Using GIS Tools to Target Health Disparities Interventions

Zachary Vernon

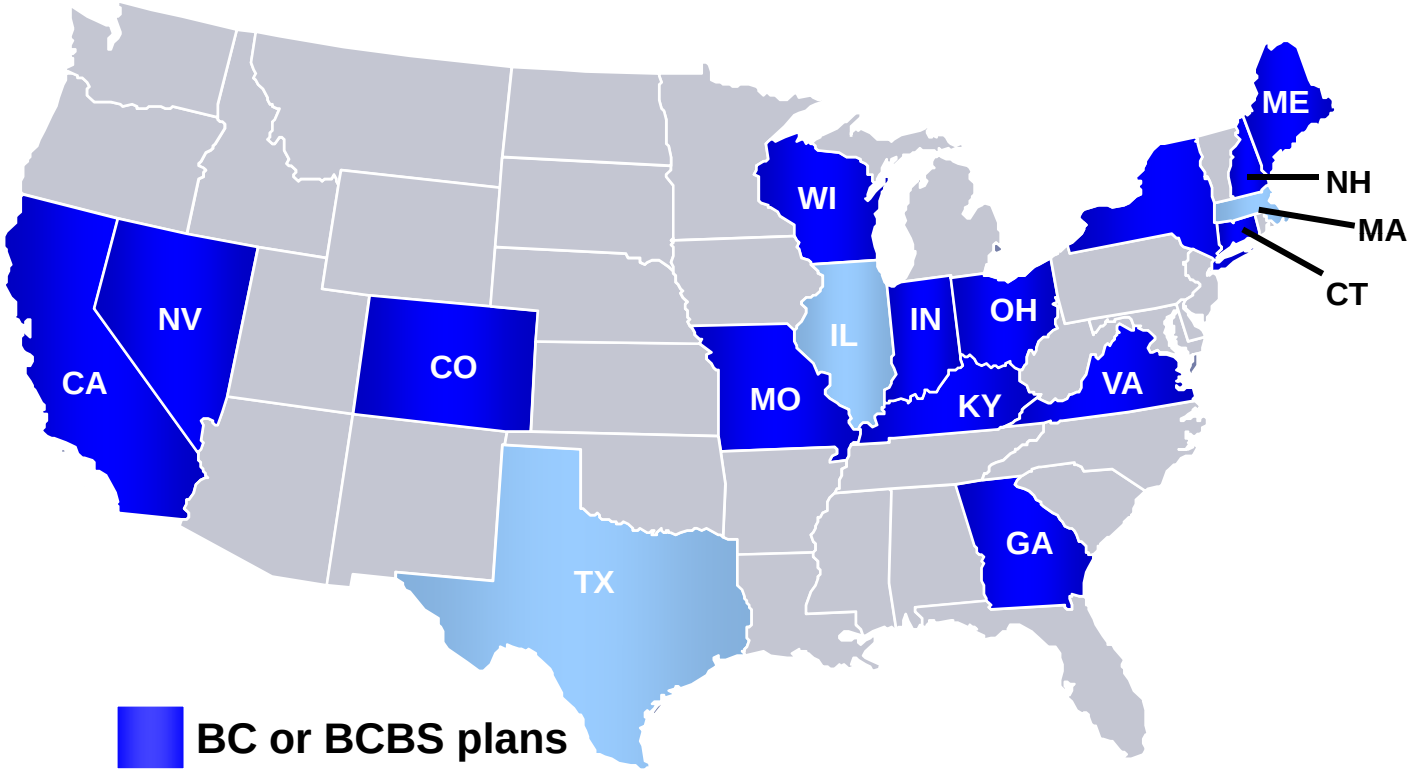
Sr. Health Information Consultant

Programs in Clinical Excellence

2008 ESRI Health GIS Conference
September 29, 2008

WellPoint, Inc.

35 Million Members Across the Country



WellPoint
nextRx
Lumenos
Health Care Enlightened
HMC
building healthier lives
HealthCore
AIM
American Imaging Management

Company Confidential | Copyright by WellPoint, Inc. | Do Not Copy Without Permission

WELLPOINT

Health Disparities – A Few Statistics

- African Americans have the highest rate of high blood pressure of all racial/ethnic groups and tend to develop it younger than others
- Compared to Whites, African Americans and Hispanic are more than twice as likely to have diabetes
- In 2002 in the United States more than 2 1/2 times as many African American newborn babies died (14.1 per 1,000 live deaths) as white newborns
- Among African American Mothers with college educations, infant mortality rates are almost 3 times higher than the rates for White mothers with the same level of education

Projected Life Expectancy from Birth - 1999

	White	Asian	Hispanic	African American
Male	74.7	80.8	77.1	68.3
Female	80.1	86.5	83.7	75.1

* Based on Data from U.S. Census

Company Confidential | Copyright by WellPoint, Inc. | Do Not Copy Without Permission

The Problem We Face

- The collection of data on race, ethnicity, and primary language of enrollees are key steps in the implementation of targeted interventions for improving overall quality and reducing racial and ethnic disparities in health and health care
- Fundamental barriers to the collection of such data exist
  - Commercial plans, which cover the majority of insured in the United States, still have limited race/ethnicity data on their enrolled members
  - Enhancements and upgrades of the technological infrastructure and databases needed to house the data are expensive investments
  - Redesign of work flows and data processes to collect the data systematically also require significant investment in human resources and training costs
  - The actual process can take years to complete, or reach a saturation point where a portion of the insured will always decline to provide such data

Company Confidential | Copyright by WellPoint, Inc. | Do Not Copy Without Permission

Addressing the Problem

WellPoint, Inc. (WellPoint) historically did not collect data in regard to the race/ethnicity of it members.

- Established the Health Disparities Analytic Unit
- Consulted with the RAND Corporation to learn more about indirect models and GIS Tools
- Developed a proxy methodology to derive indirect race and ethnicity information for members
- Used GIS Tools to along with traditional analysis techniques to plan quality interventions

Company Confidential | Copyright by WellPoint, Inc. | Do Not Copy Without Permission

What is Proxy Methodology

- At Wellpoint, we developed a statistical methodology that allows the prediction of race/ethnicity
- The methodology utilizes
  - Enrollee first and last names
  - Enrollee residential address
- First names are matched to an internally developed list of African American first names
- Surnames are matched to a series of surname list from the U.S. Census
- Through geocoding, addresses are linked to Census data based upon an enrollees neighborhood
- These factors are linked together using logistic regressions to determine the probability that a person is Asian, African American, Hispanic, or White/other

Company Confidential | Copyright by WellPoint, Inc. | Do Not Copy Without Permission

Components of Indirect Methodology

Surname Lists

- In 2007 the U.S. Census published a list of all names appearing more than 100 times in the 2000 Census
- Demographic data was also included for these names

Name	Rank	White (%)	African American (%)	Asian Pacific Islander (%)	Native American (%)	2 or More Races (%)	Hispanic (%)
SMITH	1	73.35	22.22	0.40	0.85	1.63	1.56
RODRIGUEZ	9	5.52	0.54	0.58	0.24	0.41	92.70
WILLIAMS	3	48.52	46.72	0.37	0.78	2.01	1.60
RODRIGUEZ	9	5.52	0.54	0.58	0.24	0.41	92.70
LEE	22	40.09	17.41	37.83	1.03	2.30	1.34
KIM	109	2.60	0.36	94.52	0.03	1.99	0.50

