

Innovative Predictive Modeling Approaches to Improve Quality of Care

Soyal Momin, MS, MBA

09/15/2010

The Fourth National Predictive Modeling Summit
Arlington, VA



Outline

- Maximizing the Value of Predictive Modeling
- Introduction – NCQA, HEDIS®, CAHPS®
- Business Case for Quality of Care Predictive Analytics
- Quality of Care Analytics Framework:
 - Prioritize Quality of Care Metrics – AAR Plan
 - Understand Drivers of Compliance/Response
 - HEDIS - Breast Cancer Screening
 - CAHPS - Predictors of Survey Response
 - Implement Quality of Care Improvement Efforts - OOP

Maximizing the Value of Predictive Modeling

- Understanding Population Needs
 - Risk assessment and risk adjustment
- Actuarial and Underwriting
- Care Management Member
 - Identification
 - Stratification
 - Triaging
- Care Management Program Evaluation
- **Quality of Care Improvement**

Introduction - NCQA

National Committee for Quality Assurance
(www.ncqa.org)

- Best known for assessing and reporting the quality of managed care plans across the nation
- Provides quality oversight through a voluntary accreditation program
- Accreditation well recognized by state Medicaid programs and large employers

NCQA Accreditation

- Standards
 - Quality Management and Improvement
 - Utilization Management
 - Credentialing and Re-credentialing
 - Members Rights and Responsibilities
 - Member Connections
- **HEDIS**
- **CAHPS**

Accreditation status and scoring ranges are:

- | | |
|---------------|------------|
| – Excellent | 90 – 100 |
| – Commendable | 80 – 89.99 |
| – Accredited | 65 – 79.99 |
| – Provisional | 55 – 64.99 |
| – Denied | 0 – 54.99 |

Healthcare Effectiveness Data & Information Set

- HEDIS is an **evolving** set of standard specifications for measuring health plan performance
- Originally developed by employers and the HMO group in 1991; NCQA took charge of HEDIS in 1992
- Expanded in 1996 to cover all three product lines: Commercial, Medicare, and Medicaid
- Recently expanded to include the PPO line of business
- Tool used by more than 90 percent of America's health plans to measure performance on important dimensions of care

Users of HEDIS® Data

- State regulatory agencies
- Accrediting organizations
 - National Committee for Quality Assurance (NCQA)
 - Utilization Review Accreditation Commission (URAC)
- Public and private purchasers
 - Employers
 - Medicare
 - Medicaid
- Health plans
 - RFP/RFI preparation
 - Quality improvement activities
 - Required to obtain NCQA accreditation

HEDIS® Domains

- Effectiveness of Care
- Access/Availability of Care
- Satisfaction with the Experience of Care
- Use of Services
- Cost of Care
- Health Plan Descriptive Information
- Health Plan Stability
- Informed Health Care Choices

Effectiveness of Care Measures

- 41 measures, 10 by survey
- Clinical in nature
- Focus
 - Preventive care
 - Up-to-date treatments for acute episodes of illness
 - Chronic disease care

Consumer Assessment of Healthcare Providers and Systems

- Funded and administered by the U.S. Agency for Healthcare Research and Quality (AHRQ)
- A family of standardized surveys that asks consumers and patients to report on and evaluate their experiences with health care
- Originally stood for Consumer Assessment of Health Plans Study
 - Products have evolved beyond health plans

CAHPS® Survey Products

- Ambulatory care surveys
 - CAHPS Health Plan Survey
 - Commercial
 - Medicaid
 - Medicare
 - CAHPS Clinician and Group Survey
 - Experience of Care and Health Outcomes (ECHO) Survey
 - CAHPS Dental Plan Survey
 - CAHPS American Indian Survey
 - CAHPS Home Health Care Survey

NCQA's Version of CAHPS® Health Plan Survey

- Health plans seeking NCQA accreditation must submit results from NCQA's adult version of the CAHPS Health Plan Survey (4.0H)
 - Part of the Satisfaction with Experience of Care HEDIS® domain
 - CAHPS accounts for 13% of a health plan's total accreditation score

	Accreditation
Global Rating Measures	
Rating of Personal Doctor	✓
Rating of Specialist Seen Most Often	✓
Rating of All Health Care	✓
Rating of Health Plan	✓
Composite Measures	
Getting Needed Care	✓
Getting Care Quickly	✓
How Well Doctors Communicate	✓
Customer Service	✓
Shared Decision Making	
Item-Specific Measures	
Health Promotion and Education	
Coordination of Care	
HEDIS Clinical Measures	
Advising Smokers to Quit	✓
Discussing Smoking Cessation Medications	
Discussing Smoking Cessation Strategies	

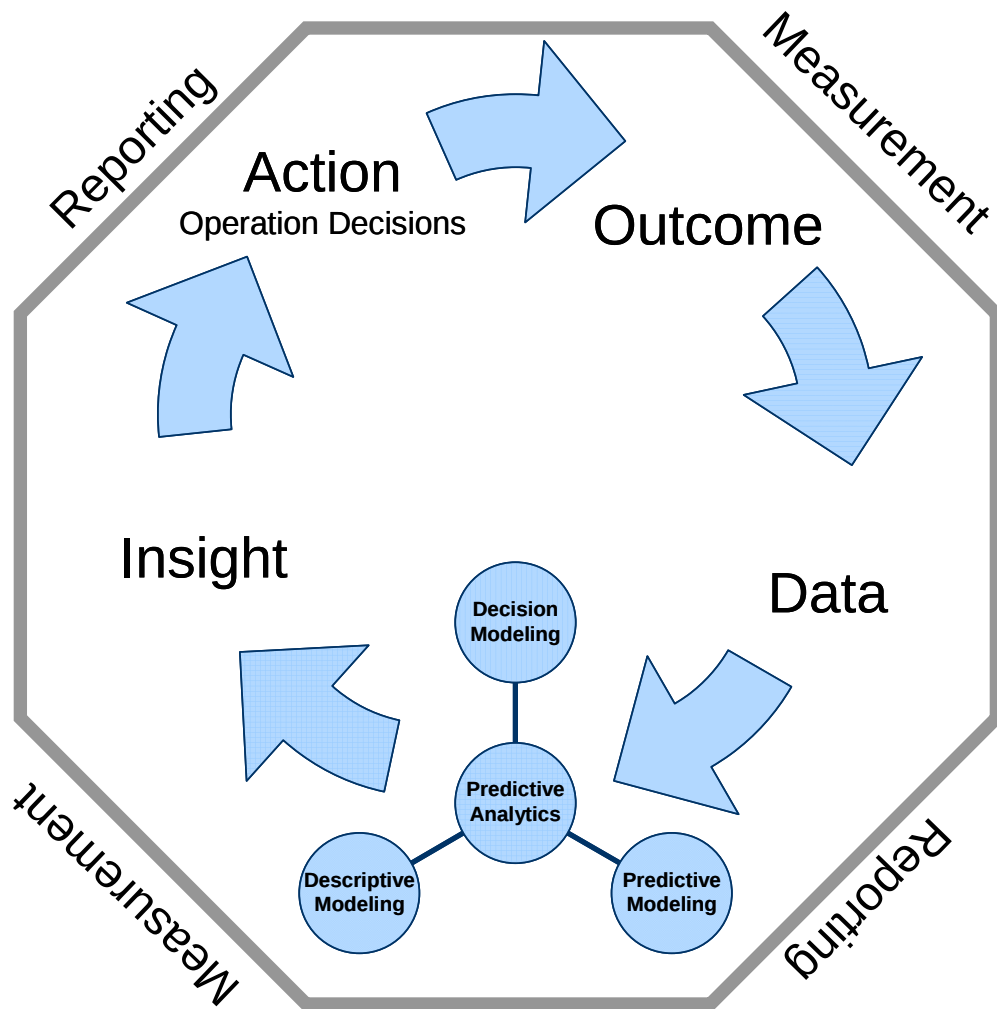
NCQA Accreditation Scoring Methodology

The accreditation score is a weighted average of the standard and performance measure scores.

The weight allocated to the performance measures has increased each year since 2008 and is targeted to be 50% by 2012.

Measure Year	Standards	Performance Measures
2007	64.3%	35.7%
2008	61.3%	38.7%
2009	57.0%	43.0%
2010	54.1%	45.9%

A Proactive Approach to Address HEDIS and CAHPS Rate Improvement



Insight derived from *analytics* that is embedded in the workflow to support *data driven quality of care improvement efforts*

Quality of Care Analytics Framework

- I. Accreditation Analytics Research (AAR) Plan
- II. Data Mining and Predictive Modeling of Quality of Care Measures
- III. Connecting Quality of Care Outreach Efforts with Care Management Efforts – Outreach Optimization Plan (OOP)

Accreditation Analytics Research (AAR) Plan

- Background:
 - Traditional intervention efforts have been both *reactive* rather than proactive and *general* rather than targeted
 - Realized a need to formalize a research and analytics strategy
- Objective:
 - To develop a proactive approach to quality improvement efforts and provide recommendations at the measure level for HEDIS and CAHPS

Accreditation Analytics Research Plan - HEDIS



Methodology

- Conduct a comparative review of the HEDIS and CAHPS 2008 and 2009 results for the measures required for accreditation in 2009
- Score each accreditation measure using established criteria
- Rank the measures in descending order by score
- Select the highest ranked measures for in-depth analysis to determine drivers of compliance/non-compliance
- Recommend areas of focus based on analysis

