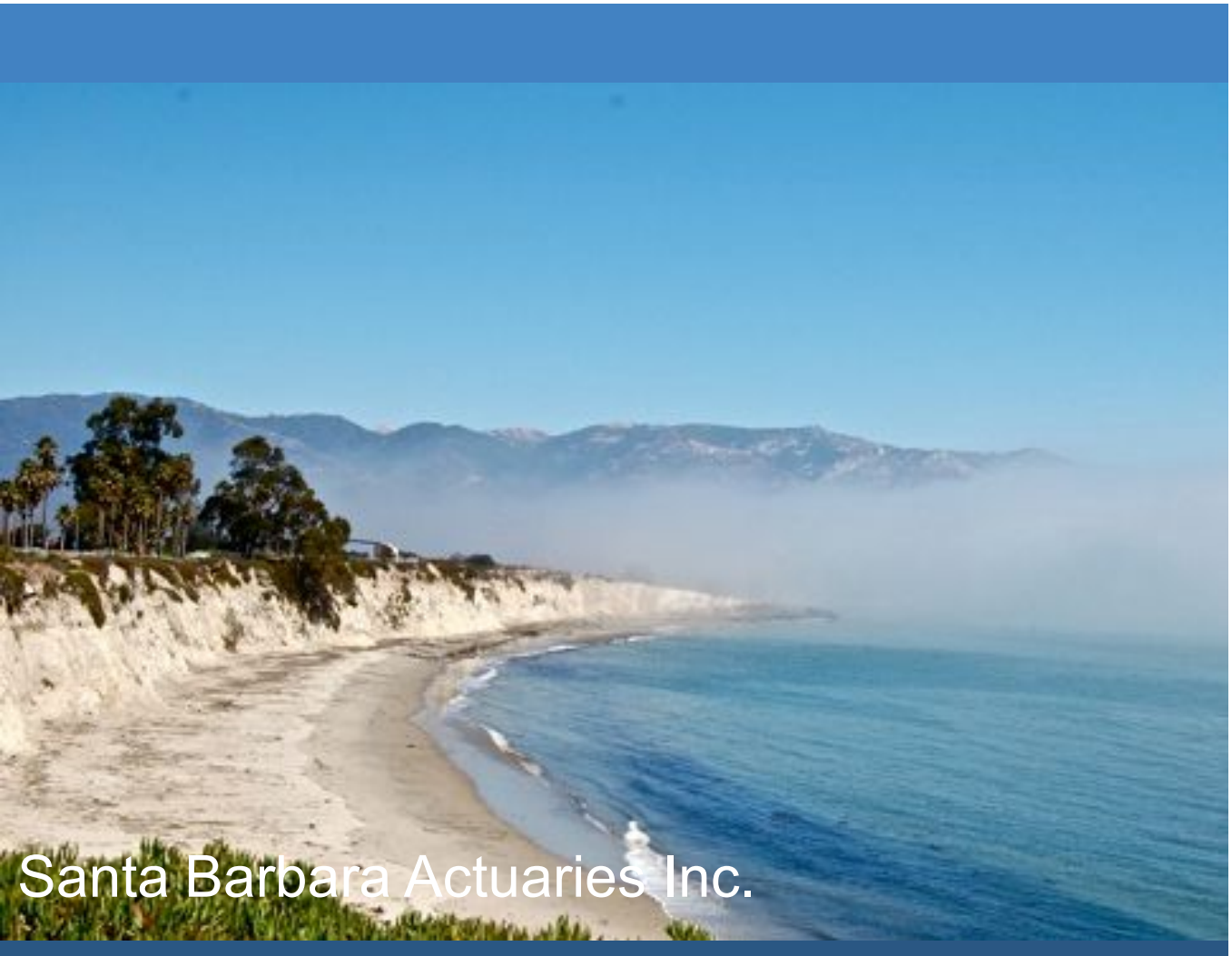
PREDICTIVE MODELING: BASICS & BEYOND

Ian Duncan FSA FIA FCIA FCA MAAA



September 2016

Agenda

1. Lecture 1: Health Risk.
2. Lecture 2: Condition and Risk Identification.
3. Lecture 3: Grouper Models.

   Break.

4. Lecture 4: Model Construction.

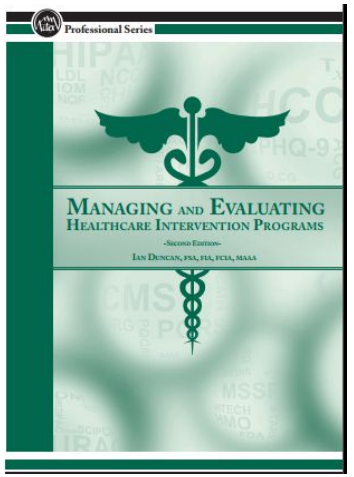
   Break

6. Applications – case studies:
   - Predictive Modeling for ACOs
   - Bundled Payments
   - End-of-life
   - Hospital Re-admissions
   - Case finding/opportunity analysis

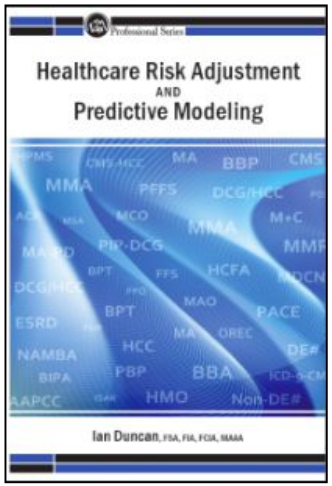
Introductions

Ian Duncan, FSA, FIA, FCIA, FCA, MAAA

- Professor of Actuarial Statistics, University of California Santa Barbara.
- President, Santa Barbara Actuaries Inc.
- Founder and former president, Solucia Consulting (now SCIO Health Analytics).
- Former board member, Massachusetts Health Insurance Connector Authority and Society of Actuaries
- Author of several books and peer-reviewed studies on healthcare management and predictive modeling.



2nd Edition 2013



What we will cover in part 1

- What is health risk?
- Typical Claims cost distributions.
- Member transitions over time.
- Traditional (Actuarial) methods of risk prediction:
  - Age/sex
  - Prior Cost
  - Prediction using Clinical Conditions.

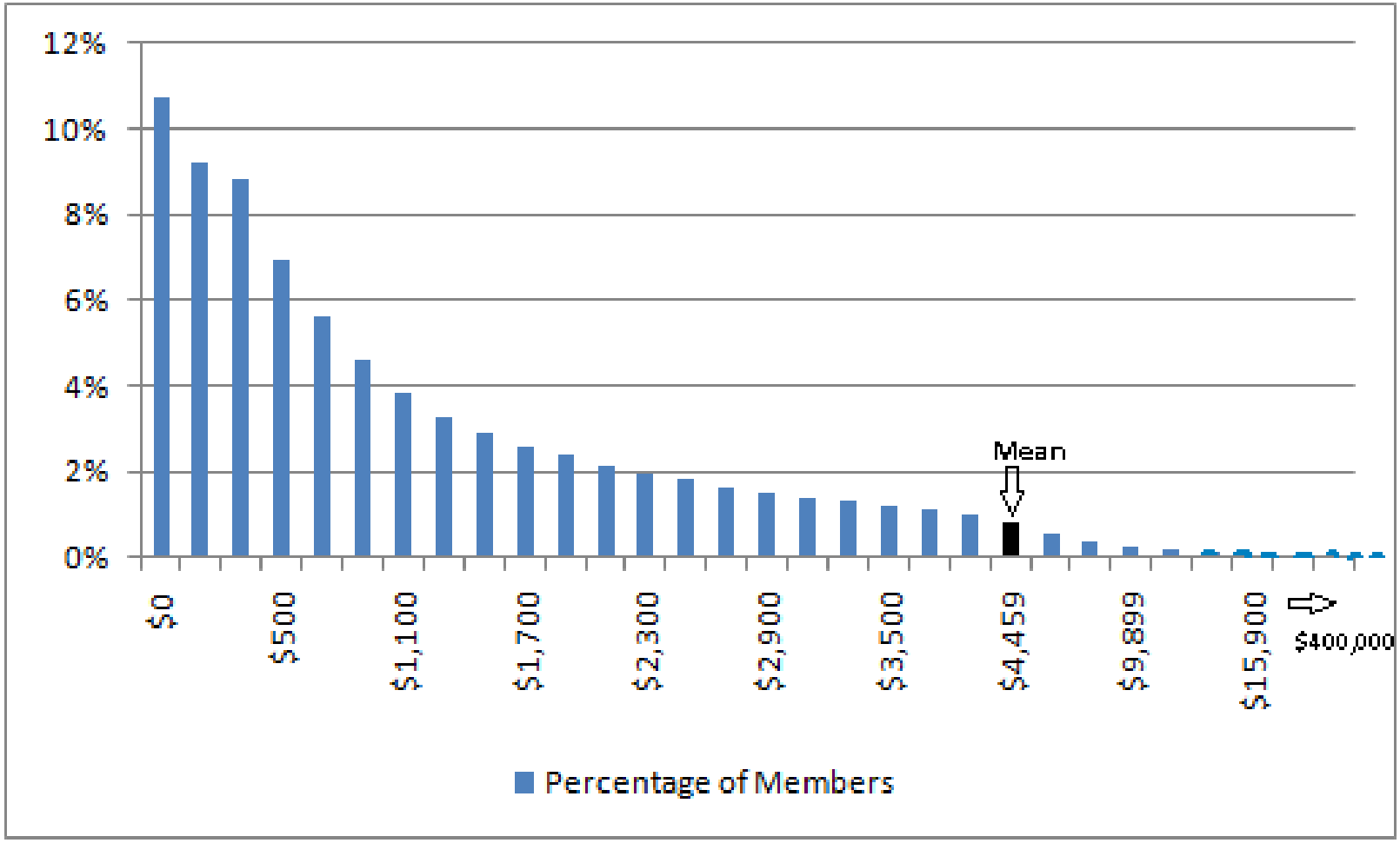
Definition of Risk

RISK = F (Loss amount; Probability)

- Another way of saying this is that Risk is a function of Frequency (of occurrences) and Severity of the occurrence.

- In healthcare, we are interested in many different states. Most frequently actuaries are interested in Financial Loss, which occurs because an event imposes a cost on an individual (or employer or other interested party). To a clinician, however, a **loss** could have a different meaning: it could be a loss of function, such as an inability to perform at a previous level or deterioration in an organ.

Typical Distribution of Population health cost



* Distribution of allowed charges within the Solucia Consulting database (multi-million member national database).

Key Concept: Member Transition

- A key concept in understanding health risk is that, while the *shape* of the distribution remains stable over time, the *composition* of the distribution changes constantly.
- Said another way: yesterday's high-cost individual isn't going to be tomorrow's.

Let's take a look at some real data.

Key Concept: Member Transition

	Baseline Year	Sequent Year		
Baseline Year Cost Group	Baseline Percentage Membership	LOW <$2,000	MODERATE $2,000-$24,999	HIGH $25,000+
LOW <$2,000	69.5%	57.4%		
			11.7%	
				0.4%
MODERATE $2,000-$24,999	28.7%	9.9%		
			17.7%	
				1.1%
HIGH $25,000+	1.8%	0.2%		
			0.9%	
				0.6%
TOTAL	100.0%	67.6%	30.3%	2.2%

Key Concept: Member Transition

	Baseline Year	Sequent Year		
Baseline Year Cost Group	Baseline Percentage Membership	LOW <$2,000	MODERATE $2,000-$24,999	HIGH $25,000+
LOW <$2,000	69.5%	57.4%		
			11.7%	
				0.4%
MODERATE $2,000-$24,999	28.7%	9.9%		
			17.7%	
				1.1%
HIGH $25,000+	1.8%	0.2%		
			0.9%	
				0.6%
TOTAL	100.0%	67.6%	30.3%	2.2%

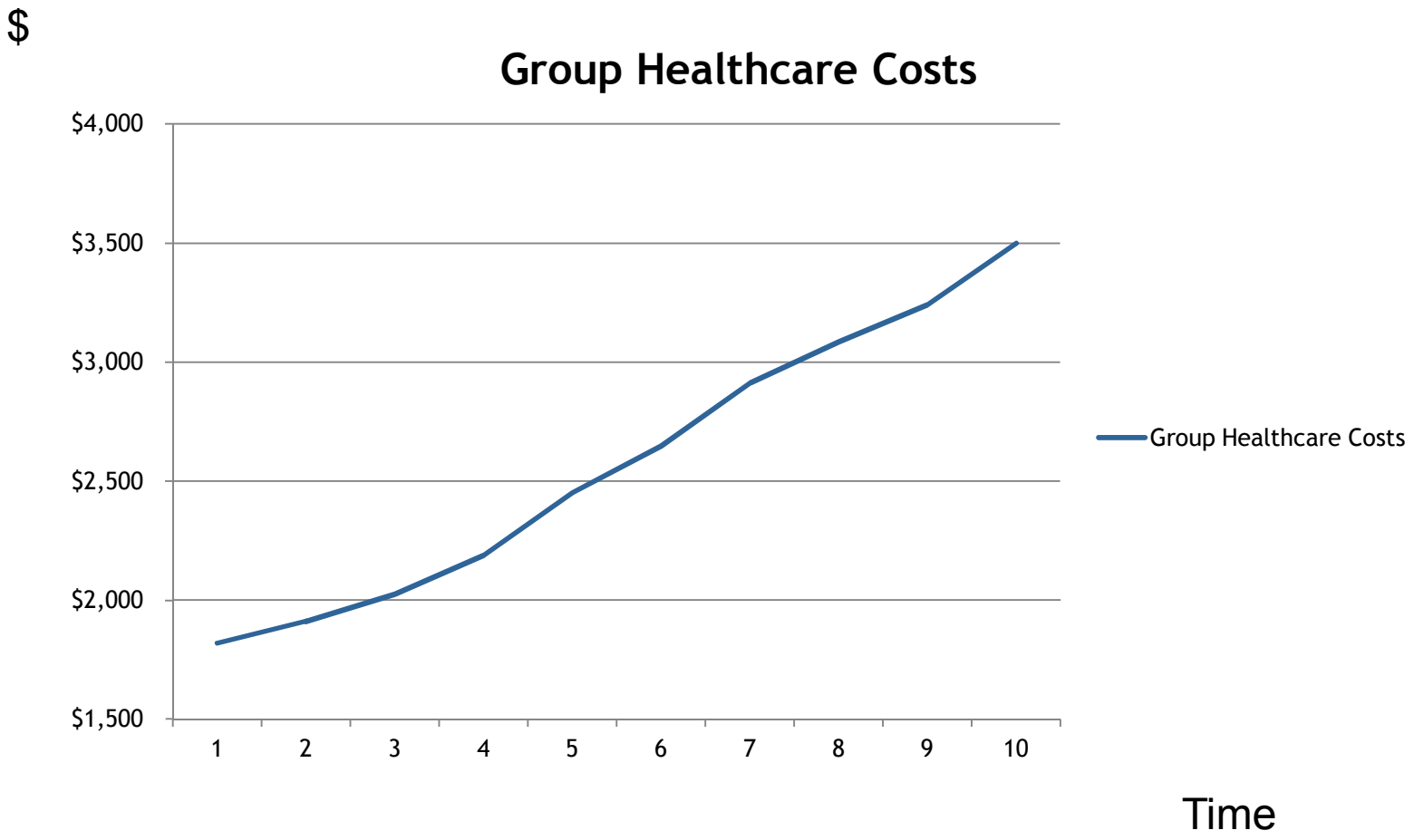
Key Concept: Member Transition

	Baseline Year	Sequent Year		
Baseline Year Cost Group	Baseline Percentage Membership	LOW <$2,000	MODERATE $2,000-$24,999	HIGH $25,000+
LOW <$2,000	69.5%	57.4%		
			11.7%	
				0.4%
MODERATE $2,000-$24,999	28.7%	9.9%		
			17.7%	
				1.1%
HIGH $25,000+	1.8%	0.2%		
			0.9%	
				0.6%
TOTAL	100.0%	67.6%	30.3%	2.2%

Key Concept: Member Transition

	Baseline Year	Sequent Year		PMPY CLAIMS
Baseline Year Cost Group	Mean Per Capita Cost	LOW <$2,000	MODERATE $2,000-$24,999	HIGH $25,000+
LOW <$2,000	$510.37	$453.24		
			$5,282.58	
				$56,166.54
MODERATE $2,000-$24,999	$6,157.06	$888.30		
			$6,803.91	
				$49,701.87
HIGH $25,000+	$55,197.12	$907.47		
			$10,435.51	
				$73,164.49
TOTAL		$518.72	$6,325.46	$57,754.19
		67.6%	30.3%	2.2%
AVERAGE	$3,090.36			$3,520.09
TREND				13.9%

Key Concept: Group costs are more stable



The Importance of Predicting Member Transition

We have seen what happens in the "real world." It is important to understand what is likely to happen to member cost, both from a financial projection perspective and from a medical management perspective.

- **Financial Management:** Health plans undertake to provide unlimited medical services to members in return for a fixed premium payment. A plan that underestimates the likely costs of health plan members can incur significant losses.
- **Healthcare Management:** An important concept for a health plan with limited resources is choose which of its members to provide medical management for; ideally, we would like to identify the 0.5% of low-cost members who will be very high cost the following year.

Traditional (actuarial) Risk Prediction

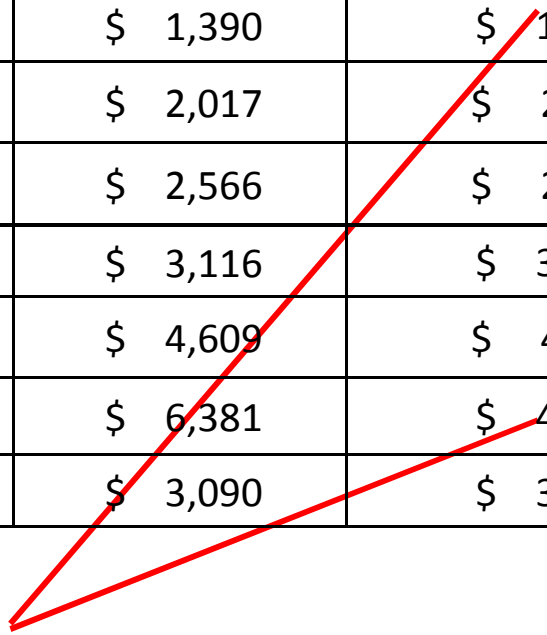
Age/Sex: although *individuals* of the same age and sex represent a range of risk profiles and costs, *groups* of individuals of the same age and sex categories follow more predictable patterns of cost. The majority of non-Government healthcare is financed by employer groups.

Relative Cost PMPY by Age/Sex			
	Male	Female	Total
< 19	$1,429	$1,351	$1,390
20-29	$1,311	$2,734	$2,017
30-39	$1,737	$3,367	$2,566
40-49	$2,547	$3,641	$3,116
50-59	$4,368	$4,842	$4,609
60-64	$6,415	$6,346	$6,381
Total	$2,754	$3,420	$3,090

Traditional (actuarial) Risk Prediction

Under the **Affordable Care Act (ACA)** traditional methods of risk management are not allowed. Rates must be uni-sex and subject to 3:1 compression. Applying these rules to our previous example, we derive:

Relative Cost PMPY by Age/Sex			
	Risk-based	**Compressed**	**Subsidy**
< 19	$ 1,390	$ 1,627	$ (237)
20-29	$ 2,017	$ 2,055	$ (38)
30-39	$ 2,566	$ 2,597	$ (31)
40-49	$ 3,116	$ 3,280	$ (164)
50-59	$ 4,609	$ 4,144	$ 465
60-64	$ 6,381	$ 4,881	$ 1,500
Total	$ 3,090	$ 3,090	$ -



3:1 Rate Compression

Typical Age/Sex Prediction (Manual Rating)

Age/Sex: Relative costs for different age/sex categories can be expressed as relative risk factors, enabling us to assess the "average" risk of an individual, or the overall (relative) risk of a population.

Relative Costs Using Age/Sex Factors					
	Male Risk Factor	**Male Number**	**Female Risk Factor**	**Female Number**	**Weighted Number**
< 19	0.46	4	0.44	12	7.12
20-29	0.42	12	0.88	19	22.00
30-39	0.56	24	1.09	21	36.33
40-49	0.82	30	1.18	24	52.92
50-59	1.41	15	1.57	12	39.99
60-64	2.08	3	2.05	1	8.29
Total	0.89	88	1.11	89	166.65
Total Membership					177
Relative age/sex factor					0.94

Based just on age/sex, this group is less risky than average.

Accuracy of Traditional Risk Prediction

Traditional (Age/Sex) risk prediction is somewhat accurate at the population level. Larger group costs are more predictable than smaller groups.

Demographic Factors as Predictors of Future Health Costs								
		Age/Sex Factors		Factor Ratio			Difference** (Predicted-Actual)	
Employer	Number of lives	Baseline	Subsequent Year	Subsequent/ Average	Predicted Cost*	Actual Cost	$	%
1	73	1.37	1.42	138%	$4,853	$23,902	**($19,049)**	-392.5%
2	478	0.74	0.76	74%	$2,590	$2,693	**($102)**	-3.9%
3	37	0.86	0.87	84%	$2,965	$1,339	$1,626	54.8%
4	371	0.95	0.97	95%	$3,331	$3,325	$6	0.2%
5	186	1.00	1.03	100%	$3,516	$3,345	$170	4.8%
6	19	1.80	1.85	180%	$6,328	$10,711	**($4,383)**	-69.3%
7	359	0.95	0.97	94%	$3,315	$3,401	**($87)**	-2.6%
8	543	0.94	0.96	93%	$3,269	$3,667	**($398)**	-12.2%
9	26	1.60	1.64	159%	$5,595	$5,181	$414	7.4%
Average		1.00	1.03	1.00	$3,520	$3,520	$ -	0.0%
Sum of absolute Differences (9 sample groups only)							$26,235	

Prior Experience adds to accuracy

To account for the variance observed in small populations, actuaries typically incorporate prior cost into the prediction, which adds to the predictive accuracy. A "credibility weighting" is used. Here is a typical formula:

$$\text{Expected Cost} = \text{Prior Year Cost} \times \text{Trend} \times Z + \text{Book of Business Cost} \times (1 - Z)$$

where $Z = \left(\frac{N}{2000}\right)^{0.5}$ and N is the number of members in the group.

Combination of Age, Sex, and Prior Cost as Predictor of Future Experience

			Cost PMPY			Difference vs. Actual	
Employer	No. of lives	Credibility Factor	Baseline	Subsequent Year Pre-dicted	Subsequent Year Actual	Difference	Difference (% of Actual)
1	73	0.19	$27,488	$9,908	$23,902	**($13,994)**	-141.2%
2	478	0.49	$1,027	$2,792	$2,693	$100	3.6%
3	37	0.14	$1,050	$2,724	$1,339	$1,385	50.9%
4	371	0.43	$2,493	$3,119	$3,325	**($205)**	-6.6%
5	186	0.30	$3,377	$3,617	$3,345	$271	7.5%
6	19	0.10	$11,352	$6,971	$10,711	**($3,739)**	-63.6%
7	359	0.42	$2,008	$2,880	$3,401	**($522)**	-18.1%
8	543	0.52	$2,598	$3,108	$3,667	**($559)**	-18.0%
9	26	0.11	$3,022	$5,350	$5,181	$169	3.2%
….	….	….	….	….	….	….	….
Average			$3,090	$3,520	$3,520	$ -	0%
Sum of absolute Differences (9 sample groups only)						$20,944	

What does Clinical information tell us about risk?

Having information about a patient's condition, particularly chronic condition(s) is potentially useful for predicting risk.

How well do standardized costs predict individual member costs?

Relative Cost PMPY by Age/Sex			
	Male	Female	Total
< 19	$1,429	$1,351	$1,390
20-29	$1,311	$2,734	$2,017
30-39	$1,737	$3,367	$2,566
40-49	$2,547	$3,641	$3,116
50-59	$4,368	$4,842	$4,609
60-64	$6,415	$6,346	$6,381
Total	$2,754	$3,420	$3,090

What does Clinical information tell us about risk?

Having information about a patient's condition, particularly chronic condition(s) is potentially useful for predicting risk.

Condition-Based Vs. Standardized Costs						
Member	Age	Sex	Condition	Actual Cost (Annual)	Standardized Cost (age/sex)	Condition-Based Cost/ Standardized Cost (%)
1	25	M	None	$863	$1,311	66%
2	55	F	None	$2,864	$4,842	59%
3	45	M	Diabetes	$5,024	$2,547	197%
4	55	F	Diabetes	$6,991	$4,842	144%
5	40	M	Diabetes and Heart conditions	$23,479	$2,547	922%
6	40	M	Heart condition	$18,185	$2,547	714%
7	40	F	Breast Cancer and other conditions	$28,904	$3,641	794%
8	60	F	Breast Cancer and other conditions	$15,935	$6,346	251%
9	50	M	Lung Cancer and other conditions	$41,709	$4,368	955%

Risk Groupers predict relative risk

Commercial Risk Groupers are available that predict relative risk based on diagnoses. Particularly helpful for small groups.

Application of Condition Based Relative Risk						
			Cost PMPY		Difference (Predicted-Actual)	
Employer	Number of lives	Relative Risk Score	Predicted	Actual	$	%
1	73	8.02	$28,214	$23,902	$4,312	15.3%
2	478	0.93	$3,260	$2,693	$568	17.4%
3	37	0.47	$1,665	$1,339	$326	19.6%
4	371	0.94	$3,300	$3,325	**($25)**	-0.8%
5	186	1.01	$3,567	$3,345	$222	6.2%
6	19	4.14	$14,560	$10,711	$3,850	26.4%
7	359	0.84	$2,970	$3,401	**($432)**	-14.5%
8	543	0.80	$2,833	$3,667	**($834)**	-29.4%
9	26	1.03	$3,631	$5,181	**($1,550)**	-42.7%
Average			$ -	0.0%	$ -	0.0%
Sum of absolute Differences (9 sample groups only)					$12,118	

Lecture 1 Recap

In this lecture we have seen:

- How Healthcare risk is a function of Amount and Probability;
- The typical, skewed distribution of healthcare costs;
- How member costs fluctuate from year-to-year, and how members move between cost levels over time;
- How standardized costs are used to predict the cost of a group, and the accuracy (or lack of accuracy) of that method;
- How predictive accuracy can be improved with information about member diagnosis.

In the next lecture we will look at other methods for estimating relative risk.

Lecture 2: Condition and Risk Identification

- At the heart of predictive modeling! How do we introduce members' diagnostic information into the identification of risk?
  - 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.