* Data from US Census Bureau - http://www.census.gov/genealogy/www/freqnames2k.html

Company Confidential | Copyright by WellPoint, Inc. | Do Not Copy Without Permission

First Name Lists

- Surnames are not as helpful in identifying African American enrollees
- Developed first name list to help
- First name lists could be developed to identify other ethnicities

Name	Rank	African American (%)
MICHAEL	1	12.80
LATOYA	27	89.69
EBONY	29	77.75

* Data developed internally by Anthem Blue Cross

Company Confidential | Copyright by WellPoint, Inc. | Do Not Copy Without Permission

Components of Indirect Methodology

Geocoding

- Determines the physical location of an address (latitude, longitude)
- Currently use a vendor for data scrubbing and geocoding
- Allows linking to Census Data for a specific census tract or block
- Other Census data such as wages, education, language, etc. can also be added

Los Angeles County, CA – Census Tract 1041.07 – Block Group 1					
Total Population	**White / Other (%)**	**African American (%)**	**Asian (%)**	**Hispanic (%)**	**Median Household Income**
552	68.5	8.7	2.4	20.4	$46,042

* Data from 2000 Census

Company Confidential | Copyright by WellPoint, Inc. | Do Not Copy Without Permission

Components of Indirect Methodology

Logistic Regression

- Combines name and location data
- Calculates probability that enrollee is White/Other, African American, Asian, and Hispanic

Member	White / Other (%)	African American (%)	Asian (%)	Hispanic (%)
A	72.5	7.4	7.7	12.4
B	3.6	0.5	2.0	93.8
C	42.4	54.7	0.9	2.0

Company Confidential | Copyright by WellPoint, Inc. | Do Not Copy Without Permission

Is Indirect Methodology Really Reliable?

Accuracy Rate – Population Aggregate

Aggregate Data

•Indirect methods have high degree of accuracy when comparing population groups

Aggregate Demographics – Predicted vs. Reported

Approach N = 192,096	Hispanic (%)	Asian (%)	Black (%)	White / Other (%)
Surname Only	46.1	6.6	7.1	40.2
Geocoding Only	41.3	7.9	11.9	39.0
WellPoint Model	52.1	7.9	14.7	25.3
Member Self – Reported Data	52.0	8.0	14.8	25.2

Company Confidential | Copyright by WellPoint, Inc. | Do Not Copy Without Permission

WELLPOINT

Accuracy Rate – Individual Member Level

Individual analysis

•Indirect methods show strong potential in identifying the R/E of individual members (86.2% of the validation set identified correctly)

Individual R/E Correctly Predicted

Approach N = 192,096	**All Members Accuracy Rate**	**Over 80% Accuracy Rate**	**Over 80% Member Count**
Surname	74.5%	90.8%	130,150
Geocoding	62.9%	84.4%	54,092
Logistic Model	86.2%	93.8%	145,009

Company Confidential | Copyright by WellPoint, Inc. | Do Not Copy Without Permission

WELLPOINT

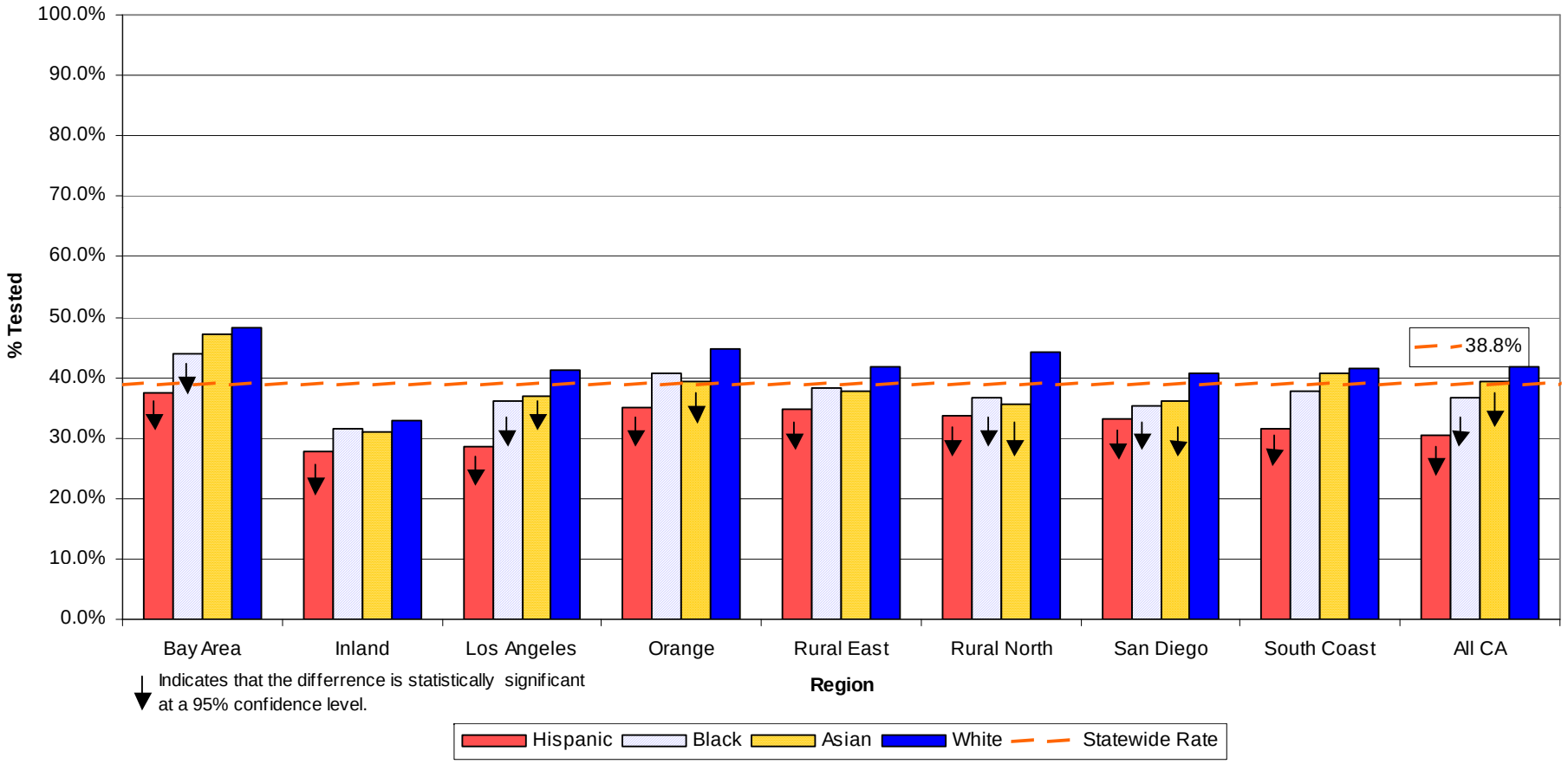
Combining Indirect Methodology & HEDIS

- HEDIS data is a good starting point for investigating health disparities
  - Data already exists
  - Focused on key health elements
- By Applying Proxy Methodology HEDIS rates can be estimated by Race/Ethnicity
- Traditional Analysis and GIS Tools allow visualization and investigation on many levels
  - Region
  - County
  - Zip code
  - Provider

Company Confidential | Copyright by WellPoint, Inc. | Do Not Copy Without Permission