HEDIS 2009 Accreditation Measures - Comparative Review

HEDIS 2009 Accreditation Measures	HEDIS 2008*		HEDIS 2009**		Change in Rate	Change in Percentile
	Adjusted Rate***	Percentile	Adjusted Rate***	Percentile		
Antidepressant Medication Management - Acute Phase	44.95%	25 th	46.83%	50 th	1.88%	↑
Antidepressant Medication Management – Continuation Phase	30.51%	25 th	29.06%	25 th	-1.45%	-----
Appropriate Treatment for Children with Pharyngitis	71.34%	75 th	72.67%	75 th	1.33%	-----
Appropriate Treatment for Children with Upper Respiratory Infection	79.65%	25 th	81.63%	25 th	1.98%	-----
Avoidance of Antibiotic Treatment in Adults with Acute Bronchitis	32.32%	75 th	32.10%	75 th	-0.22%	-----
Breast Cancer Screening	52.43%	25 th	47.57%	25 th	-4.86%	-----
Cervical Cancer Screening	64.25%	25 th	64.30%	25 th	0.05%	-----
Childhood Immunizations (Combo 2)	82.85%	75 th	83.92%	90 th	1.07%	↑
Cholesterol Management - LDL Screening Only	83.00%	75 th	85.21%	75 th	2.21%	-----
Comprehensive Diabetes Care - Eye Exams	41.73%	25 th	54.40%	50 th	12.67%	↑
Comprehensive Diabetes Care - HbA1c Testing	77.99%	25 th	81.26%	50 th	3.27%	↑
Comprehensive Diabetes Care - HbA1c Poorly Controlled (a lower rate is better performance)	41.42%	50 th	37.52%	75 th	-3.90%	↑
Comprehensive Diabetes Care - LDL-C Screening	70.08%	25 th	72.70%	25 th	4.62%	-----
Comprehensive Diabetes Care - Nephropathy Screening	68.50%	25 th	72.12%	25 th	3.62%	-----
Controlling High Blood Pressure	66.87%	90 th	63.72%	75 th	-3.15%	↓
Follow-Up After Hospitalization for Mental Illness (7 day)	33.63%	25 th	17.59%	<25 th	-16.04%	↓
Advising Smokers to Quit (CAHPS Survey Measure)	77.40%	90 th	78.86%	90 th	1.46%	-----
Persistence of Beta-Blocker Treatment After a Heart Attack	76.08%	50 th	100.00%	90 th	23.92%	↑
Prenatal Care	89.18%	75 th	83.01%	25 th	-6.17%	↓
Postpartum Care	67.65%	90 th	64.41%	50 th	-3.24%	↓
Use of Appropriate Medication for Asthma - Combined Measure	95.33%	90 th	96.17%	90 th	0.84%	-----
Use of Imaging Study for Low Back Pain	83.55%	75 th	82.78%	50 th	-0.77%	↓
Use of Spirometry Testing in the Assessment and Diagnosis of COPD	33.00%	75 th	36.11%	75 th	3.11%	-----

* Adjusted rates and percentiles computed internally

** Adjusted rates and percentiles computed by NCQA

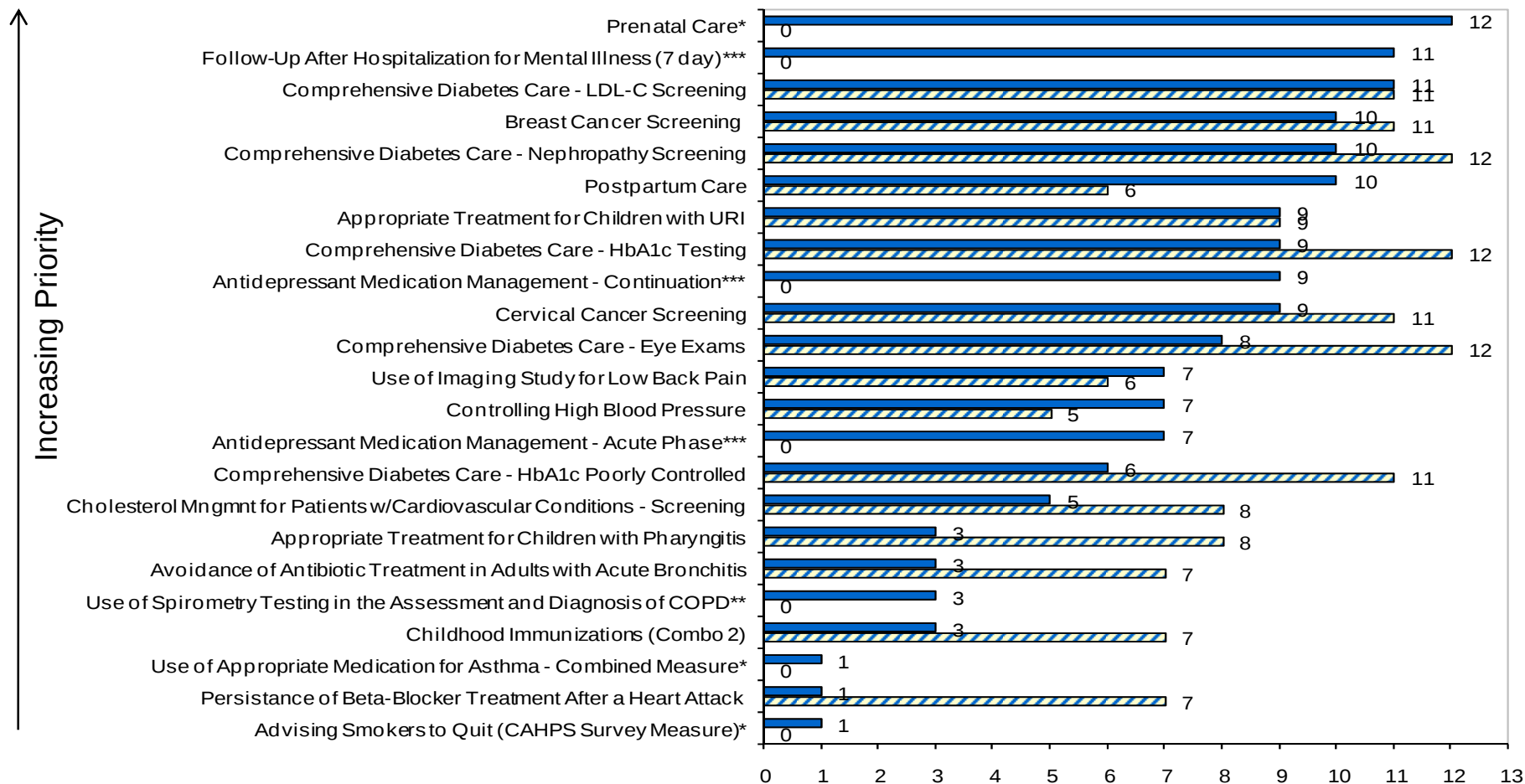
*** Adjusted for sampling and regional variations

Scoring HEDIS Measures

- Some of the criteria used when comparing HEDIS 2008 results to HEDIS 2009 results
 - Significant decrease (>2 % points) in rate = 1 point
 - Drop in percentile = 1 point
 - Current percentile status = 1-5 points (*90th=1, 75th=2, 50th=3, 25th=4, $<25th=5$ points*)
 - Nearest percentile threshold = 1-4 points
(*75th $\leftrightarrow$ 90th=1, N/A=2, 50th $\leftrightarrow$ 75th or $<25th \leftrightarrow$ 25th=3, 25th $\leftrightarrow$ 50th=4*)
 - Use of CPT II codes allowed = 1 point
 - Is this an incentive measure = 1 point
 - Is this a hybrid eligible measure = 1 point

Ranking HEDIS Measures

HEDIS 2009 Priority Measure Rankings



* 2008 Ranking=0 due to having a percentile $\geq 75^{\text{th}}$

** 2008 Ranking=0 due to being a new accreditation measure

***2008 Ranking=0 due to being a carved-out measure



Data Mining and Predictive Modeling

- Background:
 - Allows for segmentation of population into those highly likely and highly unlikely to be compliant with a guideline and apply more focused efforts
 - Enables prioritization of outreach efforts at the member level
- Objective:
 - To conduct population segmentation and determine the likelihood that a member will receive his/her recommended care as per evidenced based guidelines and then use this information in outreach deliverables

Data Mining and Predictive Modeling: A Case Study of Modeling Drivers of Compliance for Breast Cancer Screening



Defining The Problem

Breast Cancer in Tennessee

- Most common cancer in women in TN
- Accounted for 15% of all cancer cases in TN
- 4000 new cases/yr
- Second leading cause of cancer death in women in TN
- *Leading cause of death in women age 35-54 in TN
- 900 deaths are reported each year in TN
- **1 in 8 women diagnosed in their lifetime in TN**

Source: NCI, 2000 estimates

*Age-adjusted to the 2000 US standard population

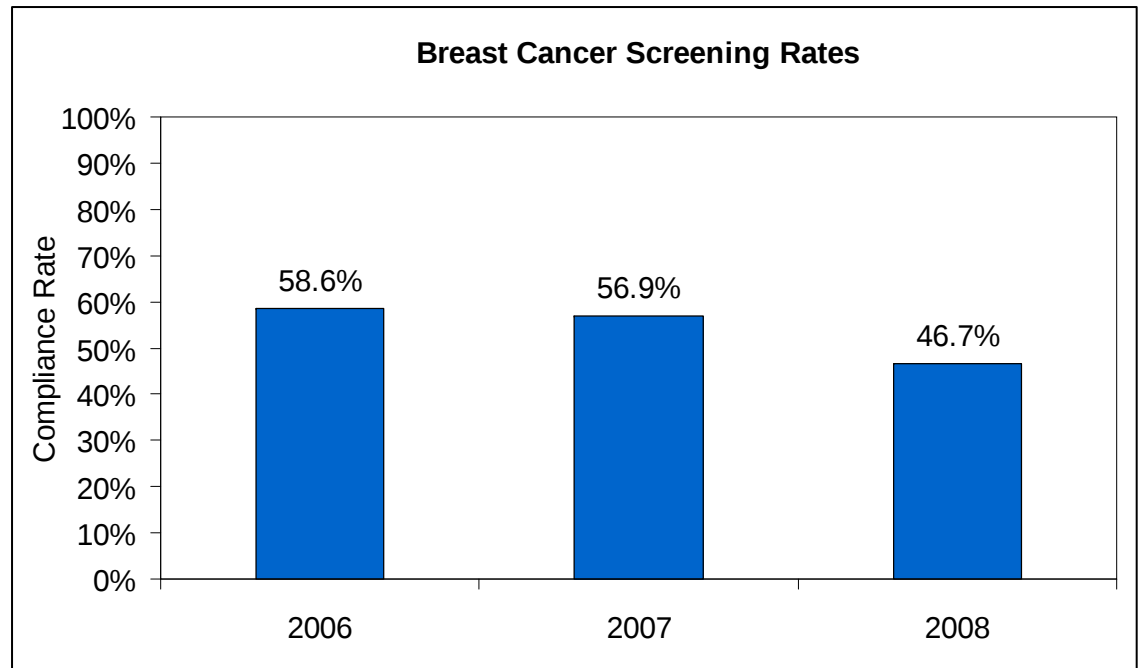
Source: US Mortality Public Use Data Tapes 2006,

Source: National Center for Health Statistics, Centers for Disease Control and Prevention, 2005.