Identification - example (Diabetes)

Diabetics can be identified in different ways:

Diagnosis type	Reliability	Practicality
Physician Referral/ Medical Records/EMRs	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)

This section discusses ICD-9 diagnoses because switch to ICD-10 is recent

Inpatient Hospital Claims – ICD-9 Claims Codes

CODE	CODE TYPE	DESCRIPTION
250.xx	ICD-9 CM	Diabetes Mellitus
357.2	ICD-9 CM	Poyneuropathy in diabetes
362.0; 362.0x	ICD-9 CM	Diabetic Retinopathy
366.41	ICD-9 CM	Diabetic Cataract
648.0 - 648.04	ICD-9 CM	Diabetes Mellitus (as other current condition in mother) classifiable elsewhere but complicating pregnancy or childbirth

Codes in the 250 series ending in 1 or 3 (e.g. 250.x3) denote Type 1 diabetes; A 2 or 4 indicates type 2.

Diabetes – additional procedure codes

CODE	CODE TYPE	DESCRIPTION
G108; G109	HCPCS	Diabetic outpatient self-management training, individual or group
J815	HCPCS	Insulin injection; per 5 units
67227	CPT4	Destruction of extensive or progressive retinopathy, one or more sessions, cryotherapy, diathermy
67228	CPT4	Destruction of extensive 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**

Algorithm Development: Diabetes Example

Not all diabetics represent the same level of risk. Different diagnosis codes help identify levels of severity.

Codes for Identification of Diabetes Severity	
Diagnosis Code (ICD-9-CM)	**Code Description**
250.0	Diabetes mellitus without mention of complication
250.1	Diabetes with ketoacidosis (complication resulting from severe insulin deficiency)
250.2	Diabetes with hyperosmolarity (hyperglycemia (high blood sugar levels) and dehydration)
250.3	Diabetes with other coma
250.4	Diabetes with renal manifestations (kidney disease and kidney function impairment)
250.5	Diabetes with ophthalmic manifestations
250.6	Diabetes with neurological manifestations (nerve damage as a result of hyperglycemia)
250.7	Diabetes with peripheral circulatory disorders
250.8	Diabetes with other specified manifestations
250.9	Diabetes with unspecified complication

Algorithm Development: Diabetes Example

Relative Costs of Members with Different Diabetes Diagnoses			
Diagnosis Code ICD-9-CM	**Description**	**Average cost PMPY**	**Relative cost**
250	A diabetes diagnosis without a fourth digit (i.e., 250 only).	$13,258	105%
250.0	Diabetes mellitus without mention of complication	$10,641	85%
250.1	Diabetes with ketoacidosis (complication resulting from severe insulin deficiency)	$16,823	134%
250.2	Diabetes with hyperosmolarity (hyperglycemia (high blood sugar levels) and dehydration)	$26,225	208%
250.3	Diabetes with other coma	$19,447	154%
250.4	Diabetes with renal manifestations (kidney disease and kidney function impairment)	$24,494	195%
250.5	Diabetes with ophthalmic manifestations	$11,834	94%
250.6	Diabetes with neurological manifestations (nerve damage as a result of hyperglycemia)	$17,511	139%
250.7	Diabetes with peripheral circulatory disorders	$19,376	154%
250.8	Diabetes with other specified manifestations	$31,323	249%
250.9	Diabetes with unspecified complication	$13,495	107%
357.2	Polyneuropathy in Diabetes	$19,799	157%
362	Other retinal disorders	$13,412	107%
366.41	Diabetic Cataract	$13,755	109%
648	Diabetes mellitus of mother complicating pregnancy childbirth or the puerperium unspecified as to episode of care	$12,099	96%
TOTAL		$12,589	100%

Algorithm Development: Diabetes Example

Which leads to a possible relative risk severity structure for diabetes:

A Possible Code Grouping System for Diabetes			
Severity Level	**Diagnosis Codes Included**	**Average Cost**	**Relative Cost**
1	250; 250.0	$10,664	85%
2	250.5; 250.9; 362; 366.41; 648	$12,492	99%
3	250.1; 250.3; 250.6; 250.7; 357.2	$18,267	145%
4	250.2; 250.4	$24,621	196%
5	250.8	$31,323	249%
	TOTAL (All diabetes codes)	$12,589	100%

Risk (Condition) Groupers

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

As we have seen with the simple diabetes example, the identification of codes can be time-consuming.

Identification Algorithms and pre-defined "Grouper" models sort-through the raw material and consolidate it into manageable like categories.

Risk Groupers (examples: DCGs; HCCs; ACGs; CRGs; CDPS) do some of this work for you. We will look at them more closely in Lecture 3.

Risk (Condition) Groupers

Risk Groupers are associated with a close relative of Predictive Modeling, Risk Adjustment. Risk Adjustment has been practiced for many years in the Medicare Advantage and many managed Medicaid programs, but not Commercial Insurance. This will change in 2014 when State-based exchanges offer small group (< 50 lives) and individual purchasers the opportunity to compare and purchase insurance from a single source.

Collectivizing insurance purchasing like this gives government the opportunity to control all premiums and to re-distribute it to participating insurers, to better match the risks that they write.

We will look at the operation of risk adjustment in the Massachusetts Exchange (Connector) later in this seminar. For now, let me try to differentiate between Risk Adjustment and Predictive Modeling.

Risk Adjustment vs. Predictive Modeling

The same techniques and tools are used for both RA and PM: both use underlying risk factors and diagnoses to predict the future utilization trajectory of at-risk members (remember Lecture 1?)

Predictive Modeling identifies the stratum into which a member may fall in the future (and the member's likely cost). Health plans may use this information to design and deliver a program aimed at changing the member's behavior and reducing the predicted utilization.

Risk Adjustment comes in two flavors: Concurrent and Prospective. Both provide a *normalized* cost, to compare with the *actual* cost of a population, providing a basis on which to move premiums between plans. Prospective Risk Adjustment is similar to Predictive Modeling: based on the member's historic risk factors a projected cost is estimated for the coming year. Concurrent Risk Adjustment looks back at the end of the year just completed and calculates a normalized cost for the member, based on the member's actual risk factors.

Risk Adjustment vs. Predictive Modeling

Risk adjustment (basic concept) revenue transfer

1. Calculate individual member risk scores.
2. Calculate weighted average risk scores for the plan (assume 1.10)
3. Calculate weighted average risk scores for the State (assume this is normalized to 1.0 for simplicity).
4. Baseline premium (average for the State): $1,000.
5. Plan net claims cost: $1,250.
6. Plan adjustment (Plan Relative Risk Score – 1)(Baseline Premium) = (0.10)($1,000) = $100.
7. Premium adjustment: + $100 = $1,100. In this case, although premium is transferred, it is insufficient to offset all additional claims.

- Under the ACA, transfers will be revenue neutral.

Risk Adjustment vs. Predictive Modeling

To understand Risk Adjustment, let's use the previous example:

Applying Risk Scores						
Mem	Age	Sex	Condition(s)	Actual Cost	Standardized Cost	Risk Score
1	45	F	None	$2,864	$4,842	0.59
2	45	M	Diabetes	$5,024	$2,547	1.97
3	40	M	Diabetes + CAD	$23,479	$2,547	9.22
4	40	M	CAD	$18,185	$2,547	7.14
5	60	F	Breast Ca. + other	$15,935	$6,346	2.51
6-25			All other members			0.59
			TOTAL	$120,000	$70,000	**1.17**

This group is expected to have costs about 1.17 times the average for a group with the same age/sex distribution. Assuming that this plan collected only 100% of the required premium (standardized cost), the Exchange authority will transfer additional premium equal to 17% of the basic premium ($11,900).

Risk Adjustment vs. Predictive Modeling

The prior example points up the need for predictive modeling, even when Risk Adjustment is applied to a population. In our example, Risk Adjustment accounted for $11,900 of the difference in claims between the standardized and actual claims. The balance (about $38,000) is due to our high-risk members experiencing claims in excess of their predicted level.

If we knew which members were likely to have high claims next year we could try to moderate that utilization, reducing the excess. We could use the risk scores generated by the Risk Adjuster model to risk-rank our members; we would then try to manage Members 3 and 4 (CAD+Diabetes and CAD), and possibly other members, if we have the resources.

The use of Risk Adjuster model is not the only way to identify high risk members, as we shall see later.

All people are not equally identifiable

An important issue with any claims-based identification algorithm is that you are *imputing,* rather than *observing* a diagnosis. Thus you are always at risk of including false positives, or excluding false negatives, from the analysis.

One consequence of using a grouper model is that you are at the mercy of the modeler's definition of diagnoses, and thus cannot control for false positives or negatives. **An important draw-back of typical Grouper models is that they assign a diagnosis based on a *single instance of the diagnosis.* They are therefore more sensitive than specific.** How does this work in practice? Let's look at some data.

Prevalence of Chronic Conditions Identified Using Different Claims Algorithms				
	Number of Claiming Events in the Year			
Condition	4 or more	3 or more	2 or more	1 or more
Asthma	2.4%	2.9%	3.9%	6.1%
Cardiovascular disease	0.8%	1.2%	1.7%	2.8%
Heart Failure	0.2%	0.2%	0.3%	0.6%
Pulmonary Disease	0.2%	0.3%	0.5%	1.0%
Diabetes	3.3%	3.7%	4.1%	4.9%
All	6.3%	7.4%	9.2%	13.1%

All people are not equally identifiable (2)

A less-rigorous algorithm will identify more people with the condition (more than twice as many in the example above). But it runs the risk of sweeping in false positives. This table shows the likelihood re-qualifying with the condition in the following year (remember that these are members with Chronic conditions that (theoretically) are permanent):

Probability that a Member Identified with Chronic Condition in Year 1 will be Identified with that Condition in Year 2				
All Chronic Conditions				
No. Claiming Events in Year 2	Number of Claiming Events in Year 1			
	4 or more	3 or more	2 or more	1 or more
4 or more	59.7%	26.3%	15.7%	7.2%
3 or more	65.8%	35.9%	22.9%	10.6%
2 or more	72.0%	47.9%	34.3%	17.2%
1 or more	78.0%	62.3%	49.9%	30.9%
Do not re-qualify	22.0%	37.7%	50.1%	69.1%

All people are not equally identifiable (3)

Why is the issue of sensitivity and specificity important?

For Predictive Modeling, we may not mind much if we use a sensitive algorithm and identify members who may not have the condition; if members are followed up by a program, their true condition can be determined.

For Risk Adjustment, however, specificity matters. Medicare's Risk Adjustment process has given rise to an industry that finds additional diagnoses (that drive the risk score). Increases in the Medicare Advantage plan's average risk score will increase its revenue. Interestingly, unlike Medicare Advantage plans, the proposed State Exchange Risk Adjustment is a zero-sum game: if another plan increases its average risk score, your plan will lose revenue. With the typical risk adjuster assigning a diagnosis based on a single instance of a diagnosis, it may be possible for aggressive plans to "game" the risk adjustment process.

Algorithm Development: Diabetes Example

Example of an identification algorithm:

Example of a Definitional Algorithm			
Disease	**Type**	**Frequency**	**Codes**
Diabetes Mellitus	Hospital Admission or ER visit with diagnosis of diabetes in any position	At least one event in a 12-month period	ICD-9 codes 250, 357.2, 362.0, 366.41, 648.0
	Professional visits with a primary or secondary diagnosis of diabetes	At least 2 visits in a twelve month period	CPT Codes in range of 99200-99499 series E & M codes or 92 series for eye visits
	Outpatient Drugs: dispensed insulin, hypoglycemic, or anti-hyperglycemic prescription drug	Three or more prescriptions in a twelve month period	Diabetes drugs (see HEDIS or similar list of drug codes).
EXCLUDE gestational diabetes.	Any (as above)	As above	648.8x

Sources of Algorithms

- NCQA – HEDIS.
- DMAA (Now CCA; Chronic definitions).
- Grouper Models.

Additionally, there has been an explosion of rules-based quality metric reporting in recent years. Just a few examples:

- PQRS (208 measures in 22 categories);
- STAR Measures (36 Medicare Advantage; 15 Medicare Part D);
- ACO quality reporting (33 measures);
- HEDIS Measures (75 measures in 8 domains).

All require risk-adjustment for their application.

Lecture 2 Recap

In this lecture we have:

- Introduced the concept of the Risk Grouper;
- Seen how diagnoses inform risk grouper and algorithm construction;
- Distinguished between Predictive Modeling and Risk Adjustment;
- Introduced the concept of Sensitivity and Specificity in algorithm construction, and seen some implications of each.

In the next lecture we will discuss grouper and algorithm construction in more depth.

Lecture 3: Grouper/Algorithm Construction

In Lecture 3, we will look at Commercial Groupers in more detail, and introduce the Society of Actuaries comparative studies of risk adjusters. We will also return to the topic of Prospective vs. Concurrent models, introduced in lecture 2.

We will also discuss the construction of a grouper algorithm for those analysts who do not want to use a commercially-available model.

We will look at Episode Groupers (a different approach to relative risk), rules-based models and models that incorporate other data sources.

Grouper Construction

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

The Society of Actuaries has published three studies that test this correlation. All are available at www.soa.org . 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. A fourth study is in preparation and should be out any day now.

Dunn DL, Rosenblatt A, Taira DA, et al. "A comparative Analysis of Methods of Health Risk Assessment." *Society of Actuaries (SOA Monograph M-HB96-1).* Oct 1996:1-88.

Cumming RB, Cameron BA, Derrick B, et al. "A Comparative Analysis of Claims-Based Methods of Health Risk Assessment for Commercial Populations". *Research study sponsored by Society of Actuaries.* 2002.

Winkelman R, Mehmud S. "A Comparative Analysis of Claims-Based Tools for Health Risk Assessment". *Society of Actuaries.* 2007 Apr:1-63.
(available at: www.soa.org/files/pdf/ risk-assessmentc.pdf).

Grouper Construction

Society of Actuaries Studies:

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.
- As we have discussed previously, "sensitive" models are open to inclusion of more false-positives.
- A well-managed patient who fails to develop a more serious form of a condition will have a lower risk score.

Commercial Groupers: SOA studies

The Society of Actuaries studies show:

1. Risk grouper modeling tools use *different algorithms* to group the source data. For example, the Symmetry models are built on episodes of care, DRGs are built on hospital episodes, while other models are built on diagnoses.
2. Similar performance among all leading risk groupers.
3. Accuracy of prediction has increased since the publication of the original study. In part, this is due to more accurate coding and the inclusion of more claims codes.
4. Risk groupers use *relatively limited data* sources (e.g. DCG and Rx Groups use ICD-9 and NDC codes but not lab results or HRA information).
5. Accuracy of **retrospective (concurrent)** models is now in the 30%-40% R^2 range. **Prospective** model accuracy is in the range of 20% to 25%.

A note about Prospective and Concurrent Models

Both have their place. Neither is perfect.

1. Concurrent models are also called Retrospective.

   The concurrent model is used to reproduce **actual historical costs.** This type of model is used for assessing relative resource use and for determining compensation to providers for services rendered because it normalizes costs across different members. Normatively, the concurrent model provides an assessment of what costs *should* have been for members, given the health conditions with which they presented in the past year. It is also used in program evaluation, which is performed once all known conditions may be identified.

2. The Prospective model predicts what costs w*ill* be for a group of members in the future. The Prospective model is predicting the unknown, because the period over which the prediction is made lies in the future. The Concurrent model, by contrast, provides an estimate of normalized costs for services that have already occurred. For prospective prediction, members with no claims receive a relative risk score component based on age/sex alone.

A note about Prospective and Concurrent Models

Concurrent models have the advantage that they represent all the known information about the member in the completed year. However, when they are used to compensate providers (for example) for managing a group of members, there is a risk to the provider that if the provider does a good job and prevents the exacerbation of the member's condition, the member risk score (and therefore the provider's compensation) will be lower than it would be if the provider does not prevent the exacerbation.

Prospective models are often used to allocate revenue to different managed care plans. The drawback to this approach is that members' prospective risk scores are based on historical data, and do not take account of developing (incident) conditions that emerge during the year.

SOA Risk Adjustment Studies - 2007

A Comparative Analysis of Claims-Based Tools for

Health Risk Assessment

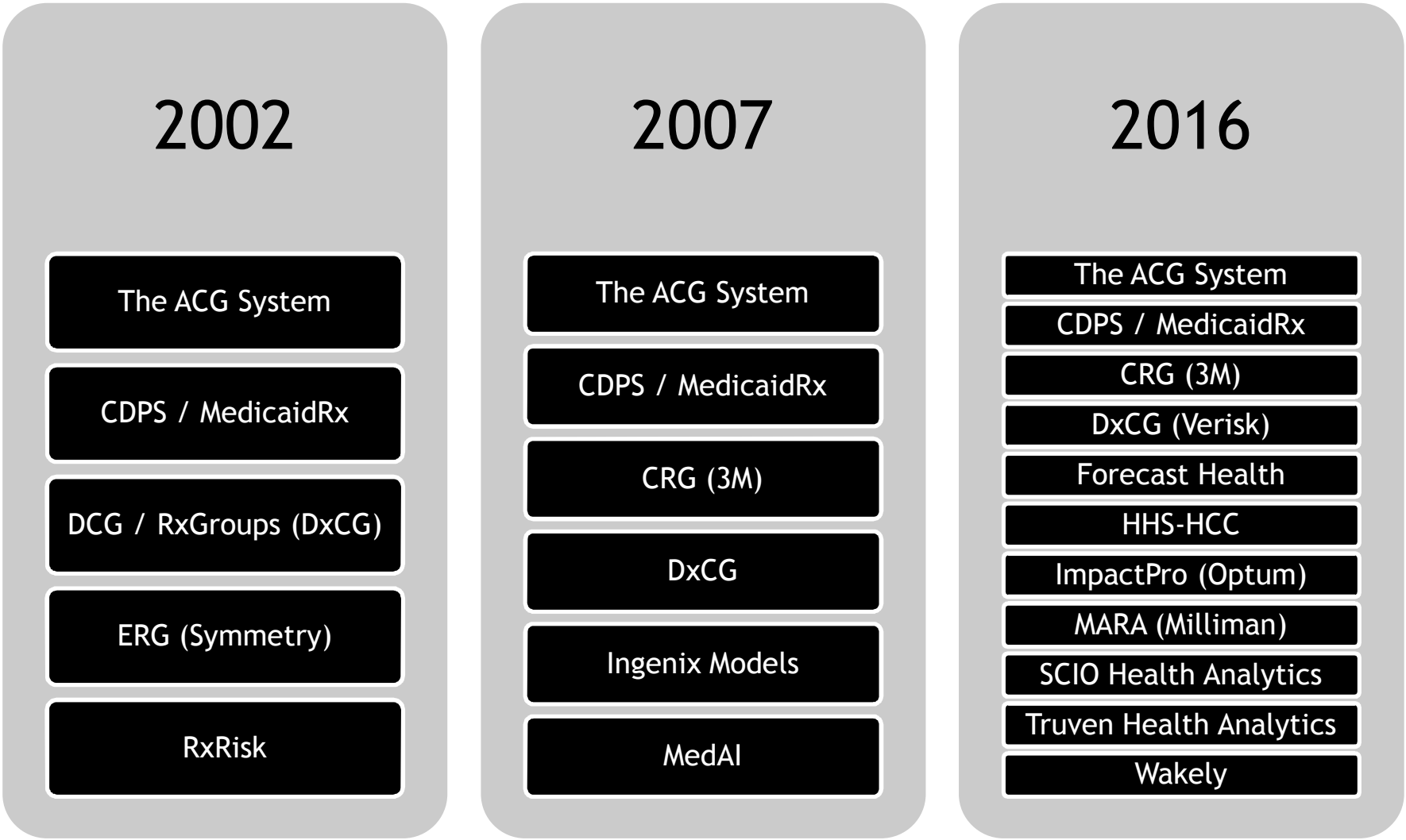
by: Ross Winkelman, FSA, Principal & Consulting Actuary, Milliman ▪ Syed Mehmud, Actuarial Assistant, Milliman
Peer reviewed by: Leigh Wachenheim, FSA, Principal & Consulting Actuary, Milliman

April 20, 2007

Available at www.soa.org

Actuaries
Risk is Opportunity.℠

Evolution of Included Models



Evolution of Predictive Accuracy

Individual Predictions, R^2

Year	Age/Sex	Retrospective Condition	Prospective Condition
1996	0.02-0.05	0.20-0.60	0.08 – 0.10
2002		0.21 – 0.47[2]	0.07 – 0.15[1]
2007		0.25 – 0.50[1]	0.12 – 0.17[1]

Notes:

1. Offered, non-lagged, non-truncated.
2. Re-calibrated, non-lagged, non-truncated

Grouper Algorithms

As an alternative to commercially-available risk groupers, analysts can develop their own models using common data mining techniques. Each method has its pros and cons:

There is a considerable amount of work involved in building algorithms from scratch, particularly when this has to be done for the entire spectrum of diseases. Adding drug or laboratory sources to the available data increases the complexity of development.

While the *development* of a model may be within the scope and resources of the analyst who is performing research, use of models for production purposes (for risk adjustment of payments to a health plan or provider groups for example) requires that a model be maintained to accommodate new codes. New medical codes are not published frequently, but new drug codes are released monthly, so a model that relies on drug codes will soon be out of date unless updated regularly.

Commercially-available clinical grouper models are used extensively for risk adjustment when a consistent model, accessible to many users, is required. Providers and plans, whose financial stability relies on payments from a payer, often require that payments be made according to a model that is available for review and validation.

Grouper Algorithms

An analyst that builds his own algorithm for risk prediction has control over several factors that are not controllable with commercial models:

1. Which codes, out of the large number of available codes to recognize. The numbers of codes and their redundancy (the same code will often be repeated numerous times in a member record) makes it essential to develop an aggregation or summarization scheme.

2. The level at which to recognize the condition. How many different levels of severity should be recognized? . The analyst will also need to determine how to handle different levels of severity of the same diagnosis: should each be included, or should the higher-severity code "trump" the lower-severity code?

3. The impact of co-morbidities. Some conditions are often found together (for example heart disease with diabetes). The analyst will need to decide whether to maintain separate conditions and then combine where appropriate, or to create combinations of conditions.

Grouper Algorithms

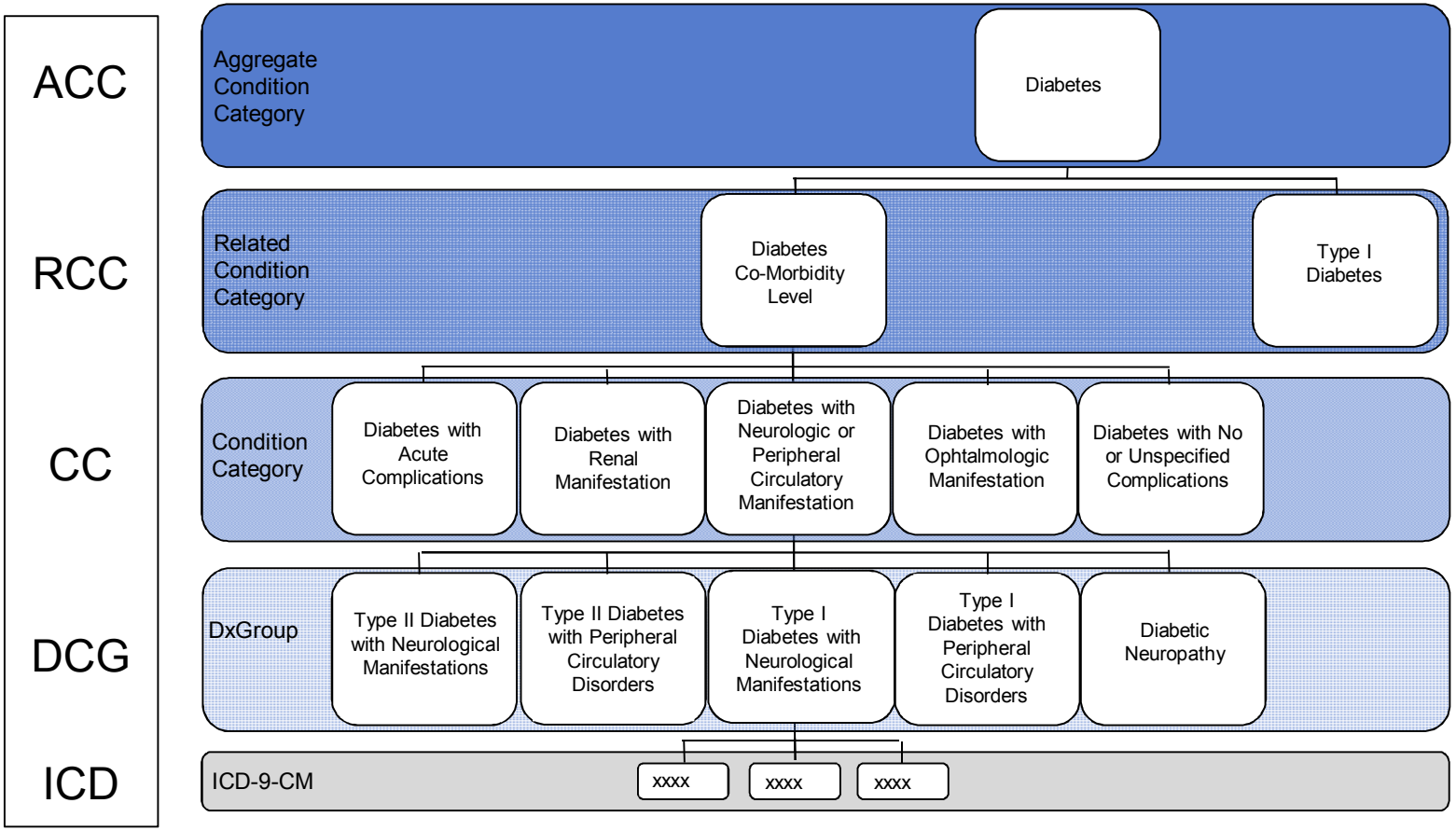
An analyst that builds his own algorithm for risk prediction has control over several factors that are not controllable with commercial models:

4. The degree of certainty with which the diagnosis has been identified (confirmatory information). (We have already seen the implication of this issue when looking at the relationship between sensitivity/specificity and the number of discrete instances of a diagnosis in the member record.) The accuracy of a diagnosis may differ based on who codes the diagnosis, for what purpose and how frequently a diagnosis code appears in the member record. The more frequently a diagnosis code appears, the more reliable the interpretation of the diagnosis. Similarly, the source of the code (hospital, physician, laboratory) will also affect the reliability of the diagnostic interpretation.

5. Data may come from different sources with a range of reliability and acquisition cost. A diagnosis in a medical record, assigned by a physician, will generally be highly reliable. Other types of data are not always available or as reliable.

