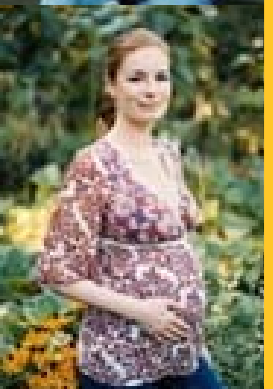


High Cost Claim Prediction for Actuarial Applications

Vincent Kane, FSA, MAAA

Research Scientist, DxCG- A Division of Urix Inc.

The Second National Predictive Modeling Summit

Washington, D.C.

September 22, 2008

Predictive Modeling vs. Risk Adjustment

- PM: Predict claims $ or stratify risk for people or groups, by any means necessary
  - Uses detailed claim-based diagnosis information and possibly procedure data, utilization data, prior costs, timing of claims, benefit provisions, lifestyle-based variables or HRA data, credit info, kitchen sink
- RA: Quantify differences in health status among populations and over time to discover illness burden
  - Picks up on differences in health status and health status alone. Risk *assessment* characterizes the relative cost differences for persons or groups, for example, using relative risk factors.

©2008 Urix, Inc.

Choice of a predictive model versus risk adjuster

- If risk-adjusting payments to providers or plans, you may not want to include prior utilization, costs or procedures.
  - Fairly assess health status, therefore, ignore diagnosis codes that are vague, difficult to audit, and gameable.
- For underwriting, care management, and stop loss or reinsurance applications, you may want to use all available predictors
  - Could recalibrate standard risk adjustment models by adding new variables, or
  - Build a predictive model from scratch for the intended application

©2008 Urix, Inc.

"High Cost Case Model" (HCCM)

- A predictive model which uses all diagnoses and pharmacy claims to prospectively find members likely to be high cost
- Based on RxGroups® and HCC clinical groupings
  - Adds proprietary variables based on prior year cost and utilization patterns
    - Blood disorders, cancers, CHF, diabetes, usual suspects
    - Extremely high cost drugs, certain injectables, etc.
  - Assumes fully run out claims
  - Does not use a lag before the prediction period

©2008 Urix, Inc.

HCCM - Model Characteristics

- Calibrated w/ Thomson MedStat Marketscan data
- Dependent variable, and therefore outcome to be predicted, are year 2 total allowable claims costs
  - A year 2 risk score is the model output
- Prospective with top coding choices
  - No top coding
  - Top coded at $250k
  - Top coded at $100k
  - Top coded at $25k

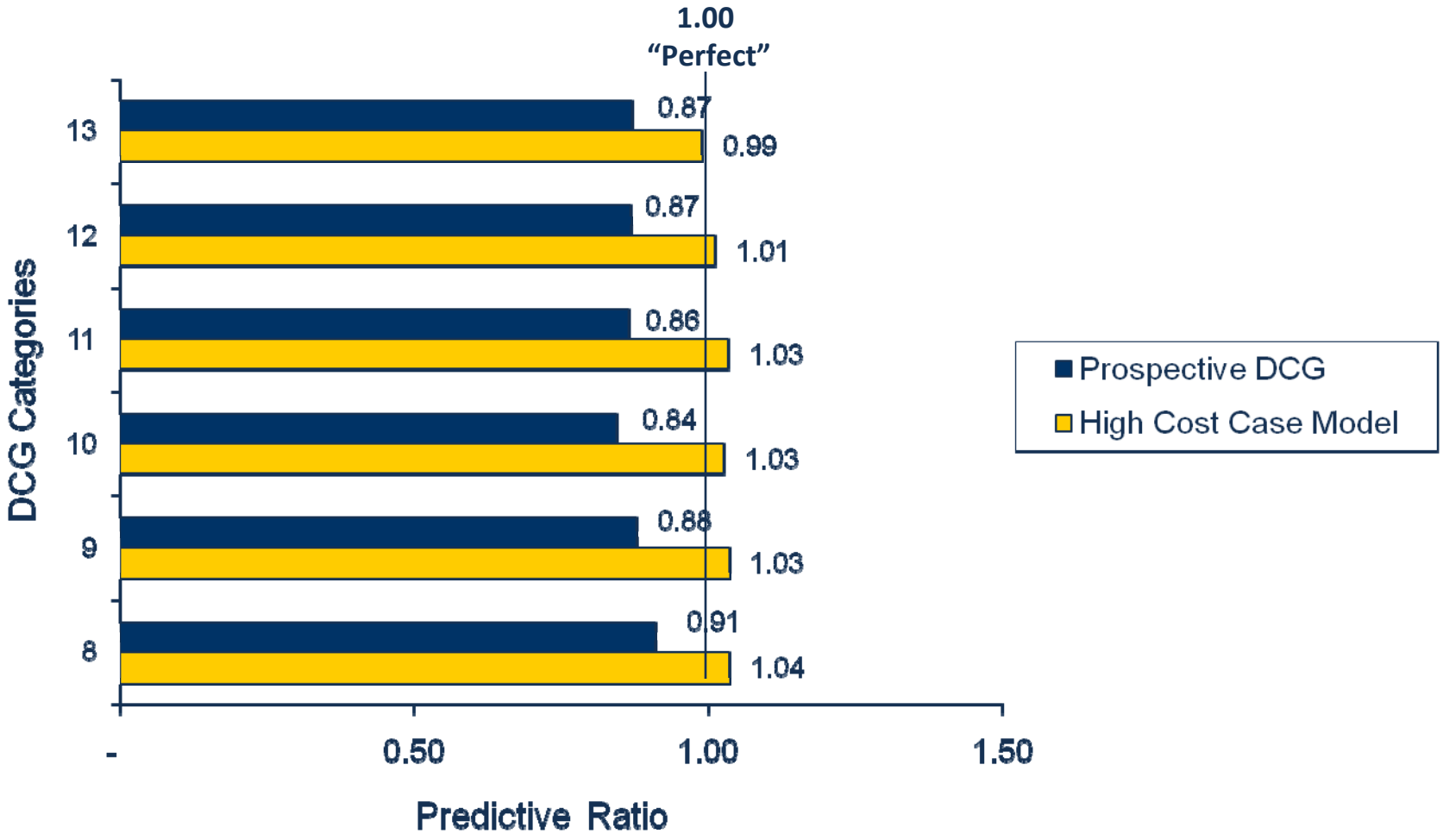
©2008 Urix, Inc.

How is HCCM Different From Prospective DCG/HCC Model?

- Uses prior costs and RxGroups® (NDC codes) as inputs
- Higher R-squared (22.1% vs 14.1%)
- Improved predictive ratios
- Performs better in top ½% and 1%
- Has a higher Positive Predictive Value (PPV) for predicting high cost patients

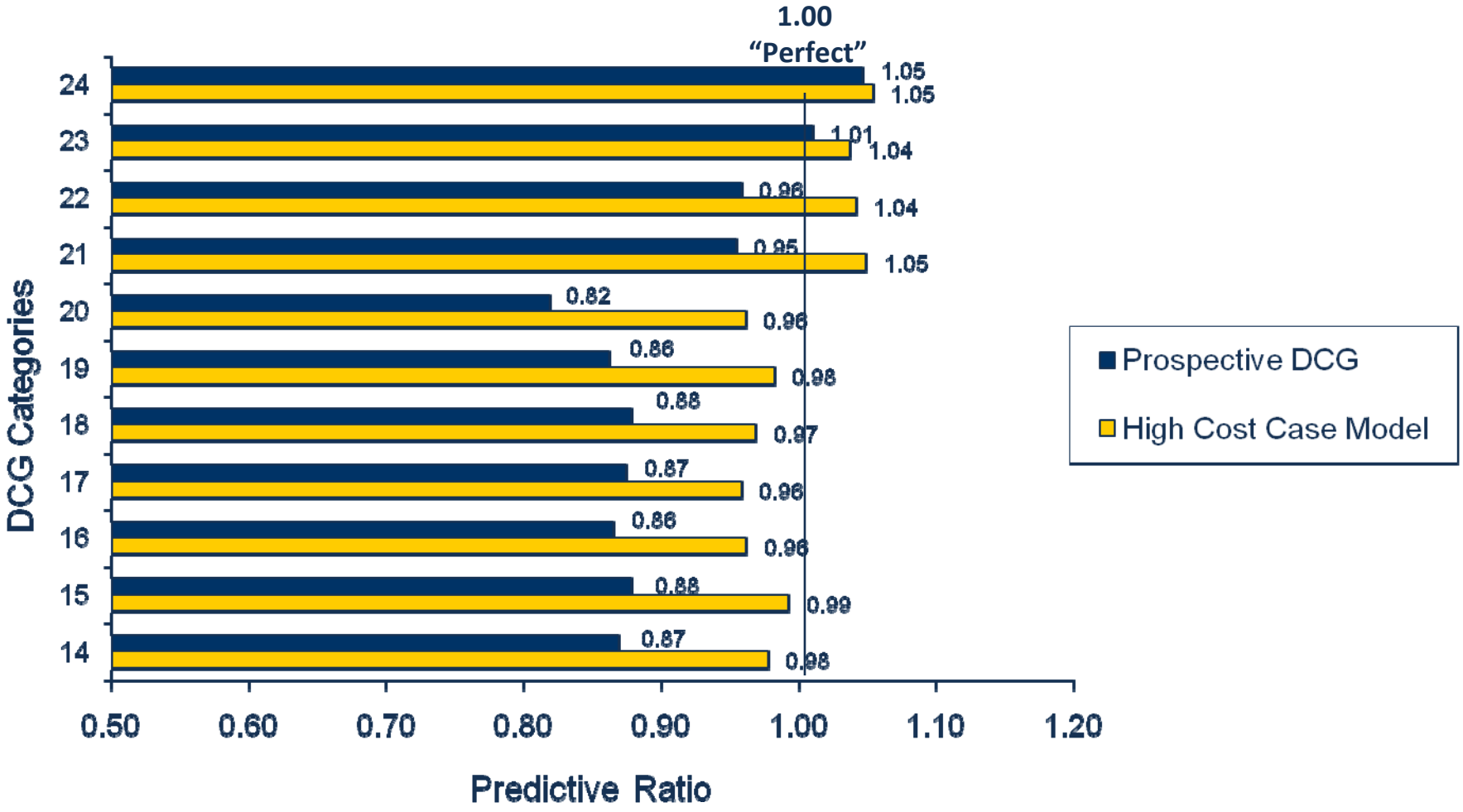
©2008 Urix, Inc.