Background – Declining Rates

- Decline in breast cancer screening rates over time
- Multiple intervention programs have been implemented

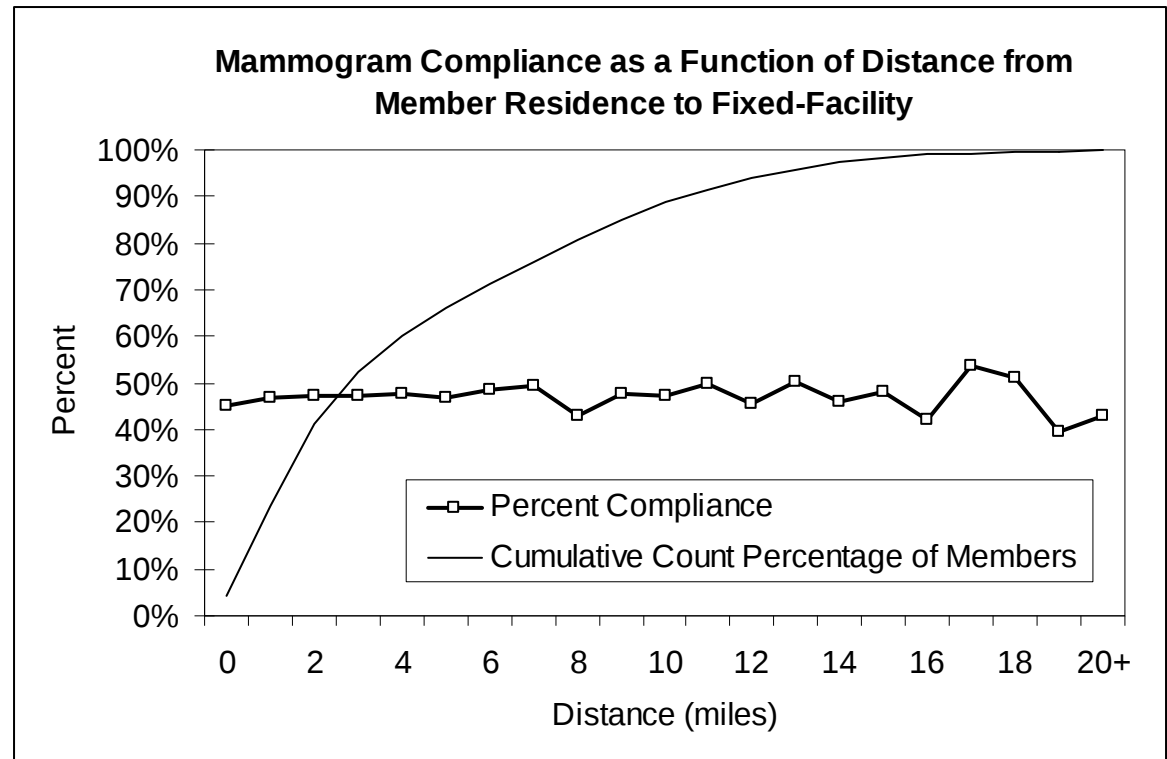


Background – Literature Search

- Mammogram screening rates are influenced by multiple factors:
 - Age, race, ethnicity, income status
 - Continuity of care with PCP (O'Malley et al. 2002)
 - Administering clinic characteristics (county hospital vs. university medical center) (Ramsey et al. 2001)
 - Logistic inconveniences (McBride 1993)
 - Suggestion from health-care professionals, perceived barriers (Champion & Menon 1997)
 - Positive views about initial screening, practice of other preventive health behaviors, knowledge of breast cancer and screening (Soler-Michel 2005)

Compliance as a Function of Distance to Facility

- Compliance rates not influenced by distance to a fixed-facility
- 98% of the study population lived within 15 miles of a fixed facility
- Average distance from home to the nearest facility was not different for compliant members versus non-compliant members



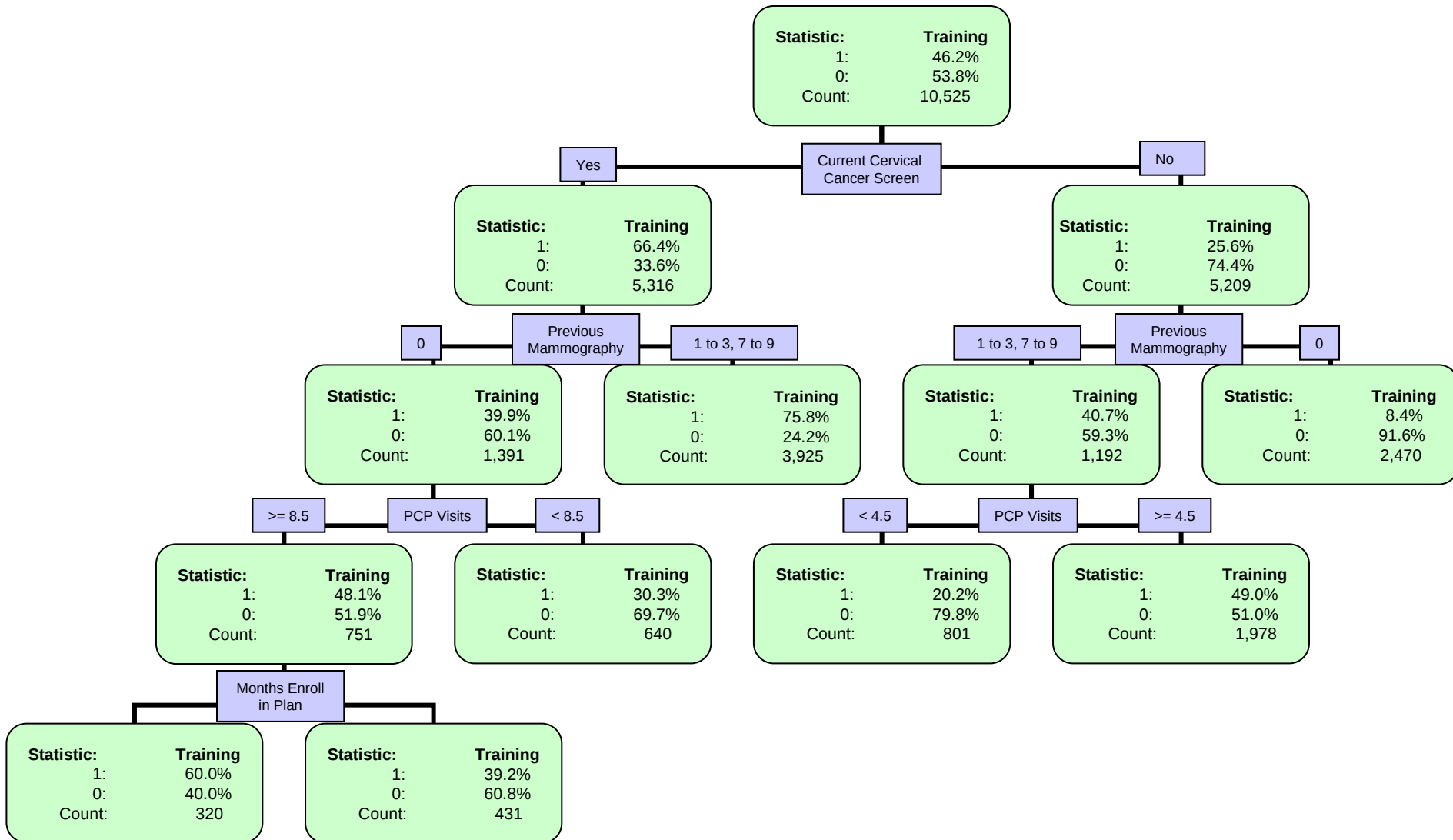
Research Objectives

- Examine mammography screening rates among adult women enrolled in 2007
- Explore driving factors of mammogram screening
- Develop propensity scores capturing likelihood of a woman getting a mammogram
 - Scores to be later used in outreach deliverables

Methods

- Target variable:
 - Compliance with most recent mammography screening as defined by HEDIS specs
- Explanatory Variables:
 - Clinical
 - Demographic
 - Compliance with other evidence-based guidelines

Results – Decision Tree



Application of Models

Creating Member Profiles and Individualized Prevention Strategies

Drivers of Compliance

1. Has a history of obtaining breast cancer screening
2. Has current cervical cancer screening
3. Visits a PCP at least 4 times a year
4. Has less than 3.5 years of enrollment
5. Not obese