Example of Grouper Construction

Grouper models are constructed in a similar fashion to that illustrated above. Below we show the hierarchical structure of the DxCG model for Diabetes:



Example of Grouper Construction

Grouper models are constructed in a similar fashion to that illustrated above. Below we show how the risk score is developed for a patient with diagnoses of Diabetes, HTN, CHF and Drug Dependence, illustrating the hierarchical and additive structure of the DxCG model:

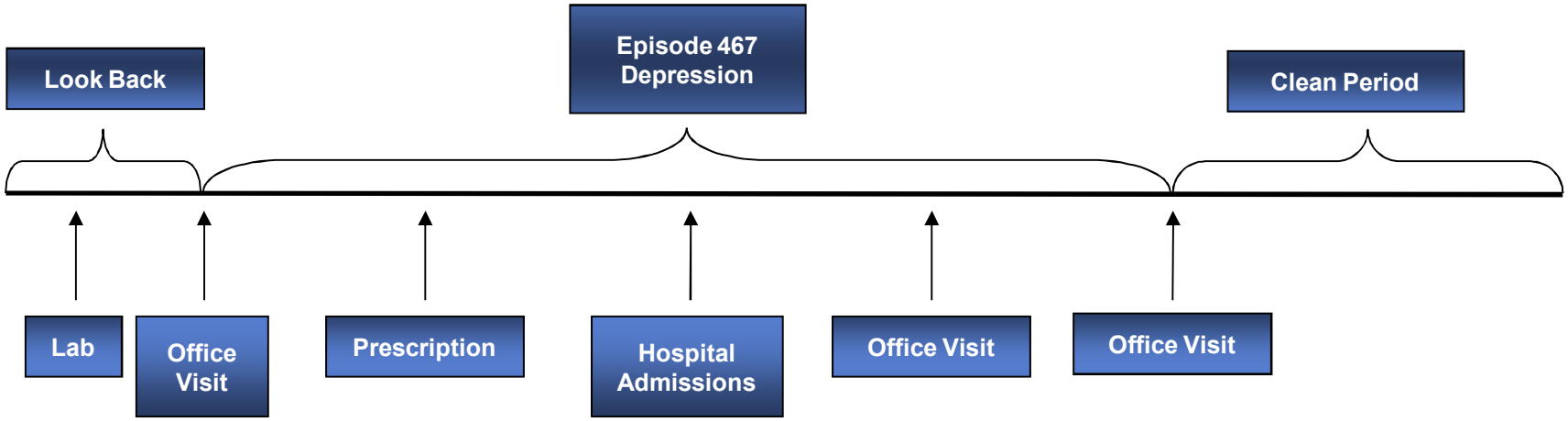
Example of Construction of a Relative Risk Score		
Condition Category	**Risk Score Contribution**	**Notes**
Diabetes with No or Unspecified Complications	0.0	Trumped by Diabetes with Renal Manifestation
Diabetes with Renal Manifestation	2.1	
Hypertension	0.0	Trumped by CHF
Congestive Heart Failure (CHF)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

Different Approach to Model Construction:

- Episode Groupers represent a different approach to model construction. Instead of a model being constructed around diagnoses, it is constructed around episodes of care. These episodes are often acute (surgery, for example) but may also be chronic or preventive.
- In the case of an acute episode, the episode grouper can be thought of as a normative tool, in that it specifies all the different services that make up the episode. For example, it may start with an office visit, followed by diagnostic testing, then inpatient admission for surgery, rehabilitation, and additional follow up visits for check-ups.
- The episode approach clearly had the advantage that it specifies the services typically ordered in the case of a specific treatment, thus enabling us to compare the utilization patterns of physicians treating the same illness. The utilization of services associated with different episodes can also be translated into a relative risk score.
- While Episode Groupers handle acute episodes well, for chronic illnesses (where there is no "episode," or the episode could be the year of treatment) the approach is very similar to condition-based models.

Episode-based Groupers

An example of an episode Group: the Symmetry Grouper.



Episode-based Groupers

Application of the Symmetry Grouper. Risk Scores are developed similarly to DxCG.

Construction of Relative Risk Scores Using ETGs						
Example: Male Aged 58						
ETG	Severity Level	Description	ERG	Description	Retrospective Risk Weights	Prospective Risk Weights
163000	2	Diabetes	2.022	Diabetes w/significant complication/co-morbidity I	0.9874	1.2810
386800	1	Congestive Heart Failure	8.043	Ischemic heart disease, heart failure, cardiomyopathy III	2.2870	2.0065
238800	3	Mood Disorder, Depression	4.033	Mood disorder, depression w/ significant cc/cb	0.8200	0.7913
473800	3	Ulcer	11.022	Other moderate cost gastroenterology II	2.3972	0.6474
666700	1	Acne	17.011	Lower cost dermatology I	0.1409	0.1023
666700	1	Acne	17.011	Lower cost dermatology I		
Demographic risk: Male 55-64						0.7331
					6.6325	5.5616

One more very useful grouper…

Drug groupers group 100,000s NDC codes into manageable therapeutic classes

Example of Therapeutic Classes Within the GPI Structure					
Group	Class	Sub Class	Group	Class	Sub Class
GROUPS 1- 16 ANTI-INFECTIVE AGENTS					
01	00	00	*PENICILLINS*		
01	10	00		Penicillin G	
01	30	00		PENICILLINASE -RESISTANT PENICILLINS	
01	50	00		AMINO PENICILLINS/BROAD SPECTRUM PENICILLINS	
01	20	00		Ampicillins	
01	40	00		EXTENDED SPECTRUM PENICILLINS	
01	99	00		*Penicillin Combinations**	
01	99	50			*Penicillin-Arninoglycoside Combinations***
01	99	40			*Penicillin-NSAIA Combinations***
02	00	00	*CEPHALOSPORINS*		
02	10	00		*Cephalosporins -1st Generation**	
02	20	00		*Cephalosporins -2nd Generation**	
02	30	00		*Cephalosporins -3rd Generation**	
02	40	00		*Cephalosporins -4th Generation**	
02	99	00		*Cephalosporin Combinations**	
03	00	00	*MACROLIDE ANTIBIOTICS*		
03	10	00		*Erythromycins**	
03	10	99			*Erythromycin Combinations***
03	20	00		*Troleandomycin**	
03	30	00		*Lincomycins**	
03	40	00		*Azithromycin**	
03	50	00		*Clarithromycin**	
03	52	00		*Dirithromycin**	
Etc.	Etc.	Etc.			

Rules-based vs. Statistical Models

What about rules-based models?

1. First, all models ultimately have to be converted to rules to apply 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, rather than statistics, determine the risk factors.

On the predictive ability of experts vs. the machine, see next slide!

Providers are not good at predicting re-admission risk

- Researchers from the VA System assessed the predictions made by
  - Physicians
  - Case managers
  - Nurses
- *"...none of the AUC values were statistically different from chance"*

Inability of Providers to Predict Unplanned Readmissions

Nazima Allaudeen, MD[1,2], Jeffrey L. Schnipper, MD, MPH[3], E. John Orav, PhD[4], Robert M. Wachter, MD[2], and Arpana R. Vidyarthi, MD[2]

[1]Department of Medicine, VA-Palo Alto Healthcare System, Palo Alto, CA, USA; [2]Division of Hospital Medicine, Department of Medicine, University of California, San Francisco, CA, USA; [3]BWH Academic Hospitalist Service and Division of General Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA; [4]Department of Biostatistics, Harvard School of Public Health, Boston, MA, USA.

BACKGROUND: Readmissions cause significant distress to patients and considerable financial costs. Identifying hospitalized patients at high risk for readmission is an important strategy in reducing readmissions. We aimed to evaluate how well physicians, case managers, and nurses can predict whether their older patients will be readmitted and to compare their predictions to a standardized risk tool (Probability of Repeat Admission, or P_{ra}).

METHODS: Patients aged ≥65 discharged from the general medical service at University of California, San Francisco Medical Center, a 550-bed tertiary care academic medical center, were eligible for enrollment over a 5-week period. At the time of discharge, the inpatient team members caring for each patient estimated the chance of unscheduled readmission within 30 days and predicted the reason for potential readmission. We also calculated the P_{ra} for each patient. We identified readmissions through electronic medical record (EMR) review and phone calls with patients/caregivers. Discrimination was determined by creating ROC curves for each provider group and the P_{ra}.

RESULTS: One hundred sixty-four patients were eligible for enrollment. Of these patients, five died during the 30-day period post-discharge. Of the remaining 159 patients, 52 patients (32.7%) were readmitted. Mean readmission predictions for the physician providers were closest to the actual readmission rate, while case managers, nurses, and the P_{ra} all overestimated readmissions. The ability to discriminate between readmissions and non-readmissions was poor for all provider groups and the P_{ra} (AUC from 0.50 for case managers to 0.59 for interns, 0.56 for P_{ra}). None of the provider groups predicted the reason for readmission with accuracy.

CONCLUSIONS: This study found (1) overall readmission rates were higher than previously reported, possibly because we employed a more thorough follow-up methodology, and (2) neither providers nor a published algorithm were able to accurately predict which patients were at highest risk of readmission. Amid increasing pressure to reduce readmission rates, hospitals do not have accurate predictive tools to guide their efforts.

Received August 27, 2010
Revised January 27, 2011
Accepted January 28, 2011
Published online March 12, 2011

KEY WORDS: readmission; unplanned; prediction.
J Gen Intern Med (26)7:771-6
DOI: 10.1007/s11606-011-1663-3

BACKGROUND

Against the background of rising concerns about both the cost and quality of American medical care, hospital readmissions have come under increasing scrutiny from both outside and within the government.[1–3] Hospital readmissions may be a marker for poor quality care, are dissatisfying for patients and families, and increase health care costs. Medicare estimates that $15 billion is spent on the 17.6% of patients who are readmitted within 30 days[4].

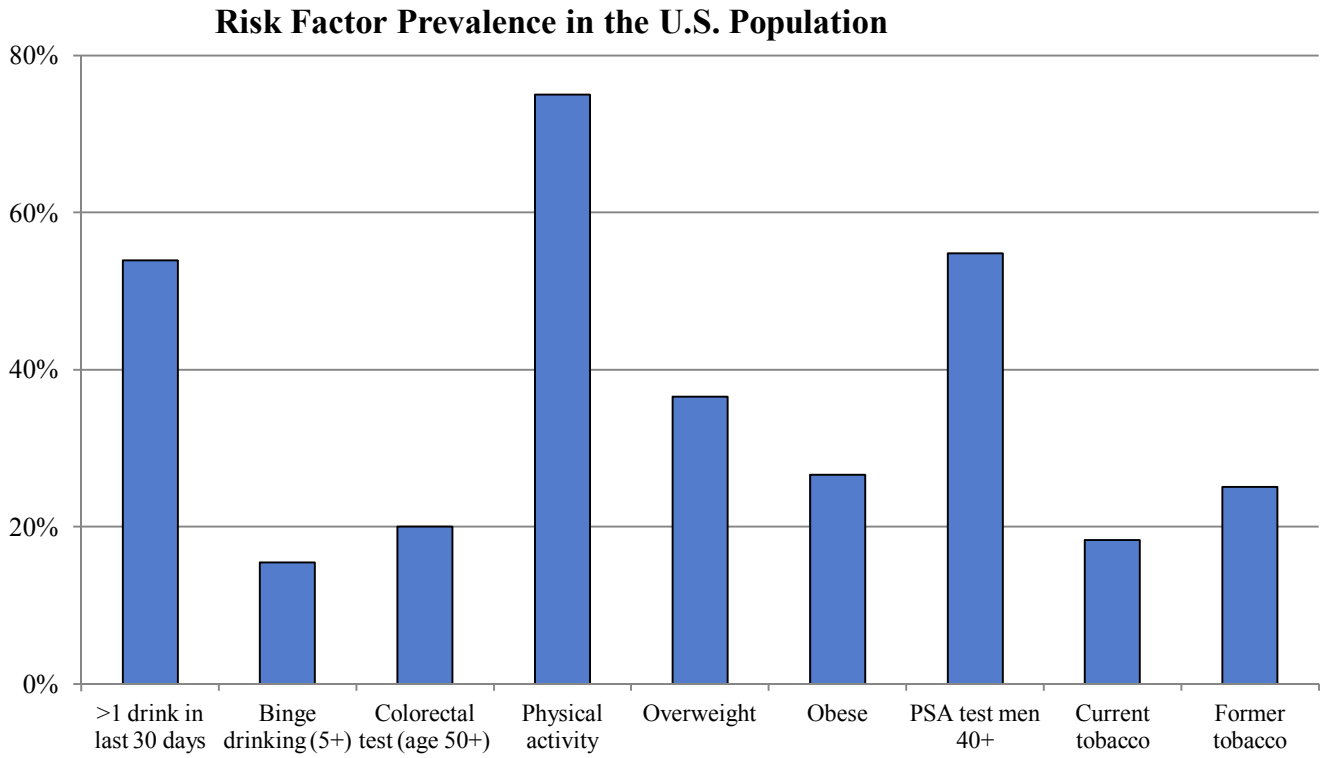
Although it would be ideal to develop interventions that improve the hospital-to-home transition for all patients, given limited resources, some have argued for targeting intense efforts—such as comprehensive discharge planning, post-discharge phone calls or home visits, and early clinic visits—towards high risk patients. However, such strategies require that we have accurate methods to identify patients at highest risk.

Anecdotal evidence suggests that inpatient providers (physicians, nurses, discharge planners) currently make informal predictions of readmission that affect discharge planning. Such predictions are not new; providers have tried to predict other outcomes, such as mortality and length of stay, in several settings (e.g., intensive care unit, emergency department), with varying success.[5–10] However, the accuracy of informal predictions of hospital readmission is unknown. Several algorithms have also been developed in recent years to predict hospital readmissions, but their use has been limited, because they require information not typically gathered during clinical care, their models are complex and difficult to use, and/or because they are not accurate. A few studies have compared providers with algorithm-based tools to predict readmission and mortality in other settings,[9] but it remains unknown how well providers' predictions of readmission for general medicine patients compare with published algorithms or how the predictions of multiple disciplines compare with one another.

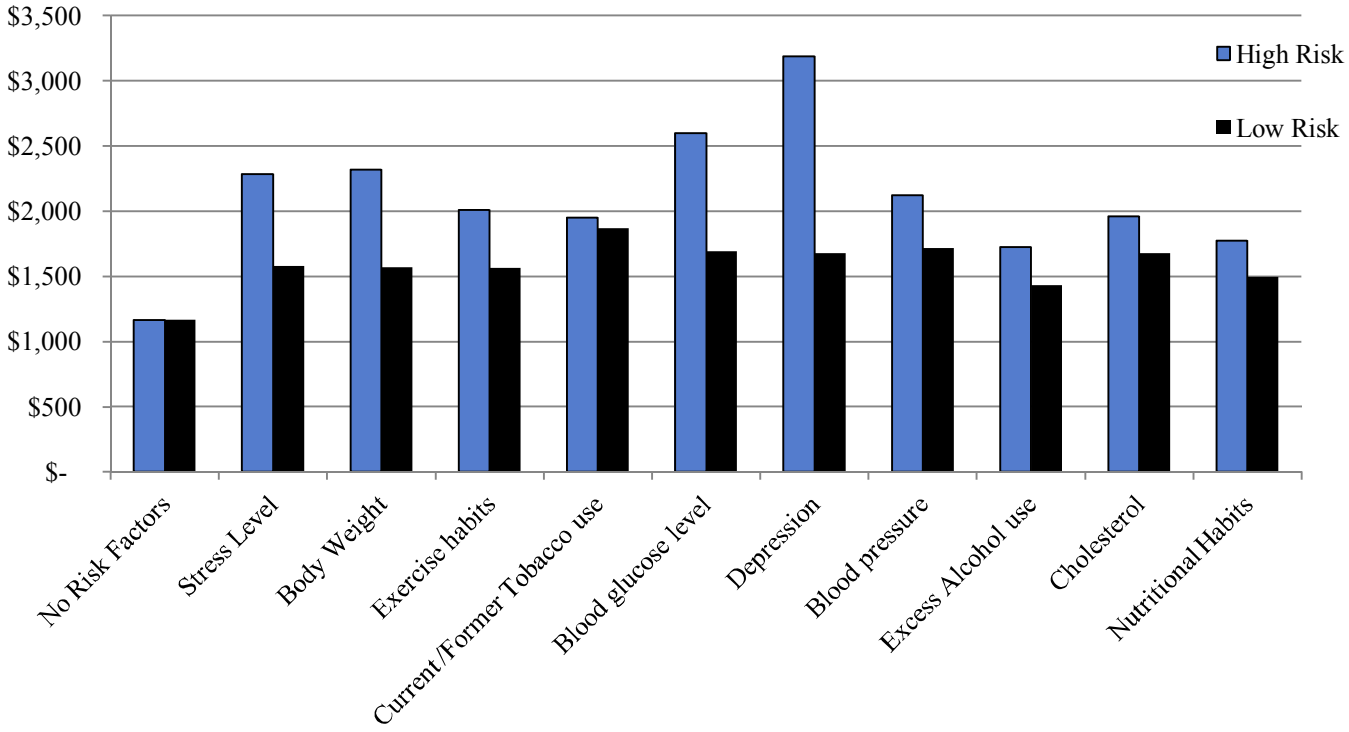
To reach the ultimate goal of preventing readmissions, identifying the highest risk patients is the first of a multistep process. Providers would next need to speculate on the reason for readmission before then targeting an effective

Allaudeen N, Schnipper JL, Orav EJ, Wachter RM, Vidyarthi AR. Inability of providers to predict unplanned readmissions. J Gen Intern Med. 2011;26(7):771-6

Don’t overlook non-condition-based Risk

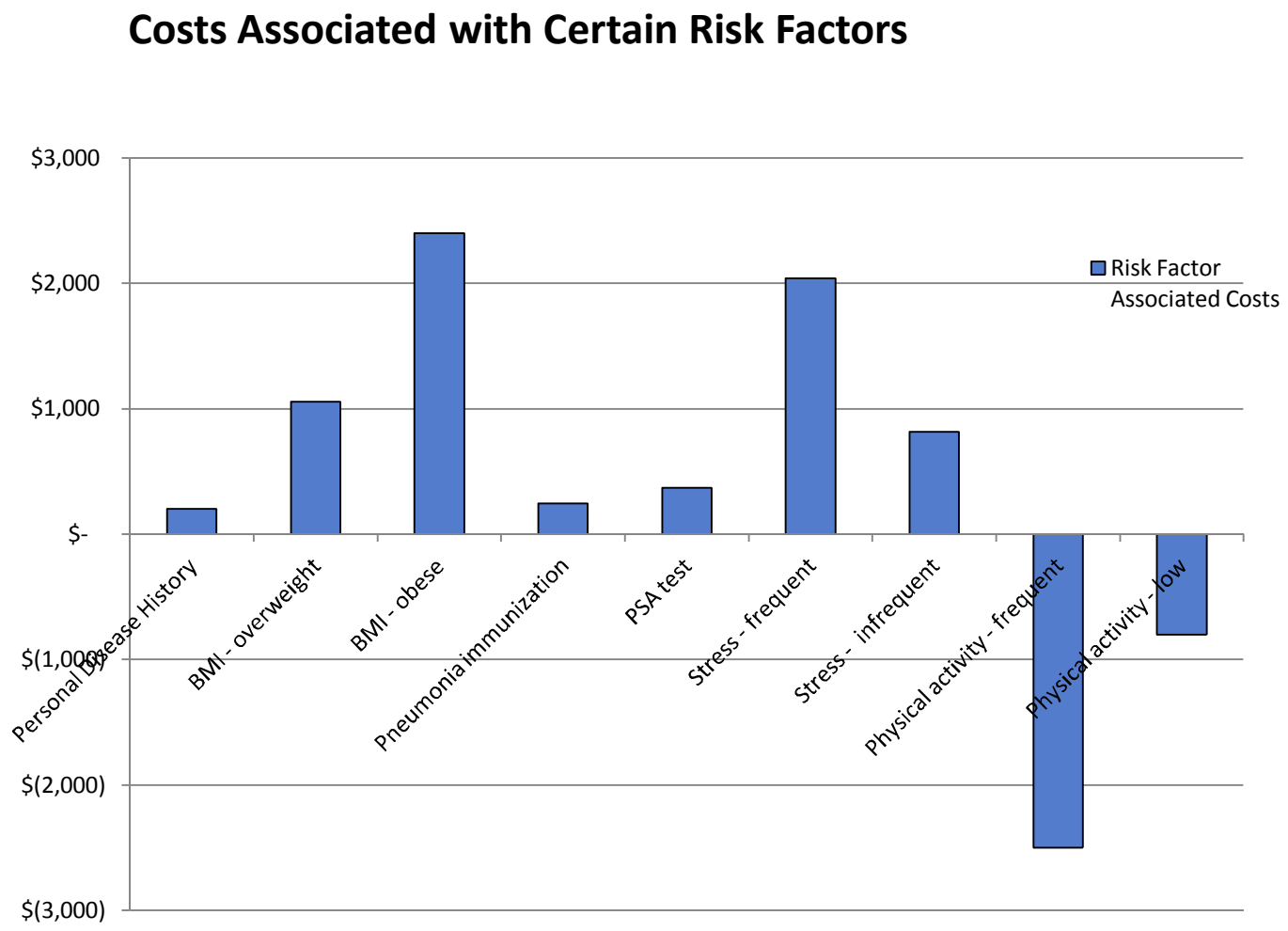


Non-condition-based Risk



Relationship Between Modifiable Health Risk Factors and Annual Health Insurance Claims

Non-condition-based Risk



Lecture 3: Re-cap

In lecture 3 we looked at Commercial Groupers in more detail, and suggested the Society of Actuaries comparative studies of risk adjusters as an unbiased source of comparative performance of different models.

We looked at Episode Groupers, rules-based models and models that incorporate other data sources.

In Lecture 4 we will look at actual model construction.

Lecture 4: Model Construction

In this lecture we will look (briefly) at model construction. This is a very significant topic, worthy of its own course. We will just skim the surface.

We will look at some common tools that are used for model building and then follow a practical example.

Developing your own model

As we have seen, there are pros and cons to using a commercially-available model. There is no reason not to develop your won predictive or risk adjustment model. Chapters 7-12 of my book address different statistical models that are encountered in this work. Examples of statistical models used frequently:

- Linear Regression. Advantage: everyone understands this.
- Generalized linear model and Logistic regression: more sophisticated models often used for healthcare data.
- Tree models: more difficult to apply operationally than regression models.
- Neural networks: black box.

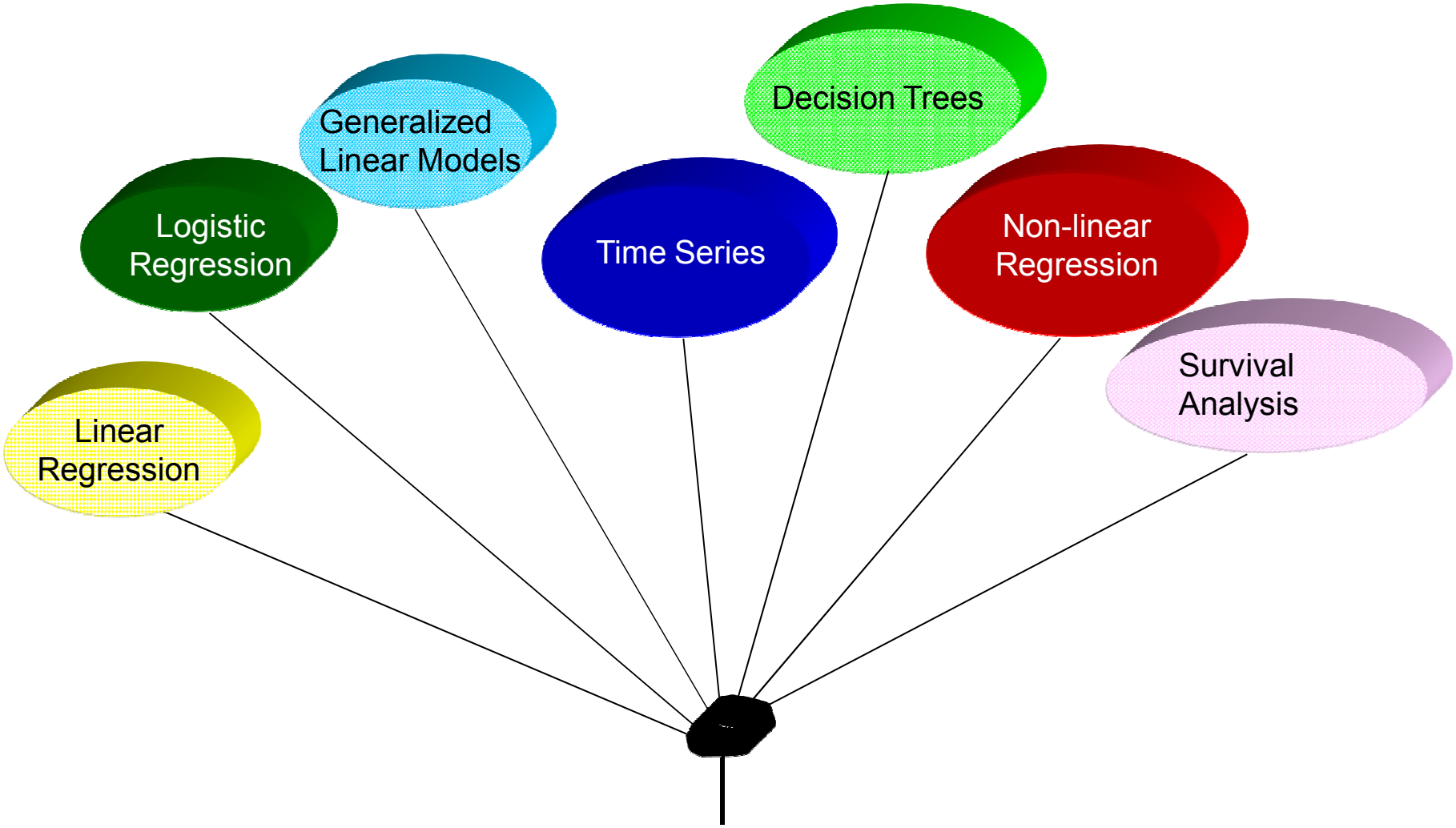
There are a couple of others less frequently encountered.