HCCM Performs Better In Low DCG Buckets and …



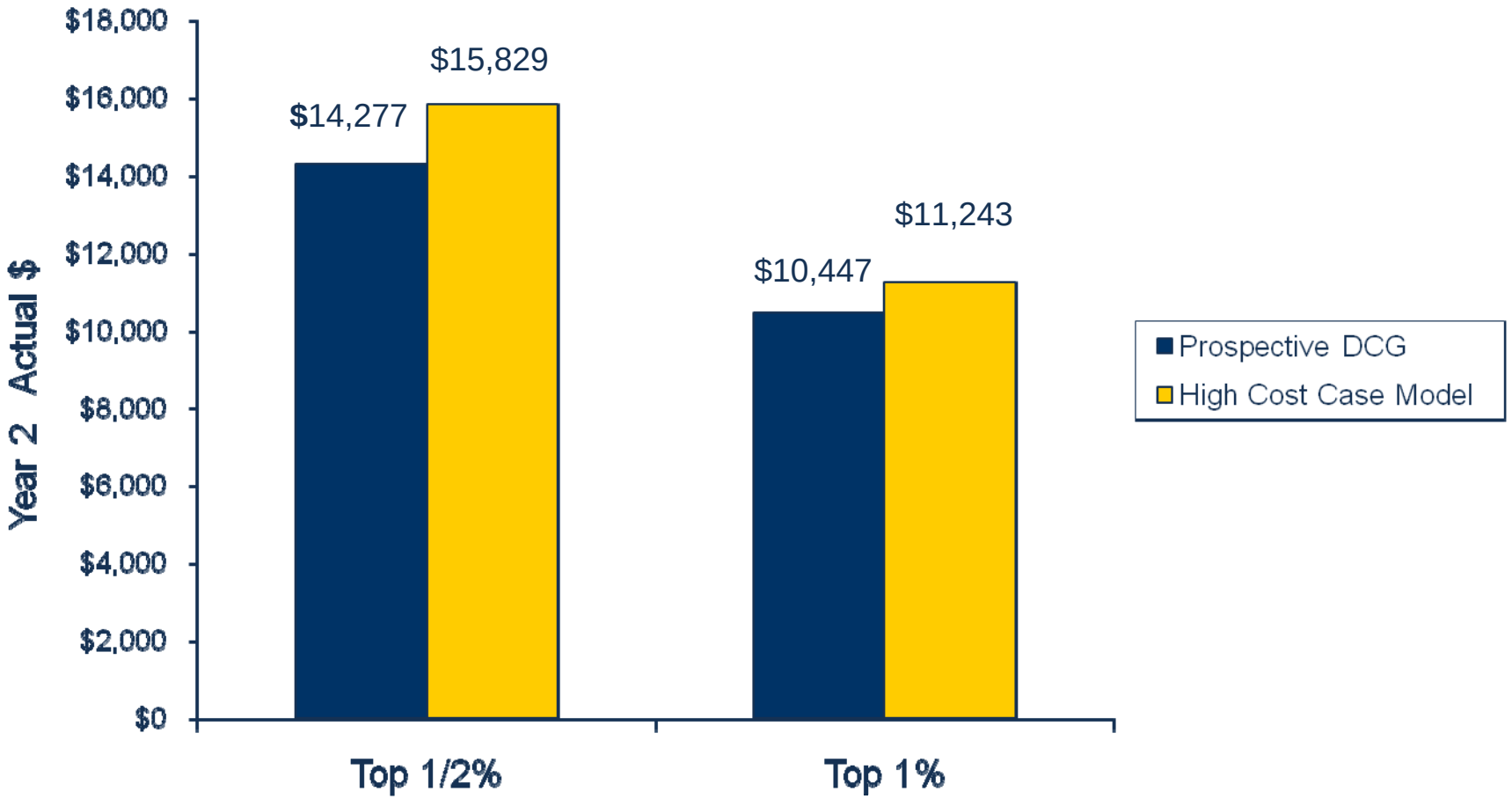
©2008 Urix, Inc.

…Performs Much Better In High DCG Buckets



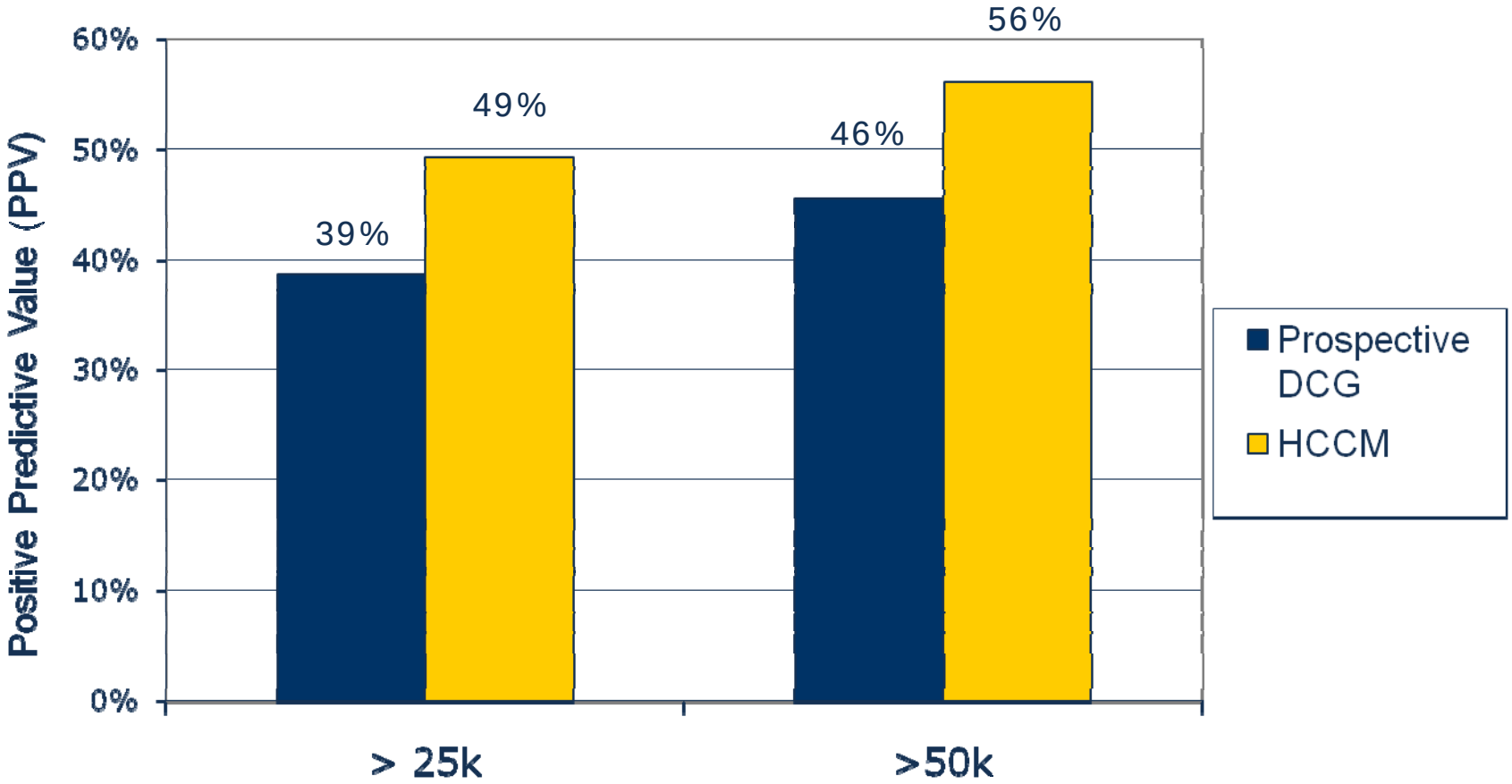
©2008 Urix, Inc.