Traditional Rate Analysis

Colorectal Cancer Screening by CA Region

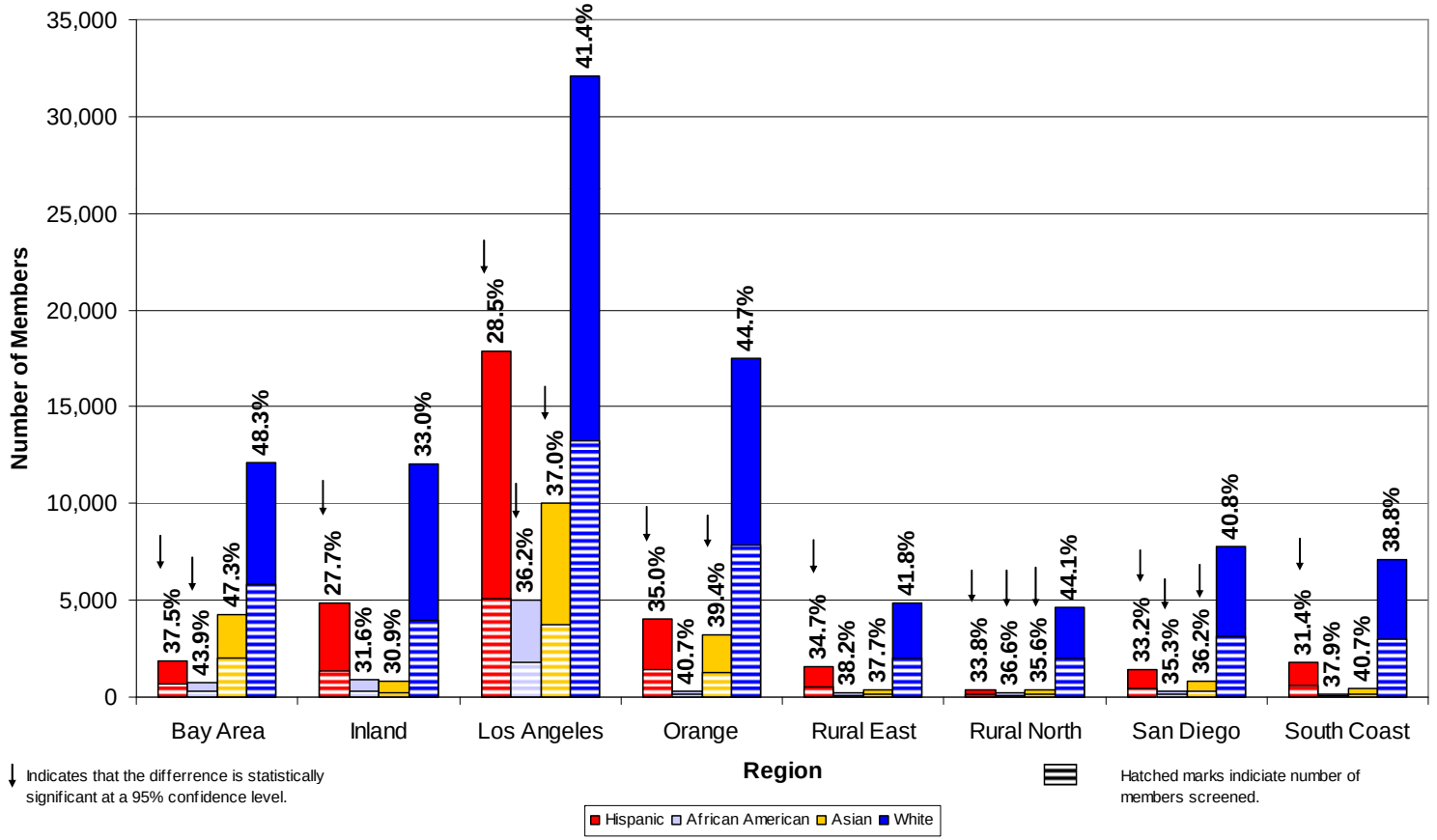


Company Confidential | Copyright by WellPoint, Inc. | Do Not Copy Without Permission

WELLPOINT

Analysis of Volume

Colorectal Cancer Screening Populations and Screening Rates by CA Region



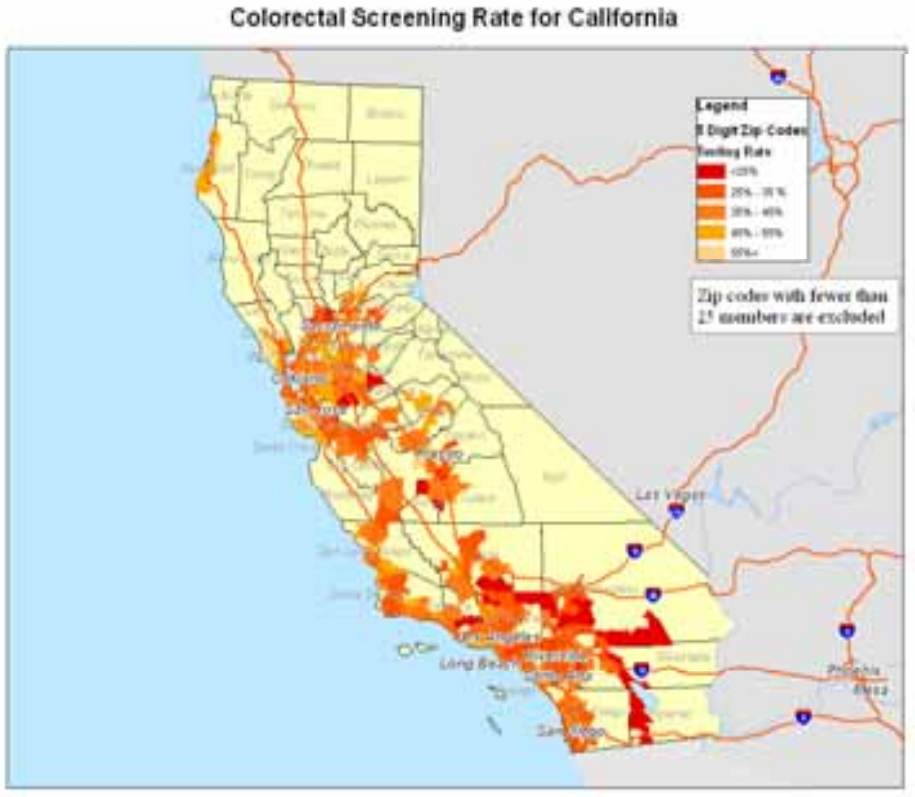
Company Confidential | Copyright by WellPoint, Inc. | Do Not Copy Without Permission

WELLPOINT

Intervention and Implementation

Applications of GIS Mapping Tools

- Geocoding has the added benefit of allowing members to be tied to Census and geographic data
- Additional, Census data can include education levels, language proficiency, income, etc.
- GIS applications facilitate visual drill down to community and individual hotspots

