RAND HEALTH

From Cost Modeling to Action

John L. Adams, Ph.D.

October 25, 2010

Overview

- **Intro to common cost models**
- **How can we turn models of costs into an understanding of “mutable” costs?**
- **How can cost modeling contribute to cost reduction?**

Costs are hard to model

- **Cost distributions are skewed (more high numbers than low numbers)**
- **Variances are pretty big compared to averages**
- **The available predictors don’t explain a large fraction of what is going on**

There are two main approaches to modeling person level costs from administrative data

- **Predict a patient's costs for the year**
  - **Mostly claims based descriptions of the patient's health status**
  - **Example: The DCGs (DxCG)**
- **Develop episodes of care and predict costs per episode**
  - **Claims are aggregated into episodes**
  - **Example: the ETGs (Symmetry)**

Predicting annual patient costs

- **Originally developed more for capitation calculations, not necessarily for case identification**
- **Most of these tools are big regression models***
- **Much of the art in developing these models is building a sensible set of claims based predictors**
- **Recently Rx data has been incorporated as well**
- **Several decisions are required:**
  - **Prospective or concurrent?**
  - **Use lagged costs as a predictor?**

***Winkelman R, Mehmud S (2007) "A Comparative Analysis of Claims-Based Tools for Health Risk Assessment," Society of Actuaries, April 20, 2007.**

Episodes of care

- The major player is the ETGs from Symmetry
- Three episode types
  - Acute
  - Chronic
  - Preventative
- The episodes a patient triggers can be used as predictors to build concurrent or prospective models

Tuning these models to your situation

- **You can/should retune these models to your population**
  - **For person-year models just rescale:**
    - **Your Population = A + B*Score**
  - **For episodes just use your own means**
- **You can use these tools to predict other things**
  - **Hospitalizations**
  - **ER use**
  - **Pharma costs**

But predictable costs are not necessarily mutable costs!*

- **The simplest example is age. Age is a useful predictor but there isn't much you can do about it.**
- **You might even improve your cost predictions by building good predictors of end of life that do not suggest actions that reduce costs**
- **Knowing who is likely to be expensive is a good thing but it is nowhere near enough**
- **You need to map the cases expected to be expensive onto actions**

***Linden A, Adams JL. Improving participant selection in disease management programmes: insights gained from propensity score stratification. *J Eval Clin Pract*. 2008;14(5):914-918.**

Cost saving strategies and cost prediction

- Case ID for intervention (e.g. disease management)
  - Find the cases predicted to be expensive
- MD profiling
  - Risk adjust the MD's cost profile for patient factors
- Program evaluation
  - Adjust for differences between those in the program and those in the comparison group
- Pay for performance
  - Adjust costs for expected costs
- Understanding cost drivers
  - See which predictors in the cost models have the biggest effects

Example: Identifying lower cost physicians

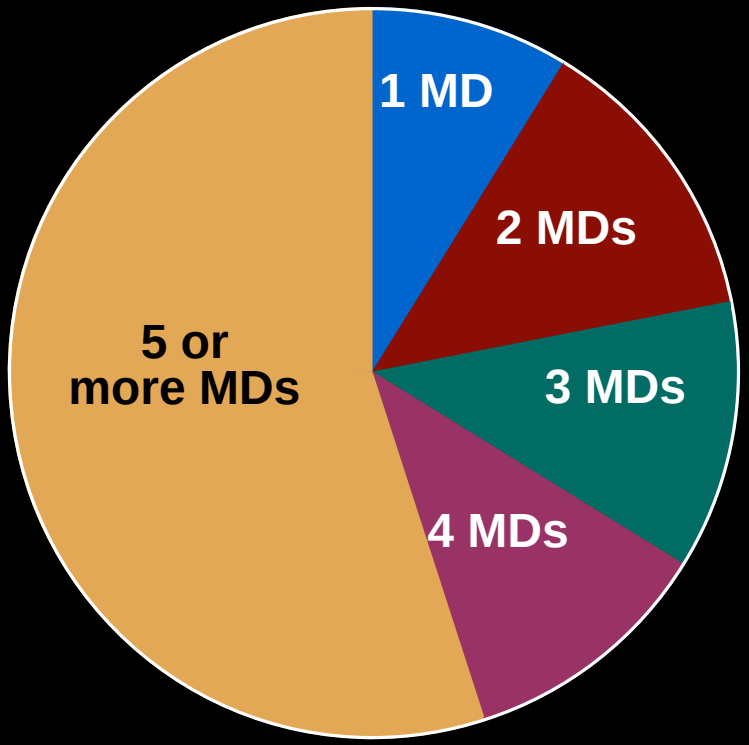
- **Identifying lower cost physicians could support several possible actions:**
  - **Bonus payments to encourage the behavior**
  - **Steering patients to lower cost MDs**
    - **Lower copays**
    - **Special recognition**
  - **Targeting higher cost MDs for feedback or other MD level interventions**
- **Episode systems are a good fit for this problem**