HCCM Finds More Expensive Individuals in Top Groups



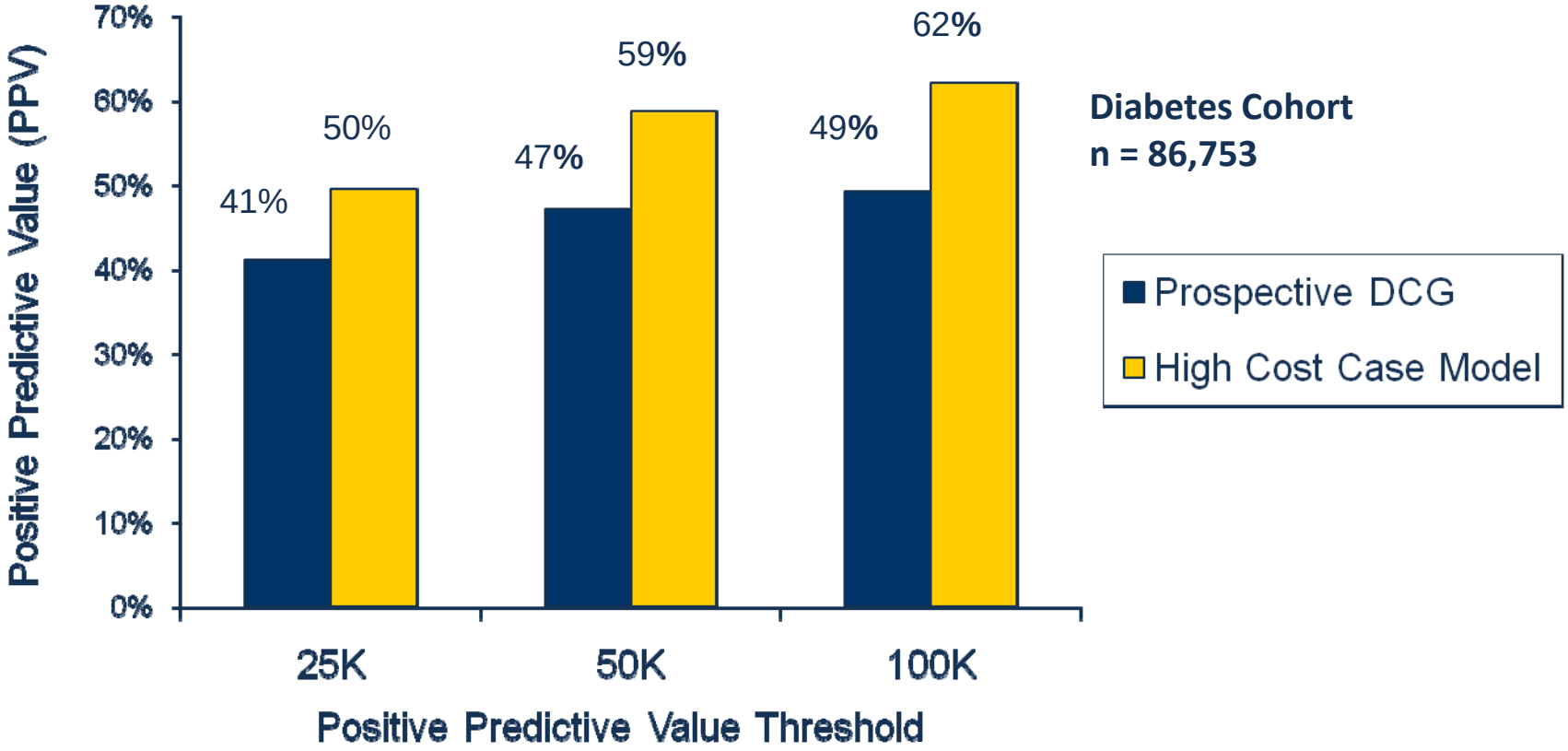
©2008 Urix, Inc.

HCCM Correctly Predicts More Expensive Individuals



©2008 Urix, Inc.

HCCM Correctly "Finds" More Cases – *PPV for Diabetic Cohort*



©2008 Urix, Inc.

Comparing HCCM with Other Means of Predicting Future Costs

- There are lots of different approaches that may be used to predict future costs
  - Age-sex
  - Prior year cost
  - Prospective DCG model
  - Prospective RxGroups model
  - Parametric methods using distributional forms
  - Two-part models
  - Other econometric models
  - Data mining techniques
  - *Combinations of methods*

©2008 Urix, Inc.

Upgrading the standard DCG-HCC model to create one type of "Combined Method"

- In the MarketScan database, DxCG created a model to simulate the combination of the traditional methods
- The recalibration combines age sex categories, the prospective DCG score and year 1 costs to predict year 2 costs
- We define this as the "Combined Method"

©2008 Urix, Inc.

“Predictive Model” performance versus standard diagnosis-based risk adjusters

	R-Squared
Prospective DCG	14.1%
Combined Method (Prospective DCG and Prior Costs)	16.5%
HCCM (no top coding)	22.1%

©2008 Urix, Inc.

Predictive performance improves with decreasing top-coding thresholds

High Cost Case Model	R-squared
No Top Coding	22.1%
$ 250k	26.6%
$ 100K	28.8%
$ 25K	31.4%

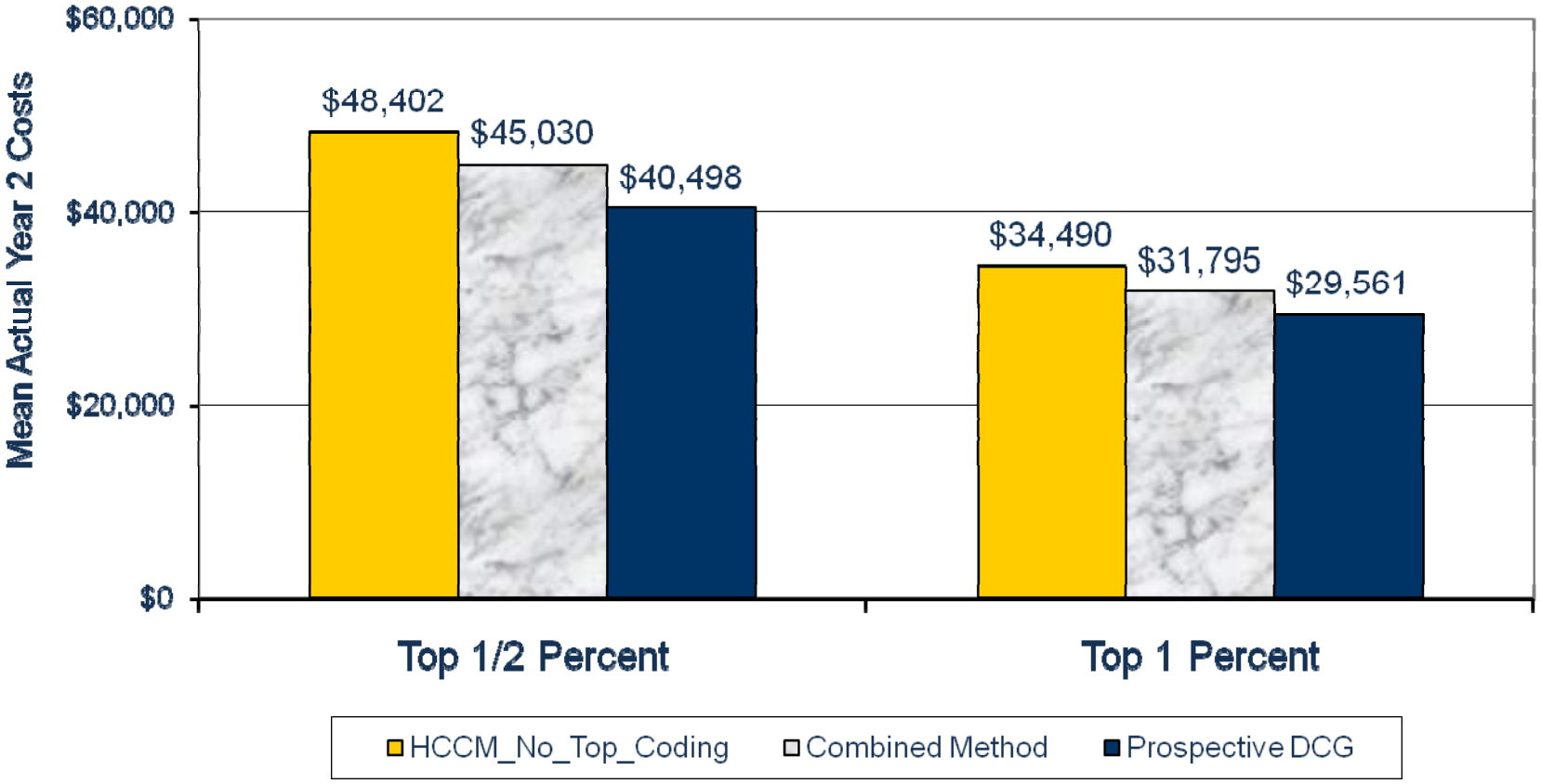
©2008 Urix, Inc.

Also possible to create “top groups” for each model

- Top groups using the prospective DCG model
  - Members who were in the top ½ percent using the prospective DCG method (N= 12,727)
  - Members who were in the top 1 percent using the prospective DCG method (N= 25,453)
- Top groups using the combined method
  - Members who were in the top ½ percent using the combined method (N= 12,727)
  - Members who were in the top 1 percent using the combined method (N= 25,453)
- Top groups using HCCM (no top coding)
  - Members who were in the top ½ percent using HCCM (N= 12,727)
  - Members who were in the top 1 percent using HCCM (N= 25,453)

