



The Johns Hopkins University's



2010 ACG International
Risk Adjustment Conference

MAY 10-12
Tucson, Arizona
Loews Ventana Canyon

ACG Predictive Modeling in Practice: Improving the Identification Process for Care Management Programs Using Cut Points

Shannon Murphy, Heather Castro & Martha Sylvia

May 12, 2010



The Johns Hopkins University's

2010 ACG International Risk Adjustment Conference

MAY 10-12
Tucson, Arizona
Loews Ventana Canyon

Outline

- Background
- Assessing Predictive Accuracy
- Determining the Optimal Cut Point
- Measuring Model Performance
- Example
- Summary and Implications



The Johns Hopkins University's



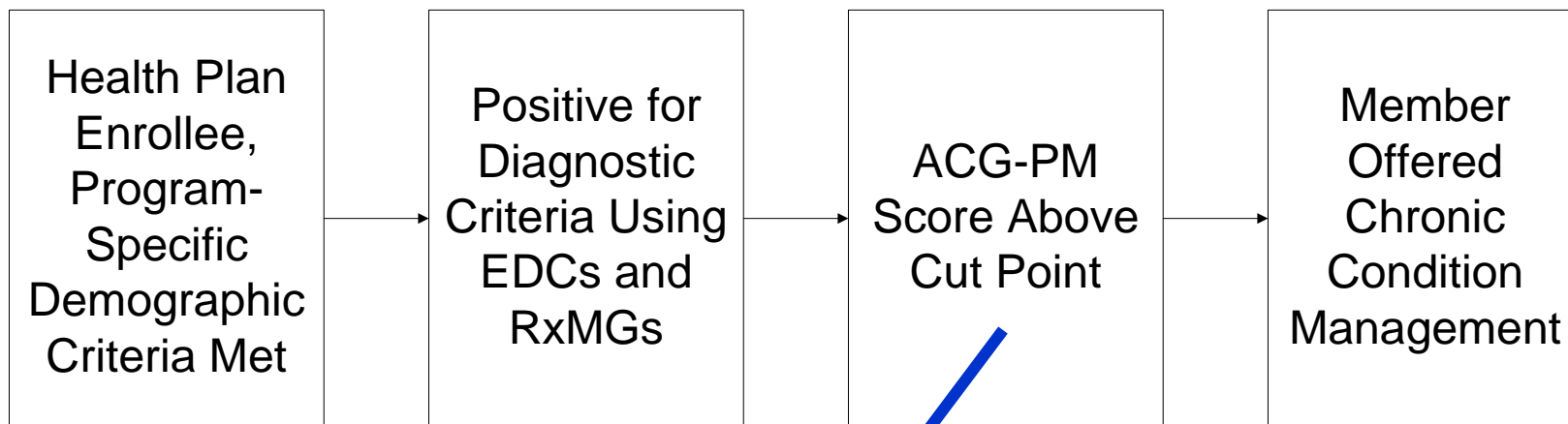
2010 ACG International
Risk Adjustment Conference

MAY 10-12
Tucson, Arizona
Loews Ventana Canyon

Background

- Direct payers of healthcare services (i.e., employers and health plans) most affected by rising healthcare costs
- Rising costs largely driven by individuals with chronic condition(s)
- Investment in population health initiatives to address needs of those with (and at-risk) for chronic illness
- Identification of appropriate members for population health initiatives is essential for program success

Background: Traditional Screening Process



- Often missing a large portion of high-cost members
- Can a different Cut Point selection process improve screening?

Assessing Predictive Accuracy: 2 x 2 Table

Prediction	Outcome		Total
	No	Yes	
N Row Pct Col Pct			
No	80 .94 .89	5 .06 .50	85 1.0 85
Yes	10 .67 .11	5 .33 .50	15 1.0 .15
Total	90 .90 1.0	10 .10 1.0	100

- For binary predictor and binary outcome
- Sensitivity: proportion of positives correctly predicted or true positive rate (tpr)
- Specificity: proportion of negatives correctly predicted or the true negative rate (tnr)
- False Positive Rate (fpr): proportion of negatives incorrectly predicted (1-Specificity)
- Positive Predictive Value (ppv): proportion of all positive predictions that were truly positive