There are a lot of details to work out

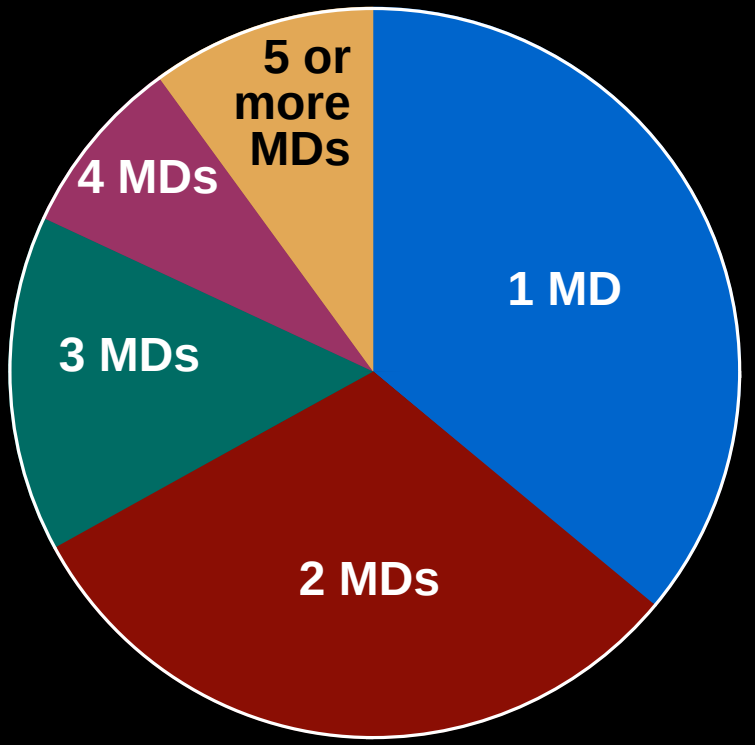
Attribution	**Which physician is assigned responsibility for which costs?**
Metric	**How is the metric constructed?**
Classification	**How are physicians assigned to categories of performance?**

Why is attribution important?

Patient



Episode



Which physician is responsible for care?

Who is responsible?

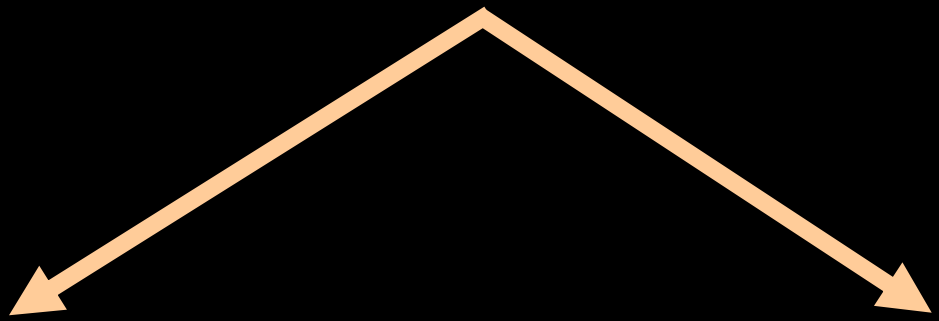
Pneumonia

Preventive care

Hysterectomy

Mrs. Smith

Signal for Assignment?