Compliant Cathy

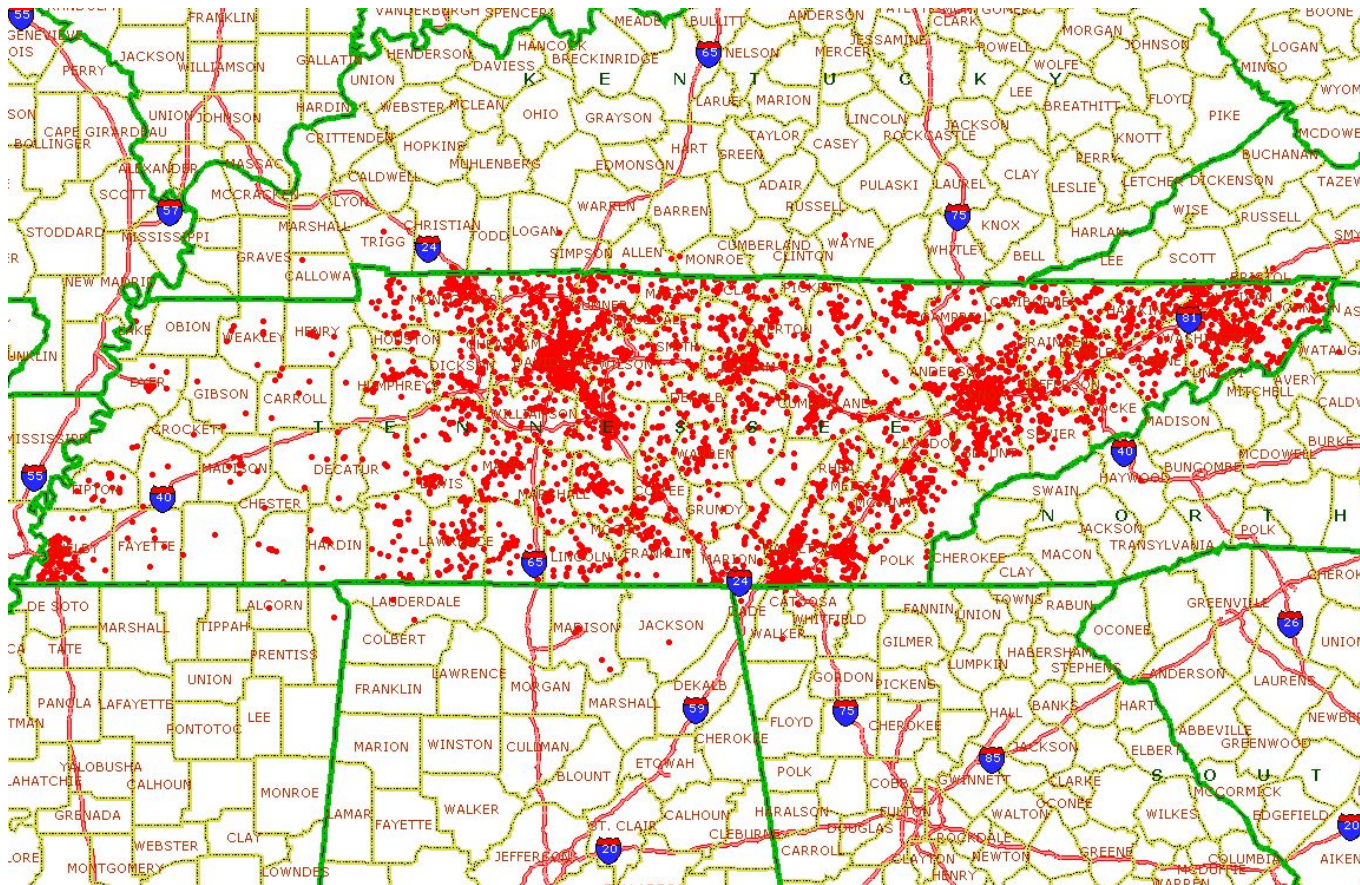


Deploying the Model – Using a GIS

- Using a Geographic Information System (GIS), locate members with a low probability of obtaining a mammography
 - Geo code member locations (latitude, longitude) using residential street address and zip code information
 - Use a spatial clustering tool to identify hot spots

Deploying the Model - Member Locations

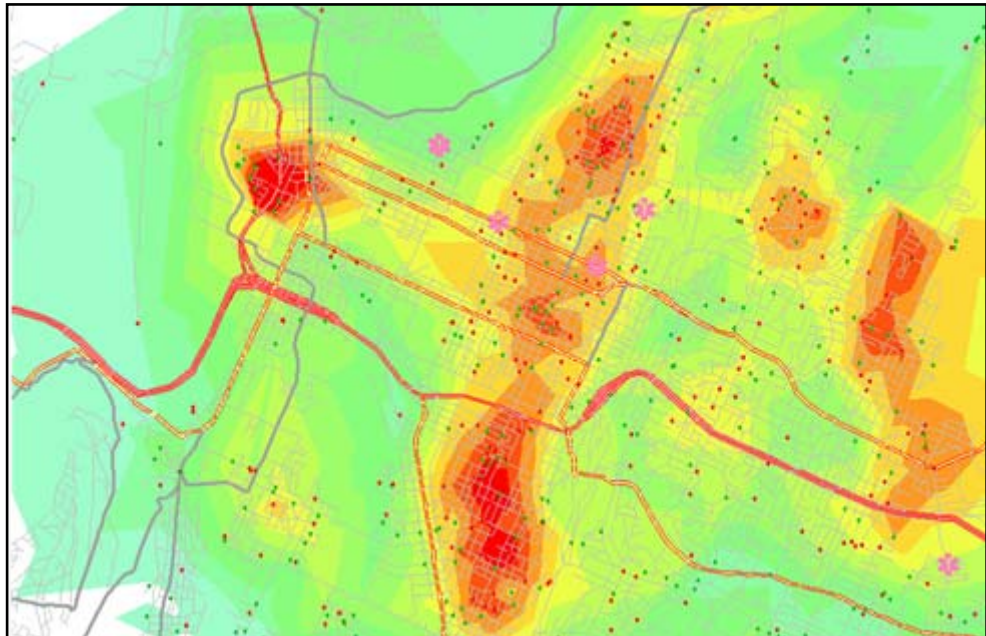
Members with a low probability of obtaining a mammogram



Deploying the Model - Neighborhood Clustering

- Using a patient-centered approach, create surface contour maps (isopleth maps) to define high densities (i.e. neighborhood clusters, or hotspots) of members with a low probability of obtaining a mammogram

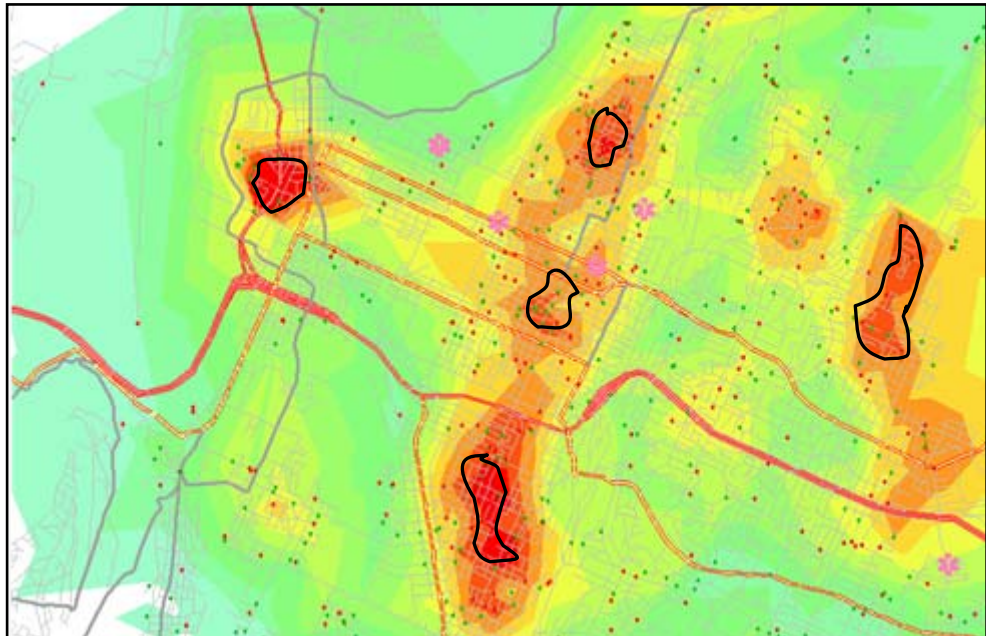
Example of Patient-Centered hotspot clustering for members with a low probability of obtaining a mammogram (red dots)



Deploying the Model - Statistically Testing of Clusters

- Using spatial statistical testing, the number of neighborhood clusters was reduced by 73% when compared to no significance testing

Example of defining statistically significant clusters (black outlines) of non-compliant members



Accreditation Analytics Research Plan - CAHPS



CAHPS 2009 Accreditation Measures - Comparative Review

CAHPS 2009 Accreditation Measures	CAHPS 2008		CAHPS 2009		Change in Score	p-value	Change in Percentile
	Adjusted Score*	Percentile	Adjusted Score*	Percentile			
Rating of Health Plan**	2.49	90 th	2.51	90 th	0.02	0.2906	-----
Rating of All Health Care	2.31	75 th	2.36	90 th	0.05	0.0393	↑
Rating of Personal Doctor	2.53	75 th	2.48	75 th	-0.05	0.0333	-----
Rating of Specialist Seen Most Often	2.66	90 th	2.51	75 th	-0.15	0.0057	↓
Getting Needed Care	2.38	75 th	2.39	75 th	0.01	0.7223	-----
Getting Care Quickly	2.46	90 th	2.43	75 th	-0.03	0.3850	↓
How Well Doctors Communicate	2.54	50 th	2.54	50 th	0.00	0.9501	-----
Customer Service	2.53	90 th	2.60	90 th	0.07	0.2055	-----

* Adjusted for sampling and regional variations. Scores are computed by NCQA.

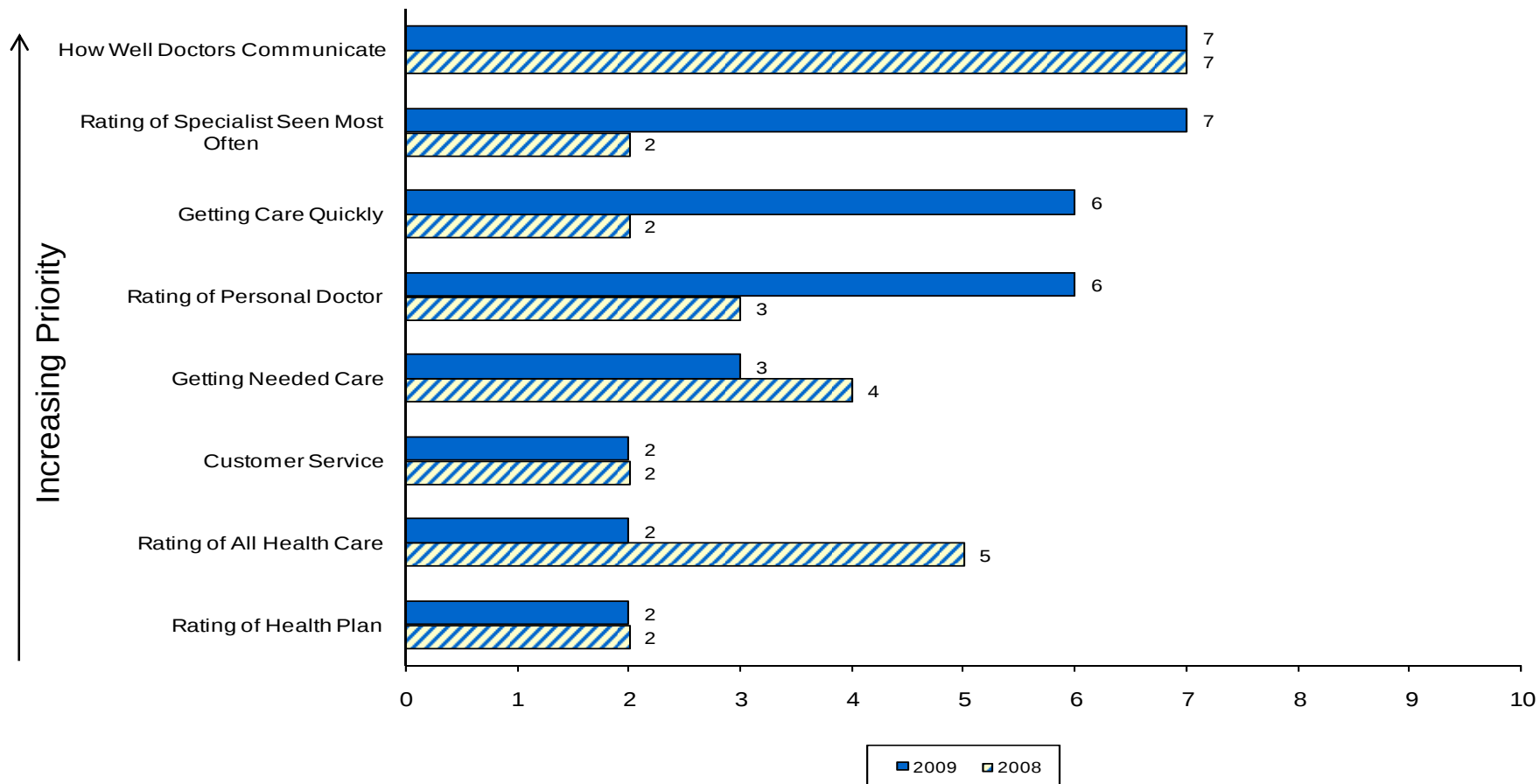
**Counts double for accreditation scoring