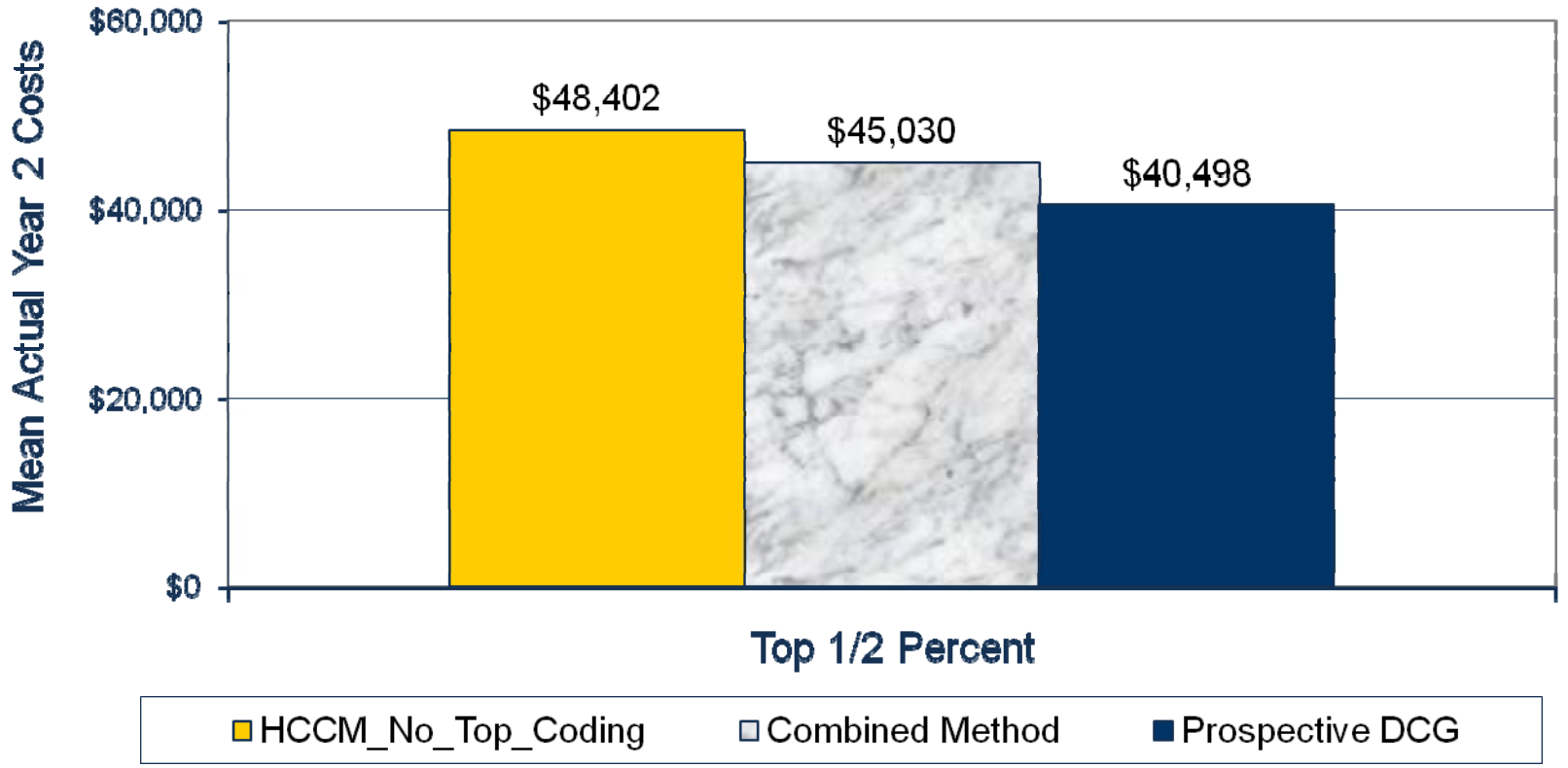
©2008 Urix, Inc.

HCCM Identifies Members With Higher Average Actual Year 2 Costs



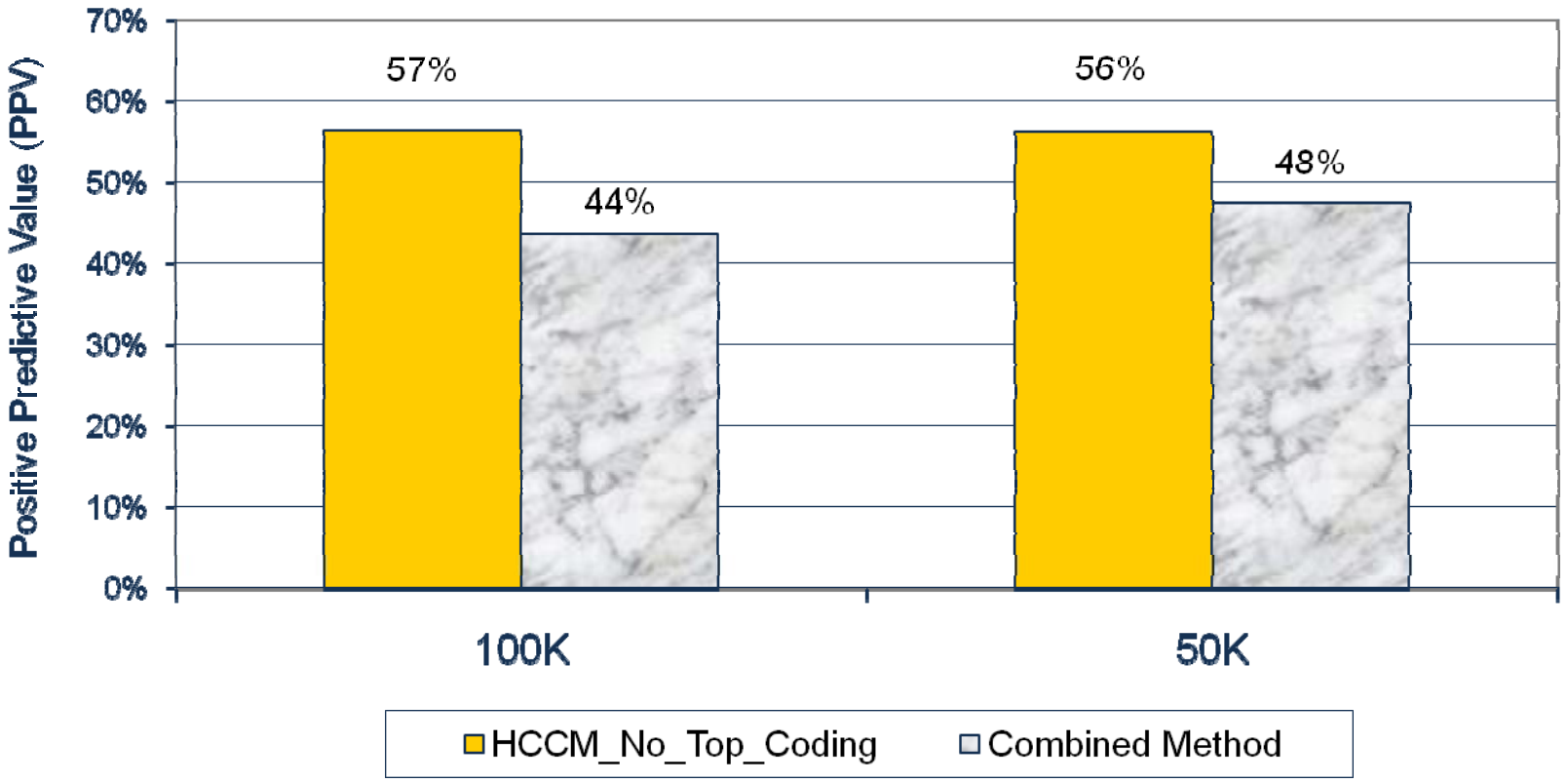
©2008 Urix, Inc.

Results for the top ½ percent group (N = 12,727)



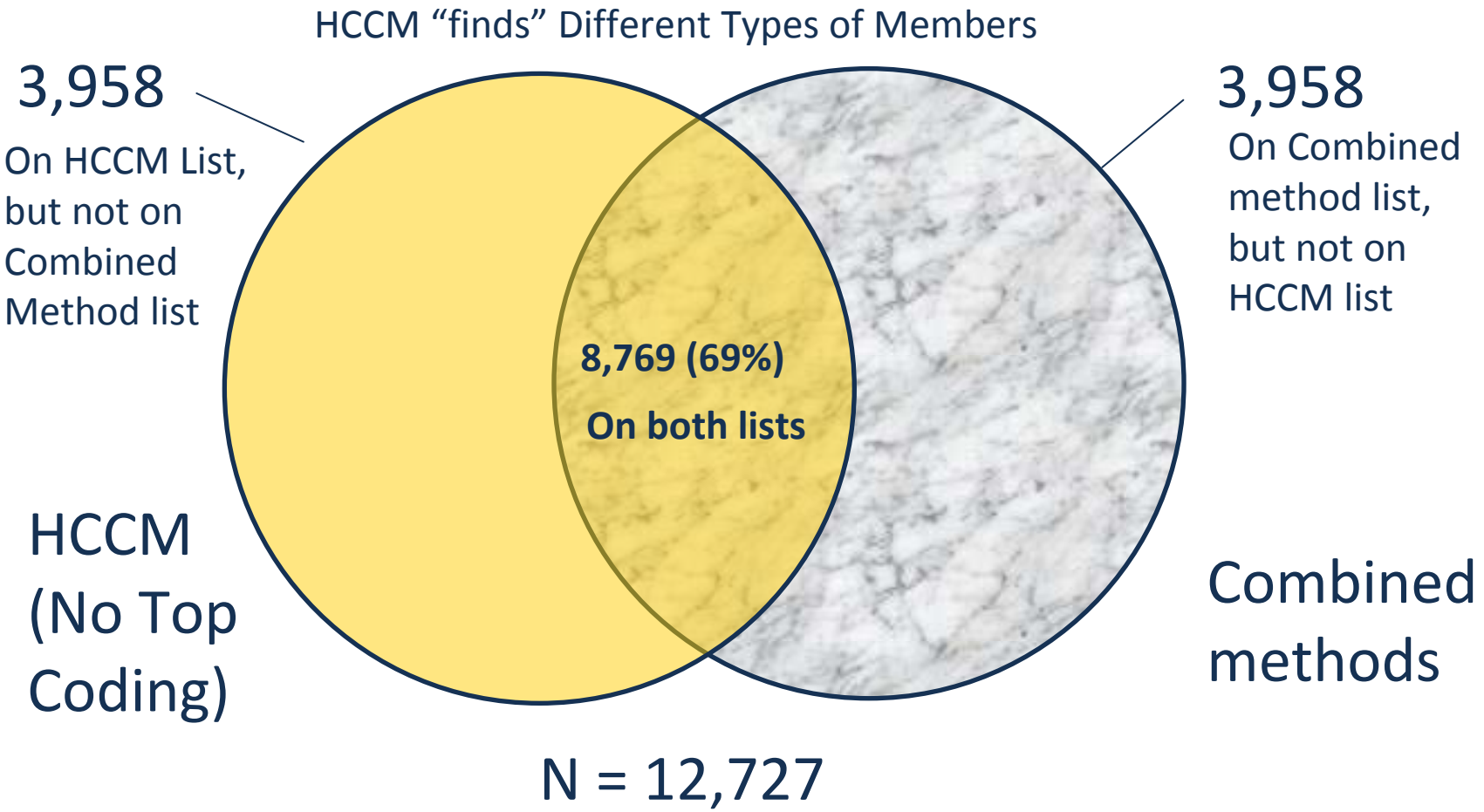
©2008 Urix, Inc.