Costs

Plurality
Majority

Visits

Plurality
Majority

Policy Implications

- Choice of attribution rules will affect categorization
- Which attribution rule to use?
  - Unfortunately no single “right” approach
- Perspective matters
  - Health plans want to include as many MDs as possible
  - Physician wants rule to reflect his or her care
- Unintended consequences also matter
  - Refuse to see certain patients?
  - Withhold care?

More details

Attribution	**Which physician is assigned responsibility for which costs?**
Metric	**How is the metric constructed?**
Classification	**How are physicians assigned to categories of performance?**

Differences or proportions?

- **Use the cost prediction model as the "expected" costs**
  - **What would my patients' have cost at other providers?**
- **Ratios:**
  - **Cost Profile = Observed/Expected**
- **Differences:**
  - **Cost Profile = Observed-Expected**

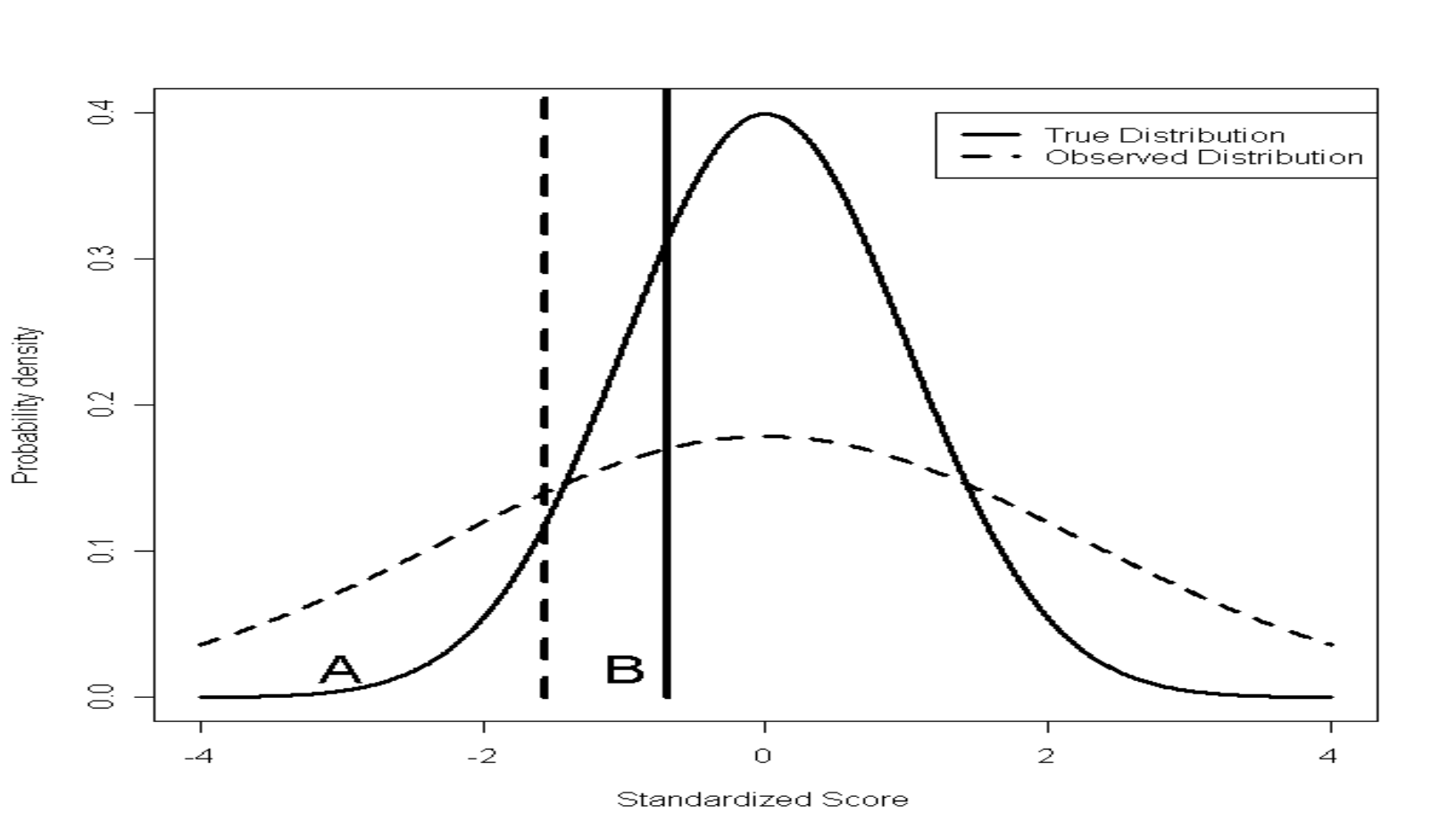
Even more details…

Attribution	**Which physician is assigned responsibility for which costs?**
Metric	**How is the metric constructed?**
Classification	**How are physicians assigned to categories of performance?**

Method 1- Empirical Cut Points

- **Pick percentiles of the observed distribution and put physicians into bins**
  - **E.g. Bottom (lowest cost) 25% of MD "high performing"**
- **Attractive because:**
  - **It is easy**
  - **It is "grading on the curve"**
  - **You can directly set the size of your "high performance" network**