Developing your own model

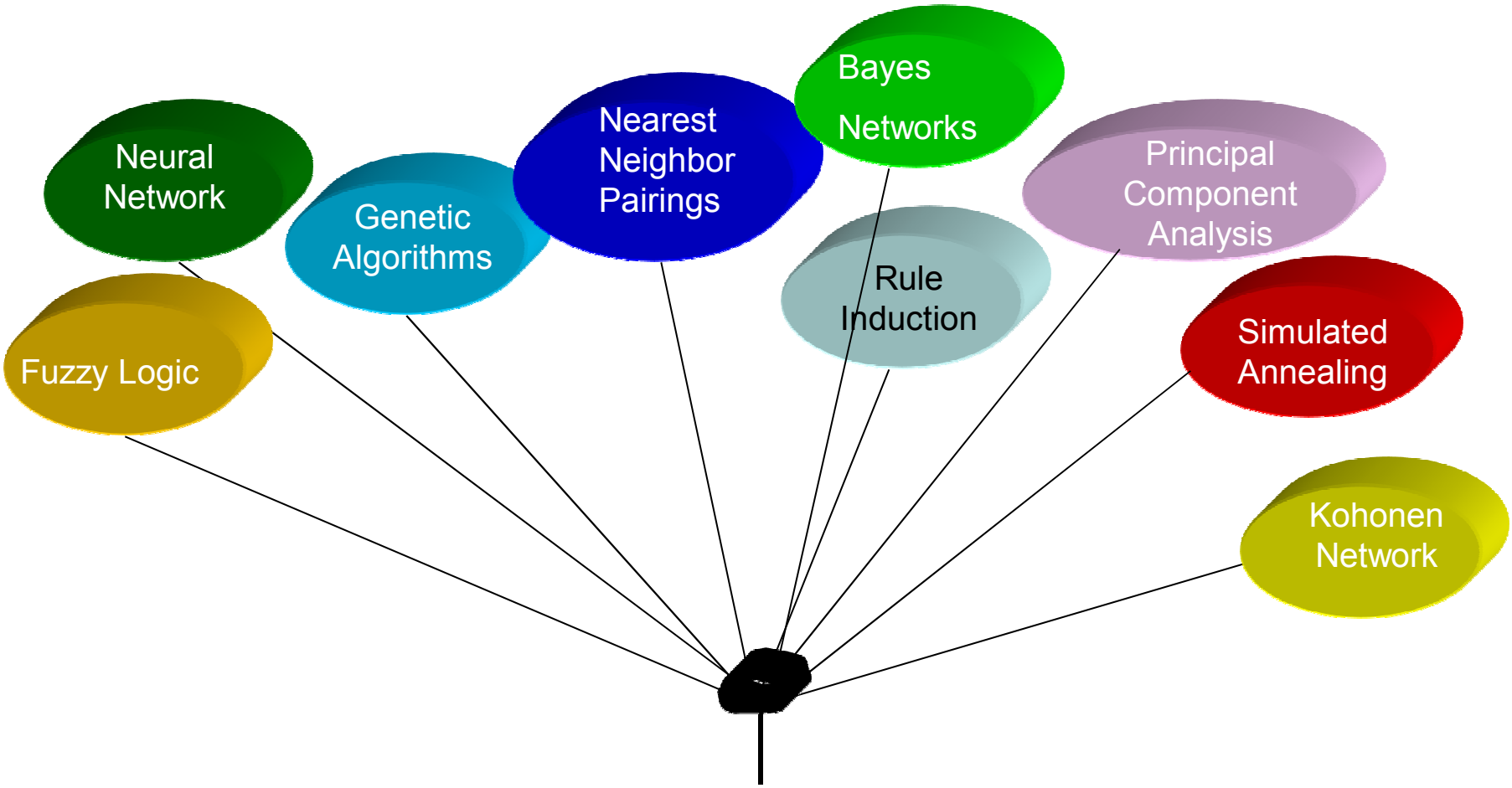
What is a Model??

- 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.). (In practical applications, individual scores are rolled-up or averaged at the population level.)
- For underwriting, we can predict either cost or risk-score. For care management, the prediction could be cost or the likelihood of an event.

Types of Statistical Models



Artificial Intelligence Models



Artificial Intelligence Models

- **Supervised learning (Predictive Modeling)**
  - Neural Network
  - Fuzzy Logic
  - Decision Trees (rule Induction)
  - K-Nearest Neighbor (KNN)
  - Etc.
- **Unsupervised Learning**
  - Association Rules (rule induction)
  - Principal components analysis (PCA)
  - Kohonen Networks, also known as Self-Organizing Maps (SOM)
  - Cluster Analysis, etc.

Optimization or Search Methods:

- Conjugate Gradient
- Genetic algorithm
- Simulated annealing
- Etc.

Developing your own model

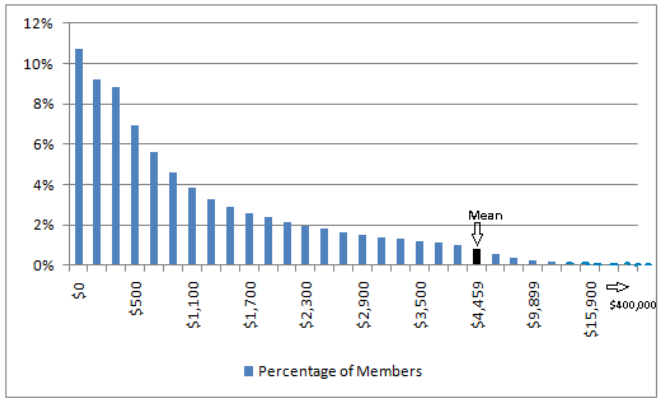
Here is a typical example (from Health Insurance):

- Can we predict future utilization (cost) of healthcare resources from current information in a member's record?
- Available data typically include:
  - Past utilization;
  - Providers that treat the patient;
  - Past and present diagnoses;
  - Past and present services (inpatient, outpatient, drugs, etc.)

Here is a snapshot of typically-available data:

Sample Observations From the Test Data Set

Member ID	Future Cost PMPM (y)	Current Cost PMPM (X_1)	Gender (X_4)
1308	35.37	27.91	0
1735	4.78	0.00	1
1954	774.28	406.58	0
2180	0.00	0.00	1
2249	180.25	131.83	1
2382	0.00	0.00	1
2914	73.70	44.48	1
2928	57.51	0.00	0
3122	175.20	211.70	1
3178	33.14	441.58	0
3199	7.03	0.00	1
3233	198.37	140.38	0
3346	325.74	406.35	1
3413	192.01	4973.6	1
3737	119.48	81.10	0
4789	1,212.85	94.17	0
4939	142.75	237.97	1
7038	814.32	738.09	1
7113	277.85	439.33	0
7203	37.00	147.14	0
Etc.			



Developing your own model

A couple of potential models come to mind:

- Time series
  - Time series could be useful for aggregate predictions; however, examination of individual-level data shows that there is considerable variation around the mean, implying that the technique may be inaccurate for individual-level prediction.
- Regression
  - Regression analysis allows us to explore associations in the data, determine which (of the many available) independent variables are associated with utilization and cost, and make projections for the future.
- We will consider some other models in a moment.

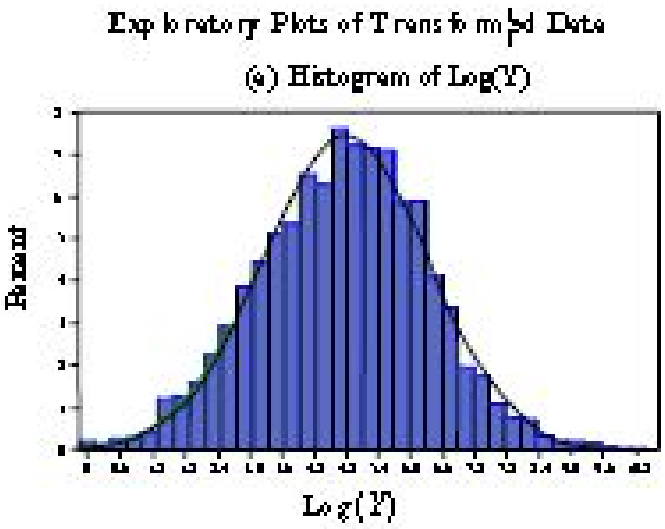
Developing your own model

Most of us are familiar with regression.

Regression *could* be applied but there are issues:

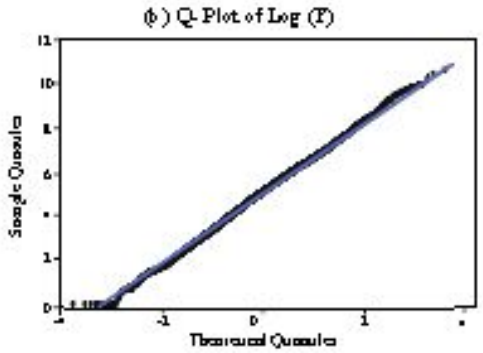
- The claims are not normally distributed;
- There is significant correlation between independent variables (multi-collinearity).

One solution is to transform the data in som form. We could take the logarithm of the claims data.



Transformed frequency distribution is approximately normally distributed

Taking the log of the dependent variable allows us to fit a straight line. (Problems at extremes though)



Developing your own model

Regular regression is inappropriate for the non-normal data often generated in healthcare.

Newer techniques exist that may be more appropriate:

- Logistic Regression, and
- Generalized Linear Models.

One way to approach the issue is to construct a 2-stage model:

- Stage 1: predicting the likelihood of a claim; and
- Stage 2: given that a member has a claim, predicting the size of the claim.

For Stage 1, a logistic regression model may be appropriate.

The logistic regression model is appropriate for categorical data (in this example, (0,1) depending on whether the member claimed).

Developing your own model

The "logit" model solves these problems:

$$\ln[p/(1-p)] = \alpha + \beta X + e$$

- p is the probability that the event Y occurs, p(Y=1); in our case, the probability of a claim.
- p/(1-p) is the "odds ratio,"
- ln[p/(1-p)] is the log odds ratio, or "logit."

Developing your own model

An example from the book: prediction of hospital admissions.

Frequency of admissions in the test dataset: 6.62%. Among predictors, current year cost, admissions, emergency room visits, PCP visits, age and member months are continuous. Other variables are binary or categorical.

After some analysis we derive the following model:

Ln (odds ratio for admission in Year 2) =

$$\beta_0 + \beta_1 \text{Admit flag in the first year} + \beta_2 \text{Male} + \beta_3 \text{Age over 64}$$
$$+ \beta_4 \text{ ER visit flag in the first year} + \beta_5 \text{PCP visits in the first year}$$

The model may be interpreted as follows: if we hold age, gender and utilization variables constant, the odds of adults with one additional PCP visit in the first year experiencing an admission are as much as:

$$\exp\left(\hat{\beta}_5\right) = \exp(0.075) = 1.078 \text{ times the odds of adults without any additional PCP visits}$$

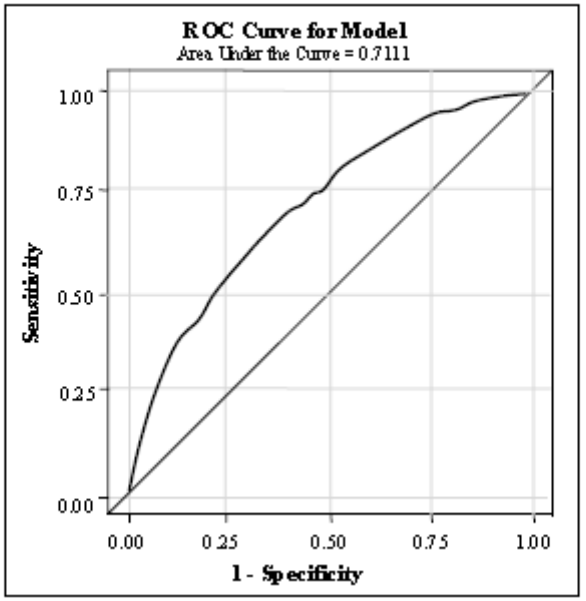
Developing your own model

Model Evaluation:

Most of us are familiar with techniques for evaluation of the accuracy of a regression model (R^2 anyone?)

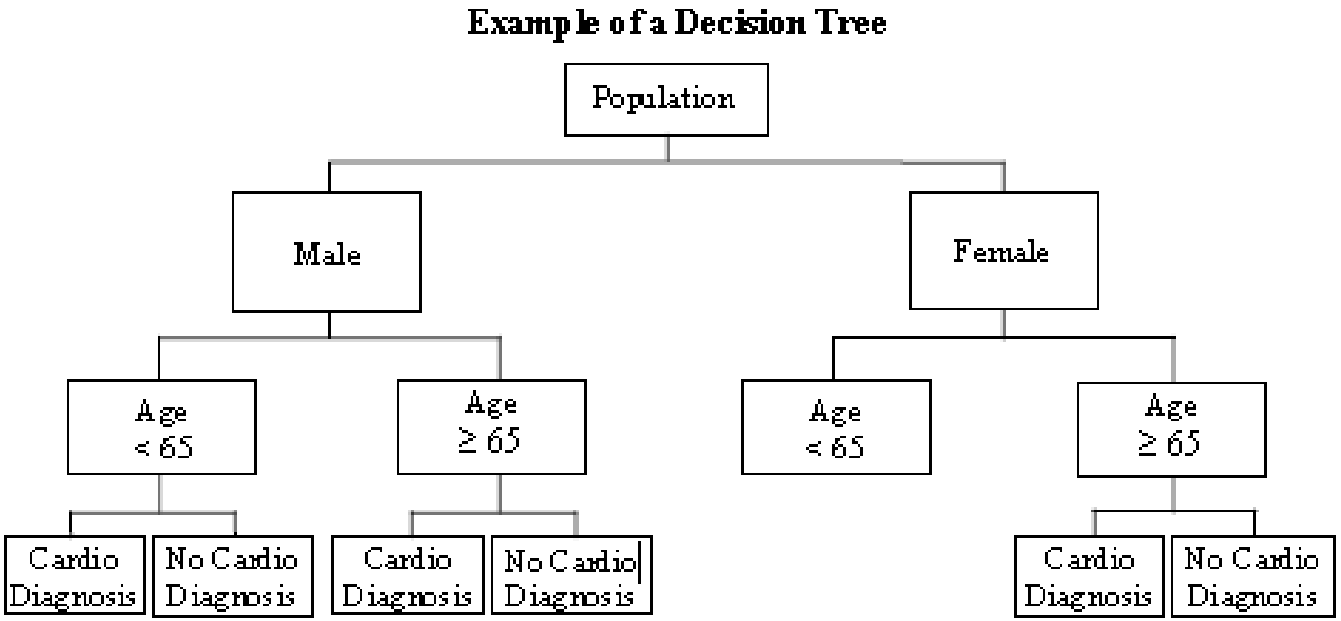
For evaluating Logistic Regression models we need other techniques. One which is useful is the Receiver Operating Curve. The area under the ROC curve, often termed the ***c*-statistic** or ***c*-index c**orresponds to the likelihood that an event will have a higher predicted Probability than an event at random.

ROC Curve for Fitted Logistic Regression Model I Using the Sample Data



Some Other Tools: Decision Trees

Here is a simple example:



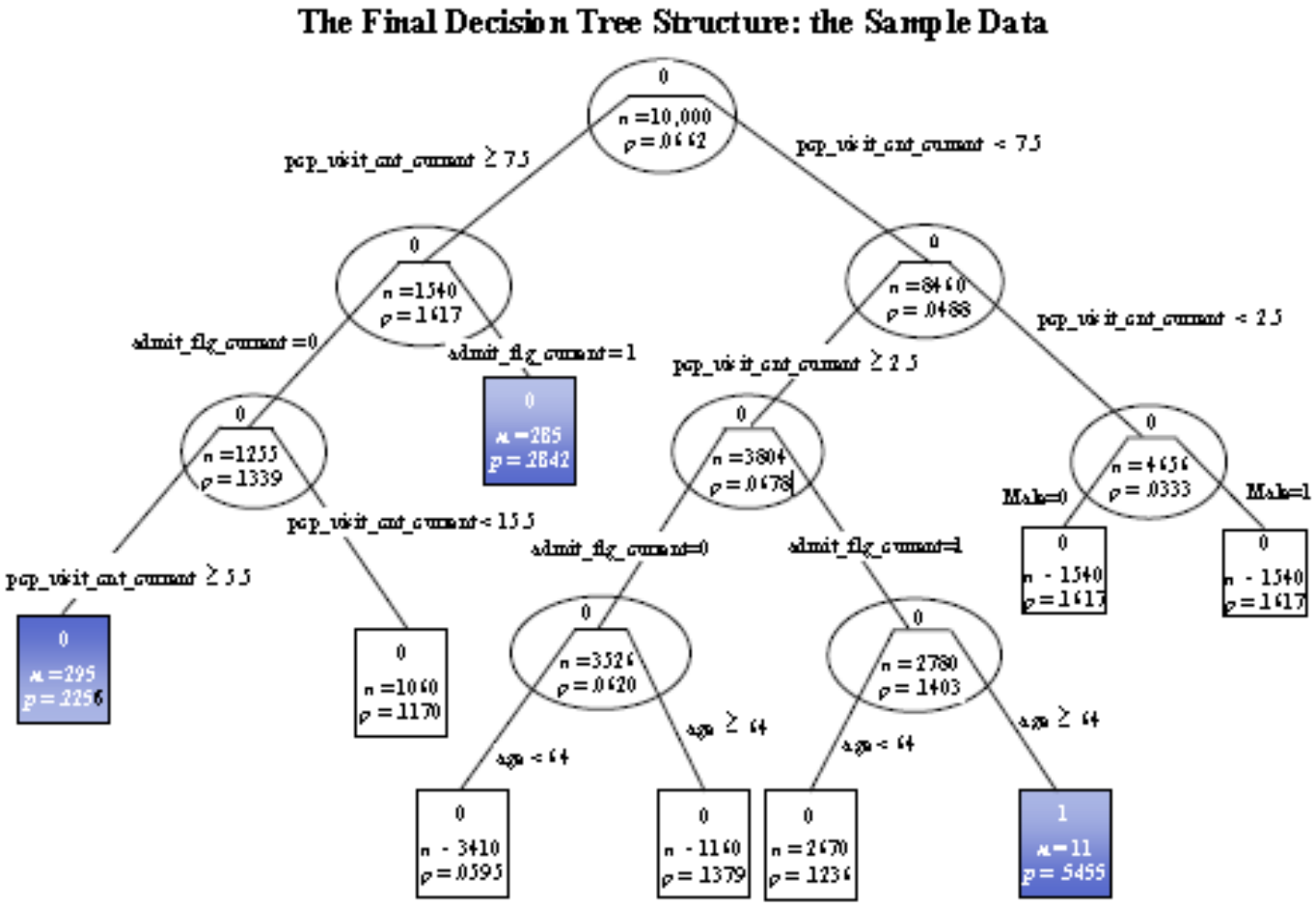
Some Other Tools: Decision Trees

Decision Trees have some advantages over other methods:

1. Tree methods are data-adaptive by giving data more freedom to choose the suitable bases that best approximate the true regression function.
2. Easy and meaningful interpretations may be extracted by tracing the splitting rules down the path to each terminal node. These combined rules help determine what has led to the event or nonevent.
3. The tree method provides a more efficient way to optimize the usage of categorical predictors, including the automatic merging of redundant levels, than regression models.
4. Trees are invariant to monotone transformations on continuous predictors.
5. The tree structure provides a natural and optimal way of grouping data, making them attractive to applications such as medical diagnosis or prognosis and credit scoring.
6. Trees excel in dealing with interactions of high order and complexity. Interactions are automatically and *implicitly* handled by the hierarchical tree architecture.
7. The tree method is *unstable* as a modeling tool, in that a small perturbation in the data could lead to dramatic changes in the final tree model. Although a single tree often does not attain satisfactory accuracy in a prediction task, multiple trees can be built and combined in model ensembles to achieve high accuracy of prediction.

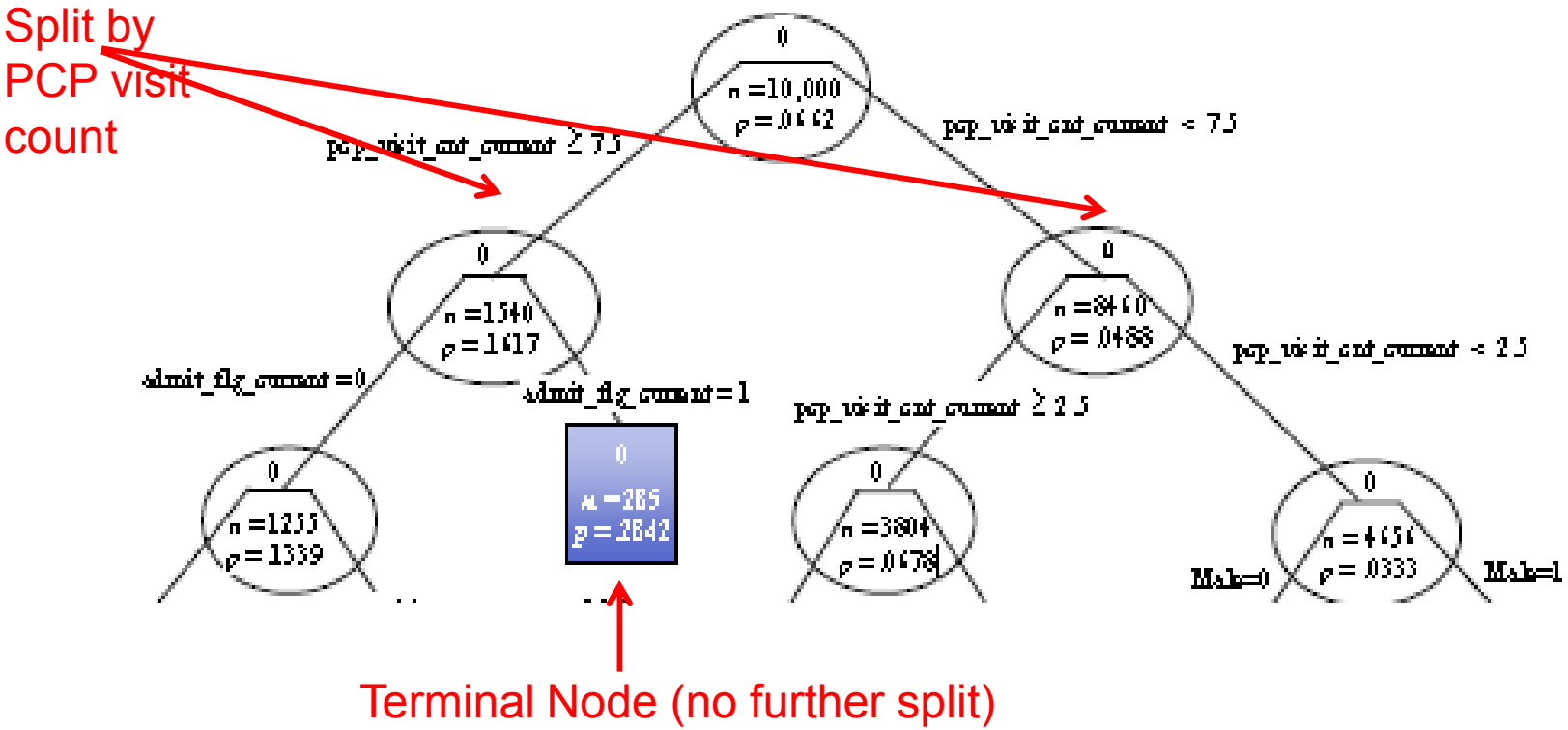
Some Other Tools: Decision Trees

Just to show what's possible, here is the hospital admission example, using decision trees:



Some Other Tools: Decision Trees

The model is hard to read, but here is a snapshot. Remember that in the logistic regression model, current year cost, admissions, emergency room visits, PCP visits and age were significant.



Some Other Tools: Neural Networks

Neural Networks

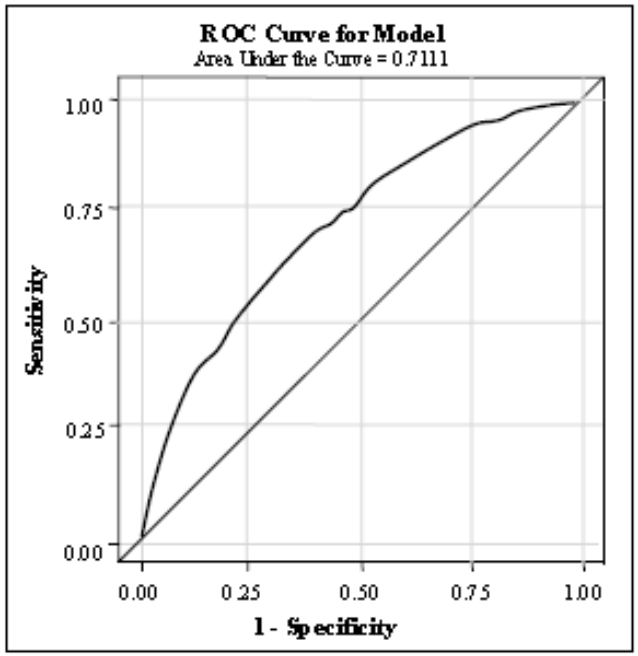
Neural networks are an example of machine-learning: instead of a single model (as in regression) the computer applies algorithms to find a series of local maxima.

Neural networks have appeal because they often fit more accurately than other models. But they have the disadvantage of not being analytical; thus for predictive modeling, for example, they do not have a form that allows us to implement them in a production environment.

Some Other Tools: Neural Networks

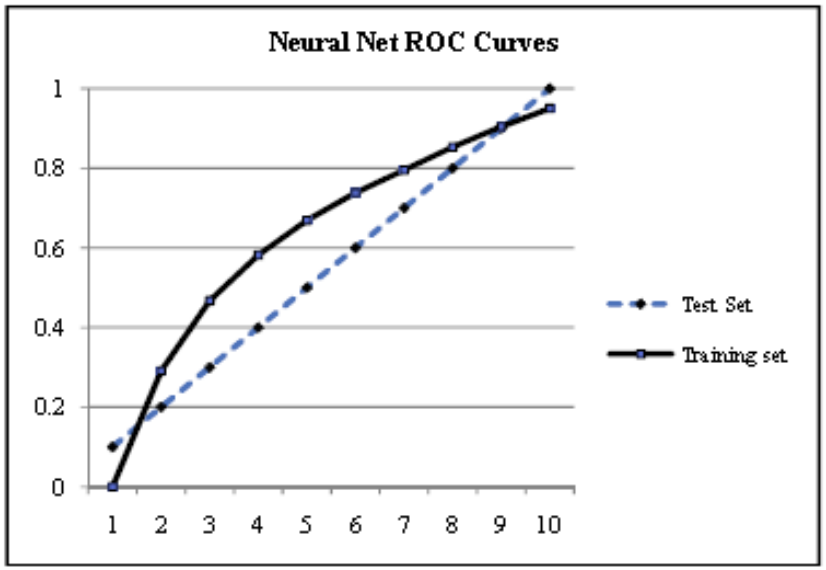
Applying a neural network model to the same data used in the hospital admission prediction problem (Logistic Regression), we get similar results:

ROC Curve for Fitted Logistic Regression Model I Using the Sample Data



C-statistic 0.71

ROC Curve from a Fitted MLP



C-statistic 0.72

Some Other Tools: Neural Networks

Unlike the different regression examples, or even the decision tree example, the underlying model is not discernible here. We cannot analytically find the likelihood of readmission in a different population.