HCCM Has a Higher PPV Compared to the Combined Method (N = 12,727)



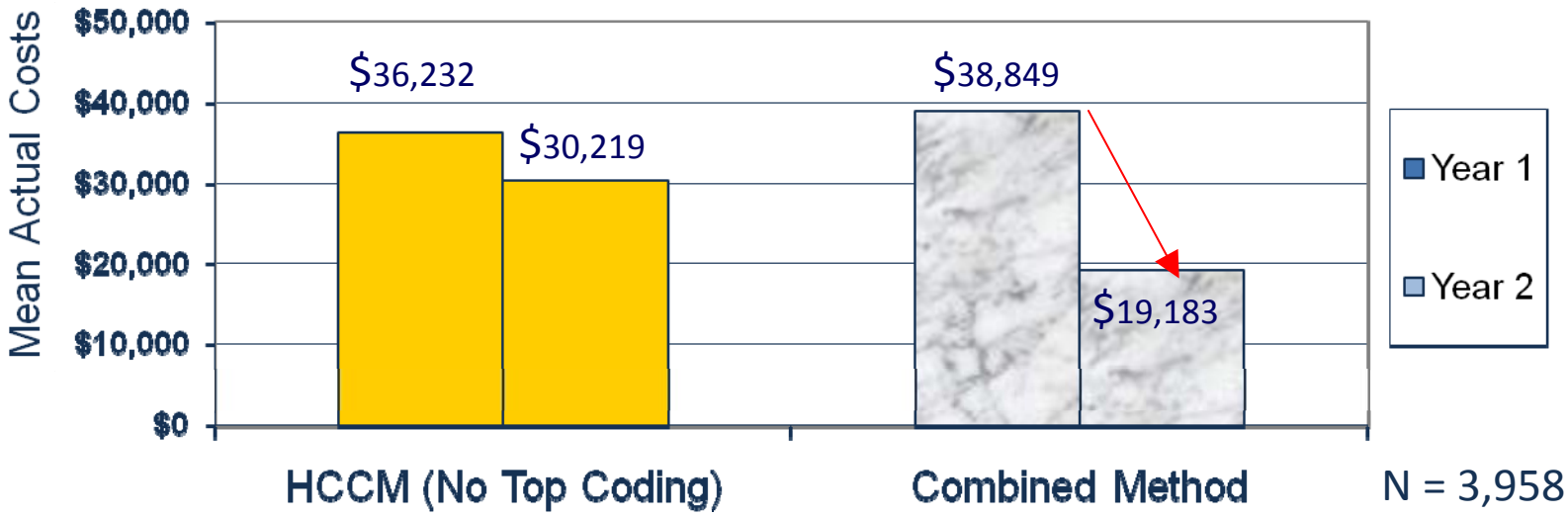
©2008 Urix, Inc.

HCCM Model Found 3,958 Individuals Not On the List from the Combined Method



©2008 Urix, Inc.

The 3,958 Non Overlapping Members Identified by the Combined Method Illustrate Regression To The Mean



Costs for the Non Overlapping 3,958 Individuals on the Combined List drop by 51% in Year 2. By contrast, the non overlapping 3,958 Individuals on the HCCM List drop by only 17% in Year 2

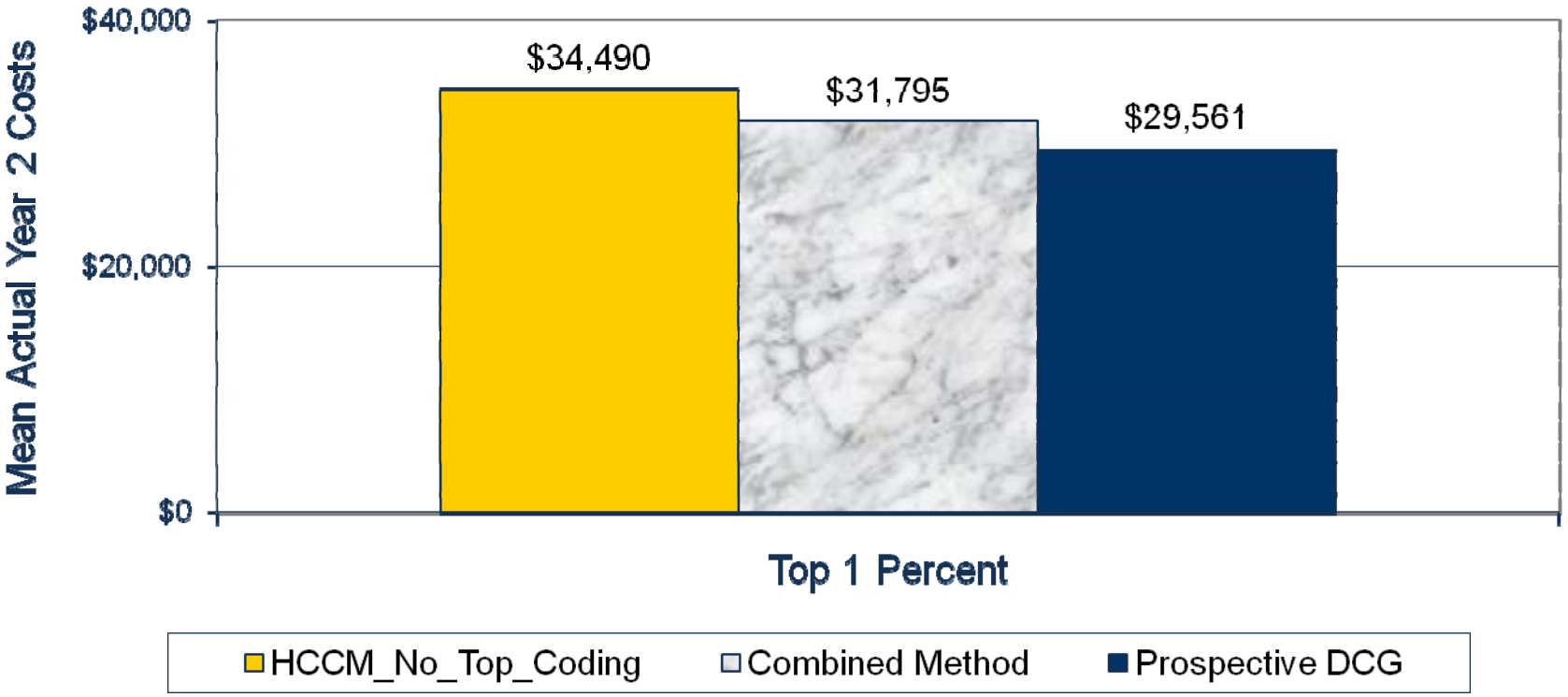
©2008 Urix, Inc.

The HCCM Model Identifies High Cost Cases Better than Traditional Methods

- 3,958 non overlapping individuals on the HCCM list had total Year 2 costs of more than $120 million
  - Average PMPY is $30,219 as shown on the previous chart
- 3,958 non overlapping individuals on the Combined method list had total Year 2 costs of $76 million
  - Average PMPY is $19,183 as shown on the previous chart