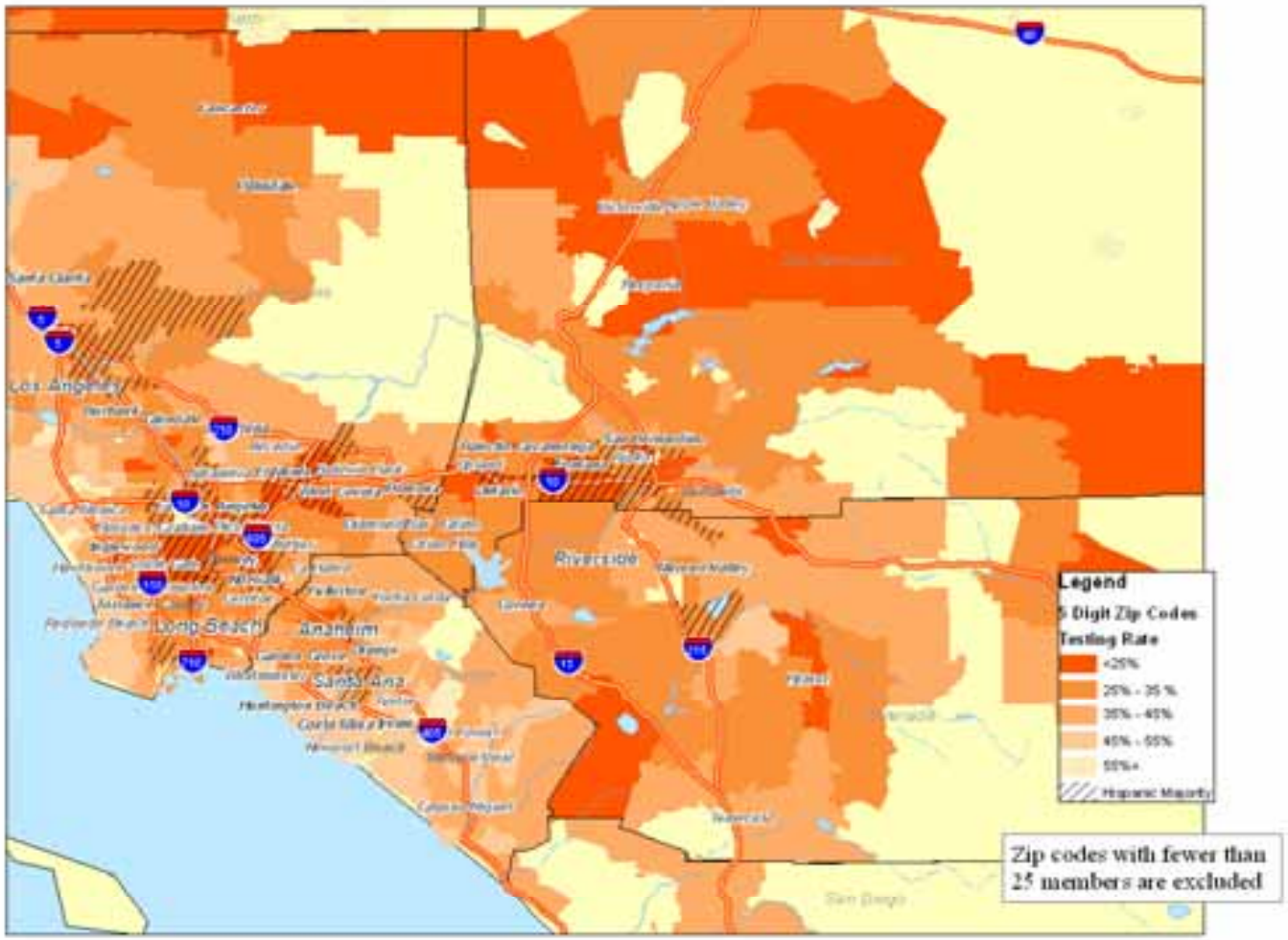


Company Confidential | Copyright by WellPoint, Inc. | Do Not Copy Without Permission

WELLPOINT

Drilldown to Region

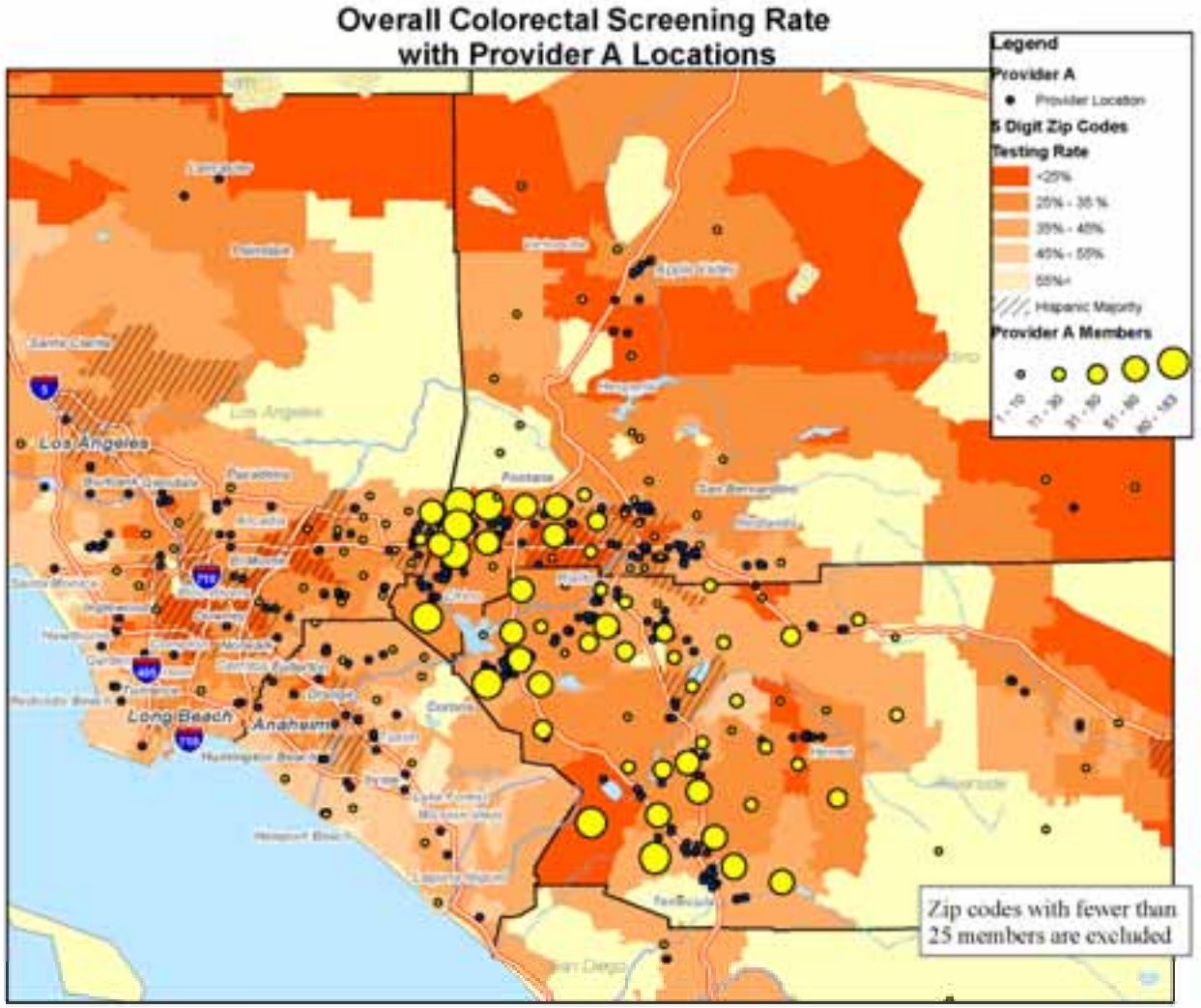
Overall Colorectal Screening Rate



Company Confidential | Copyright by WellPoint, Inc. | Do Not Copy Without Permission