In clinical models this can be problematic; clinicians like to know why patients are identified as being at high risk of re-admission.

Some Quick Comments

1. Linear Regression remains popular because it is simple, effective and practitioners understand it. Because residuals are not normally distributed, we need to use:
2. Generalized Linear Models (GLM): in these models, the linear relationship between the dependent and independent variables (basis of Linear Models) is relaxed, so the relationship can be non-linear.
3. Logistic Regression is one frequently used example of GLM in which dependence may be discrete (e.g. probability of a re-admission) rather than continuous (e.g. cost).
4. Decision Trees are a means of classifying a population using a series of structured, successive steps.
5. Non-linear regression: models relationships that are non-linear (often by transforming data).

Practical Model Building

Background

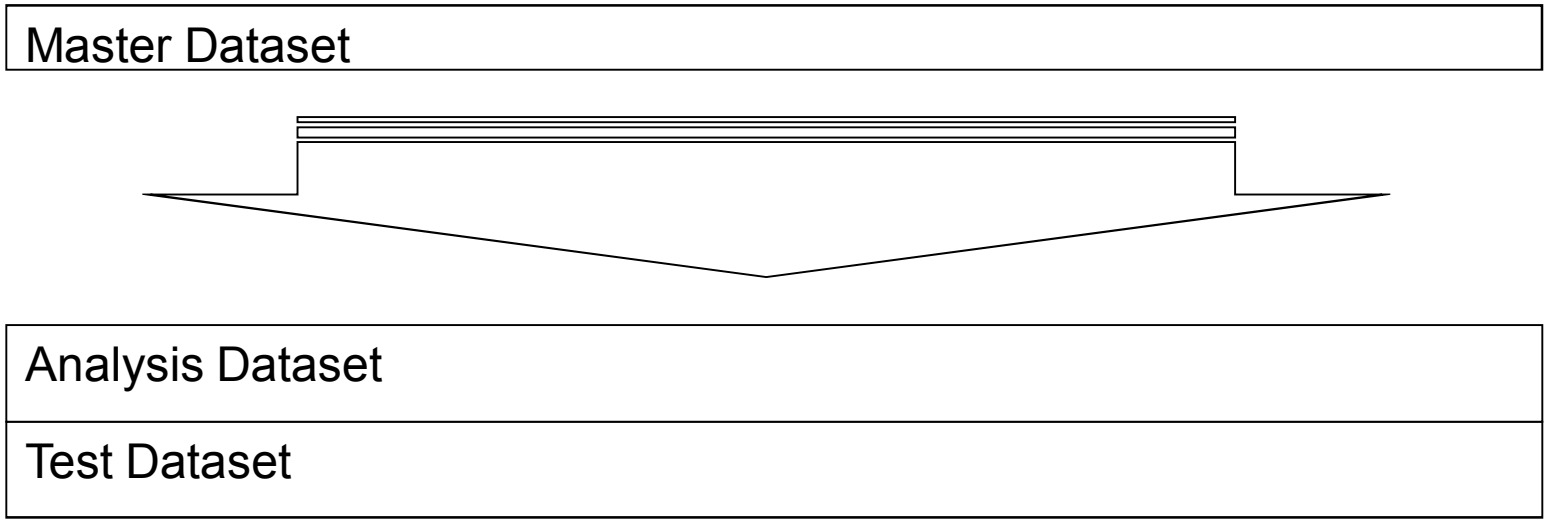
Available data for creating the risk score included the following:

- 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:

- The regression model was more intuitively understandable by end-users than other models; and
- With proper data selection and transformation, the regression was very effective, more so than the tree.

1. Split the dataset randomly into halves



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.

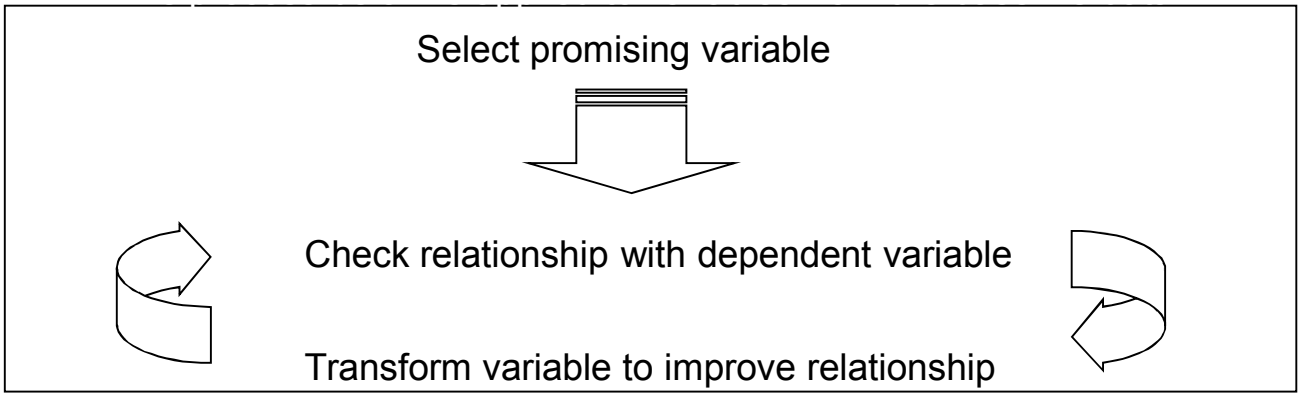
2. Build and Transform independent variables

- 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 dependent variable

- What are we trying to predict? Utilization? Cost? Likelihood of high cost?
- 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.
- So we could 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

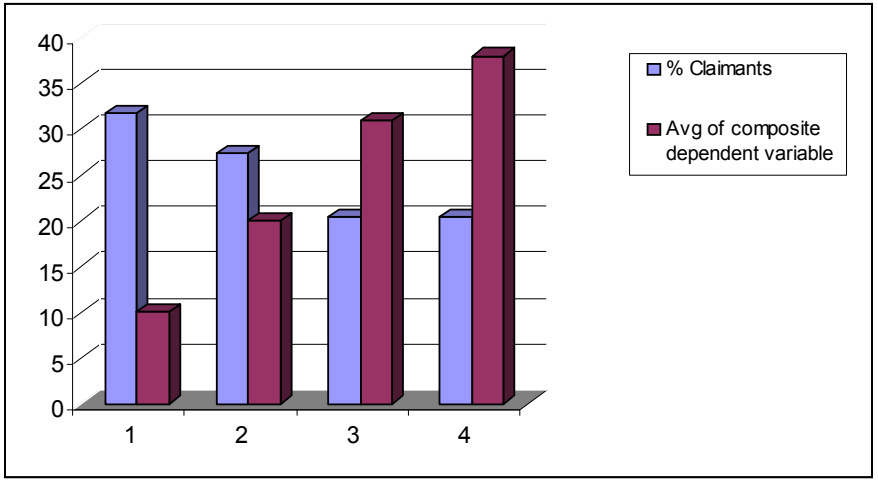


Typical transforms include.

- 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 conditions 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
- Neuroleptic drug flag
- Digoxin flag
- Number of corticosteroid drug prescriptions truncated at 2

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.

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 Regression formula:

(Female Model)

Scheduled drug prescription	358.1
NClass	414.5
MClass	157.5
Baseline cost	0.5
Diabetes Dx	1,818.9
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.

7. Results - Application in Practice

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	5k < 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

Lecture 4: Model Construction Re-cap

In this lecture we will looked at developing your own model using some typical techniques: linear and logistic regression, Decision Trees and neural networks. We also walked through an example of construction and use of a model.

In the Final Section of this Seminar we will look at a couple of case studies of applications of predictive modeling and risk adjustment.

Case Studies

In this section we will look at 4 different case studies, depending on time.

1. Predictive Modeling in ACOs.
2. Predictive Modeling for Bundled Payments.
3. Hospital Re-admissions.
4. Case management case identification and the application of Opportunity Analysis.

Example 1: Predictive Modeling for ACOs

Example 1: Predictive Modeling for ACOs

Statement of the Problem:

Financially, ACOs are provider-based organizations that are charged with managing the care of their (attributed) members. Medicare Shared Savings (MSSP) ACOs share gains relative to a projected baseline with CMS.

Although ACOs have a strong quality orientation, to save money and share in gains, the ACO needs to reduce utilization of high-utilizing members. Which members are future high utilizers? Where should the ACO apply its limited care management resources?

This is a classic Predictive Modeling/Opportunity Analysis application.

Overview - The Role of Analytics in ACOs

Analytics supports the goals of the ACO through the following processes:

Conducting opportunity analysis to identify (and then quantify) potential clinical programs;

1. Aggregating and warehousing data from multiple sources;
2. Predictive modeling/risk stratifying at the patient level for implementation of clinical programs;
3. Identifying gaps in care at the patient level;
4. Developing baseline quality measures for outcomes reporting (33 quality measures);
5. Providing ongoing reporting for program management and outcomes.

Overview – Predictive Modeling Defined

Predictive models stratify the patient population according to their likelihood of experiencing the target event. The process includes:

1. Using a similar dataset, identify all potentially correlated independent variables that predict the dependent (outcome) variable.
2. Derive scores for each patient (i.e. likelihood of experiencing the event) under numerous combinations of variables.
3. Compare the actual outcomes to the scores, to determine the scenario with the best positive predictive value. (PPV)
4. Operationalize the method for application to actual ACO data.

Develop a program to manage the targeted members.

Example 1: ACOs – Step 1

First Step: Analysis of current situation

- Where is the money spent today?
- Which high-cost patients represent an opportunity to change behavior or outcome?

For Medicare, analyze the Medicare 5% file.

Many analyses focus on *conditions.* We have found it useful to focus instead on groups of conditions (e.g. chronic) and co-morbidities.

Overall membership distribution by condition

- Members were classified into 8 hierarchical categories*:
  - Nearly 25% of members do not have an identified acute, chronic or mental health condition.
  - 35.3% of the <65 segment do not have an identified condition

Membership Distribution

Condition	<65	65+	Overall
1) Acute MH Chronic	11.0%	11.9%	11.7%
2) Acute MH	5.8%	3.6%	4.1%
3) Acute Chronic	7.3%	17.6%	15.2%
4) Acute Only	6.5%	9.1%	8.5%
5) MH Chronic	7.8%	4.3%	5.1%
6) MH Only	10.3%	3.9%	5.4%
7) Chronic Only	8.8%	15.7%	14.1%
8) EHC	7.2%	13.8%	12.2%
No Condition	35.3%	19.9%	23.6%

* Unlike typical hierarchical categories, these categories derive from type of intervention.

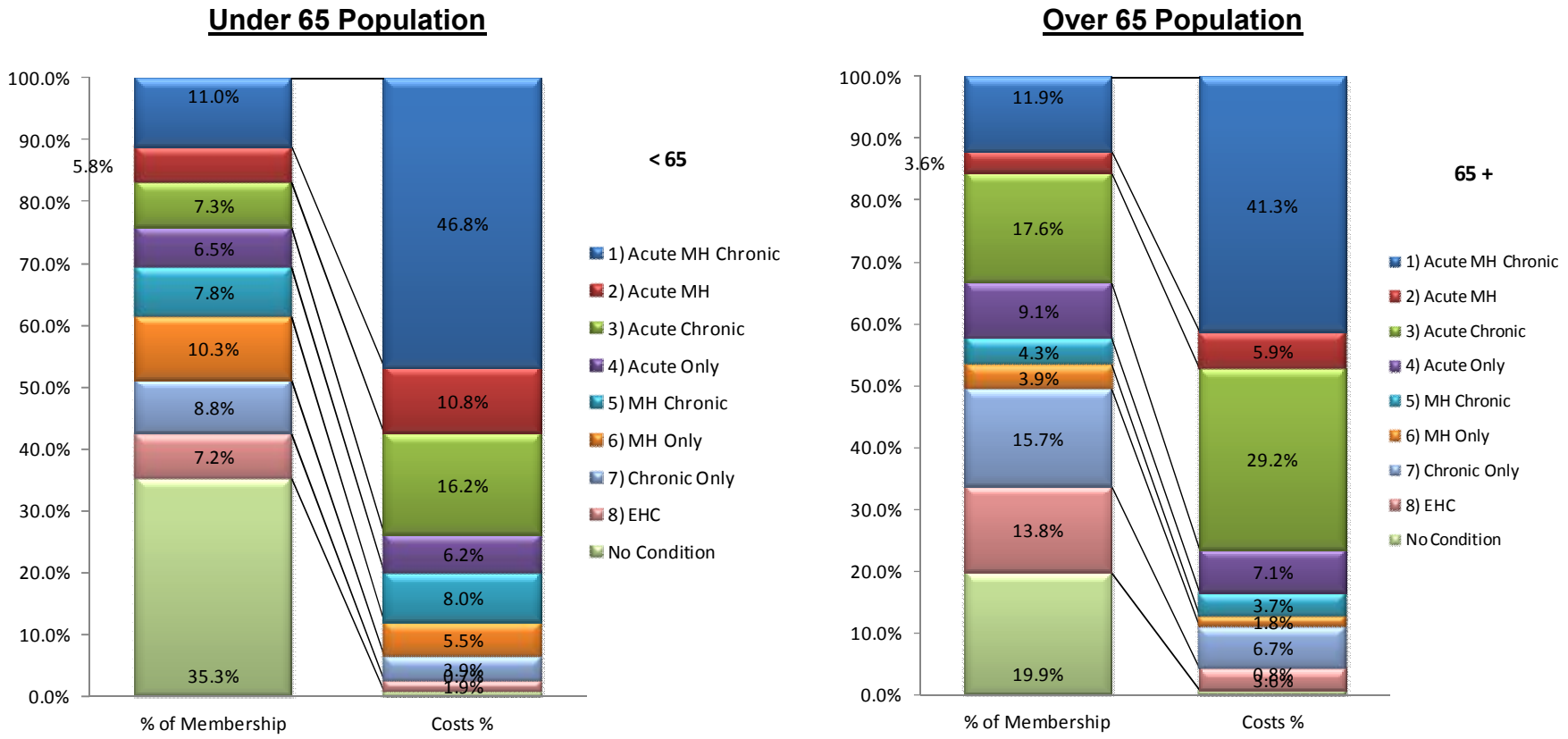
Overall cost distribution by condition

- When focusing on costs…
  - More than $4 of every $10 dollars is spent on most complex members.
  - Acute members with chronic and mental health comorbidities account for nearly 76% of all spend – care management is critical for these members

Condition	<65	65+	Overall
1) Acute MH Chronic	46.8%	41.3%	42.3%
2) Acute MH	10.8%	5.9%	6.8%
3) Acute Chronic	16.2%	29.2%	26.8%
4) Acute Only	6.2%	7.1%	6.9%
5) MH Chronic	8.0%	3.7%	4.4%
6) MH Only	5.5%	1.8%	2.4%
7) Chronic Only	3.9%	6.7%	6.2%
8) EHC	1.9%	3.6%	3.3%
No Condition	0.7%	0.8%	0.8%

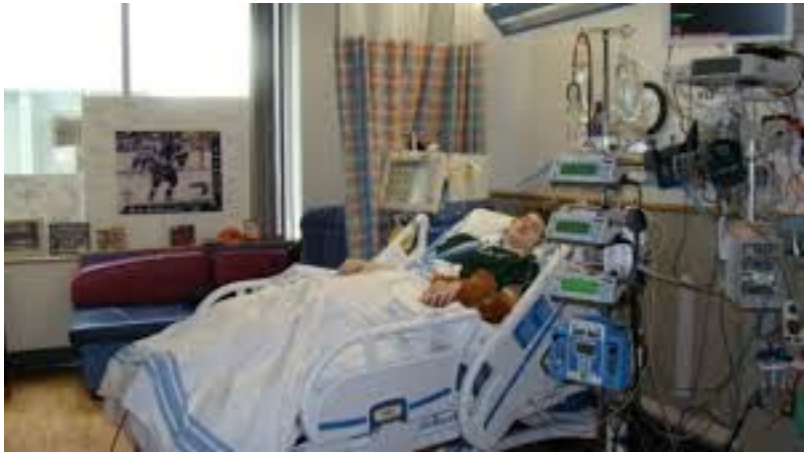
Distribution of Aged and Disabled Population

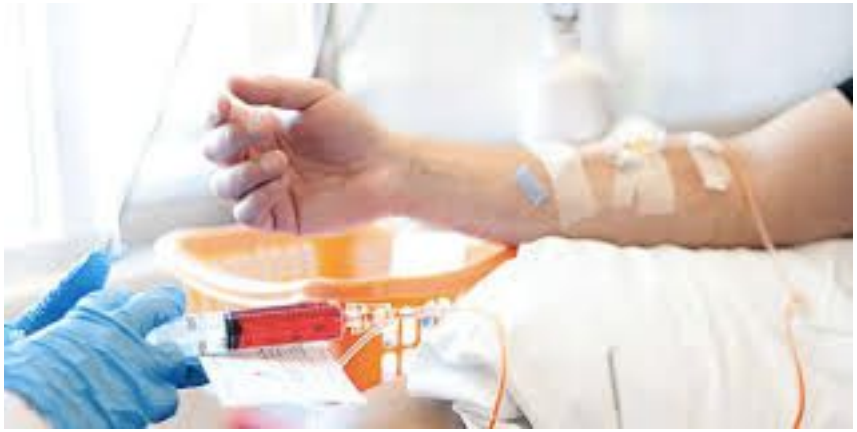
Comparison between <65 members and >65 population.



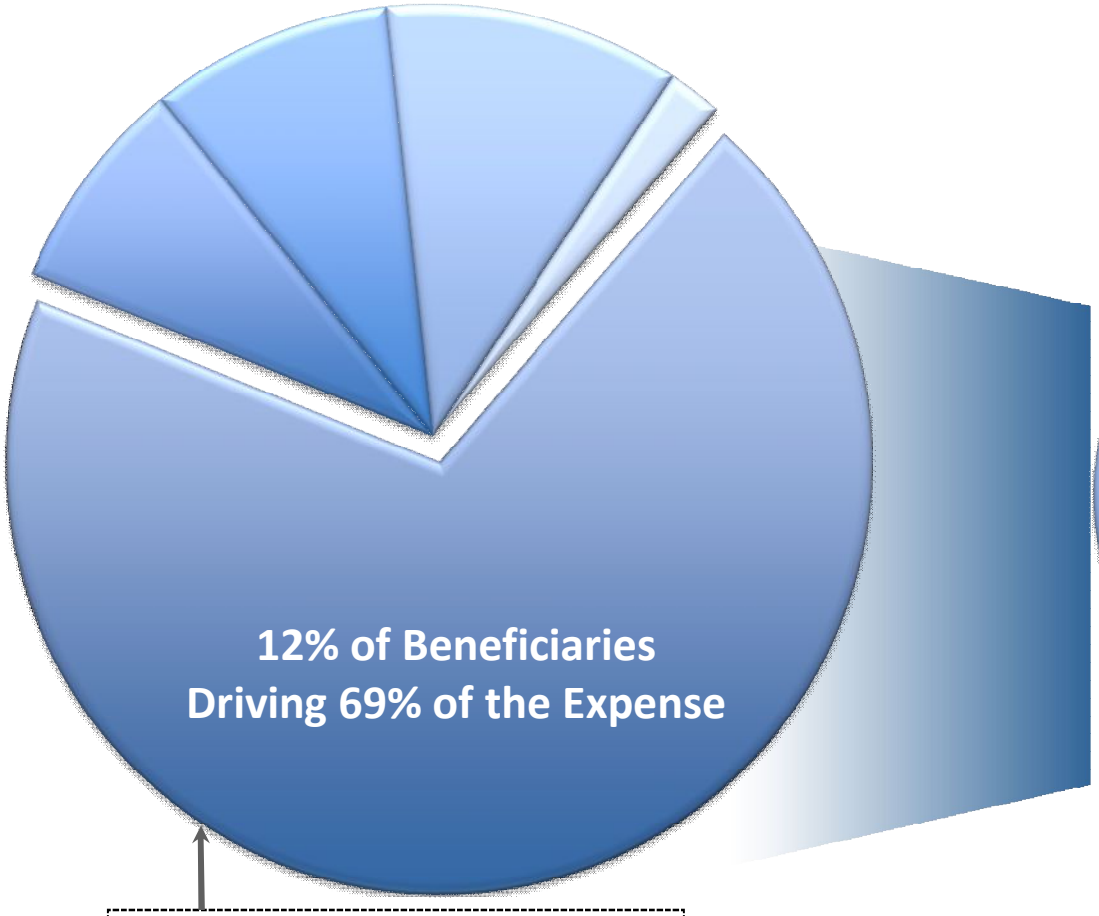
Preventing over-medicalized End-Of-Life care

Highly-Medicalized Deaths



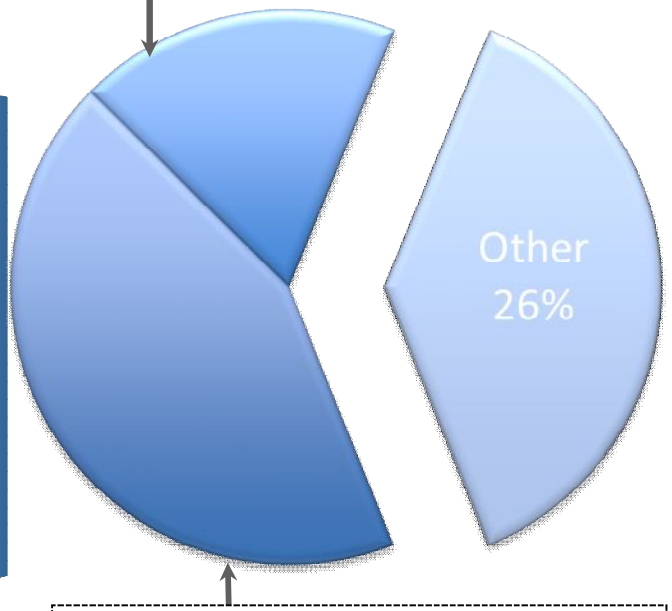


Spending at End-of-Life



5% of Medicare Beneficiaries die annually.

Second to last year of life represents 13% of the total Medicare FFS spend.



Last year of life represents ~30% of the total Medicare FFS spend.

End-of-Life (all)

The most complex members: Comparative Utilization.

Complex end of life patients have a high frequency of hospital admissions (2500 per 1000). Most of these are for medical DRGs. Also very high specialist visit frequency.

Under 65 patients are an even high-utilizing group.

	OVERALL			1) Acute MH Chronic		
	< 65	65 +	Total	< 65	65 +	Total
Average Lives	3,788	33,230	37,017	1,509	15,627	17,137
IP Admits - Overall	2,548	1,929	1,977	3,136	2,499	2,555
Average IP Length of Stay - Overall	7.8	7.2	7.7	8.5	7.6	7.7
IP 30 Days Re-Admits - Overall	800	421	426	998	566	604
Readmit % - Overall	31.4%	21.8%	21.5%	31.8%	22.6%	23.6%
IP Admits - Medical	2,243	1,682	1,680	2,691	2,154	2,201
Average IP Length of Stay - Medical	7.1	6.7	6.9	7.5	6.9	7.0
IP 30 Days Re-Admits - Medical	744	379	378	901	503	538
Readmit % - Medical	33.2%	22.5%	22.5%	33.5%	23.3%	24.4%
IP Admits - Surgical	305	247	297	445	345	354
Average IP Length of Stay - Surgical	13.1	10.6	12.1	14.5	12.1	12.4
IP 30 Days Re-Admits - Surgical	56	43	48	97	63	66
Readmit % - Surgical	18.4%	17.2%	16.1%	21.7%	18.2%	18.6%
LTC Admits	-	0.1	0.3	-	0.4	0.4
ER Visits	311	72	123	356	138	157
OP Services	17,112	15,883	12,816	15,444	15,546	15,537
PCP Visits	2,448	2,605	2,744	3,457	2,927	2,973
Specialist Visits	6,818	4,989	4,850	6,539	4,718	4,878
CT Services	3,750	2,726	2,626	4,003	3,261	3,326
MRI Services	844	482	433	718	489	509
X-Ray Services	10,352	7,916	9,365	15,796	12,251	12,563

Preventing over-medicalized End-Of-Life care

Clinical Program:

- Education for physicians and their staff on how to instigate end-of-life conversations.[3]
- Program to encourage patients to complete advance directives, consisting of materials, a helpline, and a registry.[4]
- Patient access to hospice and palliative care.
- Symptom-focused case management for very high-risk patients.[3]

↑ Population health	↓ Administrative burden	↓ Per capita cost	↑ Patient experience
Reduction in inappropriate life-sustaining treatments within 6 months of death, including a reduction in ER visits.[2]	Dedicated case managers to support physicians in caring for complex patients that are at very high risk of over-medicalized end-of-life care as defined by Barnato et al.[2]	Home-hospice care associated with significantly lower average costs ($12,434 versus $4,761 per year in 2007 dollars).[5]	Patients receiving in-home palliative care report significantly higher satisfaction and quality of life.[6]

[1] Zhang B, Wright AA, Huskamp HA, et al. Health care costs in the last week of life: associations with end-of-life conversations. *Archives of Internal Medicine.* 2009;169(5):480
[2] Barnato AE, Farrell MH, Chang CC, Lave JR, Roberts MS, Angus DC. Development and validation of hospital "end-of-life" treatment intensity measures. Medical Care. 2009;47(10):1098-1105
[3] Wright AA, Zhang B, Ray A, Mack JW, Trice E, Balboni T, et al. Associations between end-of-life discussions, patient mental health, medical care near death, and caregiver bereavement adjustment. JAMA 2008; 300(14):1665-73
[4] Nicholas L, Langa KM, Iwashyna TJ, Weir DR. Regional variation in the association between advance directives and end-of-life Medicare expenditures. JAMA 2011; 306(13):1447-53
[5] Shnoor Y, Szlaifer M, Aoberman AS, Bentur N. The cost of home hospice care for terminal patients in Israel. Am J Hosp Palliat Care. 2007 Aug-Sep;24(4):284-90
[6] Brumley R, Enguidanos S, Jamison P, Seitz R, Morgenstern N, Saito S, McIlwane J, Hillary K, Gonzalez J. Increased satisfaction with care and lower costs: results of a randomized trial of in-home palliative care. J Am Geriatr Soc. 2007 Jul;55(7):993-1000.

End of Life Predictive Model - Definition

Over-medicalized death is defined as:

- Chemotherapy for cancer patients within 14 days of death;
- Unplanned hospitalization within 30 days of death;
- More than one emergency department (ED) visit within 30 days of death
- ICU admission within 30 days of death; or
- Life-sustaining treatment within 30 days of death.