Scoring CAHPS Measures

- Some of the criteria used to identify which accreditation measures are in greatest need for quality improvement
 - Statistically significant decrease ($p < 0.10$) in score from 2008 to 2009 = 1 point
 - Drop in percentile from 2008 to 2009 = 1 point
 - Nearest percentile threshold = 1-4 points
($75^{th} \Leftrightarrow 90^{th} = 1$, $N/A = 2$, $50^{th} \Leftrightarrow 75^{th}$ or $< 25^{th} \Leftrightarrow 25^{th} = 3$, $25^{th} \Leftrightarrow 50^{th} = 4$)
 - Current percentile status = 1-5 points ($90^{th} = 1$, $75^{th} = 2$, $50^{th} = 3$, $25^{th} = 4$, $< 25^{th} = 5$ points)

Ranking CAHPS Measures

CAHPS 2009 Priority Measure Rankings



Predictors of Survey Response: A Study of Factors Influencing Response Rates to the Mailed Surveys



Defining The Problem

- Using consumer generated data to maintain & improve access to quality medical care
- Low response rates
 - Average 8-10% for Medicaid, 15-20% for Commercial
 - Financial drain
 - Non-respondent bias
 - Limited ability to publicly report the data
 - Limited ability to engage doctors as to improvement measures

Research Objective

Predict mail survey responders

- Characteristics inherent to participants
 - Demographic
 - Socio-economic
 - Benefit design (Commercial)
 - Disease burden
- Methodological techniques
 - Survey length
 - Cover letter
 - Anonymity
 - Previous respondent burden

Methods

- Models focused on predicting the likelihood of an adult responding to a mail survey
- The target variable represented the event of responding on the first mailing of a survey
- Data was used from the following surveys conducted in 2008:
 - Care Management Surveys
 - Doctor Visit Surveys
 - New Member Surveys
 - Outpatient Surveys
- Historical survey data back to 2006 was used to quantify previous survey burden

Methods

- 62 predictor variables initially considered
 - Benefit design – 41 variables (Commercial only)
 - Deductibles, coinsurance, copays, out-of-pocket max
 - Demographic – 6 variables
 - Gender, age, race, region, urbanicity, population density
 - Disease burden – 4 variables
 - Forecasted costs, prior year costs, risk index, risk category
 - Socio-economic – 4 variables
 - Education, household income, home value
 - Survey characteristics – 4 variables
 - Content, total pages, total questions, anonymity
 - Account characteristics – 2 variables (Commercial only)
 - Fully/self-insured, account type
 - Survey burden – previous 2 years

Methods

- Variable screening:
 - Variables with uniform or nearly uniform distributions were excluded
 - Variables with more than 30 percent missing values were excluded
 - Continuous variables with positively skewed distributions were normalized using the log transformation
 - Categorical variables were collapsed so each level contained approximately an equal number of cases

Methods

- Final variables used in building models:

	Commercial	Medicaid
Benefit Design	Out-of-pocket max Overall in-network coinsurance % Total health and Rx deductible for member	
Disease Burden	Insurance risk category	Insurance risk category Forecasted costs
Demographic	Age Gender Population w/in 1 mile Urbanicity	Age Gender Urbanicity Customer Service Area
Socio-demographic	%BS or higher Median HH income	%BS or higher Median HH income
Survey characteristics	Total surveys received in past 2 years	Survey type Total questions on survey Total surveys received in past 2 years

Methods

- Responding to a mail survey can be considered a rare event
- Unless non-responders are sampled down, models will have very low sensitivity in detecting a responder
- 3 different predictive models were constructed
 - Decision Tree
 - Autonomous tree growth
 - Logistic Regression
 - Stepwise selection (entry p-value=0.2, exit p-value=0.05)
 - Neural Network

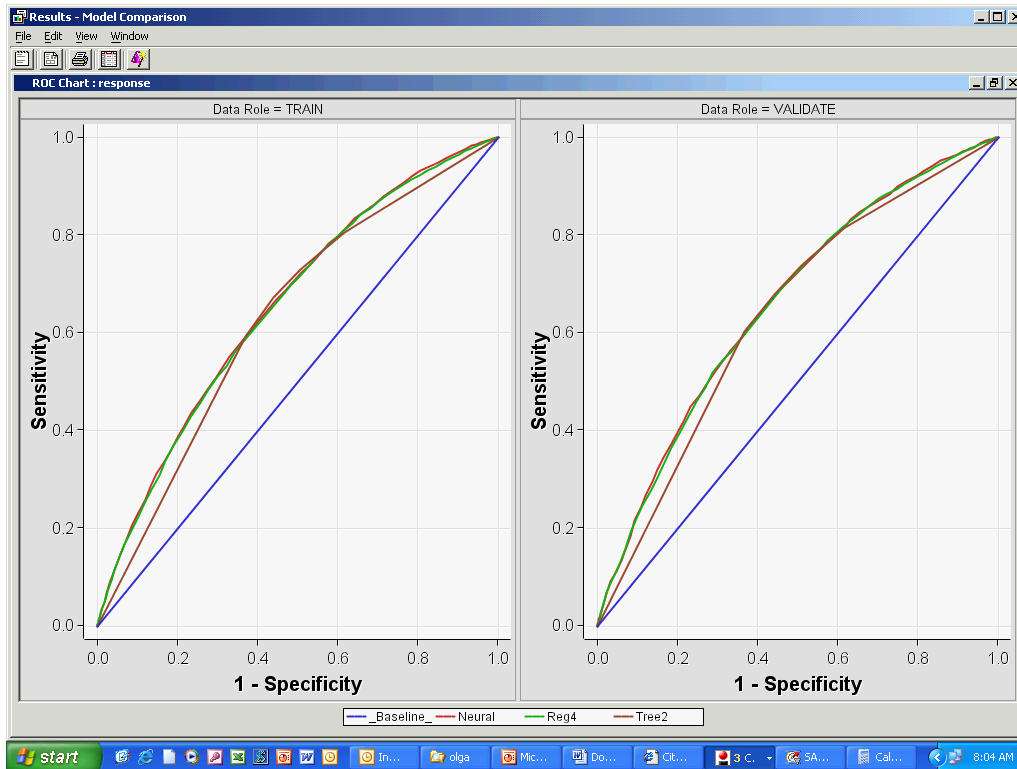
Data was split into training (70%) and validation (30%)

Statistical Terminology

- Positive Predictive Value (PPV) is the proportion of members responding to a survey who were predicted to respond
- Negative Predictive Value (NPV) is the proportion of members not responding to a survey who were **not** predicted to respond
- Misclassification Rate describes how accurate a model is in predicting the disposition of an outcome (e.g. survey response)
- Receiver Operating Curve (ROC) describes how well a model discriminates between positive (e.g. survey responder) and negative (e.g. nonresponder) assignments

Results

- Commercial
 - ROC curves



- Assessment/validation

	Tree	Regression	Neural Net
Misclassification Rate	0.38	0.39	0.39

Results

- Assessment on the scoring dataset (11,428 cases)
 - Logistic Regression
 - PPV = 33%
 - NPV = 82%
 - Overall correctly classified = 61%
 - Decision Tree
 - PPV = 33%
 - NPV = 84%
 - Overall correctly classified = 59%
 - Neural Network
 - PPV = 34%
 - NPV = 82%
 - Overall correctly classified = 63%

Impact

- Assume a sample of 10,000 members
 - Not utilizing predictive model
 - All 10,000 members get mailed a survey
 - Low estimate – 1,500 returned (15% response rate)
 - High estimate – 2,000 returned (20% response rate)
 - Utilizing predictive model
 - 4,000 members get mailed a survey (model flagged 40% of the scoring cases)
 - 1,320 are returned (model had 33% PPV)
 - 6,000 members get a telephone call
 - 1,500 cooperate (telephone surveys average 25% response rate)
 - 2,820 total surveys completed, 28% final response rate
 - **Liberal – 1,320 more surveys completed, 13% gain in response rate**
 - **Conservative – 820 more surveys completed, 8% gain in response rate**

Outreach Optimization Plan (OOP) - Geographic Prioritization for Outreach and Care Management Quality Improvement Efforts