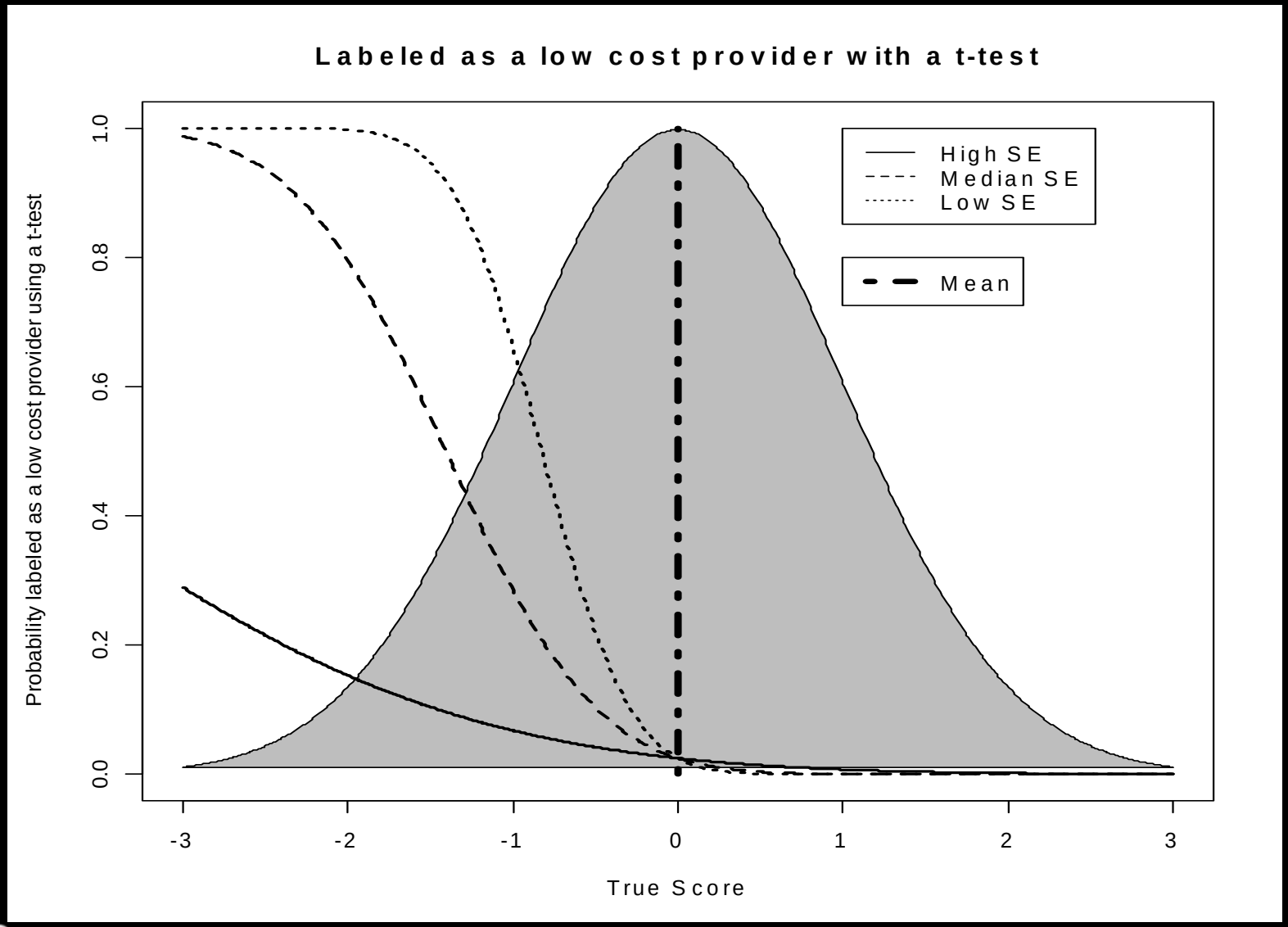
Noisy Cut Points Are A Problem



Method 2 – Statistical testing vs. the mean

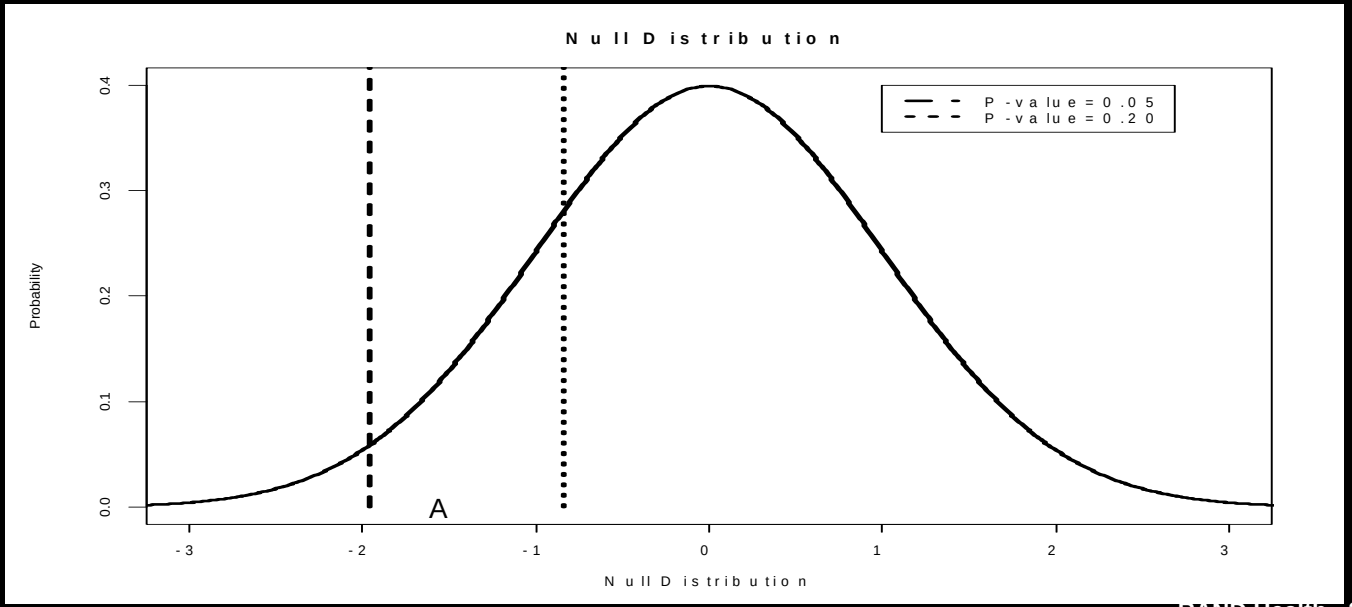
- First you need a standard error
  - $$SE(PROFILE) = \sqrt{\sum \text{var}(Observed)} \Big/ \sum Expected$$
  - Plug in population quantities
  - A "null hypothesis" style SE
- Then you test against the mean:
  - $$t = \frac{PROFILE - mean(PROFILES)}{SE(PROFILE)}$$