· Ho, T. H., Barbera, L., Saskin, R., Lu, H., Neville, B. A., & Earle, C. C. (2011). Trends in the aggressiveness of end-of-life cancer care in the universal health care system of Ontario, Canada. *J Clin Oncol, 29*(12), 1587-1591. doi:10.1200/JCO.2010.31.9897. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3082976/pdf/zlj1587.pdf

· Earle, C. C., Park, E. R., Lai, B., Weeks, J. C., Ayanian, J. Z., & Block, S. (2003). Identifying potential indicators of the quality of end-of-life cancer care from administrative data. Journal of Clinical Oncology, 21(6), 1133-1138. doi: 10.1200/jco.2003.03.059 Retrieved from http://jco.ascopubs.org/content/21/6/1133.long

End of Life Predictive Model - Scoring

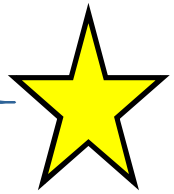
- An EOL risk score is calculated for each member.
- Risk scores range in value from 0.0-1.0.
- Model is based on the following member attributes (121 in all):
  - Age and gender;
  - Race;
  - Region
  - Clinical Grouper Flags (65 HCCs);
  - Baseline admission count(s)
  - Baseline readmission count(s)
  - Baseline ER visit count(s)
  - Baseline admission via ER indicator
  - Baseline dollars spent for healthcare resources

Conditions and Attributes that Add Most to Scores

1. Acute Myocardial Infarction
2. Acute Leukemia
3. Craniotomy with major device implant
4. Cardio-Respiratory Failure & Shock
5. Metastatic Cancer & Acute Leukemia
6. Lung, Upper Digestive Tract and Other Severe Cancers
7. Septicemia or Severe Sepsis
8. Number of Admissions

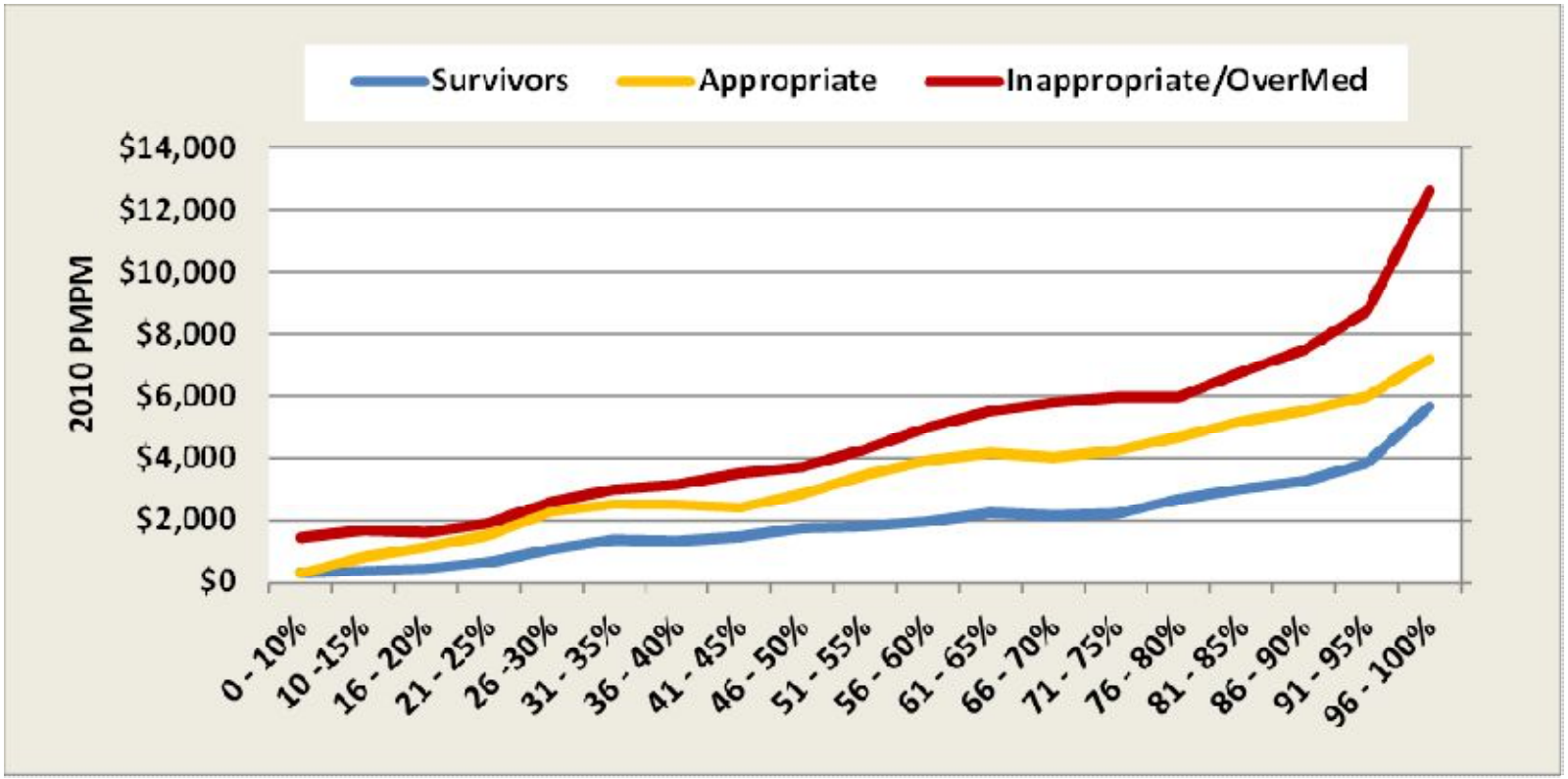
Opportunity as seen from Medicare 5% Database

Medicare Patients and Deaths (based on 50% of the 5% file)			
Categories	**Members**	**% of Total Population**	**PMPM**
Survivors	819,189	92.0%	$684.80
Deceased	71,059	8.0%	$4,323.73
Appropriate	22,989	2.6%	$2,249.62
Inappropriate	**9,832**	**1.1%**	**$3,433.30**
OverMedicalized	**38,238**	**4.3%**	**$5,797.08**
Total	890,248	100.0%	$975.26



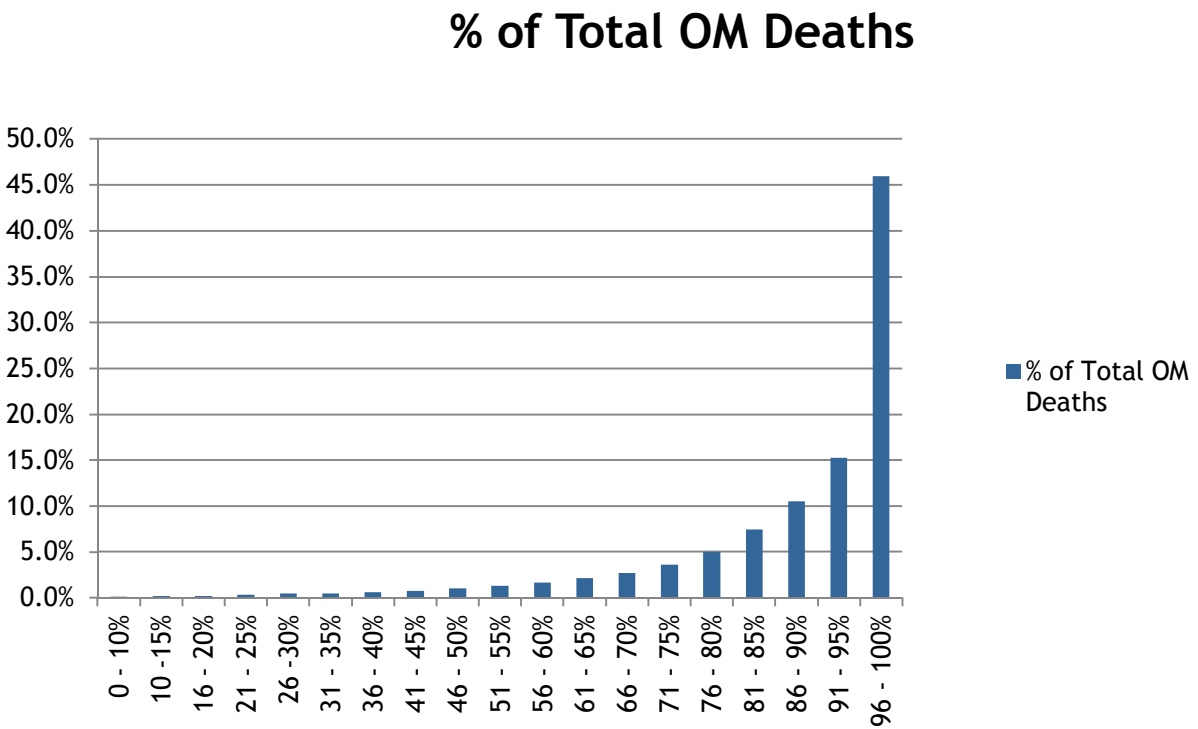
The difference between over-medicalized and appropriate death represents a financial and clinical opportunity. (Inappropriate death also represents an opportunity, although a smaller one).

Member costs by category and risk score



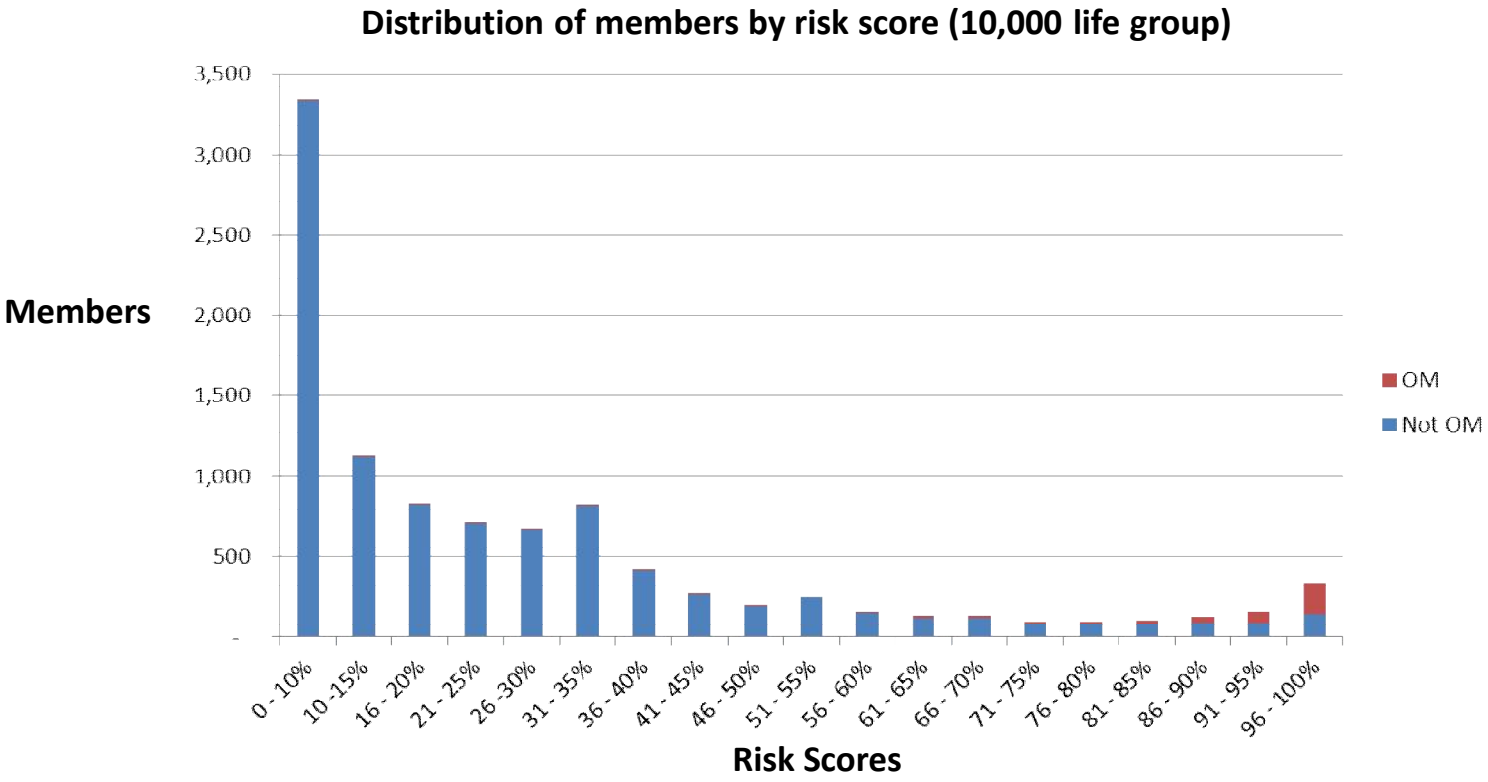
The PMPMs for members in each category vary across the bands of risk scores. The difference in the costs between those that experience overmedicalized deaths versus those that experience appropriate deaths is greatest in members with risk scores >.95.

Performance of Model on Medicare 5% Database



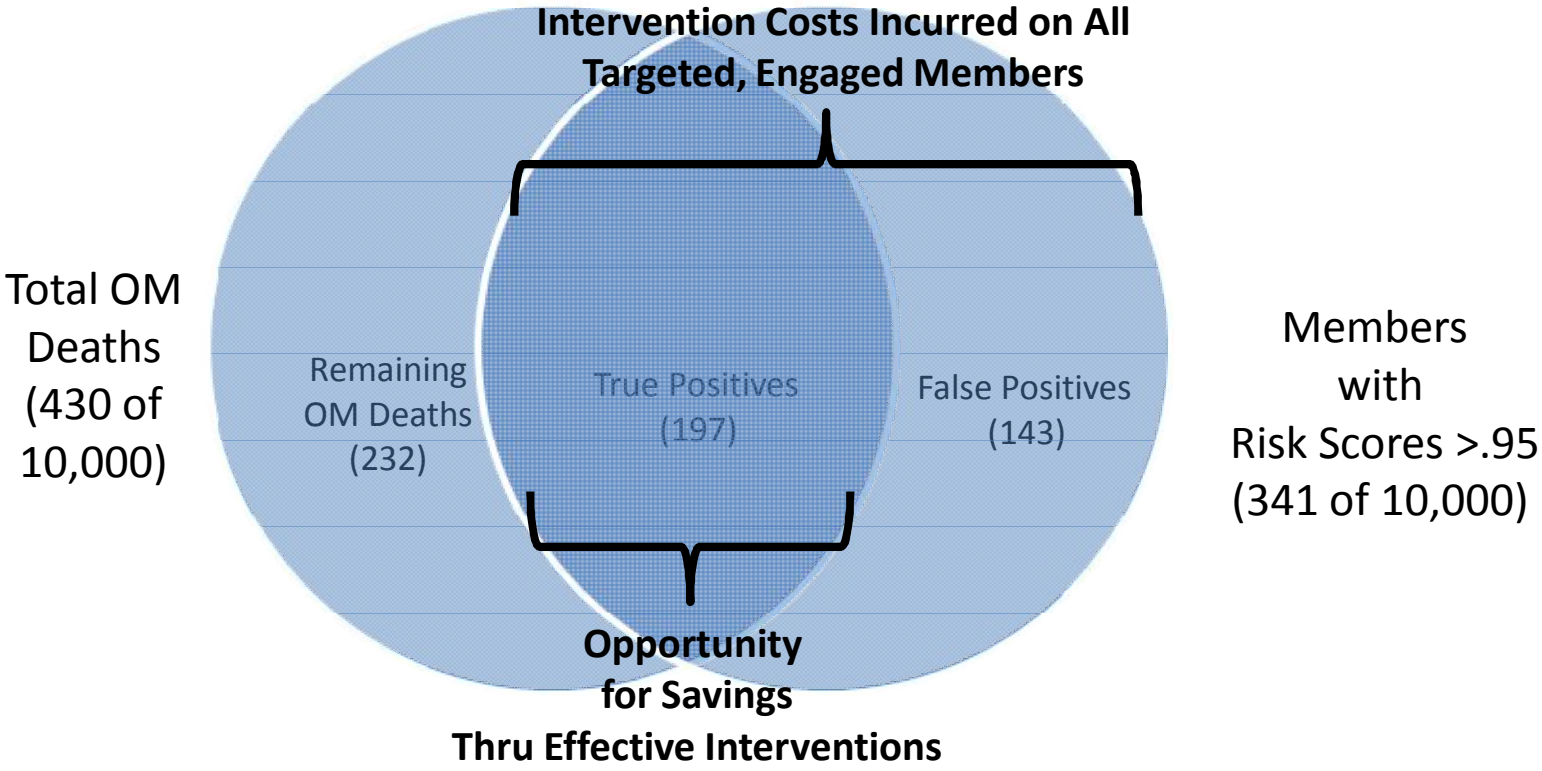
Out of a 10,000 attributed life group, we would expect 430 overmedicalized deaths (4.3%). Based on our model, approximately 46% of these members will have risk scores >.95.

Performance of Model on Medicare 5% Database



Out of a 10,000 attributed life group, we would expect 341 members to have risk scores >.95. Of these members, we expect 197 (57.9%) to be "true positives"; that is, these are the members that represent an opportunity to avoid an overmedicalized death.

End of Life Predictive Model - Targeting



The risk score "cut off" point is determined by evaluating the number of total members above a given risk score with the number of "true positives" found in that group. We will incur intervention costs on all members with risk scores above the cut-off, but only have the opportunity to generate savings on the "true positives" within that group.

Financial Scenario at 95% Risk Score Threshold

Based on Members with Risk Scores >.95	
# of Members (out of 10,000)	341
% of Members (out of 10,000)	3%
Over Medicalized Sensitivity	46.0%
PPV (OM Deaths)	57.9%
# of True Positives (out of 10,000)	197
# of False Positives (out of 10,000)	143
Estimated Gross Savings	
# of True Positives (a)	197
Engagement Rate (b)	40%
Effectiveness Rate (c)	50%
Potential Savings per True Positive (d), (1)	$ 15,981
Estimated Gross Savings (a x b x c x d)	**$630,853**
Estimated Net Savings	
# of Members with p>.95 (e)	341
Engagement Rate (b)	0%
Cost of Case Management (f)	$ 940.67
Total Cost (e x b x f)	$128,234
Net Savings/(Costs)	**$502,619**

(1) Difference in costs between OM death and appropriate death, over 6.5 months (PMPM*6.5).

Example 2: Bundled Payments

- CMS (CMMI) has offered hospitals and other providers the opportunity to receive bundled payments for episodes of care, rather than current fee-for-service reimbursement. Other payers are also interested.
- Under the CMMI Bundled Payments initiative, CMS would link payments for multiple services patients receive during an episode of care. For example, instead of a surgical procedure generating multiple claims from multiple providers, the entire team is compensated with a "bundled" payment that provides incentives to deliver health care services more efficiently while maintaining or improving quality of care. Proposers will have flexibility to determine reimbursement of different providers from the bundled payment.
- The expectation is that bundled payments will align incentives for providers – hospitals, post acute care providers, doctors, and other practitioners– encouraging them to partner closely across all specialties and settings that a patient may encounter to improve the patient's experience of care during a hospital stay in an acute care hospital, and during post-discharge recovery.

Example 2: Bundled Payments

- From the perspective of predictive modeling, the Bundled Payments initiative is no different than other applications: it requires modeling to predict frequency and severity in a population and methods for identifying and mitigating risk.
- The CMS bundled payment initiative allowed proposers to select their own episodes and define those procedures that they wished to exclude.
- Providers have focused on inclusiveness (finding as many DRGs as possible to include in the bundle) to maximize revenue. However, this runs the risk of heterogeneity: including all sorts of different patients, at different risk levels and with potentially variable outcomes in the bundle.
- Providers address the variability/unpredictability problem by excluding many down-stream procedures, re-admissions for complications etc. While this is *a* solution to the problem of variability in outcome, it is not ideal, reducing the provider's risk to the primary admission only and destroying the incentive to coordinate care and reduce risk.

Example 2: Bundled Payments

Optimally, how would we organize a bundled payment initiative?

- Choose the condition or procedure for which bundling is being considered. Generally these tend to be acute episodes, although some bundlers have also undertaken bundling of chronic conditions.
- Ideally the episode or condition should be one for which clear treatment protocols and evidence-based guidelines exist.
- Once the conditions or procedures are identified, investigate what the data tell us:
  - Frequency of different procedures
  - Severity
  - Outliers
  - If the episode is admission-based, what is the ideal pre- and post-admission period for the inclusion of services?
  - If protocols/guidelines are available, to what extent does actual practice mirror best-practice?

Example 2: Bundled Payments

Optimally, how would we organize a bundled payment initiative?

- Develop a risk adjusted tiered bundle payment that ensures provider is appropriately compensated in caring for sicker patients. (Ideally, the Insurer should assume population risk (frequency and severity) not the provider.
- The number of tiers, and the identification of risk factors for tiering patients will be an empirical exercise.
- Additionally, determine claims distributions in the "tail" to inform the outlier/reinsurance discussion.
- Develop a predictive model for prospectively assigning patients at inception/diagnosis by risk tier.
- Determine the price per bundle to be charged to the payer.

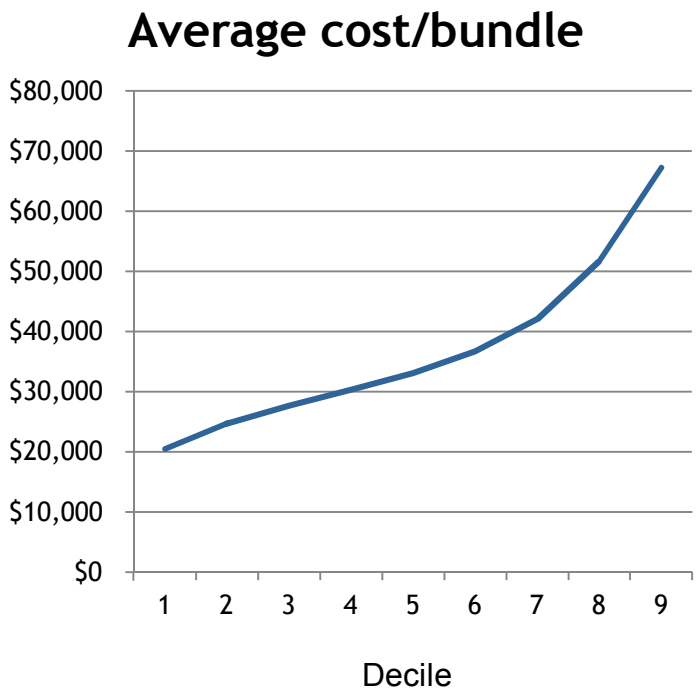
The bundler's reimbursement will be determined by the number of patients accepted at each tier, the price per bundle per tier, and reimbursement (if any) for outliers.

Example 2: Bundled Payments

Payment data indicate that for this state, the reimbursements for the bundle of activities varied from a low of $20,000 to a high of more than $600,000.

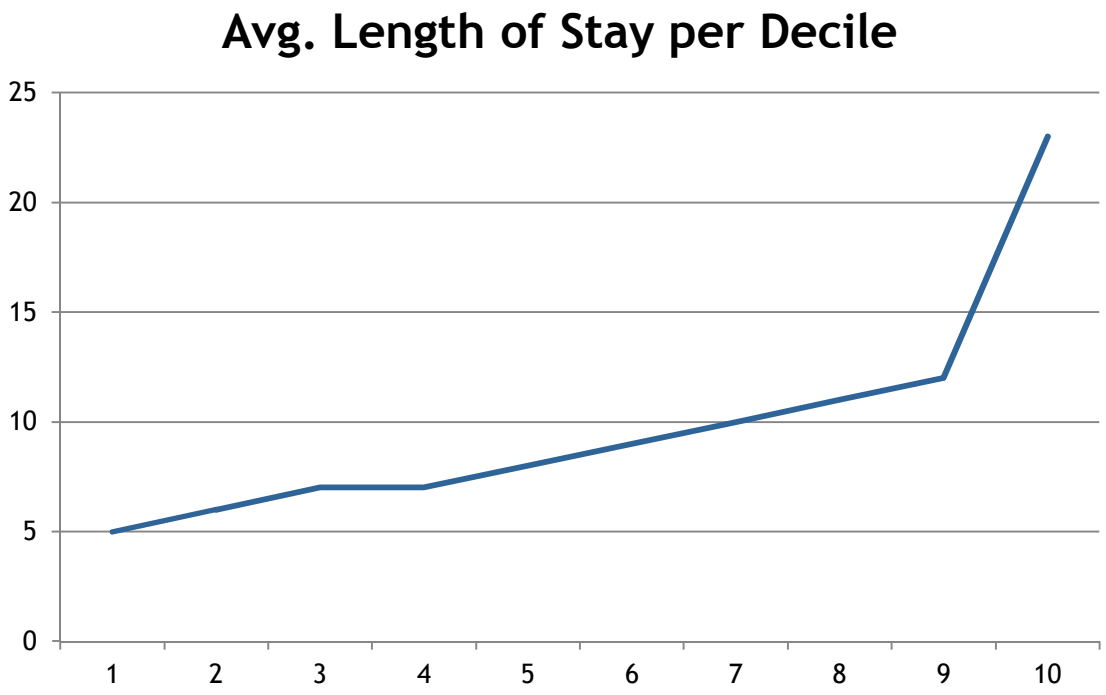
This range underscores the importance of identifying a patient's risk prospectively and of assigning the patient to the appropriate stratum for reimbursement.

The risk adjustment process assigns patients prospectively to the appropriate stratum; providers are then reimbursed for the bundle of services at the rate appropriate for the patient's stratum.