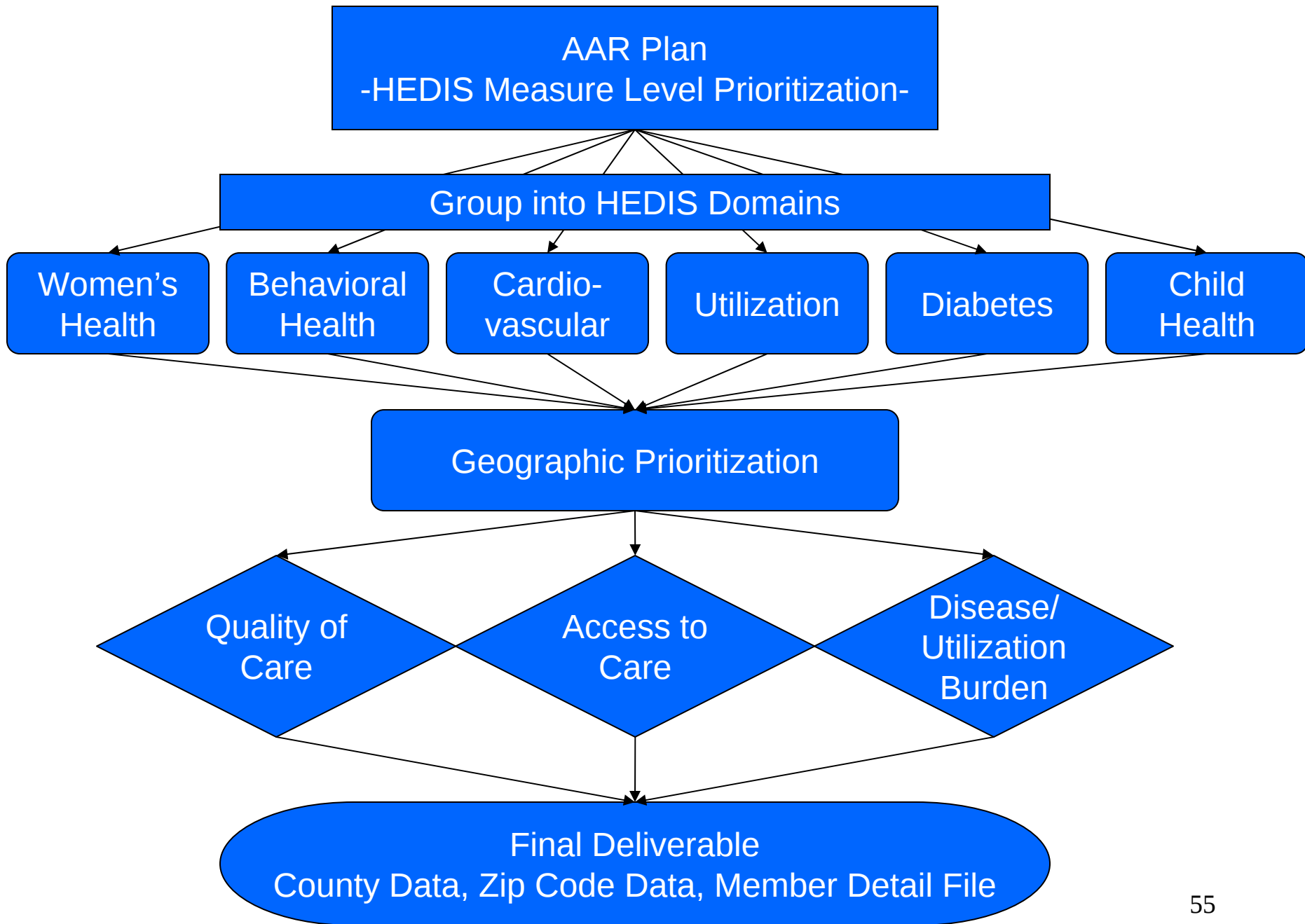


Outreach Optimization Plan (OOP)

- Background:
 - AAR Plan did not address which counties are in most need of intervention
 - Measures identified as priority in the AAR plan span across broad disease areas and impact entire population
 - Factors such as disease and utilization burden, access to care, and opportunities for gaps in care improvement (gaps in care burden) need to be considered in prioritizing intervention counties
- Objective:
 - To develop an analytical deliverable that strategically directs outreach and care management quality improvement efforts to specific areas of the state most in need

Categorizing High Priority Measures

- Measures identified as high priority from the AAR plan were categorized into 6 domains:
 1. Behavioral Health
 2. Cardiovascular Health
 3. Children Health
 4. Comprehensive Diabetes Care
 5. Women Health
 6. Misuse treatment in adults



Methods - Develop Prioritization Ranking System

- Rank counties based on 3 categories:
 1. Quality of Care - Number of gaps in care
 2. Access to Care - Barriers to compliance
 3. Disease / Utilization burden – Prevalence, ER and Admissions
- Multiple ranking factors are examined within each of these 3 categories
- Each county/zip will be scored for each ranking factor
- Counties/zips with higher scores will be higher priority
- Create individual disease category scores and an overall score for multiple optimization options

Methods – Quality of Care

- Use monthly data to address gaps in care
- Prioritize by total number of gaps
- Align measures with HEDIS

Outreach Prioritization: Measure List			
Measure Domain	HEDIS Domain	Report Measure	Measure Description
Behavioral Health	AMM	DepChronic	Antidepressant medication: effective continuation phase treatment
	AMM	DepAcute	Antidepressant medication: effective acute phase treatment
	FUH	DepOVPostIP	Inpatient admission: Office visit within 30 days of admission
	FUH	DepOVPostER	Outpatient emergency room visit: Office visit within 30 days of visit
	FUH	SchzOVPostIP	Follow up within 30 days of IP discharge
	FUH	SchzOVPostER	Follow up within 30 days of ER visit
Cardiovascular Health	CMC	CADLDLScrn	LDL-C screening performed
	PBH	MIBBPersist	Beta blocker persistence after heart attack
Children's Health	AWC	WCTeen12_21	One preventive care physical exam per year for ages 12 to 21
	CIS	HepBVacc	Hepatitis B: 3 vaccinations or Hepatitis B disease or carrier
	CIS	DTaPVacc	DTaP: 4 vaccinations
	CIS	HiBVacc	HiB: 3 vaccinations
	CIS	PneumoVacc	Pneumococcal Conjugate: 4 vaccinations
	CIS	VZVacc	Varicella (VZV): 1 vaccination or evidence of disease
	CIS	IPVacc	Polio (IPV): 3 vaccinations
	CIS	MMRVacc	Measles, Mumps, Rubella (MMR): 3 vaccinations
Comprehensive Diabetes Care	CIS	RotaVacc	Rotavirus: 3 vaccinations
	CDC	Eye Exam	Eye exam (retinal) performed
	CDC	HbA1c	Hemoglobin A1c (HbA1c) testing
	CDC	Nephropathy	Medical attention for nephropathy: screening or evidence of nephropathy
Women's Health	CDC	LDL	LDL-C screening performed
	BCS	Breast	Breast cancer screening: Women 40-69 years
	CCS	Cervical	Cervical cancer screening: Pap test within the previous 2 years
Misuse Treatment in Adults	CHL	Chlamydia	Chlamydia screening, women age 16 to 25
	LBP	LBPXray	Appropriate use of imaging studies
Pediatric Asthma	ASM	AsthmaRx	Persistent asthma: appropriately prescribed medication
	NONE	AsthmaStrd	Inhaled corticosteroid

Methods – Access to Care

- Use pre-established specialty to measure assignments from P4P criteria
 - Assignments were based on evidence of at least 500 episodes being allocated to a particular specialty within a 12 month period
 - Applicable health care providers (HCP) were compared to 3 different populations:
 - Ratio of study members (i.e. members eligible for gaps) to HCP
 - Ratio of overall population members to HCP
 - Ratio of pertinent census population to HCP (e.g. for diabetes, population aged 18-75)
 - Apply a ½ weighting factor to control for counties with low numbers
 - counties with few members were being weighted too heavily

Methods – Disease / Utilization Burden

- Disease prevalence
 - Diabetes, CHF, CAD, Asthma, etc.
- Rank based on the following:
 - Count of members with disease/condition
 - ER visits **due** to the diseases
 - Inpatient stays **due** to the diseases

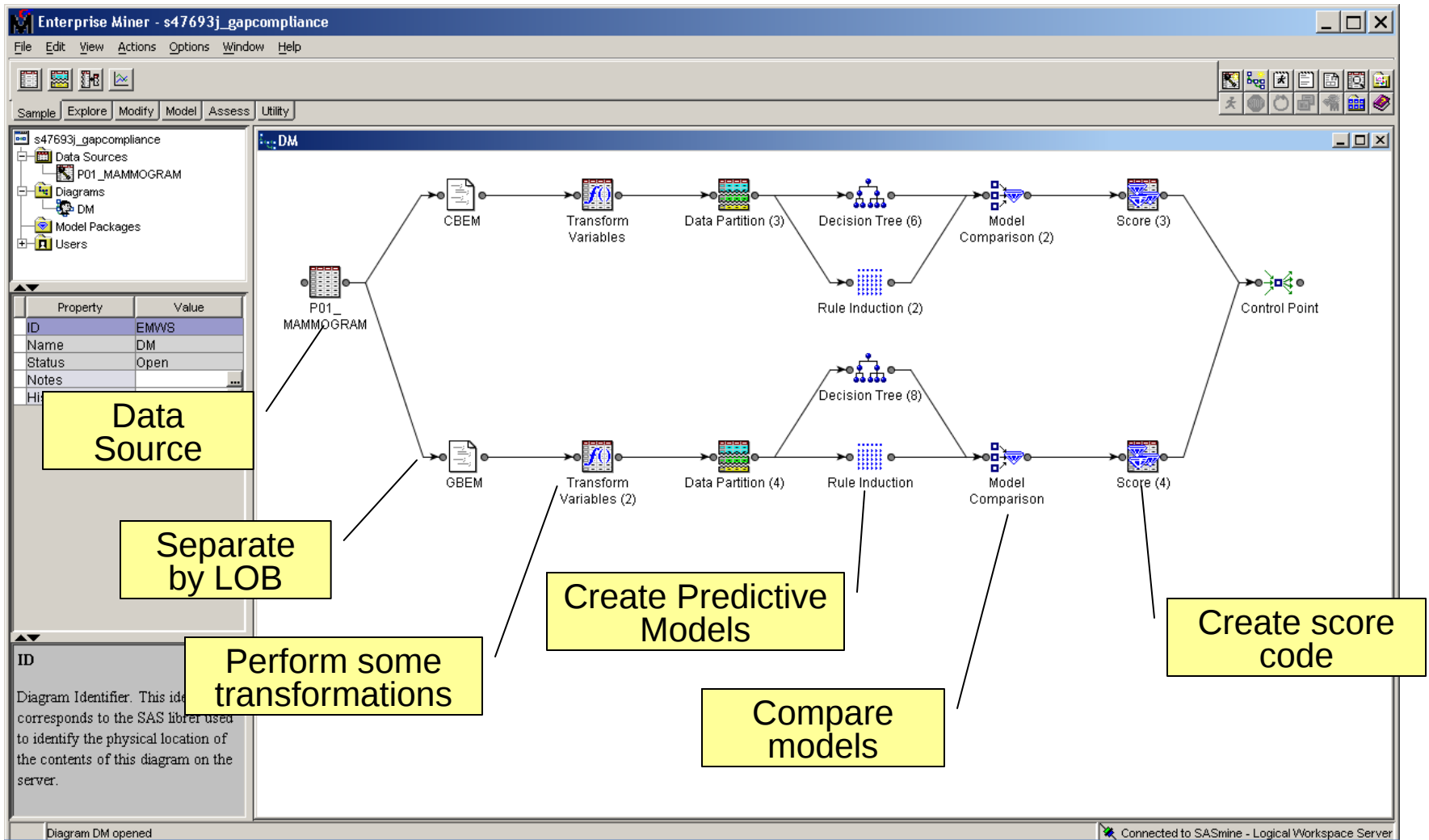
Using Advanced Data Mining Techniques to Predict Adherence to Evidence-Based Guidelines