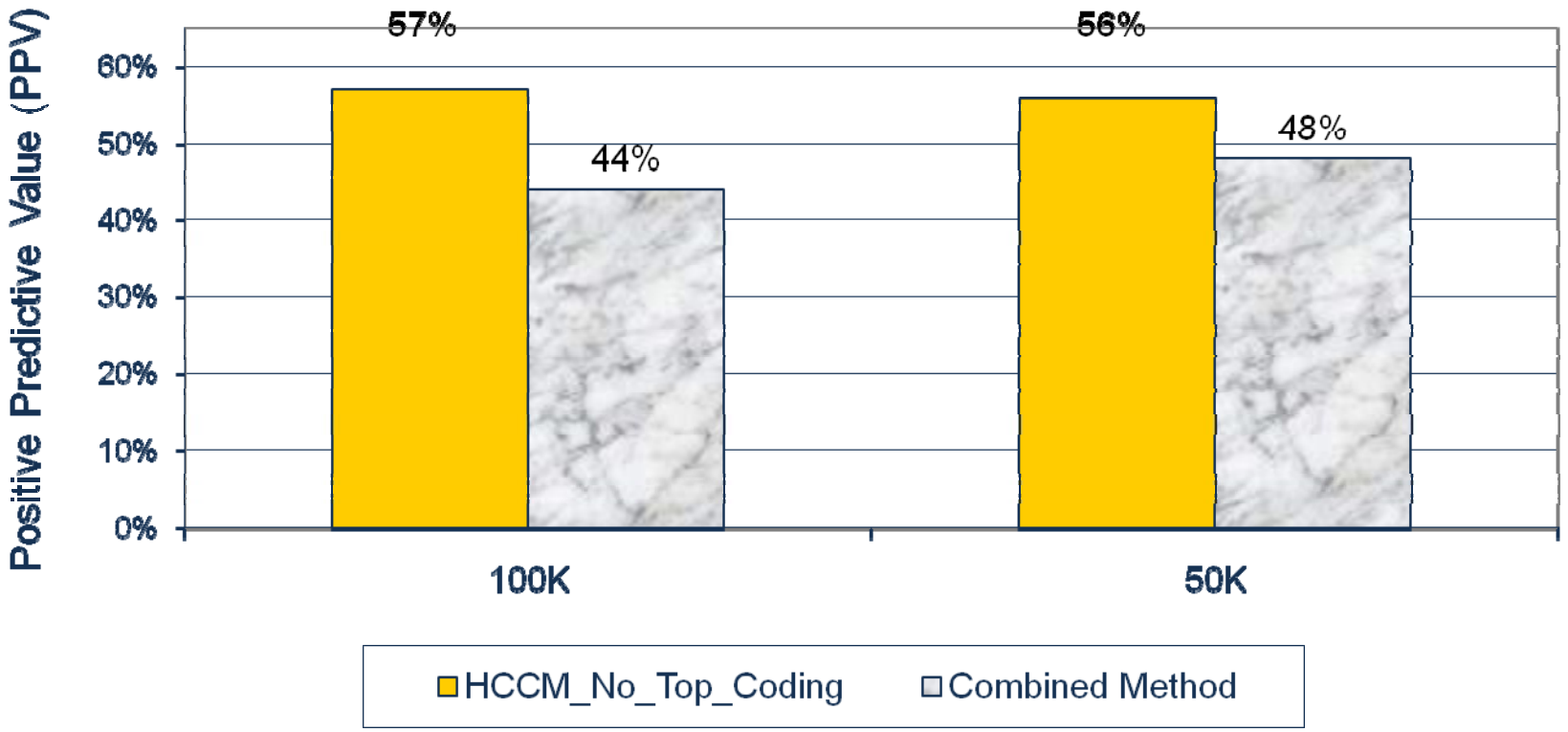
©2008 Urix, Inc.

Results for the top 1 percent group (N=25,453)



©2008 Urix, Inc.

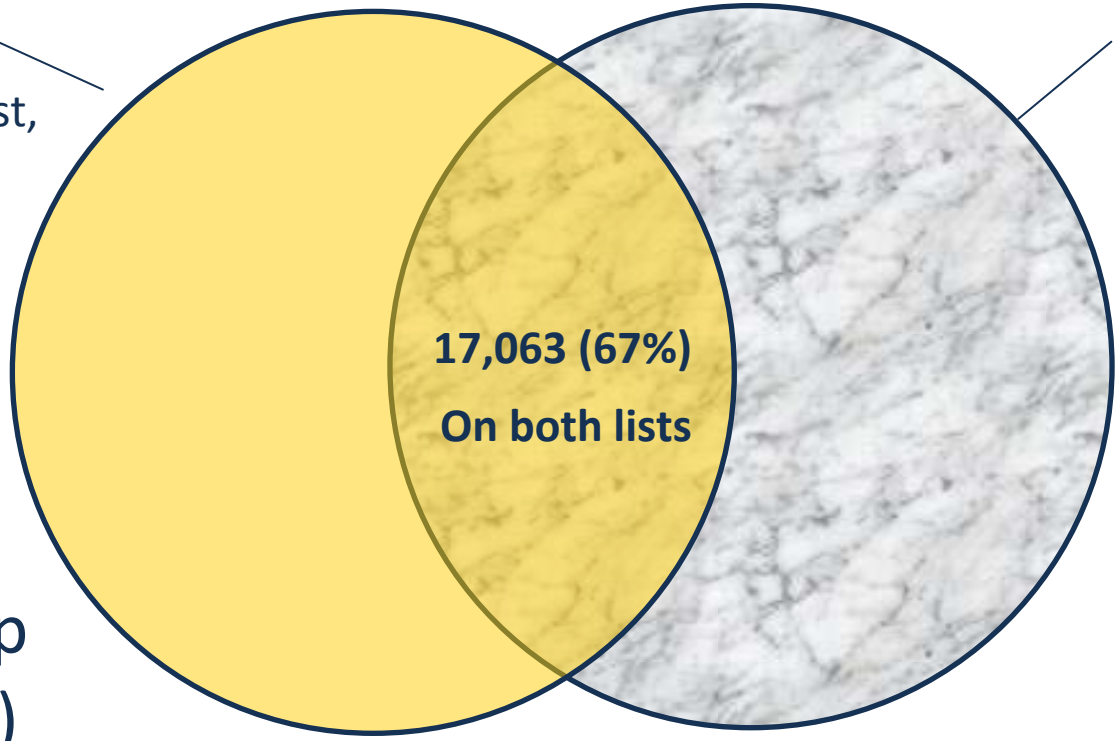
HCCM Has a Higher PPV Compared to the Combined Method (N = 25,453)



©2008 Urix, Inc.

HCCM Model Found 8,390 Individuals Not On the List from the Combined Method

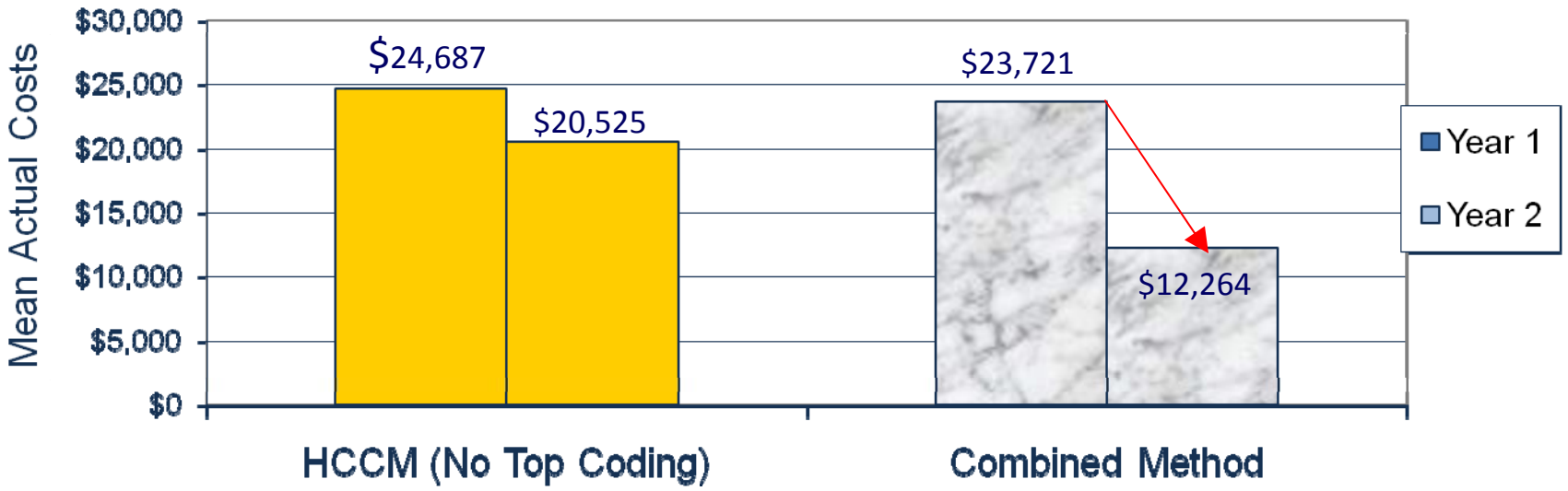
HCCM "finds" Different Types of Members



N = 25,453

©2008 Urix, Inc.

The 8,390 Non Overlapping Members Identified by the Combined Method Illustrate Regression To The Mean



Costs for the Non Overlapping 8,390 Individuals on the Combined List drop by 48% in Year 2. By contrast, the non overlapping 8,390 Individuals on the HCCM List drop by only 17% in Year 2

©2008 Urix, Inc.

The HCCM Model Identifies High Cost Cases Better than Traditional Methods

- 8,390 non overlapping individuals on the HCCM list had total Year 2 costs of more than $172 million
  - Average PMPY is $20,525 as shown on the previous chart
- 8,390 non overlapping individuals on the Combined method list had total Year 2 costs of $103 million
  - Average PMPY is $12,264 as shown on the previous chart

©2008 Urix, Inc.

How are the members in the top groups different?

- Randomly sampled 100,000 lives from Marketscan data set for 2005 and 2006
- Sorted the population using three different methods using 2005 as baseline
  - By High Cost Case Model risk score
  - By Prospective All-Encounter DCG-HCC score
  - By 2005 total allowable claims dollars
- Created 1% top-groups for each method (1,000)

©2008 Urix, Inc.

How are the members in the top groups different?

Top 1% Groups

	% Female	Avg. Age	Hospitalizations 2005	Hospitalizations 2006	Emergency Room 2005	Emergency Room 2006
HCCM	55%	50.1	1.1	0.66	1.3	0.93
Prospective DCG	50%	51.3	1.2	0.67	1.2	0.87
Prior Costs	50%	49.5	1.5	0.61	1.5	0.92

	HCCM Prevalence 2005	HCCM Prevalence 2006		Prosp. DCG Prevalence 2005	Prosp. DCG Prevalence 2006		Prior Cost Prevalence 2005	Prior Cost Prevalence 2006	
Diabetes	277	240	-13%	357	304	-15%	244	200	-18%
CVD	86	49	-43%	101	66	-35%	94	60	-36%
CHF	122	106	-13%	178	140	-21%	118	92	-22%
COPD	113	80	-29%	146	96	-34%	110	57	-48%
VD	132	88	-33%	160	98	-39%	138	72	-48%
CAD	192	156	-19%	231	171	-26%	251	183	-27%
RF	118	91	-23%	176	121	-31%	80	58	-28%
Respiratory	318	224	-30%	345	222	-36%	296	179	-40%

©2008 Urix, Inc.