WELLPOINT

Provider Locations



Company Confidential | Copyright by WellPoint, Inc. | Do Not Copy Without Permission

WELLPOINT

Provider Dashboard

Provider Dashboard
Colorectal Screening

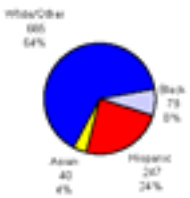
Provider & Region Provider A (Inland)

Rating	Overall	Disparity
	Above Average	Yes

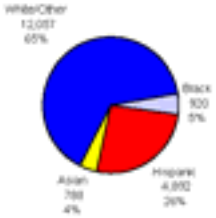
Provider R / E Mix (Number of Patients)

Black	Hispanic	Asian	White/Other	Total
79	247	40	665	1,031

Estimated Racial / Ethnic Demographics
Colorectal Screenings
Provider A
(Inland)

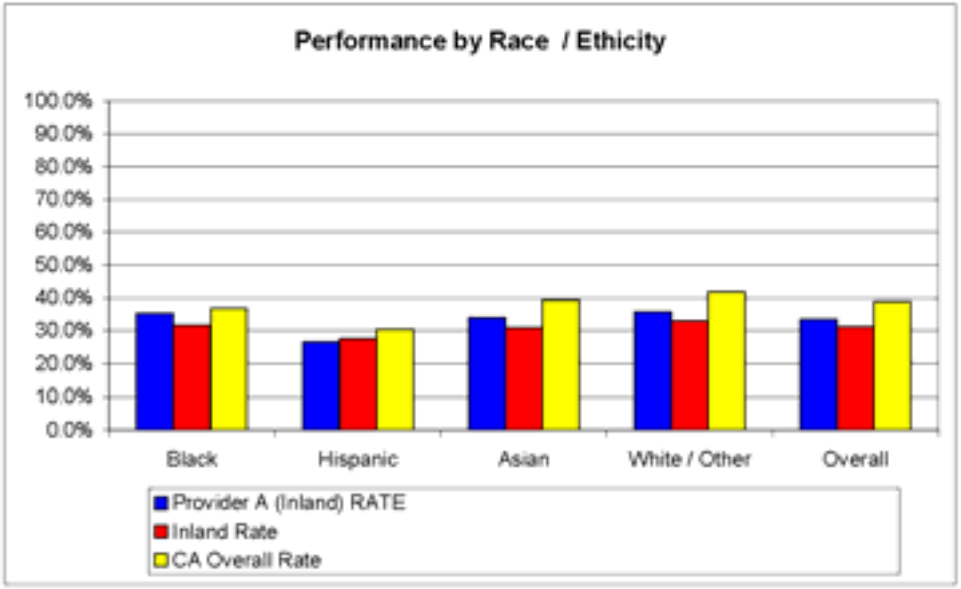


Estimated Racial / Ethnic Demographics
Colorectal Screenings
Inland



Overall Performance	
CA Overall Rate	38.8%
Inland Rate	31.1%
Provider A (Inland) RATE	33.4%
2.3% Above Average Inland Rate	

Performance By R / E	Black	Hispanic	Asian	White / Other	Overall
CA Overall Rate	36.6%	30.2%	39.9%	41.9%	38.8%
Inland Rate	31.6%	27.7%	30.9%	33.0%	31.1%
Provider A (Inland) RATE		35.2%	26.6%	34.0%	35.7%
Difference vs. Inland Rate	3.1%	-1.1%	3.1%	2.3%	2.3%
Difference vs. White / Other Treated by Provider	-0.8%	-0.1%	-1.1%		
Disparity vs. White / Other Treated by Provider (95% Confidence)	No	Yes	No		

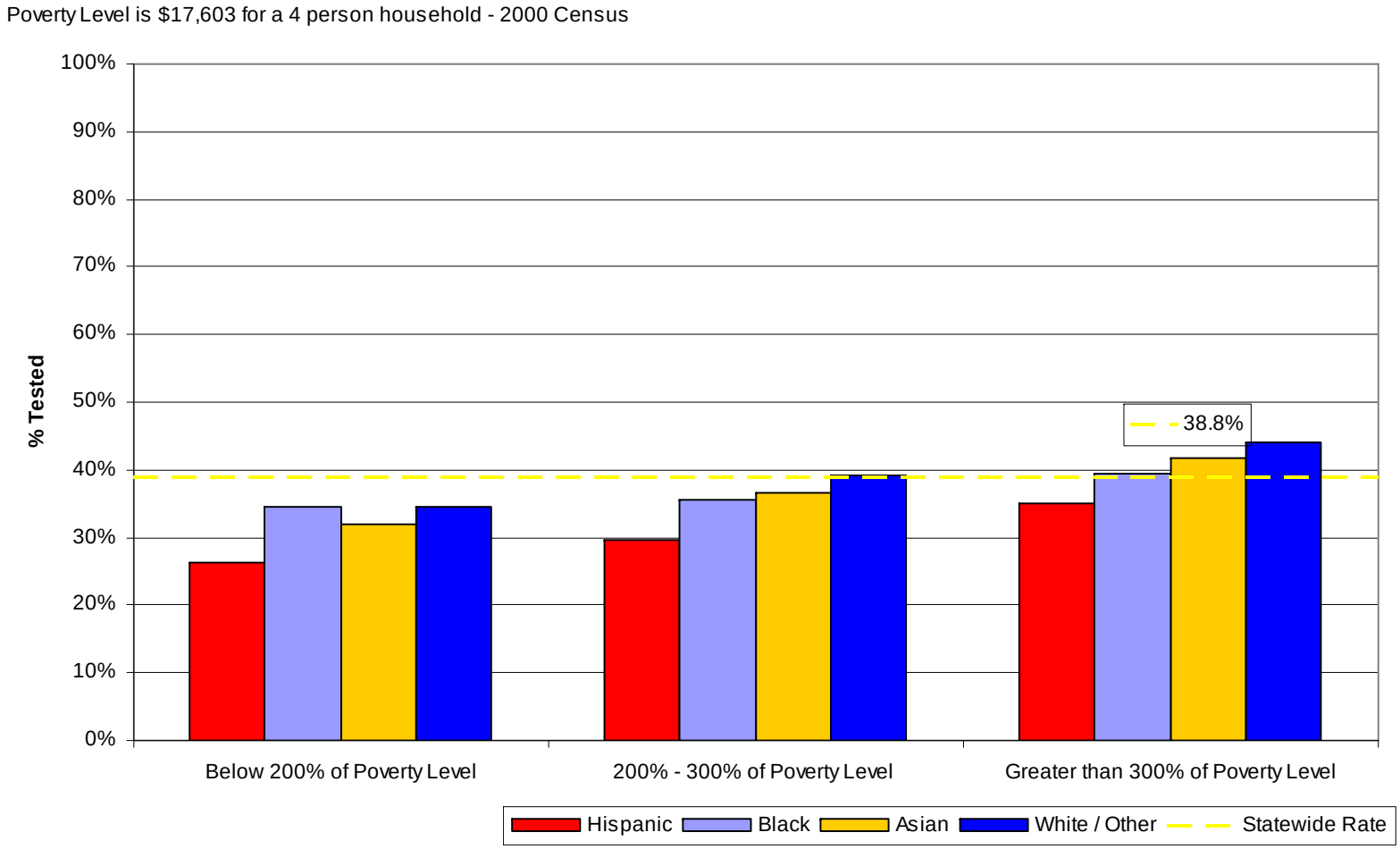


Company Confidential | Copyright by WellPoint, Inc. | Do Not Copy Without Permission