The Johns Hopkins University's

 2010 ACG International
Risk Adjustment Conference

MAY 10-12 
Tucson, Arizona
Loews Ventana Canyon

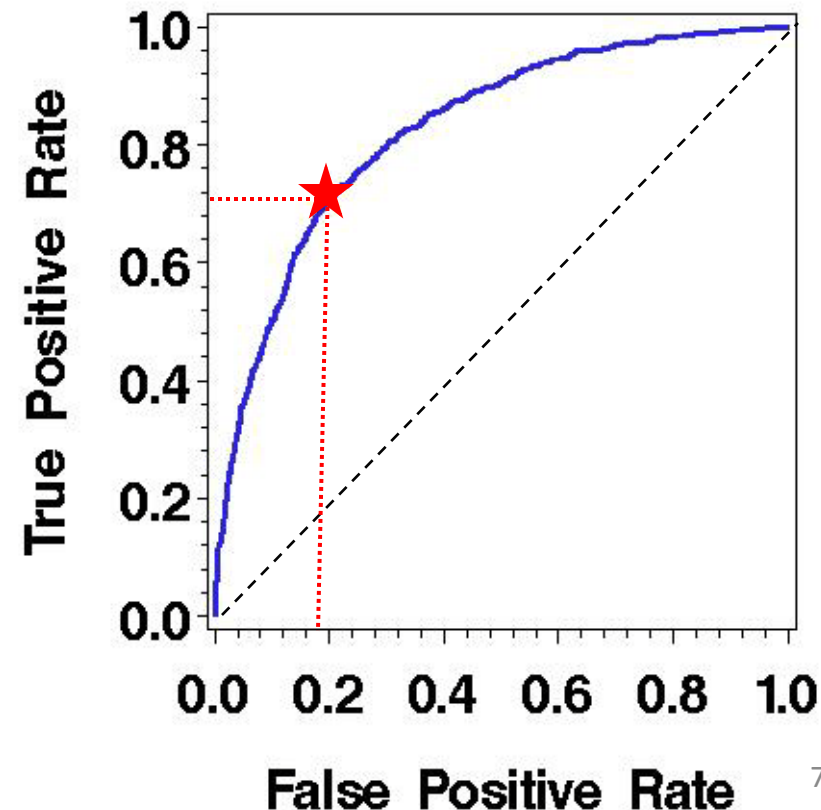
Assessing Predictive Accuracy: Logistic Regression

- For binary or continuous predictors and binary outcome
- Can produce the 2x2 table with binary predictor
- C-statistic: concordance probability, or the estimated probability of making a correct prediction (0 to 1)

Assessing Predictive Accuracy: ROC Curve

Receiver Operating Characteristic (ROC) Curve

- For continuous predictors
- True positive rate (tpr) = Sensitivity
- False positive rate (fpr) = $1 - \text{Specificity}$
- Area under the curve (AUC): probability of making a correct prediction (~ c-statistic)



Methods for Determining Optimal Cut Point(s): Create Dataset

- 2 sets of ACG scores with experience periods spanning 2 consecutive 12-month periods
- ACG Risk Predictors (Year 1)
 - ❖ ACG predictor (e.g. Unscaled PRI)
 - ❖ Disease flags, if applicable (based on EDCs and RxMGs)
- High Cost Outcome (Year 2)
 - ❖ Indicator of high cost (based on total cost)
- Merge year 1 and year 2 data – one record per member

Methods for Determining Optimal Cut Point(s): Compute ROC Curve

- Calculate the tpr (sensitivity) and fpr (1-specificity) obtained when using each possible risk score as the cut point

In SAS, run Proc Logistic to get ROC curve dataset

```
proc logistic data=dataset_name ;
  model high_cost (event="Yes")=ACG_score / outroc=ROC_Dataset;
run;
```

- Create ROC curve by plotting the results for each score (fpr,tpr) - optional

```
symbol1 v=dot i=join;
axis1 label=('False Positive Rate') order=0 to 1 by .2;
axis2 label=('True Positive Rate') order=0 to 1 by .2;
proc gplot data=ROC_Dataset;
  plot _sensit_*_1mspec_ / haxis=axis1 vaxis=axis2;
run; quit;
```