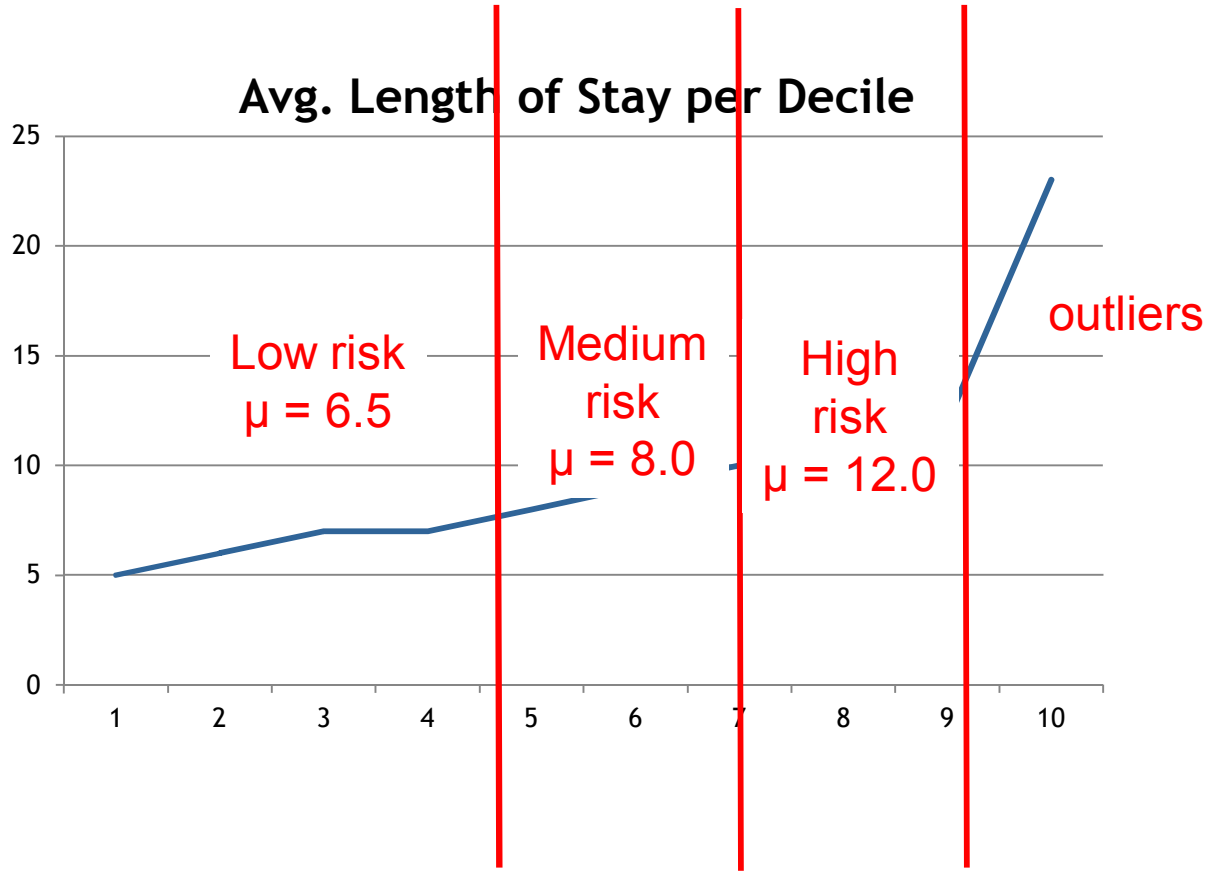
Example 2: Bundled Payments

Looking at average length of stay by decile helps inform the stratification and outlier strategy.



Example 2: Bundled Payments

One possible grouping structure:



Example 2: Bundled Payments

Developing a Model for Bundled Payments

The bundled payment model assigns patients to a risk stratum based on commonly-available risk factors. RISK FACTORS MUST BE COMMONLY-AVAILABLE PRIOR TO ADMISSION.

The ideal risk-adjusted bundled payment model will minimize intra-stratum variation, while maximizing variation among strata. Said another way, the analysis must identify the minimum number of strata that recognize legitimate variation in payments, without introducing an overly complex number of strata.

A key to developing the model was to identify risk factors that were correlated with higher payments. A patient presenting for surgery will then be placed into one of several strata based on his/her risk factors. The assignment to a particular stratum must be based on the types of data available when the patient is scheduled for surgery. Thus, while it is possible to construct sophisticated bundle algorithms, a practical model must be tailored to available payer (or analytics provider) data.

Example 2: Bundled Payments

To develop the model, we used as a measure of relative patient risk the ratio of average payment per episode to the aggregate average payment. (See Distribution of Relative Risk Scores). This relative score for each episode was used as an explanatory variable to predict future payment. A wide range of other independent variables was tested, including member condition, demographic, socio-economic status, and lab and blood screenings.

Variables selected for the model were:

- Age
- Sex
- Socioeconomic status
- Diabetes diagnosis
- Metabolic syndrome diagnosis
- Cancer diagnosis
- Mental Health/Psychiatric diagnoses
- Pulmonary disease diagnosis
- Renal disease diagnosis
- CNS diagnoses diagnosis
- Substance abuse treatment/diagnosis
- Presence of other procedures
- Previous inpatient claim for osteoarthritis

Decile	Risk Score
10	0.490
20	0.593
30	0.662
40	0.728
50	0.793
60	0.880
70	1.007
80	1.240
90	1.613
100	14.726

Example 2: Bundled Payments

The key output of the model is the predicted average payment, which was generated by applying the model weights to each patient's explanatory variable values. The predicted and actual average costs were reasonably close for risk score bands 1 and 2, indicating that the model is likely to perform well as a predictor of cost for patients in those bands. The model somewhat under-predicted costs for members in risk band 3.

Risk Score Bands, Actual and Predicted Payments

Risk Score Band	Number of Members	Risk Score Mean	Actual Average Payment, Truncated	Predicted Average Payment
1 (≤ 0.9)	965	0.80	$29,291	$29,814
2 (0.9-1.2)	1,842	1.00	$36,689	$37,660
3 (> 1.2)	650	1.40	$57,053	$53,524

Example 2: Bundled Payments

It is important to understand how this type of bundling could affect a participating hospital. This table illustrates how a specific hospital's current reimbursement would be affected if it participated in bundling using the predicted risk score and reimbursement by stratum in the prior slide.

Distribution of Episodes and Reimbursement at the Sample Hospital

Risk Score Band	Number of Members	Risk Score Mean	Actual Average Payment Truncated	Predicted Average Payment	Variance
1 (≤ 0.9)	130	0.78	$33,981	$29,814	-$4,167
2 (0.9-1.2)	249	1.00	$40,736	$37,660	-$3,076
3 (>1.2)	123	1.53	$62,187	$53,524	-$8,663

Example 2: Bundled Payments

Overall the model indicates that our sample hospital will receive less reimbursement under the bundling model than under the current fee-for-service model. This provides an incentive for the hospital to increase its efficiency and ensure evidence-based practice.

Actual and Predicted Reimbursement for the Sample Hospital

	Sample Hospital		Ratio
Risk Score Band	**Actual Total Payment**	**Predicted Total Payment**	**Predicted/ Actual**
1 (≤ 0.9)	$4,417,485	$3,875,855	88%
2 (0.9-1.2)	$10,143,202	$9,377,372	92%
3 (>1.2)	$7,649,011	$6,583,457	86%
ALL	$22,209,698	$19,836,684	89%

Example 2: Bundled Payments

Risk Stratification Model Application

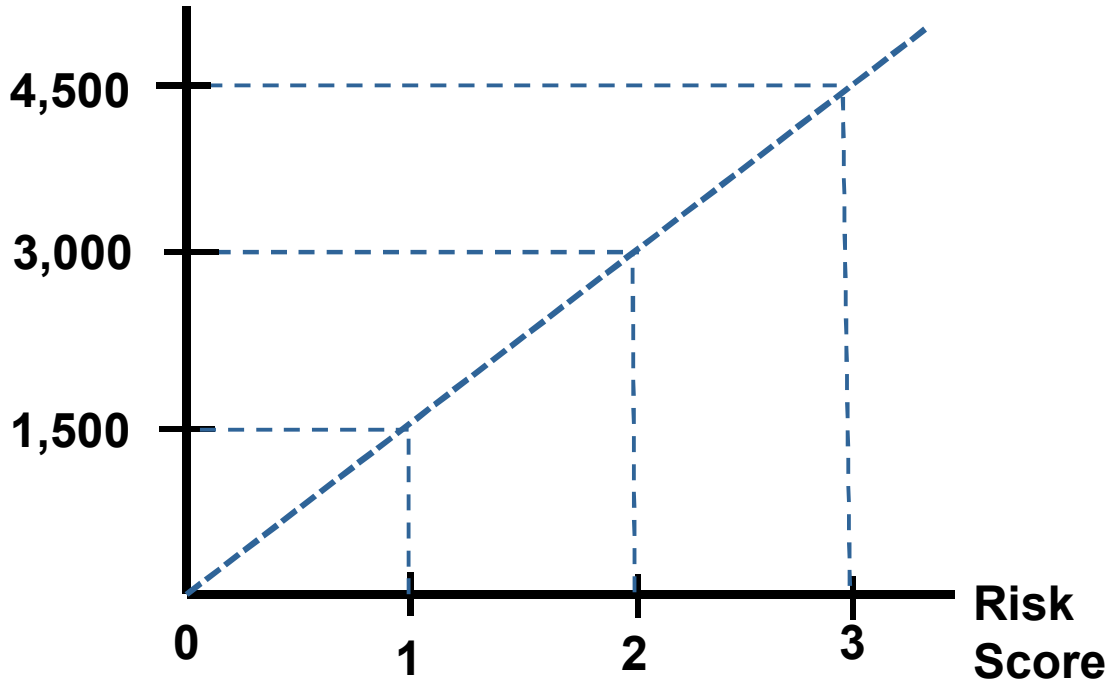
The bundle in the Sample Hospital appears to be composed of higher-cost providers. Although providers will determine the price they bid for the entire bundle, this number must fall below the payer's actual average payment to be attractive. In this example there is an opportunity for the hospital to negotiate more cost-effect pre- and post-surgery arrangements.

Descriptions	ALL	Sample Hospital
Number of Episodes	**3,457**	**502**
Payment per Episode	$ 38,453	$ 44,242
Avg. Pre Surgery Payment	$ 6,203	$ 8,352
Avg. Peri-Surgery Payment	$ 28,338	$ 31,165
Avg. Post-Surgery Payment	$ 3,912	$ 4,726

Example 3: Healthcare Reform and Exchanges

Underlying Risk Adjustment Hypothesis

The hypothesis underlying the transfer of funds is that Risk and Cost are correlated. On the next slide we see how Risk Adjustment is supposed to work.



Underlying Risk Adjustment Hypothesis

Scenario 1: Baseline Case – Uniform Cost/Risk

	State	Plan A	Plan B
Actuarial value	0.700	0.700	0.700
Risk Score	1.000	0.918	1.082
Rating Factor	1.952	1.952	1.952
Induced Demand	1.020	1.020	1.020
Geographic cost factor	1.000	1.000	1.000
Premium	$ 489.82	$ 489.82	$ 489.82
Risk x Ind. Demand x Geog.	1.020	0.936	1.104
Normalized ( )	1.000	0.918	1.082
Normalized ( )	1.000	1.000	1.000
	$ 000		
	-	$ (40.17)	$ 40.17

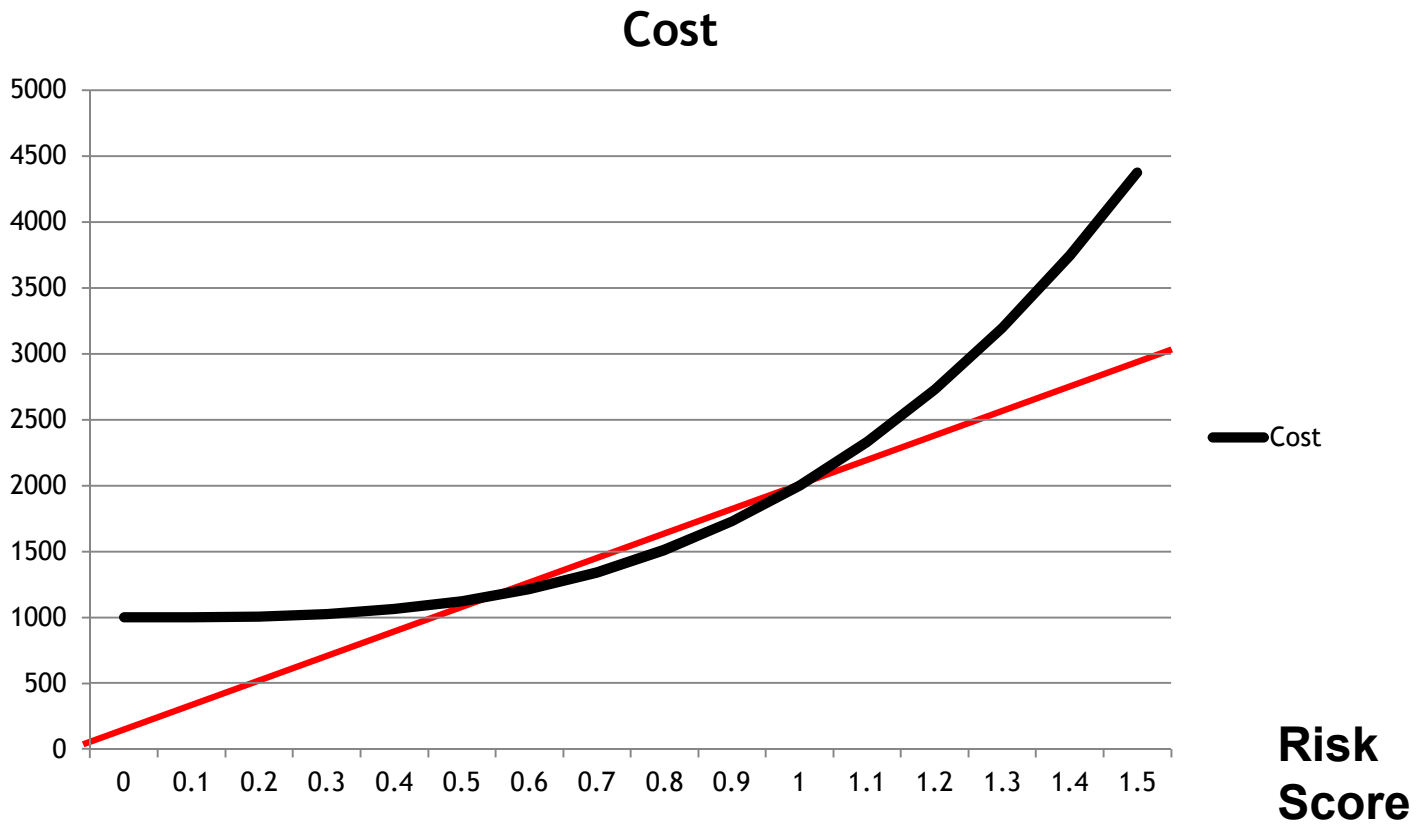
Plan Financials	Before Transfer	After Transfer	After Transfer
Members	2,000	1,000	1,000
Premium*	$ 11,755,680	$ 5,877,840	$ 5,877,840
Claims ($489.82 * Risk Factor)	$ 11,755,680	$ 5,395,857	$ 6,359,823
Gain	$ 000	$ 481,983	$ (481,983)
Funds Transfer	$ 000	(481,983)	481,983
Net income	$ 000	$ -	$ -

* assumed net for simplicity

Underlying Risk Adjustment Hypothesis

What if Cost and Risk are not perfectly correlated?

What if, instead, the relationship is non-linear?



Underlying Risk Adjustment Hypothesis

Scenario 2: Non-linear cost/risk relationship			
	State	Plan A	Plan B
Actuarial value	0.700	0.700	0.700
Risk Score	1.000	0.950	1.050
Rating Factor	1.952	1.952	1.952
Induced Demand	1.020	1.020	1.020
Geographic cost factor	1.000	1.000	1.000
Premium	$ 489.82	$ 489.82	$ 489.82
Risk x Ind. Demand x Geog.	1.020	0.969	1.071
Normalized ( )	1.000	0.950	1.050
Normalized ( )	1.000	1.000	1.000
	$ -	$ (24.49)	$ 24.49
Plan Financials	Before Transfer	After Transfer	After Transfer
Members	2,000	1,000	1,000
Premium*	$ 11,755,680	$ 5,877,840	$ 5,877,840
Claims ($489.82 * Adjustment)	$ 11,799,764	$ 5,458,677	$ 6,341,087
Gain	$ (44,084)	$ 419,163	$ (463,247)
Funds Transfer		(293,892)	293,892
Net income	$ (44,084)	$ 125,271	$ (169,355)
		2.1%	-2.9%
* assumed net for simplicity			

Analysis of Commercial Dataset

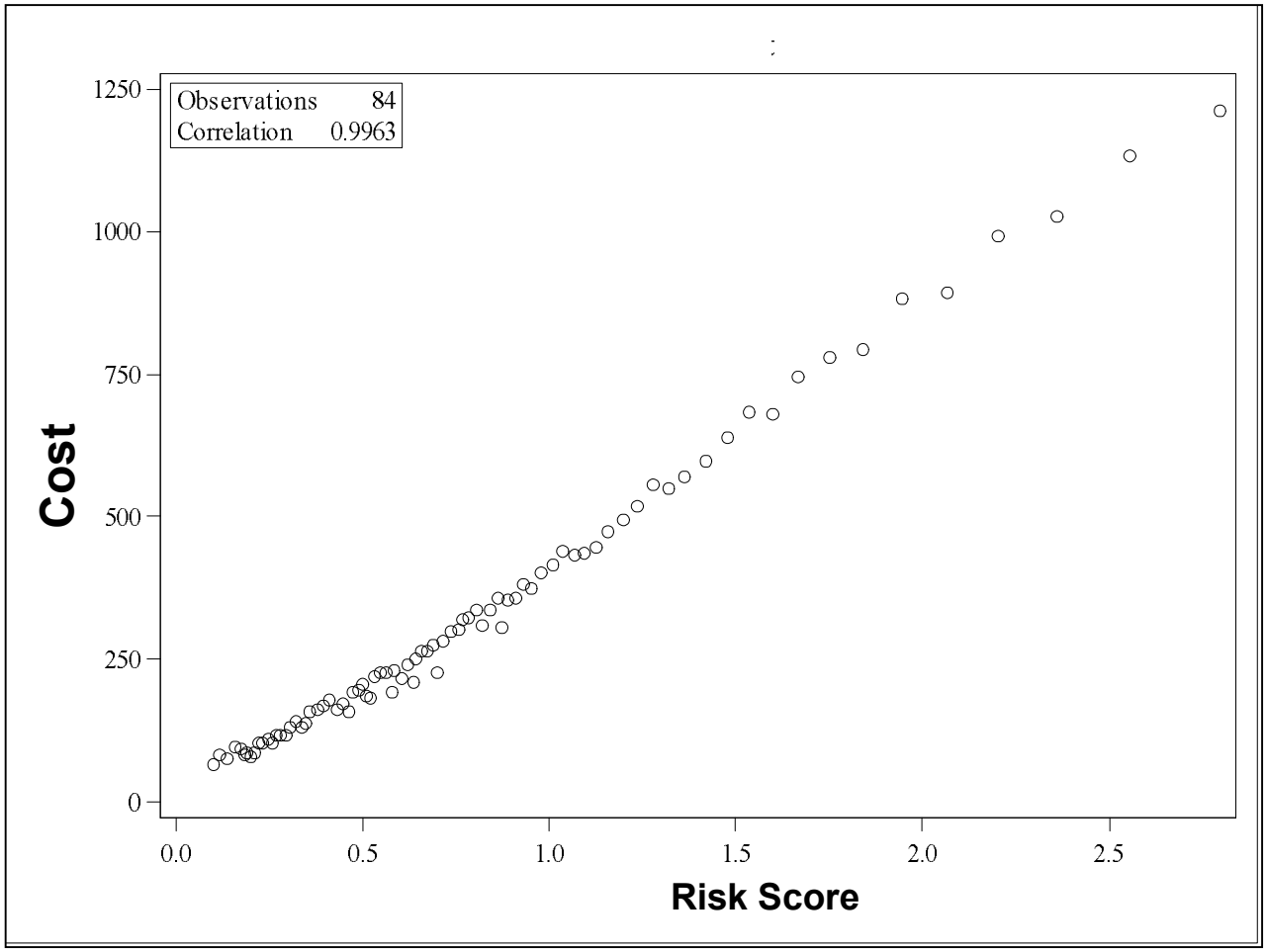
Leveraged SOA's latest Risk Adjustment study.

Risk Scores of one vendor, relative to cost.

Joint work with Prof. Sreenu Konda, Univ. of Illinois Chicago

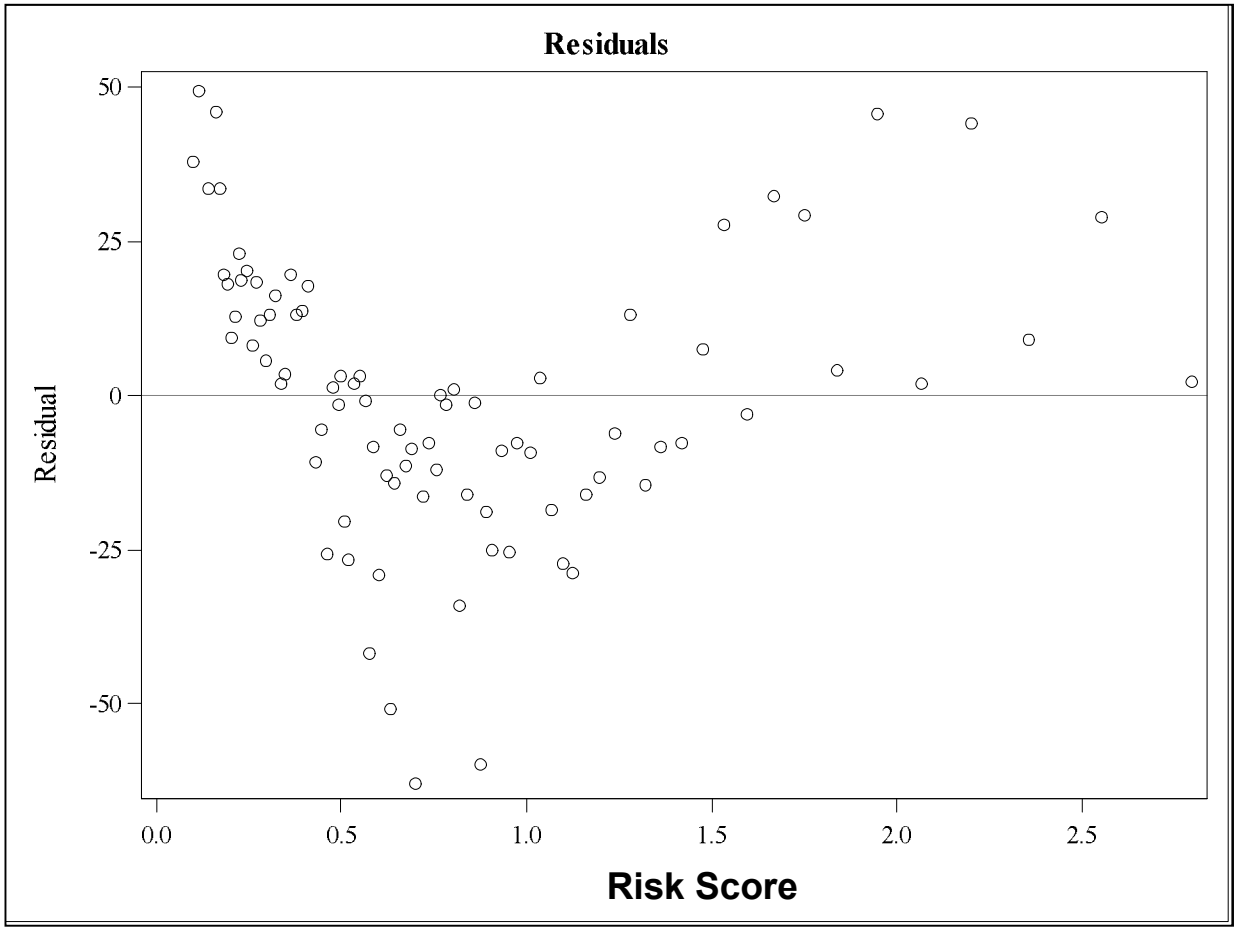
Analysis of Commercial Dataset

Superficially, Cost and Risk are highly correlated



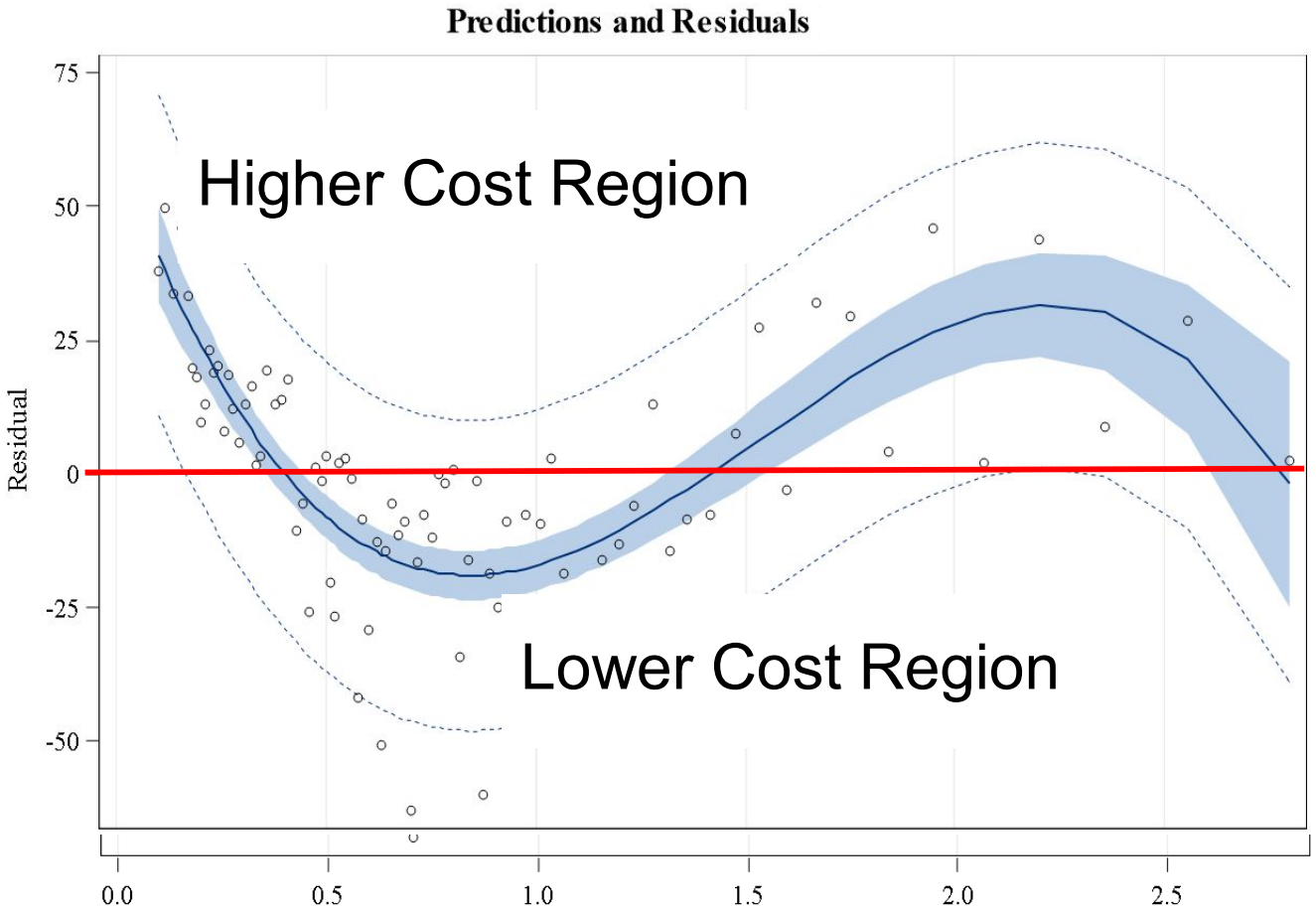
Analysis of Commercial Dataset

Analysis of Residuals shows a different result



Analysis of Commercial Dataset

Best Fit Cubic Model



Analysis of Commercial Dataset

It is possible that the distortion effect could occur, and health plans could be under- or over-compensated, depending on a plan's relative risk score.