WELLPOINT

Geocoding – Provides a Link to Data

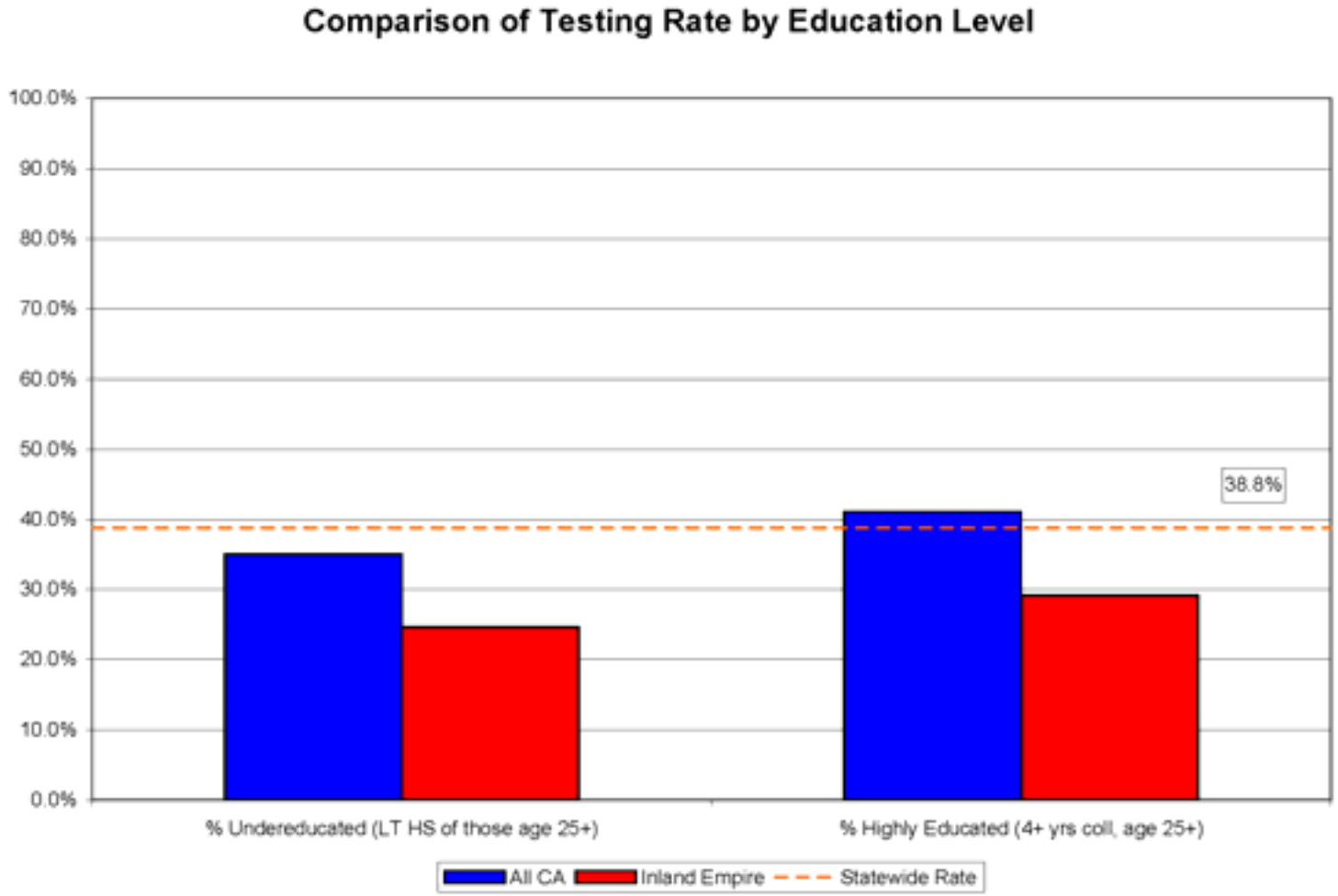
Colorectal Screening Rates by Poverty Level and Race



Company Confidential | Copyright by WellPoint, Inc. | Do Not Copy Without Permission