Aggregated Condition Category Descriptor	HCCM Prevalence	Prosp. DCG Prevalence	Prior Cost Prevalence
ACC001: Infectious and Parasitic	272	270	241
ACC002: Malignant Neoplasm	340	397	297
ACC003: Benign/In Situ/Uncertain Neoplasm	133	116	125
ACC004: Diabetes	277	357	244
ACC005: Nutritional and Metabolic	553	571	564
ACC006: Liver	134	157	115
ACC007: Gastrointestinal	464	463	470
ACC008: Musculoskeletal and Connective Tissue	580	545	590
ACC009: Hematological	354	409	343
ACC010: Cognitive Disorders	54	57	60
ACC011: Substance Abuse	89	100	115
ACC012: Mental	242	205	230
ACC013: Developmental Disability	15	15	24
ACC014: Neurological	248	249	226
ACC015: Cardio-Respiratory Arrest	87	122	117
ACC016: Heart	532	600	601
ACC017: Cerebro-Vascular	86	101	94
ACC018: Vascular	229	262	237
ACC019: Lung	420	452	391
ACC020: Eyes	240	253	230
ACC021: Ears, Nose and Throat	392	365	374
ACC022: Urinary System	365	433	303
ACC023: Genital System	202	194	203
ACC024: Pregnancy Related	12	12	22
ACC025: Skin and Subcutaneous	333	337	316
ACC026: Injury, Poisoning, Complications	433	428	487
ACC027: Symptoms, Signs and Ill-Defined Conditions	802	809	828
ACC028: Neonates	2	9	5
ACC029: Transplants, Openings, Other V-Codes	79	106	77
ACC030: Screening / History	805	799	847

©2008 Urix, Inc.

Aggregated RxGroup Category Descriptor	HCCM Prevalence	Prosp. DCG Prevalence	Prior Cost Prevalence
ARXG001: Analgesics/anti-inflammatories	772	699	798
ARXG002: Anti-hyperlipidemics	373	385	359
ARXG003: Anti-infectives	813	785	784
ARXG004: Coagulants and Anticoagulants	237	238	309
ARXG005: Biologicals	186	169	145
ARXG006: Cardiovascular	619	649	621
ARXG007: Neurological agents	679	588	653
ARXG008: Dermatologicals	372	318	310
ARXG009: EENT preparations	275	236	232
ARXG010: Endocrine/metabolic agents	574	480	488
ARXG011: Diabetes drugs	254	227	202
ARXG012: Pulmonary drugs	247	302	220
ARXG013: GI drugs	665	605	596
ARXG014: Genitourinary agents	241	215	205
ARXG015: Immunologic agents	132	107	61
ARXG016: Nutritionals	301	314	254
ARXG017: Upper respiratory agents	397	339	342
ARXG018: Additional groups	324	346	272

©2008 Urix, Inc.

When to use the High Cost Case Model

- When a plan needs to identify the top ½ percent or top 1% of cases expected to be high cost
  - Care management
- When the business problem is:
  - Identifying cases that are going to be catastrophic (high cost) for the plan
    - Pricing, Underwriting
  - Understanding how many and what kinds of stop loss cases are likely to occur (e.g. in a self-insured account)
  - Understanding if there are excess risk coverage or reinsurance considerations

©2008 Urix, Inc.

Recommended Uses of HCCM Top Coding Choices

- "No top coding" – for budgeting and projecting *total* costs
- $250K and $100K - when predicting costs below these attachment points
- $25k - for use by forecasting actuaries and also disease management professionals
  - Model has the best PPV for predicting those likely to exceed $25k
- HCCM top coding options (250K, 100K and 25K) simulate the impact of reinsurance or stop loss at those levels
  - Top coded models have improved predictive accuracy (as measured by R^2)

©2008 Urix, Inc.

Applications of high cost claim prediction

- More accurate predictions for individuals & groups
- Group by disease, and then rank
  - DM program involvement
- Rank groups or identify groups with higher concentrations of expected high cost claims
  - Rank by expected year 2 cost
  - Monitoring accounts
- Pooling charges in underwriting or self-insured pricing
- Simulation of reinsurance arrangements or risk pools
  - Better estimate the right tail of the claims distribution

©2008 Urix, Inc.

Reinsurance Considerations

- American Re HealthCare (now Munich Re) gave a user conference presentation in 2004 on high cost claim prediction
  - Evaluated several types of models for predicting high cost claims
    - 2-Part Prospective DCG model with simple recalibration
    - 2-Part Prospective DCG model with “total” recalibration
    - Age-sex tables
    - Prior Costs
    - Claims distributions (e.g., Log-normal, discrete continuance tables)

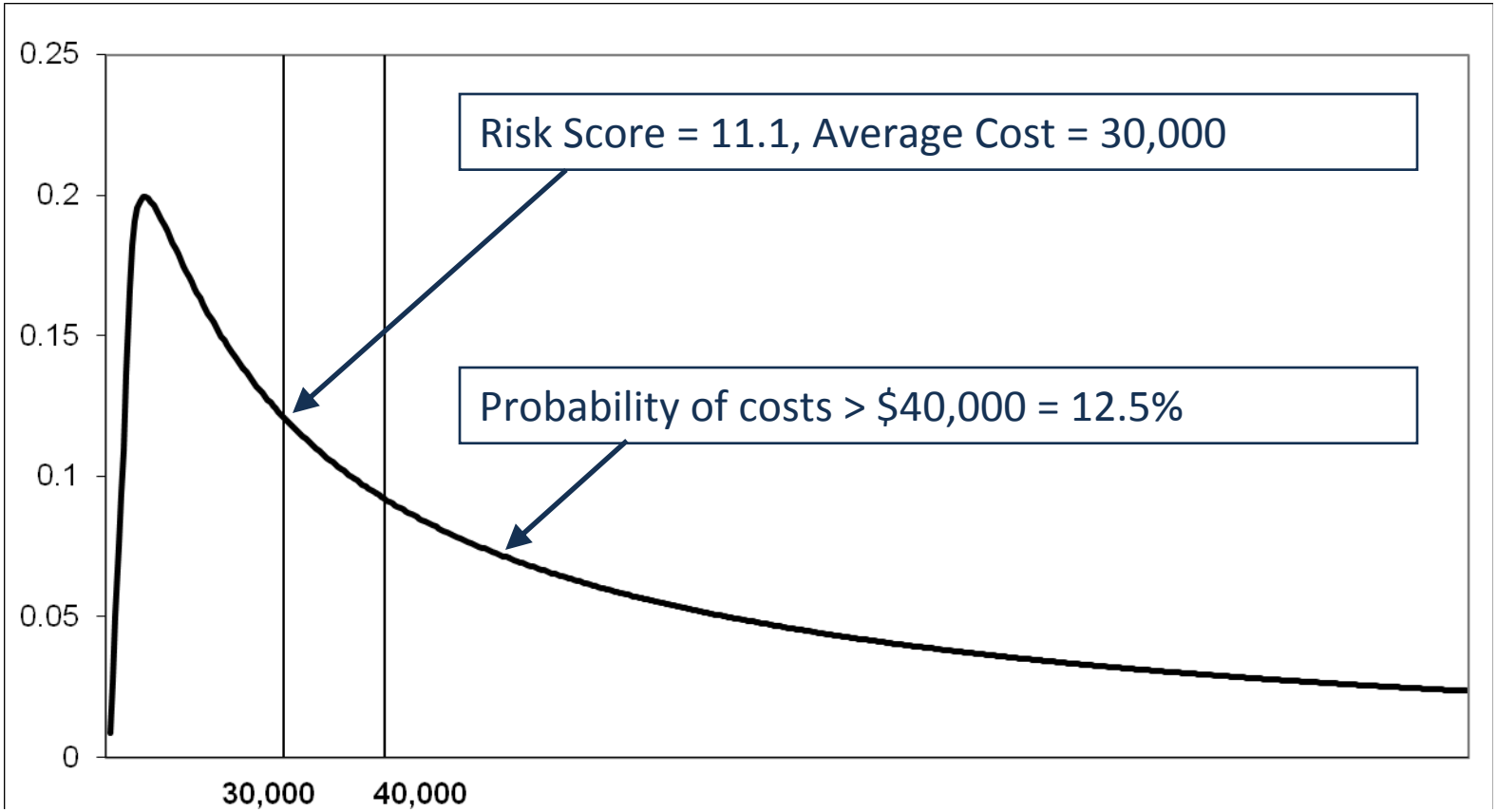
©2008 Urix, Inc.

Reinsurance Considerations (cont'd)

- Risk scores for non-top-coded model reflect total costs
  - You can look at the prevalence of risk scores that would put you over the stop loss threshold (by multiplying by population's average cost)
  - You can look at the prevalence of actual year 2 claims over the stop loss threshold
  - There will be a disconnect!

©2008 Urix, Inc.