Some comments about how tests work



Potential concern: Not enough outliers

- Some policy applications require a sufficient fraction of physicians to be labeled as high performing
  - Statistical testing – 12.9% are low cost
  - Cut-points – 25% are low cost
- If you need to increase the number of MDs, use a higher p-value



Comparing the two methods

- **Cut-points**
  - **Top 25% = high cost**
  - **Bottom 25% = low cost**
- **Statistical testing**
  - **Significantly higher than average (p<0.05) = high cost**
  - **Significantly lower than average (p=<0.05) = low cost**
- **Data used here is described in:**

Adams JL, Mehrotra A, Thomas JW, McGlynn EA. Physician Cost Profiling — Reliability and Risk of Misclassification. *N Engl J Med* 2010;362:1014-21.

30% of MDs are classified differently across the two methods

	T-test (p=0.05)		
Cut-Points	**Low Cost**	**Average Cost**	**High Cost**
Low Cost Bottom 25%	**11%**	**14%**	**0**
Average Cost	**2%**	**47%**	**1%**
High Cost Top 25%	**0**	**13%**	**12%**

Pros and Cons

- Empirical cut-points
  - Pros
    - Gives the big standard error (small sample size) providers a chance to be flagged as good (mostly by mistake)
    - Easy to explain
  - Cons
    - Can be very noisy
    - Lots of misclassification for small sample size providers
    - It may not help to be a low SE provider
- Statistical testing
  - Pros
    - Reduces the number of providers flagged as above/below average by chance
    - Conforms to typical medical evidence standards
  - Cons
    - May not fill a high performance network (without a relaxed statistical standard)
    - May be harder for some purchasers to implement
    - Harder for most people to understand

A warning: MDs may be incorrectly classified

1. **Unfair to low-cost physician**
   - **Low-cost physician is labeled as average**
2. **Undermines impact of profiling**
   - **Average physician is labeled as low cost**
3. **Unfair to patient in a high-deductible plan**
   - **Physician labeled low-cost is actually not**

High Rates of Misclassification Found Across Specialties

Specialty	Low cost MD labeled average	MD labeled low-cost is actually average	Overall Misclassification
Internal medicine	77%	50%	25%
Family practice	52%	39%	21%
OB-GYN	29%	36%	17%
Cardiology	37%	40%	20%
Psychiatry	61%	48%	24%

Improves One Type of Misclassification…

	MD labeled low cost is actually average	
Specialty	**Percentile cut-off**	**Statistical testing**
Internal medicine	**50%**	**5%**
Family practice	**39%**	**5%**
OB-GYN	**36%**	**4%**
Cardiology	**40%**	**6%**
Psychiatry	**48%**	**11%**

…At the Cost of Another Type

	Low cost MD is labeled as average	
Specialty	**Percentile cut-off**	**Statistical testing**
Internal medicine	**77%**	**81%**
Family practice	**52%**	**83%**
OB-GYN	**29%**	**80%**
Cardiology	**37%**	**85%**
Psychiatry	**61%**	**90%**

Policy Implications

- **Patients and MDs may not receive a useful signal from profiles given the rates of misclassification observed**
- **“Too much” misclassification depends on perspective**

Where To Go From Here?

- **Essential to find a way to improve quality and manage costs (as well as ensure access)**
- **There are probably no painless ways to do this**
- **Ideally these decisions will involve all stakeholders -- but getting something for nothing probably isn't one of the options**
  - **So, what is each group willing to give up to ensure value and sustainability in the health system?**
- **Transparent, participatory processes are critical for moving forward**

The Language of Misclassification

		TRUE		
		Low Cost (Positive)	Average Cost (Negative)	
OBSERVED	Low Cost (Positive)	True Positive (TP)	False Positive (FP)	Positive Predictive Value TP/(TP+FP)
	Average Cost (Negative)	False Negative (FN)	True Negative (TN)	Negative Predictive Value TN/(FN+TN)
		Sensitivity TP/ (TP+FN)	Specificity TN/(FP+TN)	