WELLPOINT

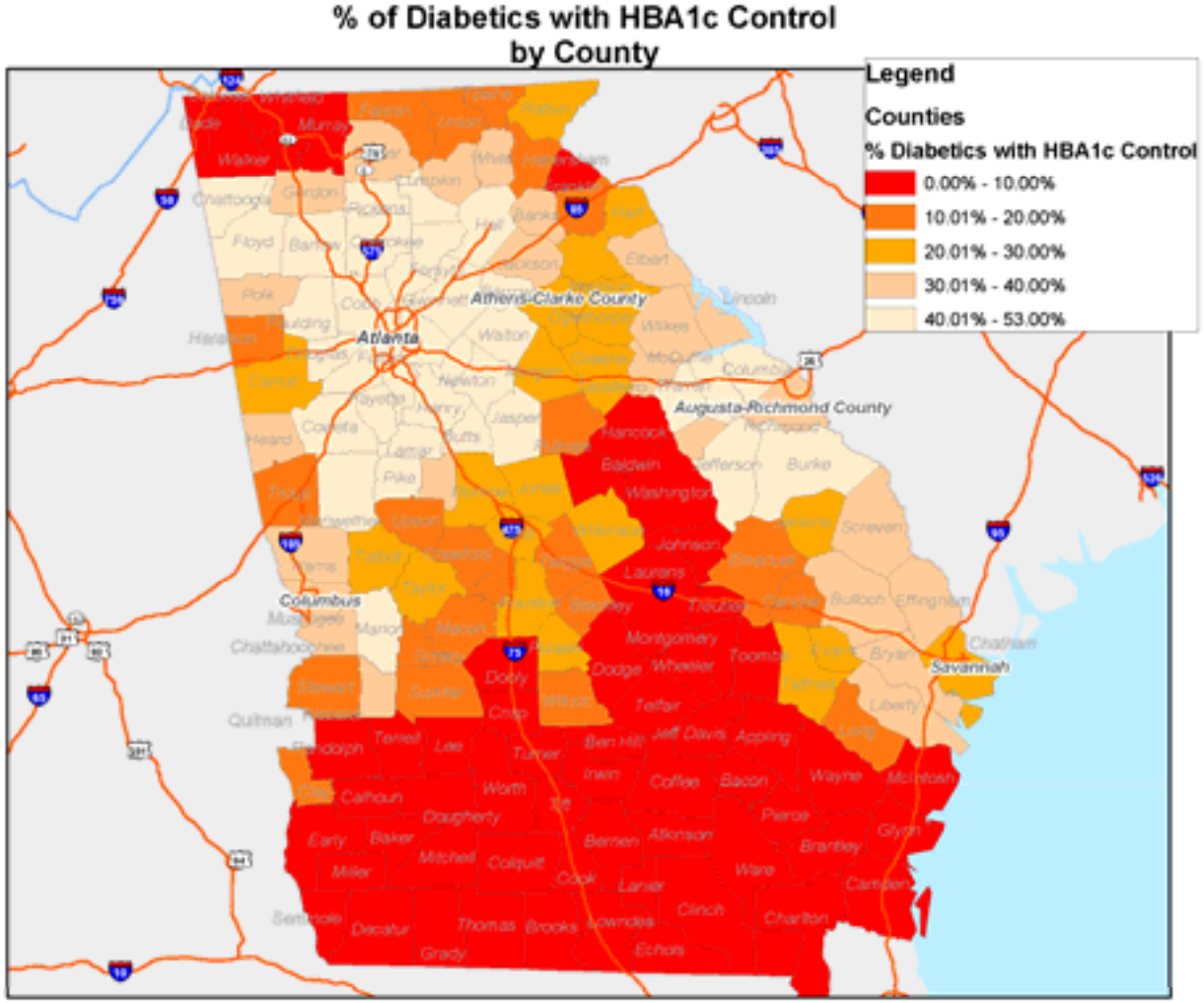
Census Data Provides Insight



Company Confidential | Copyright by WellPoint, Inc. | Do Not Copy Without Permission

WELLPOINT

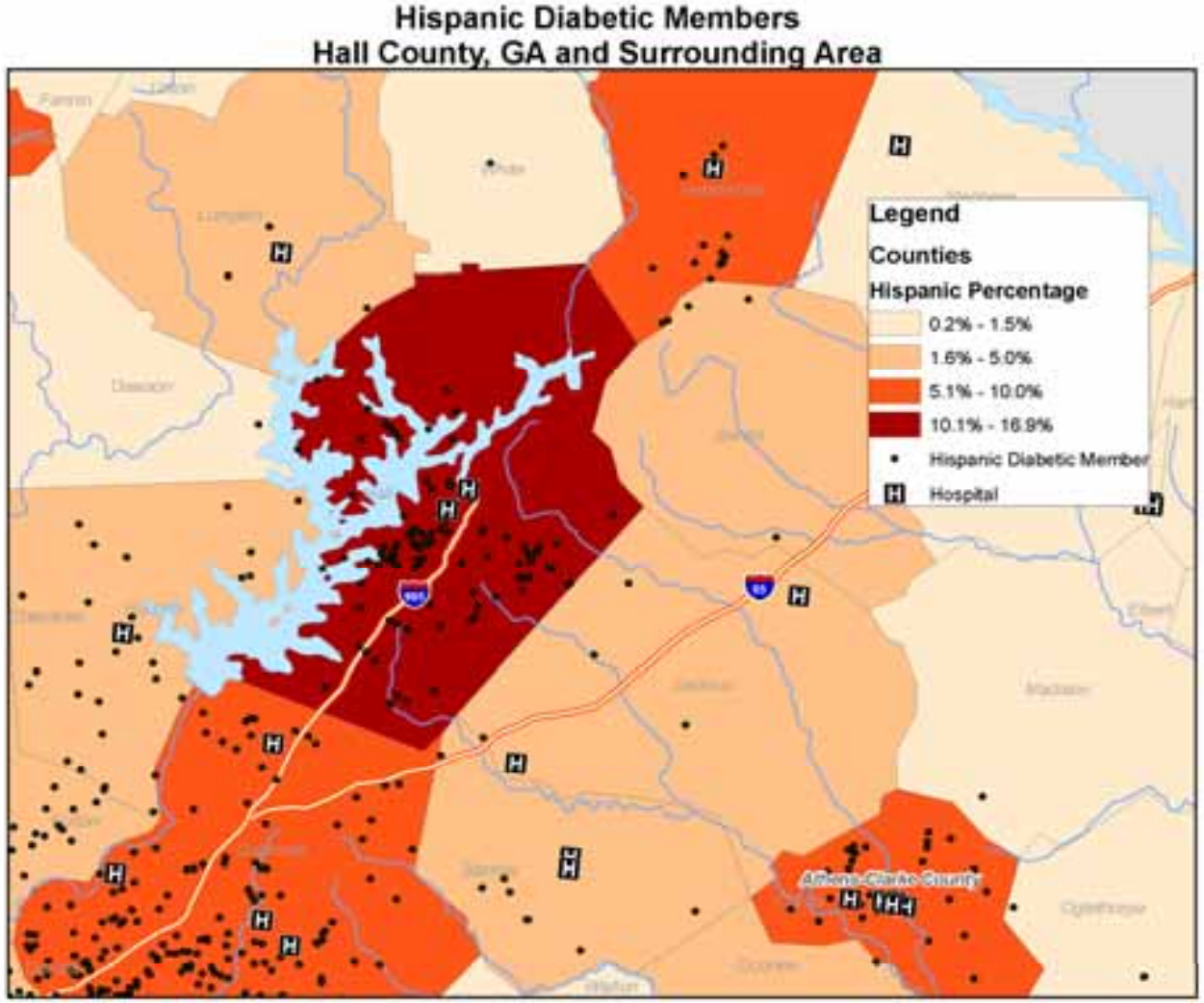
Determining Best Locations for Telemedicine



Company Confidential | Copyright by WellPoint, Inc. | Do Not Copy Without Permission

WELLPOINT

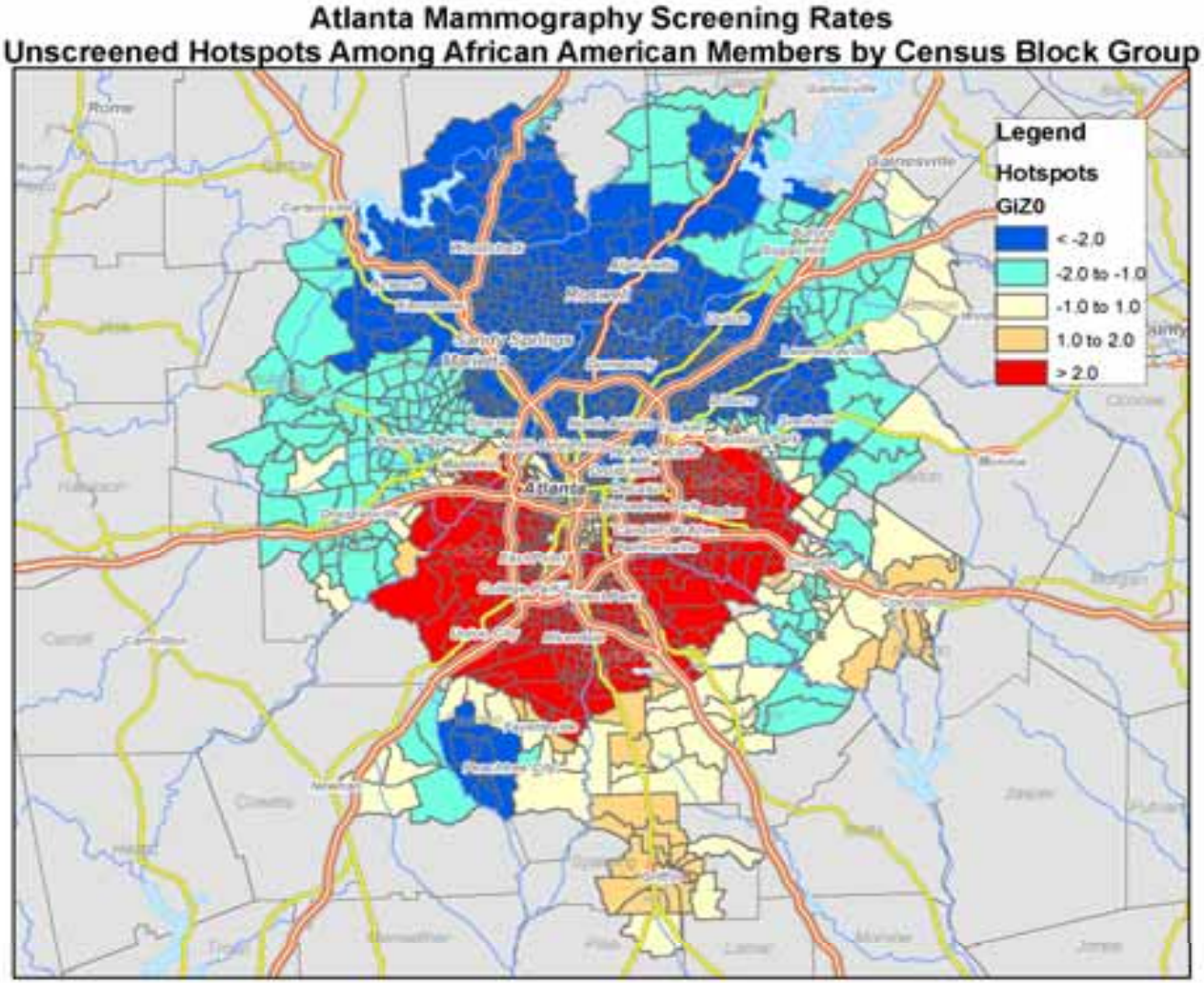
Hall County Georgia – Hispanic Diabetics