Example 4: Predicting Hospital Re-admissions

Example 4: Predicting Hospital Re-admissions

Purpose

- 2012- Centers for Medicare and Medicaid Services reduced Medicare payments for hospitals with excess readmissions.
- Conditions: Heart Attack, Heart Failure, Pneumonia, Hip/Knee Replacement, Chronic Obstructive Pulmonary Disease.

What are excess readmissions?

- Admissions 30 days after discharge.
- Patients with select conditions (heart attack, heart failure, pneumonia, knee/hip replacement, COPD).

Calculated by a ratio:

$$\frac{\textit{No. Predicted Readmissions}}{\textit{No. Expected Readmissions}}$$

https://www.medicare.gov/hospitalcompare/readmission-reduction-program.html

Example 4: Predicting Hospital Re-admissions

Objective

- Create models that can be used to predict the risk of readmission
- Understand causes of readmissions
- Vulnerable Groups
  - Are there specific age, race or gender groups that are at a higher risk of being admitted?

Example 4: Predicting Hospital Re-admissions

Most Common Approach: LACE

- L: Length of stay
- A: Acuity (ER vs. elective)
- C: Co-morbidity (Charlson)
- E: # ED Visits

Table 1: Components of the LACE index.

Variable	Value	Points
Length of stay, days	<1	0
	1	1
	2	2
	3	3
	4–6	4
	7–13	5
	≥14	7
Acute (emergent) admission	Yes	3
Charlson comorbidity index score	0	0
	1	1
	2	2
	3	3
	≥4	5
Emergency department visits during previous 6 months	0	0
	1	1
	2	2
	3	3
	≥4	4

The LACE score is calculated by summing the points of the above 4 variables.

Example 4: Predicting Hospital Re-admissions

Why is LACE popular?

- LACE ranges 0-19:
  - Low risk: 0-4
  - Moderate risk: 5-9
  - High risk: >=10
- Predict early death and urgent readmission.
- Paper tool, uses existing resources.
- Easy to use in daily workflow.

Example 4: Predicting Hospital Re-admissions

Why should hospitals not rely solely on LACE?

- Assumption: valid to use on different hospitals' populations.
- NOT clinical data.
- Accuracy of the score (c-statistic) is .72
- Does not account for specific information on the patients (e.g.: race, age, sex…)

"Until the LACE index is externally validated with primary data, we recommend that it be used for outcomes research and quality assurance rather than in decision-making for individual patients."

Van Walraven C, Dhalla IA, Bell C, et al. *Derivation and Validation of an Index to Predict Early Death or Unplanned Readmission After Discharge From Hospital to the Community.* CMAJ 2010; 182: 551-557.

Example 4: Predicting Hospital Re-admissions

Data Summary for Total Patients

Variable	Type	Summary
Race	Category	White: 75%
		Hispanic: 19%
		Asian: 2%
		Black: 2%
		Native American, Hawaiian/Pac. Island, Other & Unknown: 2%
DRG Class	Category	DRG Medical: 50%
		DRG Surgical: 43%
		DRG Ungroup: 7%
Gender	Category	Female: 59%
		Male: 41%

Example 4: Predicting Hospital Re-admissions

Variable	Type	Summary
Admit From Type	Category	Emergency: 43%
		Pre Admit: 36%
		Observation: 15%
		Pre Clinic, Clinic, and SDC & Other: 6%
Readmission	Category	No Readmission: 93%
		Readmission: 7%
Length of Stay (days)	Numeric	Min:0
		Median:3
		Mean: 4.04
		Max: 239
AGE	Numeric	Min: 15
		Mean: 58
		Max: 112

Example 4: Predicting Hospital Re-admissions

Variable	Type	Summary
ED Visits in 2010	Numeric	Min: 0
		Mean: 0.16
		Max: 43
ED Visits in 2011	Numeric	Min: 0
		Mean: 0.16
		Max: 41
ED Visits in 2012	Numeric	Min: 0
		Mean: 0.18
		Max: 38
ED Visits in 2013	Numeric	Min:0
		Mean: 0.18
		Max:38
ED Visits in 2014	Numeric	Min: 0
		Mean: 0.18
		Max:38

Example 4: Predicting Hospital Re-admissions

- **CDPS Risk Score**
  - Diagnostic-based risk model that uses ICD-9 codes to assess risk.
  - Provides a summary measure of the burden of illness.
- **LACE Index**
  - Length of stay
  - Number of Charlson comorbidity index conditions.
  - Acuity of admission.
  - Number of ED visits in previous 6 months.

Variable	Type	Summary
CDPS Risk Score	Numeric	Min: 0.14
		Mean: 3.24
		Max: 29.85
LACE Index	Numeric	Min: 1
		Mean: 5.87
		Max: 19

Example 4: Predicting Hospital Re-admissions

Logistic Regression

- Model the probability of an event occurring depending on the values of the independent variables.
- Estimate the probability that an event occurs for a random selected observation versus the probability that event does not occur.
- Odds =Pr(Occurring)/Pr(Not Occurring)
- Odds Ratio: a ratio between two odds
- Odds ratio for a variable represents how the odds change with a 1 unit increase in that variable holding all other variables constant.

Example 4: Predicting Hospital Re-admissions

Summary of the model

Variable	Coefficient	Odds Ratio	95% Confidence Interval
Intercept	-3.130	0.044	(.032,0.059)
Sex Male (vs. Female)	0.079	1.072	(1.004,1.145)
Race Black (vs. Asian)	0.198	1.219	(0.885,1.686)
Race Hispanic (vs. Asian)	0.299	1.348	(1.054,1.749)
Race White (vs. Asian)	0.101	1.106	(0.873,1.423)
Race Other (vs. Asian)	-0.410	0.664	(0.409,1.047)

Example 4: Predicting Hospital Re-admissions

Summary of the model (contd.)

Variables	Coefficient	Odds Ratio	95% Confidence Interval
Intercept	-3.130	0.044	(.032,0.059)
Admission From ED (vs. No Admission From ED)	0.420	1.522	(1.403,1.653)
DRG Surgical (vs. DRG Medical)	-0.761	0.467	(0.429,0.508)
DRG Ungroup (vs. DRG Medical)	0.128	1.137	(1.021,1.263)
LACE Low (vs. LACE High)	-1.157	0.314	(0.270,0.365)
LACE Moderate (vs. LACE High)	-0.240	0.786	(0.723,0.855)

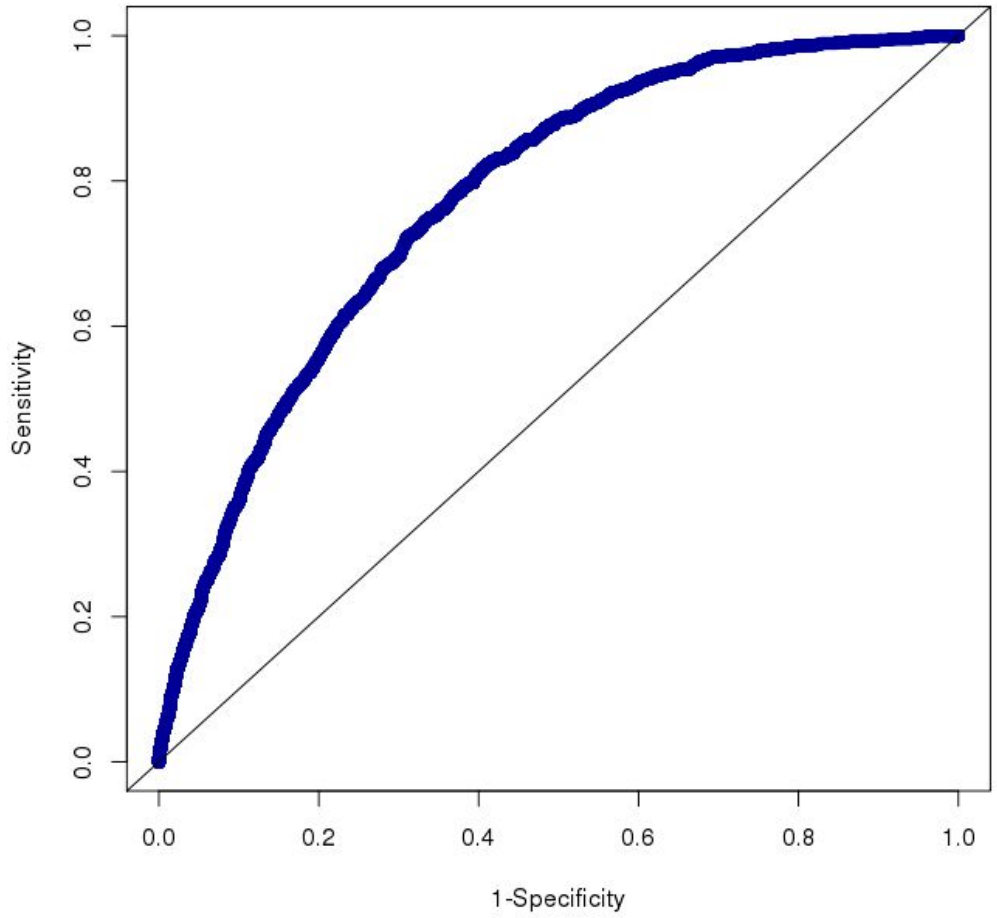
Example 4: Predicting Hospital Re-admissions

Summary of the model (contd.)

Variables	Coefficient	Odds Ratio	95% Confidence Interval
Intercept	-3.130	0.044	(.032,0.059)
Age (at 65)	0.003	1.208	(1.211,1.211)
CDPS Risk Score	0.101	1.107	(1.096,1.118)
Length of Stay	0.014	1.014	(1.009,1.019)
ED visits in 2010	0.069	1.072	(1.050,1.093)
ED visits in 2011	0.093	1.098	(1.073,1.123)
ED visits in 2012	0.106	1.112	(1.090,1.135)
ED visits in 2013	0.081	1.085	(1.061,1.108)
ED visits in 2014	0.075	1.078	(1.057,1.100)

Example 4: Predicting Hospital Re-admissions

ROC Curve



Example 4: Predicting Hospital Re-admissions

Financial Modeling (example)

	ALL PATIENTS				MEDICARE PATIENTS			
Decile	Number in decile	Mean Prediction within Quantile	Actual Readmissions	Predicted Readmissions	Number in decile	Mean Prediction within Quantile	Actual Readmissions	Predicted Readmissions
0-10	1,611	0.0092	8.0	14.7	666	0.0092	6.0	6.1
10-20	1,611	0.0112	11.0	18.1	666	0.0114	4.0	7.6
20-30	1,611	0.0177	20.0	28.5	666	0.0185	15.0	12.3
30-40	1,611	0.0248	48.0	39.9	666	0.0255	15.0	17.0
40-50	1,611	0.0359	68.0	57.8	666	0.0364	22.0	24.2
50-60	1,611	0.0569	94.0	91.7	666	0.0568	48.0	37.8
60-70	1,611	0.0822	130.0	132.5	666	0.0821	60.0	54.7
70-80	1,611	0.1025	164.0	165.1	666	0.1032	61.0	68.7
80-90	1,611	0.1339	230.0	215.7	666	0.1366	98.0	91.0
90-100	1,611	0.2351	360.0	378.7	666	0.2319	157.0	154.4
	16,110		1,133.0	1,142.8	6,660		486.0	473.9

Example 4: Predicting Hospital Re-admissions

Financial Modeling (example)

Population Decile	No. Patients	Expected Re-admits	90% Enrolled	Cost/ mgd	15% Avoided Re-admit	20% Avoided Re-admit	15% Revenue Reduction	20% Revenue Reduction	15% Penalty Reduction	20% Penalty Reduction
1	700	140	630	$ 315,000	21	28	$ 252,000	$ 336,000		
2	700	105	630	$ 315,000	15.75	21	$ 189,000	$ 252,000		
3	700	84	630	$ 315,000	12.6	16.8	$ 151,200	$ 201,600		
4	700	63	630	$ 315,000	9.45	12.6	$ 113,400	$ 151,200		
5	700	63	630	$ 315,000	9.45	12.6	$ 113,400	$ 151,200		
6	700	56	630	$ 315,000	8.4	11.2	$ 100,800	$ 134,400		
7	700	52.5	630	$ 315,000	7.875	10.5	$ 94,500	$ 126,000		
8	700	49	630	$ 315,000	7.35	9.8	$ 88,200	$ 117,600		
9	700	45.5	630	$ 315,000	6.825	9.1	$ 81,900	$ 109,200		
10	700	35	630	$ 315,000	5.25	7	$ 63,000	$ 84,000		
	7,000	693	6,300	$3,150,000	104	139	$ 1,247,400	$ 1,663,200		

Example 5: Case Finding and Opportunity Analysis

Example 5: Case Finding and Opportunity Analysis

A recent article about the application of predictive models to the "Triple Aim:"

- Lower cost
- Better Quality Care
- Higher patient satisfaction.

QUALITY & GOVERNANCE

By Gerald Lewis, Heather Kirkham, Ian Duncan, and Rhenu Vaithianathan

How Health Systems Could Avert 'Triple Fail' Events That Are Harmful, Are Costly, And Result In Poor Patient Satisfaction

ABSTRACT Health care systems in many countries are using the "Triple Aim"—to improve patients' experience of care, to advance population health, and to lower per capita costs—as a focus for improving quality. Population strategies for addressing the Triple Aim are becoming increasingly prevalent in developed countries, but ultimately success will also require targeting specific subgroups and individuals. Certain events, which we call "Triple Fail" events, constitute a simultaneous failure to meet all three Triple Aim goals. The risk of experiencing different Triple Fail events varies widely across people. We argue that by stratifying populations according to each person's risk and anticipated response to an intervention, health systems could more effectively target different preventive interventions at particular risk strata. In this article we describe how such an approach could be planned and operationalized. Policy makers should consider using this stratified approach to reduce the incidence of Triple Fail events, thereby improving outcomes, enhancing patient experience, and lowering costs.

The "Triple Aim" of health care is to improve individual patients' experiences of care, advance population health, and reduce per capita health care costs.[1] A central tenet of the Triple Aim is to restructure care in ways that lead to improvements across all three of these goals.

The Institute for Healthcare Improvement has worked with organizations in many countries to implement populationwide interventions to foster the Triple Aim.[2] Examples are programs that encourage self-management of chronic conditions,[3] promote e-mail communication between patients and physicians,[4] and encourage greater use of primary care.[5]

Other organizations have adopted a more targeted approach to achieving the Triple Aim. For example, a Commonwealth Fund case study found examples of organizations that were focusing on improving access and care for individual patients who had low incomes, were uninsured, or had complex chronic conditions.[6] Indeed, several authors have argued that success will require both population health and individually focused strategies, such as those employed by Geisinger Health System in Flint, Michigan.[7,8,9] For example, Geisinger increased its primary care capacity (that is, a population approach) and offered health navigators to its high-risk patients (that is, a targeted approach).

The objectives of this article are to propose a third, stratified approach to tackling the Triple Aim and to explore some of the ethical challenges that this new approach presents. The stratified approach to the Triple Aim involves identifying and prioritizing subpopulations according to their risk of experiencing health encounter failure—what we call "Triple Fail" events—and according to their likelihood of benefiting from preventive care.[10]

We define a *Triple Fail event* as a health outcome

Example 5: Case Finding and Opportunity Analysis

Traditionally, predictive modeling for case-finding has aimed at identifying high-risk individuals. This is frequently done by applying a commercial risk-adjuster model and intervening on the highest-risk score population. The following procedures are common:

Model 1: Run a predictive model and stratify members according to their predictive risk score. One potential draw-back of this approach is the high prevalence at the top of that list of members who (although high risk) are minimally intervenable. Even if an intervenability algorithm is applied to the entire population, the resulting list will consist of a mix of members with different conditions, issues and needs.

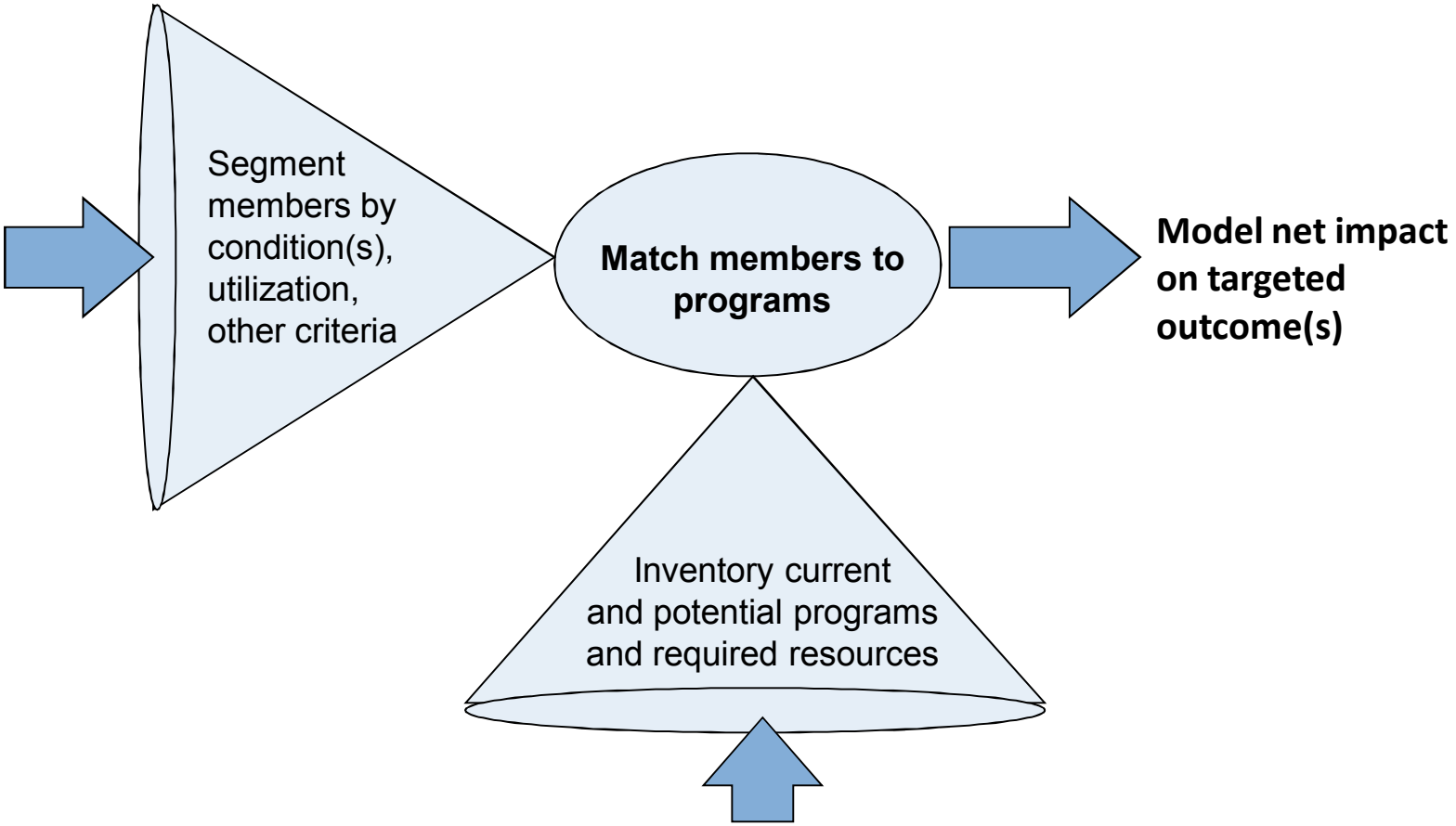
Model 2: Model 2 is a condition-specific model. For simplicity, Program planners frequently want to focus on members with a specific condition, say diabetes. This approach has the benefit of addressing the member heterogeneity inherent in Model 1. However, the high prevalence of co-morbidities in the high-risk population requires that any program targeted at a condition population will ultimately have to be sufficiently broad to address all conditions of the population. Moreover, it is often the interplay between comorbidities that drives the complexity and its associated costs, so a focus on one disease may well miss this greater opportunity.

Model 3: a rules-based approach is often used in case management programs. In this model, program managers determine a set of rules to identify target patients for management. Sometimes the rules are condition-specific; sometimes a financial threshold is used (for example, $50,000). Depending on how the rules are determined, high-opportunity members may or may not be targeted for intervention.

Example 5: Case Finding and Opportunity Analysis

- Opportunity Analysis is designed to address a number of the shortcomings associated with the typical models. It maintains the stratification of Model 1 but adds the element of intervenability by assigning lower priority to those patients with conditions that are less amenable to an intervention program.
- As a general rule, Opportunity Analysis avoids disease-specific programs in favor of programs that target members with common risk profiles (for example, all chronic members or all members at end of life) although there may be a few notable exceptions (for example end-stage kidney disease or some specific preference-sensitive conditions).
- Opportunity Analysis requires research and understanding of the targeting, operation and outcomes of programs that have been implemented in similar populations.
- Finally, Opportunity Analysis takes into consideration the economics of programs: the cost of the intervention vs. the expected reduction in utilization that each individual can be expected to contribute.

Example 5: Case Finding and Opportunity Analysis



Example 5: Case Finding and Opportunity Analysis

A summarization that we have found helpful is to group patients not be risk or condition but by class of condition; this helps with relative intervenability and in the design of programs. Note the disproportionate cost of the "Episodic, Mental Health and Chronic" group, whose costs are about three and one-half times their numbers. Almost half of the total cost of the population is concentrated in this, the most complex segment of members. Unfortunately, this is also the most difficult segment to design programs for or to manage. But any population program that avoids addressing the needs of the most complex patients will be doomed to financial failure.

Condition Category	Population %	Cost %
Episodic, Mental Health and Chronic	12%	42%
Episodic and Mental Health	4%	7%
Episodic and Chronic	15%	27%
Episodic only	9%	7%
Mental Health and Chronic	5%	5%
Chronic only	14%	6%
Mental Health only	5%	2%
Emerging Conditions	12%	3%
None	24%	1%
	100%	100%

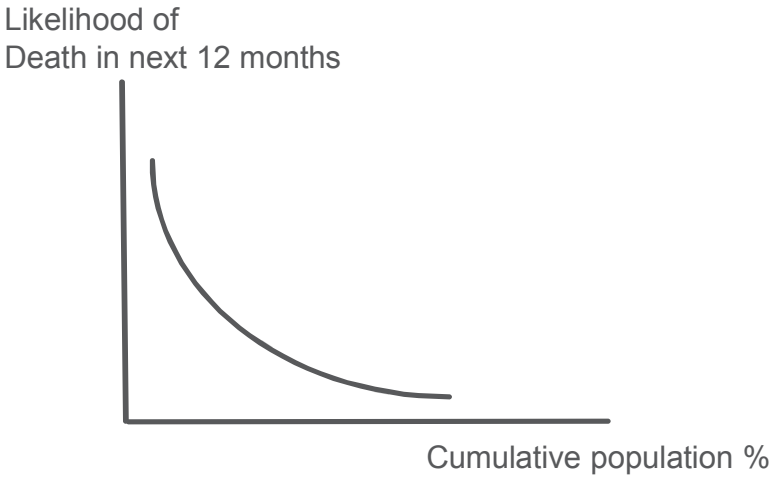
Example 5: Case Finding and Opportunity Analysis

Considerable analytics can be performed to try to identify classes of patients who represent both high cost and an opportunity to change the outcome. One such class consists of patients within 6 to 12 months of end of life. This class consumes considerable resources, without much affecting the ultimate outcome.

Service Category	Complex Population	End-of-life Sub-Population
Inpatient Admissions	1,000 per 1000	2,500 per 1000
IP Admissions Medical	750 per 1000	2,150 per 1000
IP Admissions Surgical	250 per 1000	350 per 1000
IP 30 day readmissions	18%	25%
30-day Readmits Medical	20%	25%
30-day Readmits Surgical	12%	20%
Emergency Room visits	1,200 per 1000	200 per 1000
Specialist visits	6,500 per 1000	5,000 per 1000
MRI services	400 per 1000	500 per 1000
Primary care visits	4,500 per 1000	3,000 per 1000

Example 5: Case Finding and Opportunity Analysis

Fortunately, there is plenty of peer-reviewed literature about, and good outcomes from programs that aim to manage patients at the end of life. We developed a program and constructed a predictive model to predict those patients at risk of death in the next 6 to 12 months. Not all patients, however, are targets for the program. We combine the predictive model and program with an economic model, ranking members by likelihood death.



Example 5: Case Finding and Opportunity Analysis

Combining the predicted probability of the event, the likelihood of the member engaging, and the expected outcome (earlier transfer to hospice, for example) with the cost of intervention, we are able to determine which members represent an economic return for a particular program.

Depth of Dive Target

ROI Target

Member #	Dive %	Intervention Cost	Net-Net Opportunity	Cum Savings	Cum ROI
001	0.1%	$500	$5,000	5,000	10.0
002	0.2%	$500	$4,800	9,800	9.8
003	0.3%	$500	$4,300	14,100	9.4
…	….	$500		100,000	2.0
		$500	$0		
		$500	-$100	…	…
		$500	-$500		
1000	100%	…	…		

Max. Absolute Savings

Selected references

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

- Duncan, I. Healthcare Risk Adjustment and Predictive Modeling. 2011 Actex Publication.
- 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.
- Duncan I, Loginov M and Ludkovski M. *Testing Alternative Regression Frameworks for Predictive Modeling of Health Care Costs.* North American Actuarial Journal January 2016. DOI: 10.1080/10920277.2015.111491.

Selected references (contd.)

- Winkelman R. and S. Ahmed. *A comparative analysis of Claims Based Methods of health risk assessment of 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
- Duncan I: *Managing and Evaluating Healthcare Intervention Programs* (2013). Actex Publications. www.actexmadriver.com
- Lewis, GL, Duncan I, Kirkham H, Vaithianiathan R: *How Health Systems Could Avert "Triple Fail" Events that are Harmful, Are Costly and Result in Poor Patient Satisfaction."* Health Affairs 32 No. 4 (2013) 669-676.
- Duncan I: "*Mining Health Claims Data for Assessing Patient Risk*" In *Data Mining: Foundations and Intelligent Paradigms* Volume 3: *Medical, Health, Social, Biological and other Applications.* D.E. Holmes and L. Jain (eds.), Springer Verlag 2012.

Ian Duncan FSA FIA FCIA MAAA
Adj. Assoc. Professor
Dept. of Statistics & Applied Probability
University of California, Santa Barbara

(860) 614-3295
duncan@pstat.ucsb.edu
iduncan@sbactuaries.com