Methods

- Build predictive models
 - Decision Trees
 - Rule Induction
- Assess models based on performance criteria (e.g. PPV)
- Incorporate likelihood scores in Outreach Optimization Plan deliverables
- Likelihood score can be used to prioritize member engagement

Example of Data Mining Flow



Process Flow

- Process flow to prioritize outreach efforts include:
 - **Step:1** - Identify the county(ies) most in need, using maps and/or disease burden/utilization, number of gaps, and access to care
 - **Step:2** - Drill down to the zip codes within the county(ies) identified in step #1 to get detailed information
 - **Step:3** - Extract the member level information for the counties and/or zip codes identified from steps 1 and 2

Outreach Deliverable

e.g. Tab 1 Deliverable – County Level Data

County	Members with Diabetes	Disease Burden/Utilization					Eye Exam		HbA1c Testing		Rankings			
		Incidence per 1,000	IP Stays	IP Days	ER Visits	Office Visits	# Non-Compliant	% Compliance	# Non-Compliant	% Compliance	Compliance	Disease	Access to Care	Overall
A	2,649	66.6	203	903	338	18,882	1,767	33.3%	648	75.5%	1	1	17	1
B	1,083	88.4	65	343	213	7,310	796	26.5%	288	73.4%	3	2	61	2
C	1,105	93.7	43	215	198	6,776	809	26.8%	314	71.1%	4	3	51	3
D	89	100.8	2	18	17	623	68	23.6%	18	79.1%	5	4	91	4
E	650	113.2	33	125	163	4,690	439	32.5%	153	76.5%	6	5	41	5
F	128	114.6	1	2	7	649	69	46.1%	31	75.8%	33	37	2	6
G	352	130.8	9	31	53	3,032	237	32.7%	77	78.1%	8	7	14	7
H	115	117.6	6	30	17	954	88	23.5%	22	80.9%	37	18	4	8

•County A is worst overall county

e.g. Tab 2 Deliverable – Zip Code Level

County	ZIP	Members with Diabetes	Disease Burden/Utilization					Eye Exam		HbA1c		Rankings			
			Incidence per 1,000	IP Stays	IP Days	ER Visits	Office Visits	# Non-Compliant	% Compliance	# Non-Compliant	% Compliance	Compliance	Disease	Access to Care	Overall
A	12345	9	61.2	0	0	0	44	7	22.2%	4	55.6%	252	572	351	410
A	12346	5	69.4	0	0	0	38	3	40.0%	1	80.0%	533	557	484	583
A	12347	76	79.8	13	54	26	672	49	35.5%	14	81.6%	65	14	77	20
A	12348	5	106.4	0	0	1	57	2	60.0%	0	100.0%	592	415	163	483
A	12349	52	124.7	1	3	3	292	39	25.0%	10	80.8%	88	165	184	113
A	12350	1	35.7	0	0	0	7	1	0.0%	0	100.0%	564	620	542	625
A	12351	83	95.2	1	2	10	524	49	41.0%	21	74.7%	40	128	75	70
A	12352	2	52.6	0	0	0	3	2	0.0%	1	50.0%	451	605	305	572
A	12353	18	62.3	4	16	4	129	13	27.8%	0	100.0%	304	94	47	135
A	12354	3	120.0	0	0	0	16	2	33.3%	1	66.7%	451	434	357	488
A	12355	4	117.6	1	4	0	22	2	50.0%	1	75.0%	438	248	330	372
A	12356	1	83.3	0	0	0	1	1	0.0%	0	100.0%	598	543	444	595
A	12357	11	71.0	0	0	1	82	9	18.2%	2	81.8%	369	501	126	409
A	12358	39	75.1	2	4	11	372	25	35.9%	9	76.9%	129	126	19	100
A	12359	3	103.4	0	0	0	13	1	66.7%	0	100.0%	511	490	375	531

Outreach Deliverable

e.g. Tab 2 Deliverable – Zip Code Level

Location		Members with Diabetes	Disease Burden/Utilization					Eye Exam		HbA1c		Rankings			
County	ZIP		Incidence per 1,000	IP Stays	IP Days	ER Visits	Office Visits	# Non-Compliant	% Compliance	# Non-Compliant	% Compliance				
A	12345	9	61.2	0	0	0	44	7	22.2%	4	55.6%				
A	12346	5	69.4	0	0	0	38	3	40.0%	1	80.0%				
A	12347	76	79.8	13	54	26	672	49	35.5%	14	81.6%				
A	12348	5	106.4	0	0	1	57	2	60.0%	0	100.0%				
A	12349	52	124.7	1	3	3	292	39	25.0%	10	80.8%				
A	12350	1	35.7	0	0	0	7	1	0.0%	0	100.0%				
A	12351	83	95.2	1	2	10	524	49	41.0%	21	74.7%				
A	12352	2	52.6	0	0	0	3	2	0.0%	1	50.0%				
A	12353	18	62.3	4	16	4	129	13	27.8%	0	100.0%				
A	12354	3	120.0	0	0	0	16	2	33.3%	1	66.7%				
A	12355	4	117.6	1	4	0	22	2	50.0%	1	75.0%				
A	12356	1	83.3	0	0	0	1	1	0.0%	0	100.0%	598	543	444	595
A	12357	11	71.0	0	0	1	82	9	18.2%	2	81.8%	369	501	126	409
A	12358	39	75.1	2	4	11	372	25	35.9%	9	76.9%	129	126	19	100
A	12359	3	103.4	0	0	0	13	1	66.7%	0	100.0%	511	490	375	531

•E.g. Now that Zip Code 12345 of County A has been identified as highest priority, use Tab 3 auto-filters to drill-down into the member detail

Member #	First Name	Last Name	...	Address	ZIP code	...	Provider Last Name	Provider First Name	EyeExam	HBA1C	NephScrn	LDL	EyeExam Probability	HBA1C Probability	NephScrn Probability	LDL Probability	...
12345678	Hope	Nukes	...	...	12345	...	...	...	Compliant	Compliant	Compliant	Compliant	Compliant	Compliant	Compliant	Compliant	...
12345679	Chris	Jeffery	...	...	12345	...	...	...	Non-Compliant	Compliant	Compliant	Compliant	Non-Compliant	Compliant	Compliant	Compliant	...
12345680	Harold	Timms	...	...	12345	...	...	...	Non-Compliant	Compliant	Compliant	Compliant	Compliant	Compliant	Compliant	Compliant	...
12345681	Ian	Moots	...	...	12345	...	...	...	Compliant	Compliant	Compliant	Compliant	Compliant	Compliant	Compliant	Compliant	...
12345682	Ibis	Sevens	...	...	12345	...	...	...	Non-Compliant	Compliant	Compliant	Compliant	Compliant	Compliant	Compliant	Compliant	...
12345683	Heather	Hampton	...	...	12345	...	...	...	Non-Compliant	Compliant	Compliant	Compliant	Compliant	Compliant	Compliant	Non-Compliant	...

Predictive model outcome

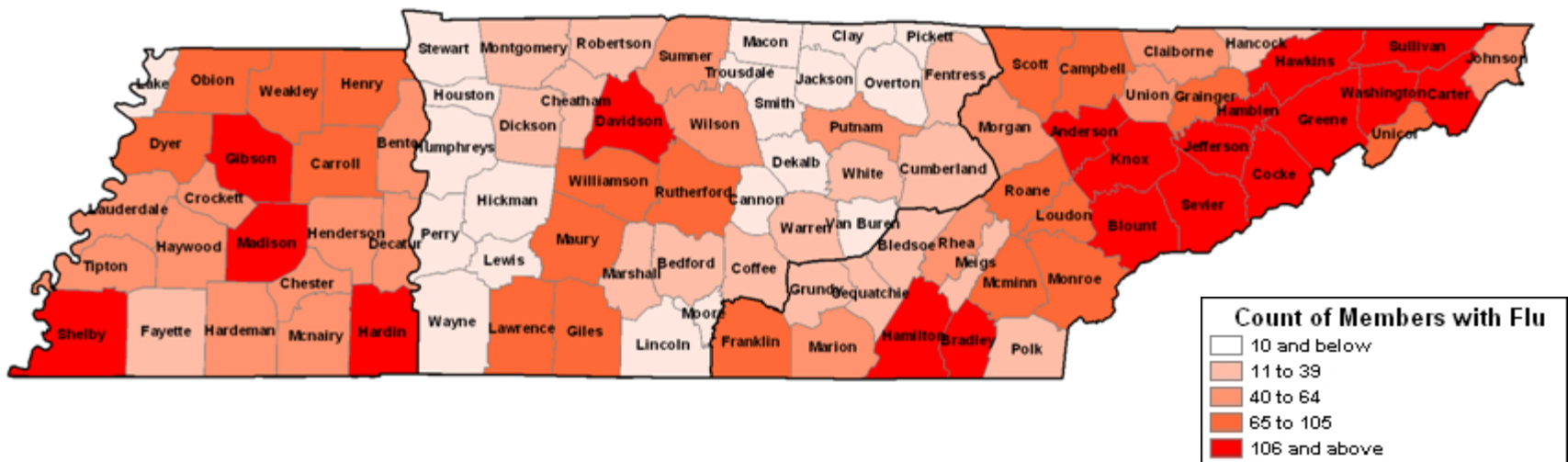
Member Detail includes mbr#, name, contact info, assigned PCP, compliant status with all measures, LOB, Stratification Tier, IP/ER flag, CM/DM flag, HEDIS continuous enrollment indicator, Propensity Score if applicable