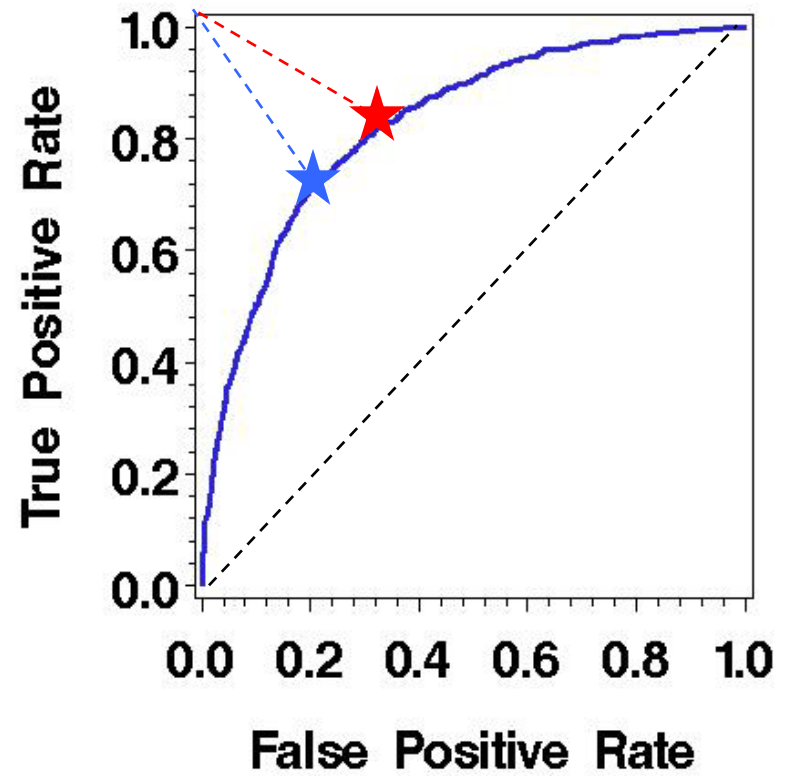
Methods for Determining Optimal Cut Point(s): Calculate Cut Point

- Calculate the distance from each point (fpr,tpr) to the “perfect” coordinate (0,1)
- Distance formula: A (0,1) to B (fpr,tpr)

$$d_{AB} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

$$d_{AB} = \sqrt{(0 - fpr)^2 + (1 - tpr)^2}$$

- Select the point with the shortest distance to “perfect”





The Johns Hopkins University's



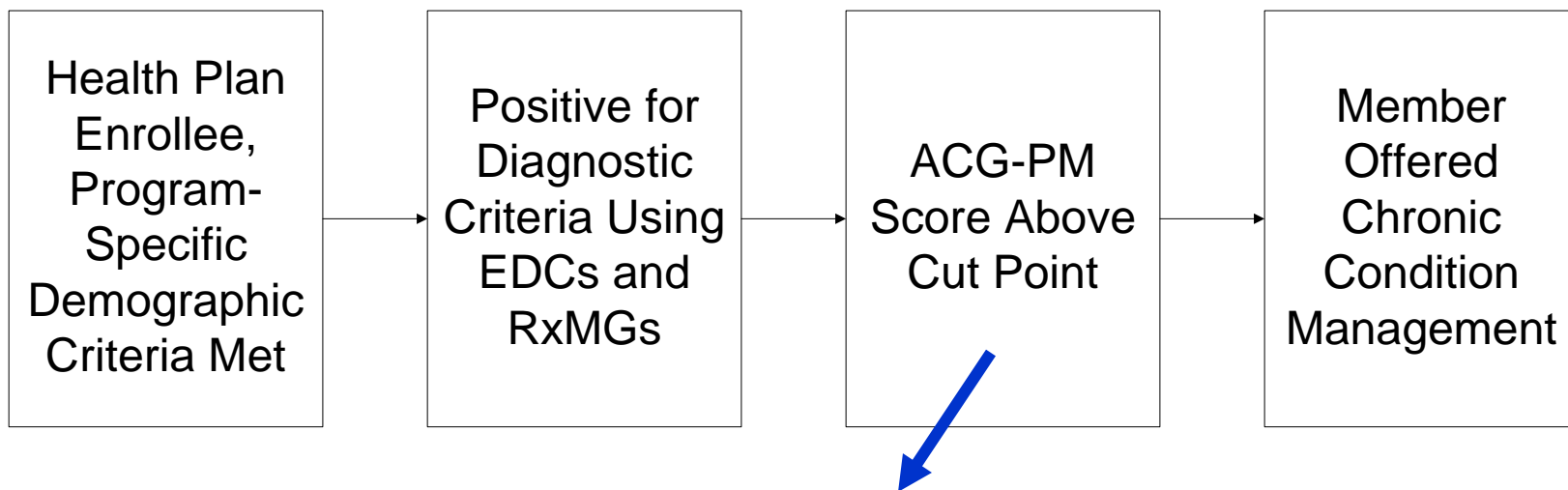
2010 ACG International
Risk Adjustment Conference