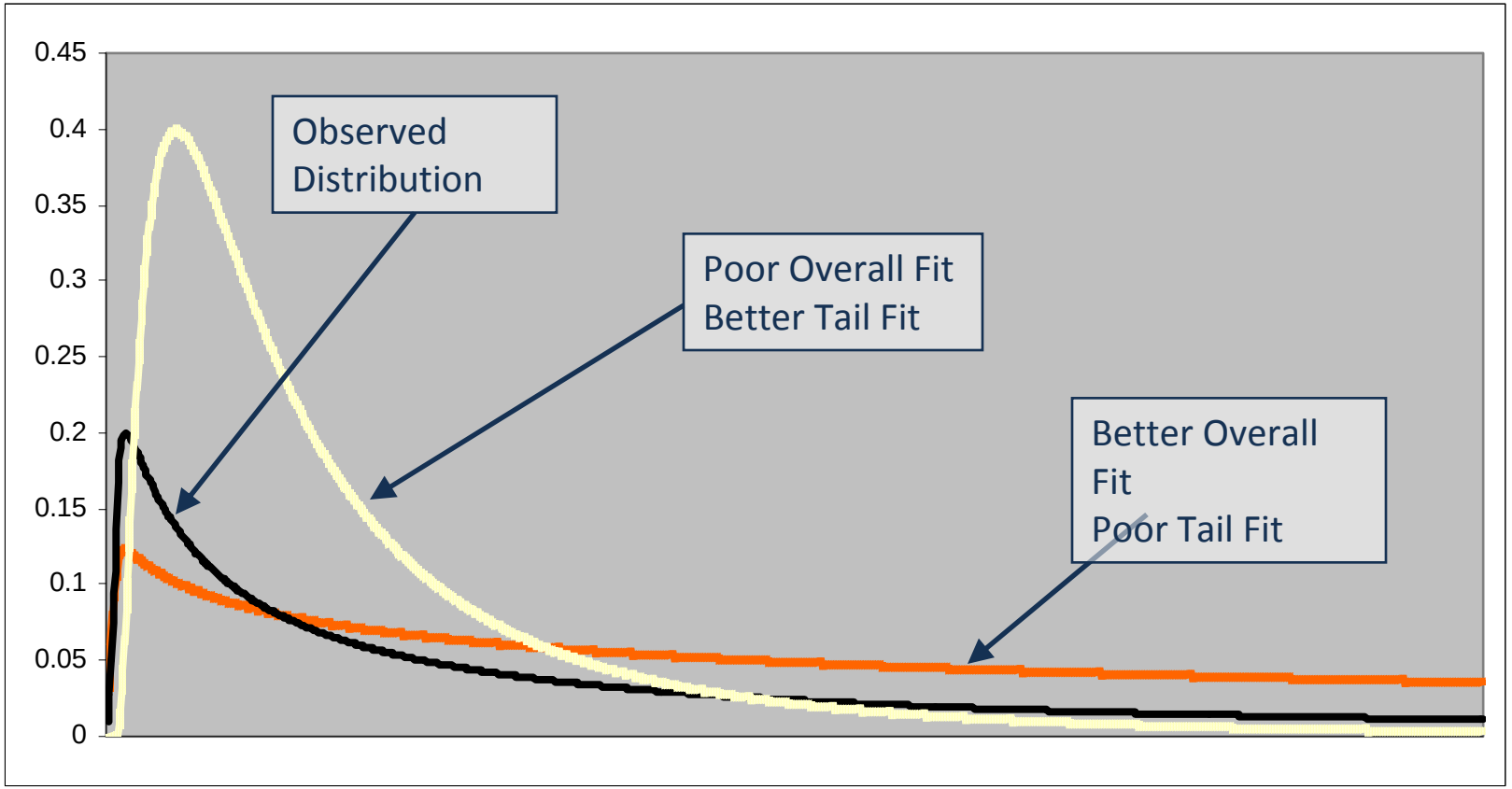
Reinsurance Considerations (cont'd)



From American Re "Using DxCG for Stop Loss and Reinsurance Pricing", 2004 DxCG User Conference Presentation

©2008 Urix, Inc.

Reinsurance Considerations (cont'd)



From American Re "Using DxCG for Stop Loss and Reinsurance Pricing", 2004 DxCG User Conference Presentation

©2008 Urix, Inc.

American Re retrospective study- methodology

- Methods evaluated:
  - 2-part recalibrations (all HCCs, limited set)
  - Claims distributions based on scores (best fit overall, best fit for top 50%)
  - Age-sex factors
  - Prior year costs
- Looked at ability to identify high cost claimants, excess loss PMPM and grouped R-Squared

©2008 Urix, Inc.

American Re retrospective study- findings

- High cost claim identification
  - Diagnostic models superior in finding high cost claims at all stop loss thresholds
  - Those that the prior cost method successfully identified as high cost had higher excess claims
- PMPM Excess Loss
  - Recalibrated model with limited HCCs was best
  - Prior cost and DxCG raw predictions were equivalent
  - Recalibrated “All HCCs” did not perform well as others

©2008 Urix, Inc.

American Re retrospective study- findings (cont'd)

- Group pricing (PM versus standard methods)
  - Standard methods are age-sex or prior cost
  - Age-sex always worse than diagnostic models
  - Small to mid-size groups (<250): Diagnostic better than prior costs alone (all thresholds)
    - Diagnostic model more limited at $250K threshold

©2008 Urix, Inc.

American Re retrospective study- findings (cont'd)

- Group pricing (within class of PM)
  - At lower thresholds, recalibrated "All HCCs" better
    - Limited HCCs and distributional models equivalent
  - At $100K threshold, recalibrate "All HCCs" model and distributional models equivalent
  - At $250K threshold, the distributional models were better than either of the recalibrated models, though predictive performance was not very strong

©2008 Urix, Inc.

Reinsurance Pooling Scheme

- Large, self-insured employer with national PPO and many Business Units (BUs) each accountable for own healthcare financials
- Corporate decided to “risk-adjust” and bill BUs premiums adjusted to their population
  - Risk premium proxies for Aggregate Stop Loss
  - Billed premiums reconciled with actual claims
  - “Recoveries” paid from Corporate pool, with desired outcome that loss ratios approach 100%

©2008 Urix, Inc.

Pooling Scheme for Self-Insured Employer with Several Business Units

Without Diagnosis-Based Risk Adjustment

Pooled PPO Claim PMPM **$250.00**

Business Unit:	**Corporate**	**XYZ Co.**	**ABC Co.**	**Total**
Membership	45	455	1,500	2,000
Average Age	48	42	36	38
% Male	70%	25%	70%	60%
Demographic Factor Adjustment	**1.15**	**1.05**	**0.98**	**1.00**
(Normalized)				
Business Unit Expected PMPM	$287.50	$262.50	$245.08	$250.00
Risk Pooling Charge as % of Claims	10.0%	7.0%	5.0%	5.6%
Charged Risk Premium PMPM	$28.75	$18.38	$12.25	$14.02
Total Risk Premium Dollars	$15,525	$100,328	$220,575	**$336,428**
Actual Incurred PPO Claims PMPM	$183.75	$367.50	$232.75	$262.30
Initial Business Unit Loss Ratio	63.9%	140.0%	95.0%	104.9%
Actual minus Expected Claims				
PMPM	($103.75)	$105.00	($12.33)	$12.30
Annual Dollars	($56,025)	$573,300	($222,000)	$295,275
Recoveries collected from Pool	$0	**$336,428**	$0	$336,428
Net Owed to the Reinsurance Pool	$0	$236,873	($222,000)	$14,873
Final Business Unit Loss Ratio	63.9%	116.5%	95.0%	

©2008 Urix, Inc.

Without Diagnosis-Based Risk Adjustment

Final Pool Accounting

Starting value	$0
Risk premium collected	$336,428
Recoveries paid to units	($336,428)
Collected from "losers"	$236,873
Paid to "winners"	($222,000)
Ending value	**$14,873**

©2008 Urix, Inc.

With Diagnosis-Based Risk Adjustment				
Pooled PPO Claim PMPM	**$250.00**			
Business Unit:	**Corporate**	**XYZ Co.**	**ABC Co.**	**Total**
Membership	45	455	1,500	2,000
Average Age	48	42	36	38
% Male	70%	25%	70%	60%
Demographic Factor Adjustment *(Normalized)*	1.15	1.05	0.98	1.00
"Hidden Health Information"	*Low incidence of chronic disease for this age group*	*High incidence of diabetes, heart disease and associated comorbidities*	*Young, healthy and invincible*	
Unit's Average Relative Risk Score *(Prospective, Normalized)*	0.85	1.40	0.88	1.00

©2008 Urix, Inc.

With Diagnosis-Based Risk Adjustment				
Pooled PPO Claim PMPM	**$250.00**			
Business Unit:	**Corporate**	**XYZ Co.**	**ABC Co.**	**Total**
Membership	45	455	1,500	2,000
Average Age	48	42	36	38
% Male	70%	25%	70%	60%
Unit's Average Relative Risk Score	0.85	1.40	0.88	1.00
(Prospective, Normalized)				
Business Unit Risk-Adjusted PMPM	$212.50	$350.00	$220.79	$250.00
Risk Pooling Charge as % of Claims	10.0%	7.0%	5.0%	5.7%
Charged Risk Premium PMPM	$21.25	$24.50	$11.04	$14.33
Total Risk Premium Dollars	$11,475	$133,770	$198,713	**$343,958**
Actual Incurred PPO Claims PMPM	$183.75	$367.50	$232.75	$262.30
Initial Business Unit Loss Ratio	86.5%	105.0%	105.4%	104.9%
Actual minus Expected Claims				
PMPM	($28.75)	$17.50	$11.96	$12.30
Annual Dollars	($15,525)	$95,550	$215,250	$295,275
Recoveries collected from Pool	$0	**$95,550**	**$215,250**	$310,800
Net Owed to the Reinsurance Pool	$0	$0	$0	$0
Final Business Unit Loss Ratio	86.5%	100.0%	100.0%	

©2008 Urix, Inc.

With Diagnosis-Based Risk Adjustment

Final Pool Accounting

Starting value	$0
Risk premium collected	$343,958
Recoveries paid to units	($310,800)
Collected from "losers"	$0
Paid to "winners"	$0
Ending value	**$33,157**

©2008 Urix, Inc.

Any Questions?

Vincent Kane, FSA, MAAA
Research Scientist
DxCG – A Division of Urix, Inc.
vincent.kane@dxcg.com
www.dxcg.com

©2008 Urix, Inc.