Geographic Prioritization: A Case Study - Flu Initiative



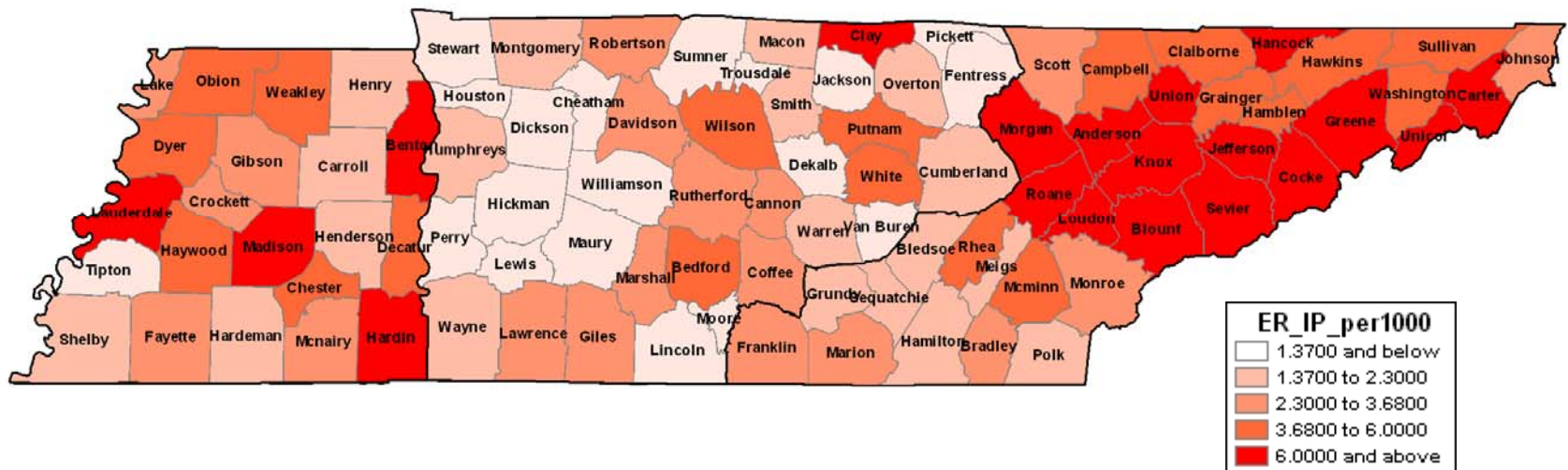
Count of Members with Flu

- **Count of Members** = Number of unique members with a flu diagnosis count for report period: June 2008 – May 2009



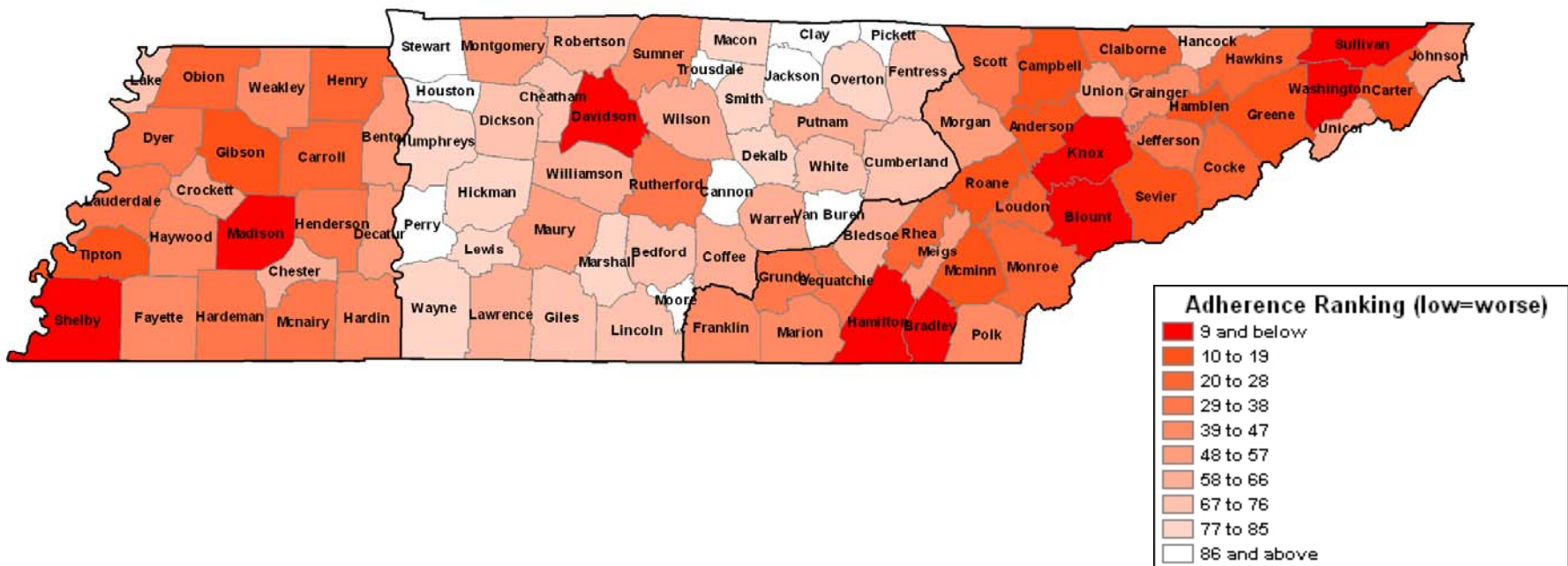
Utilization Burden Related to Flu

- **Utilization Burden** = Number of ER visits and Inpatient Stays per 1000 related to flu diagnosis for report period: June 2008 – May 2009



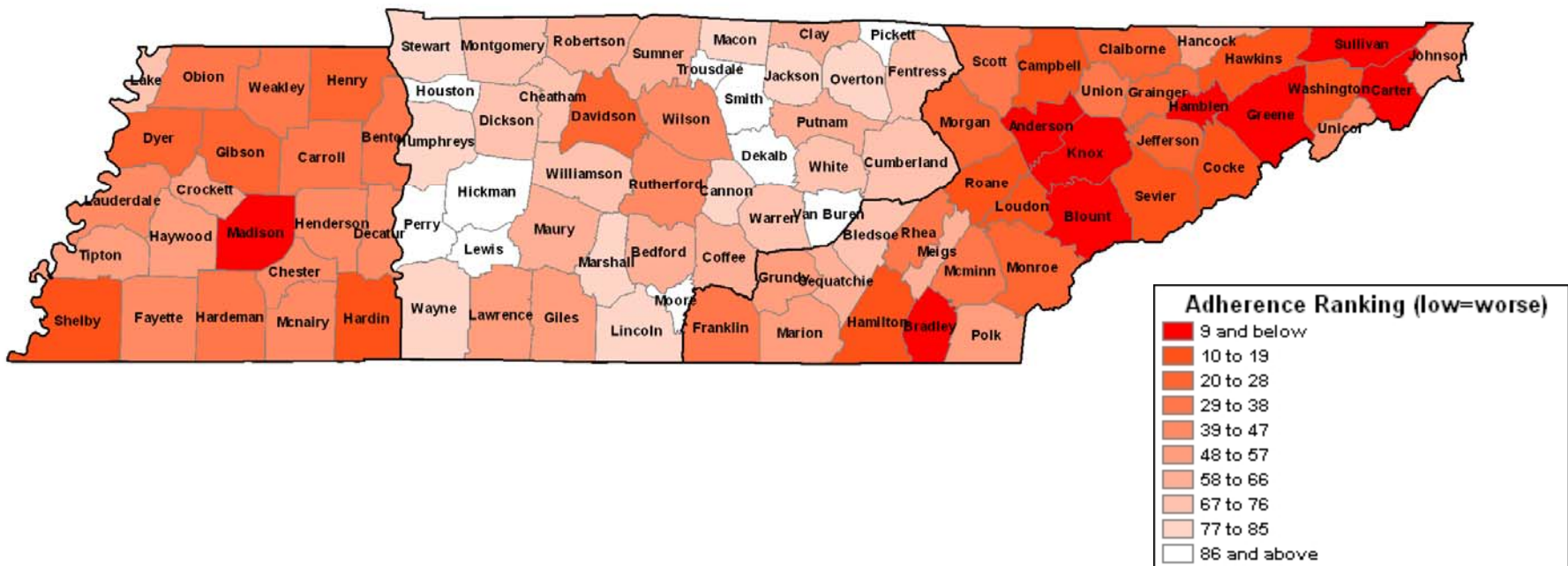
Adherence to Flu Vaccines

- Adherence Ranking** = Rank of County from 1 to 95 where 1 is worst related to overall adherence to flu vaccines for report period: June 2008 – May 2009



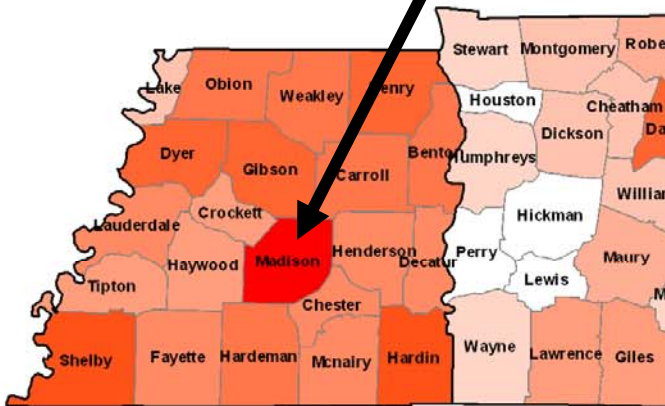
Create Focus Areas

- **Overall Focus Areas:** Combine flu prevalence, utilization burden and adherence rankings to determine priority counties



Create Focus

- Overall Focus Areas: utilization burden and determine priority of



Name	MADISON TN
MbrCount	14242
Inpatient_Stays	1
Inpatient_Days	2
ER_Visits	89
Office_Visits	306
Unique_Members_with_Flu	289
[num_Non-Compliant_ChildFluVac]	1629
num_Eligible_ChildFluVac	1908
pcnt_Compliant_ChildFluVac	0.1462
[pcnt_of_All_Non-Compliant_ChildFluVac]	0.0318
[num_Non-Compliant_PreSrFluVac]	422
num_Eligible_PreSrFluVac	497
pcnt_Compliant_PreSrFluVac	0.1509
[pcnt_of_All_Non-Compliant_PreSrFluVac]	0.0196
[num_Non-Compliant_SeniorFluVac]	16
num_Eligible_SeniorFluVac	16
pcnt_Compliant_SeniorFluVac	0.0000
[pcnt_of_All_Non-Compliant_SeniorFluVac]	0.0189
[num_Total_Non-Compliant_AllFlu]	2067
num_Total_Eligible_AllFlu	2421
pcnt_Compliant_AllFlu	0.1462
[pcnt_of_All_Non-Compliant_AllFlu]	0.0281
Inpatient_Stays_per_1000Mbrs	0.0702
ER_Visits_per_1000mbrs	6.2491
Flu_Prevalence	20.2921
ChildFluVac	4
PreSrFluVac	7
SeniorFluVac	8
AllFlu	4
Inpatient_Stays_per_1000MbrsRnk	26
ER_Visits_per_1000mbrs1	15
Flu_Prevalence1	35
MbrCountwithFlu	7
Avg_All_Ranks	5



(low=worse)