MAY 10-12
Tucson, Arizona
Loews Ventana Canyon

Measuring Model Performance: Compare Summary Statistics

- Compare AUC, c-statistic, sensitivity, specificity, PPV, and cut points
- Determination of best results depends on organization goals, resources, etc.

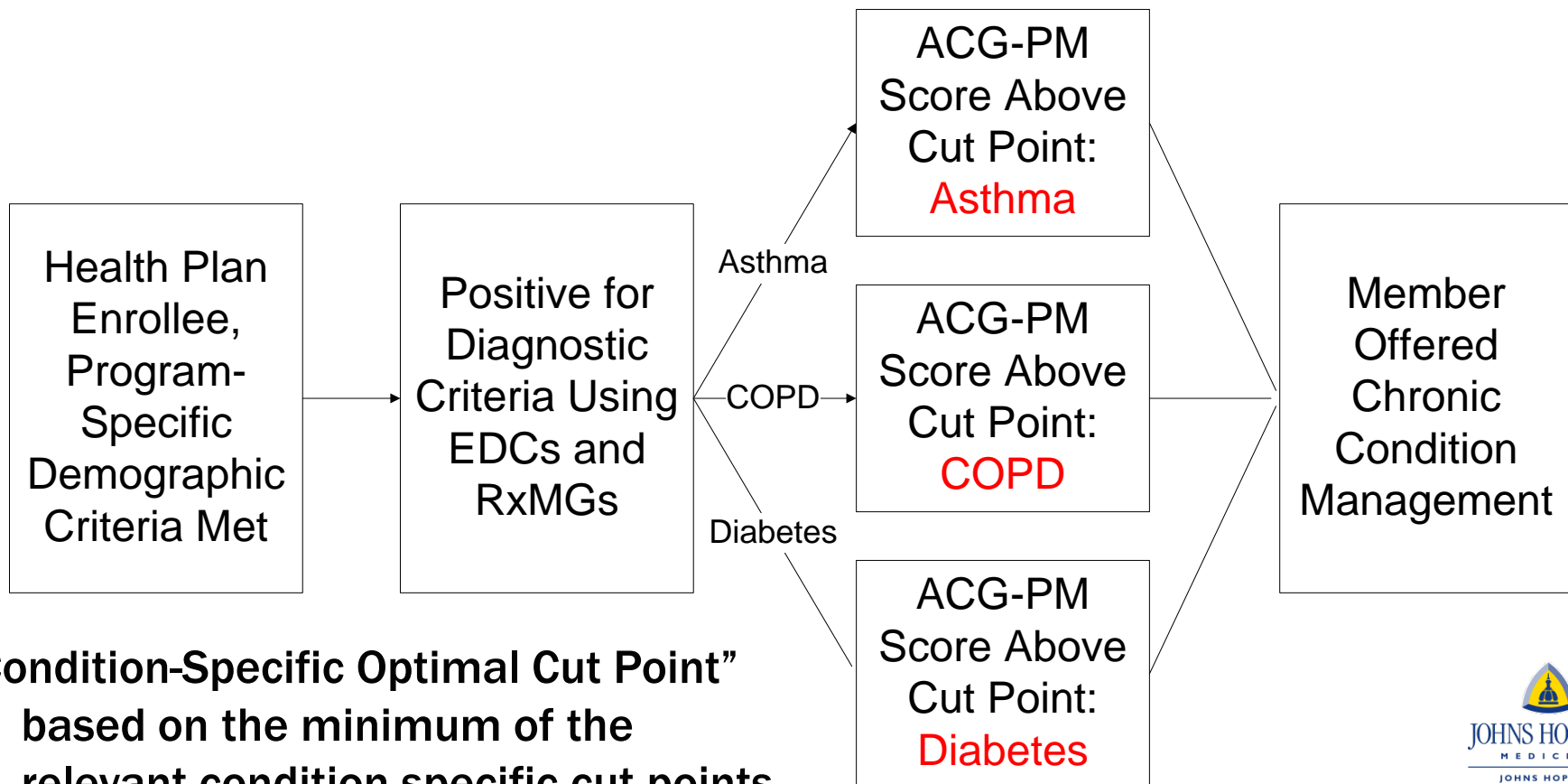
Example of Uniform Screening



■ Single Cut Point selection options:

- ❖ “Arbitrary”: Based on capacity and cumulative distribution of risk scores (e.g. 5% of members had Unscaled PRI score >13.87)
- ❖ “Optimal”: Based on a formula for calculating the optimal cut point

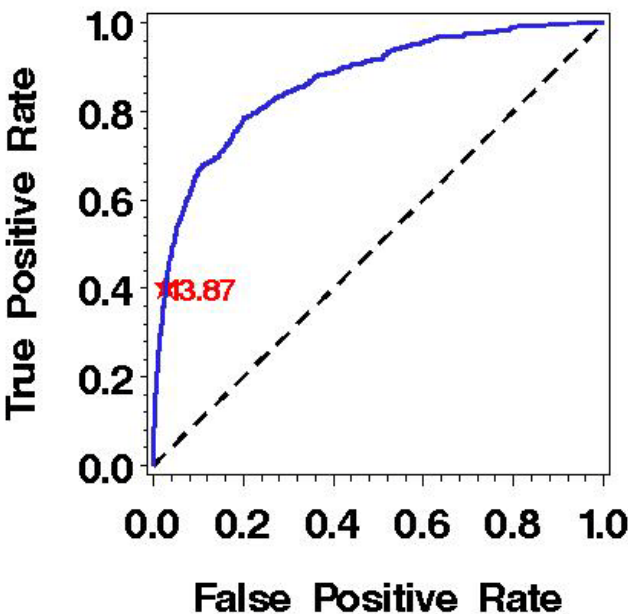
Measuring Model Performance: Example of Condition-Specific Screening



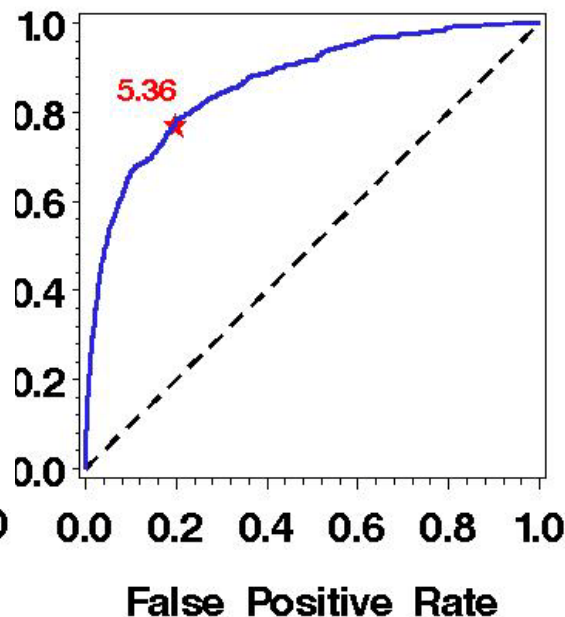
“Condition-Specific Optimal Cut Point”
based on the minimum of the
relevant condition-specific cut points