Company Confidential | Copyright by WellPoint, Inc. | Do Not Copy Without Permission

WELLPOINT

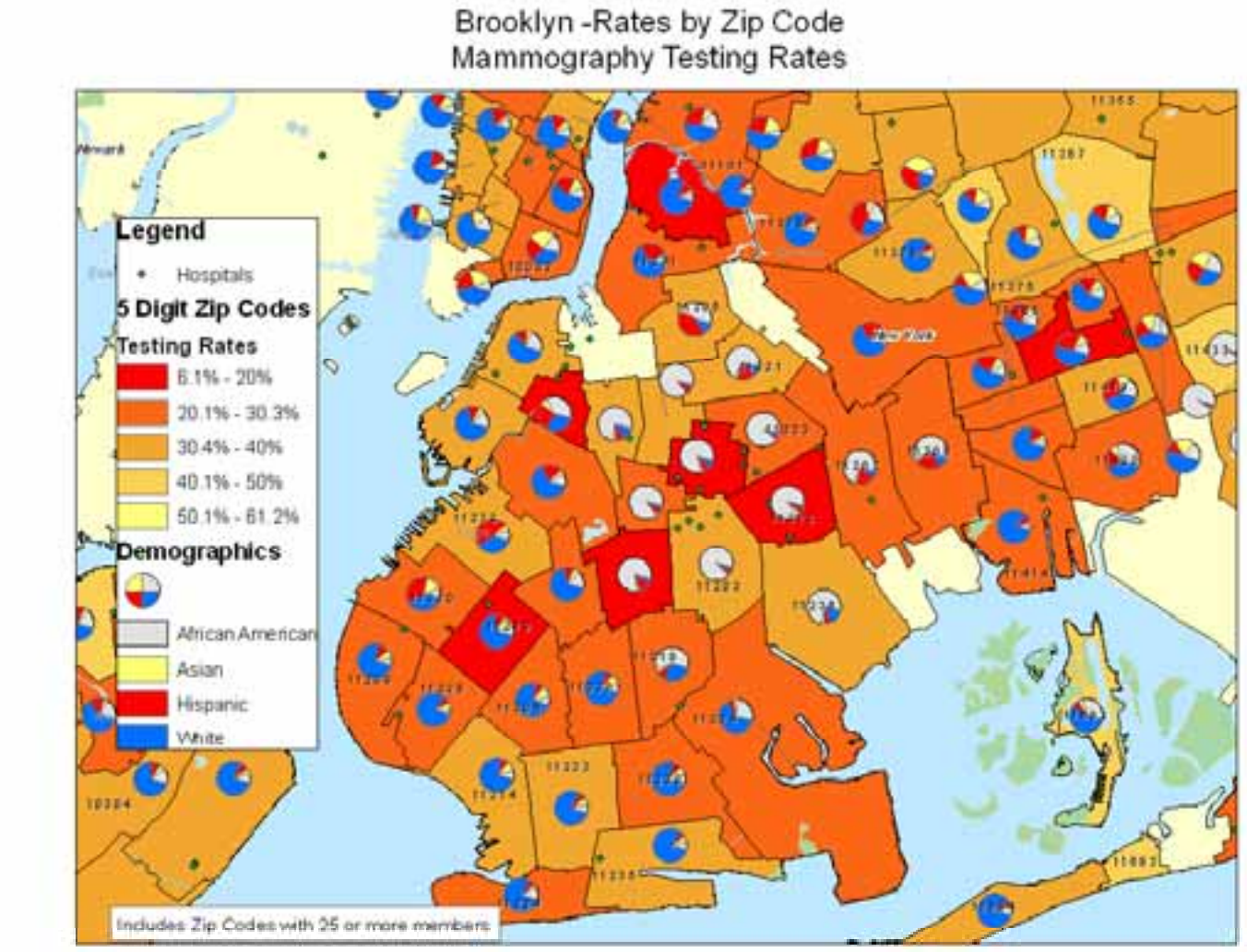
Atlanta – Locating Unscreened Hot Spots



Company Confidential | Copyright by WellPoint, Inc. | Do Not Copy Without Permission

WELLPOINT

Brooklyn – Combining Demographics and Rates



Company Confidential | Copyright by WellPoint, Inc. | Do Not Copy Without Permission

WELLPOINT

How is the Analysis Used

- The logistic model of indirect race/ethnicity data methodology has proven itself to be flexible and applicable to address a wide range of project and business support needs
  - Used to identify hotspots of unscreened members
  - Study provider access for minority members
  - Identify minority members for culturally/linguistically appropriate health screening reminders and health education materials
  - Determine member threshold language needs to meet regulatory requirements

Company Confidential | Copyright by WellPoint, Inc. | Do Not Copy Without Permission

Using the Analysis Outside Wellpoint

- Reports and maps generated using the indirect race/ethnicity data process
  - Promote collaboration with its network medical groups
  - Promote partnering with external health advocacy organization like the American Cancer Society
  - Encourage open dialogues with elective government officials on the issue of health disparities and jointly work towards collaborative quality initiatives to reduce the gap in care

Company Confidential | Copyright by WellPoint, Inc. | Do Not Copy Without Permission

Looking toward the future

- Advancements in predictive abilities
  - Improved individual identification
  - Further granularity beyond major races (African-American, Hispanic, Asian, and White / Other) to major ethnicities (i.e. Chinese, Korean, etc.)
  - Language spoken predictions
  - Cultural assimilation
- Methodology Advancements
  - Explore use of updated name lists
  - Further explore the use of data mining tools such as tree models, neural networks, and other machine learning
- What are other industries doing?

Company Confidential | Copyright by WellPoint, Inc. | Do Not Copy Without Permission