Cut Points Based on Each Selection Approach

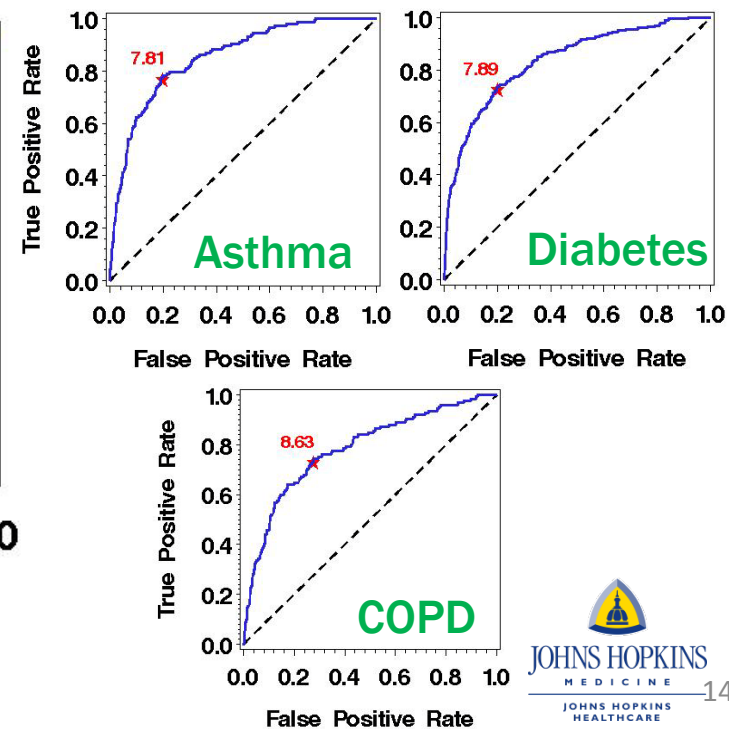
Single Arbitrary



Single Optimal



Condition-Specific Optimal



Accuracy Achieved with Each Cut Point Approach

Condition	Cut Point	Sensitivity	Specificity	PPV	C-Statistic
Single Arbitrary Cut Point	13.87	44.7	95.0	41.5	69.9
Single Optimal Cut Point	5.36	82.0	69.3	17.4	75.6
Condition-Specific Optimal Cut Point	Multiple	70.8	83.7	25.5	77.2



Impact of Cut Point Approaches on Screened Members

Condition	Condition-Specific Optimal Cut Point	# With Condition	# (%) With Condition Above Cut Point		
			Single Arbitrary Cut Point	Single Optimal Cut Point	Condition-Specific Optimal Cut Point
Hypertension	7.89	3,824	375 (9.8%)	1,575 (41.2%)	899 (23.5%)
Depression	7.88	3,250	317 (9.8%)	1,216 (37.4%)	683 (21.0%)
Diabetes	7.89	2,543	246 (9.7%)	1,059 (41.6%)	602 (23.7%)
Asthma	7.81	1,868	191 (10.2%)	770 (41.2%)	406 (21.7%)
Arthritis	7.88	1,851	172 (9.3%)	825 (44.6%)	402 (21.7%)
COPD	8.63	1,206	195 (16.2%)	662 (54.9%)	330 (27.4%)
Malignancies	10.59	672	150 (22.3%)	428 (63.7%)	206 (30.7%)
CHF	11.01	613	151 (24.6%)	440 (71.8%)	180 (29.4%)
Total (Any Condition)		6,459	509 (7.9%)	2,225 (34.4%)	1,310 (20.3%)

Summary

- **To calculate optimal cut point:**
 - ❖ Create dataset with 2 sets of ACG scores spanning 2 consecutive years
 - ❖ Calculate the false positive rate (fpr) and true positive rate (tpr) for each possible ACG score
 - ❖ Calculate the distance of each score to “perfect” using the distance formula
 - ❖ Select the point with the shortest distance to perfect
- **To assess model performance:**
 - ❖ Use the 2x2 table, logistic regression and/or ROC curve to calculate summary statistics for accuracy of model
 - ❖ Compare AUC/c-statistic, sensitivity, specificity, ppv



The Johns Hopkins University's



2010 ACG International
Risk Adjustment Conference

MAY 10-12
Tucson, Arizona
Loews Ventana Canyon

Implications

- Optimal cut points should be calculated
- Optimal cut points vary by condition/disease
- May result in a higher number of identified members

Explore whether condition-specific screening algorithms could enhance predictive accuracy



Contact Information

Shannon Murphy, M.A.

Biostatistician

Johns Hopkins HealthCare LLC

6704 Curtis Court

Glen Burnie, MD 21060

smurphy@jhhc.com