



# NHIA Over-Identifies Hispanics in Areas with a Low Proportion of Hispanic Americans

Carin Perkins, PhD and  
Sally Bushouse, DVM, PhD  
Minnesota Cancer Surveillance System

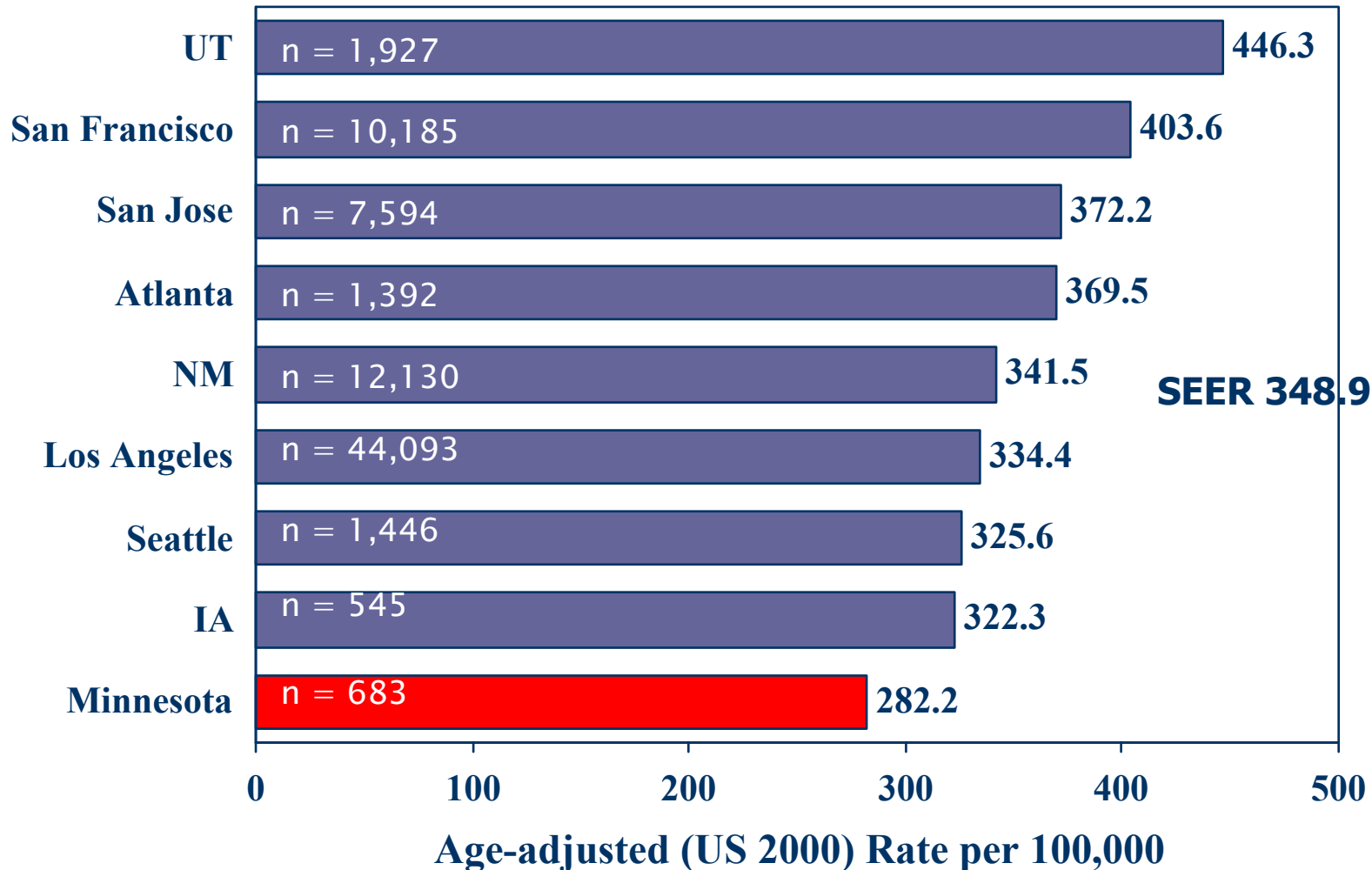
# Hispanic Cancer Rates in Minnesota (1)

- ◆ MCSS began population-based data collection in 1988, didn't collect race/ethnicity until 1995 with funding from NPCR.
- ◆ Internal and external audits indicate MCSS is highly complete, accurate, and timely.
- ◆ 2000 Census enumerated 143,382 Hispanics in Minnesota (2.9% of Minnesota population).

# Hispanic Cancer Rates in Minnesota (2)

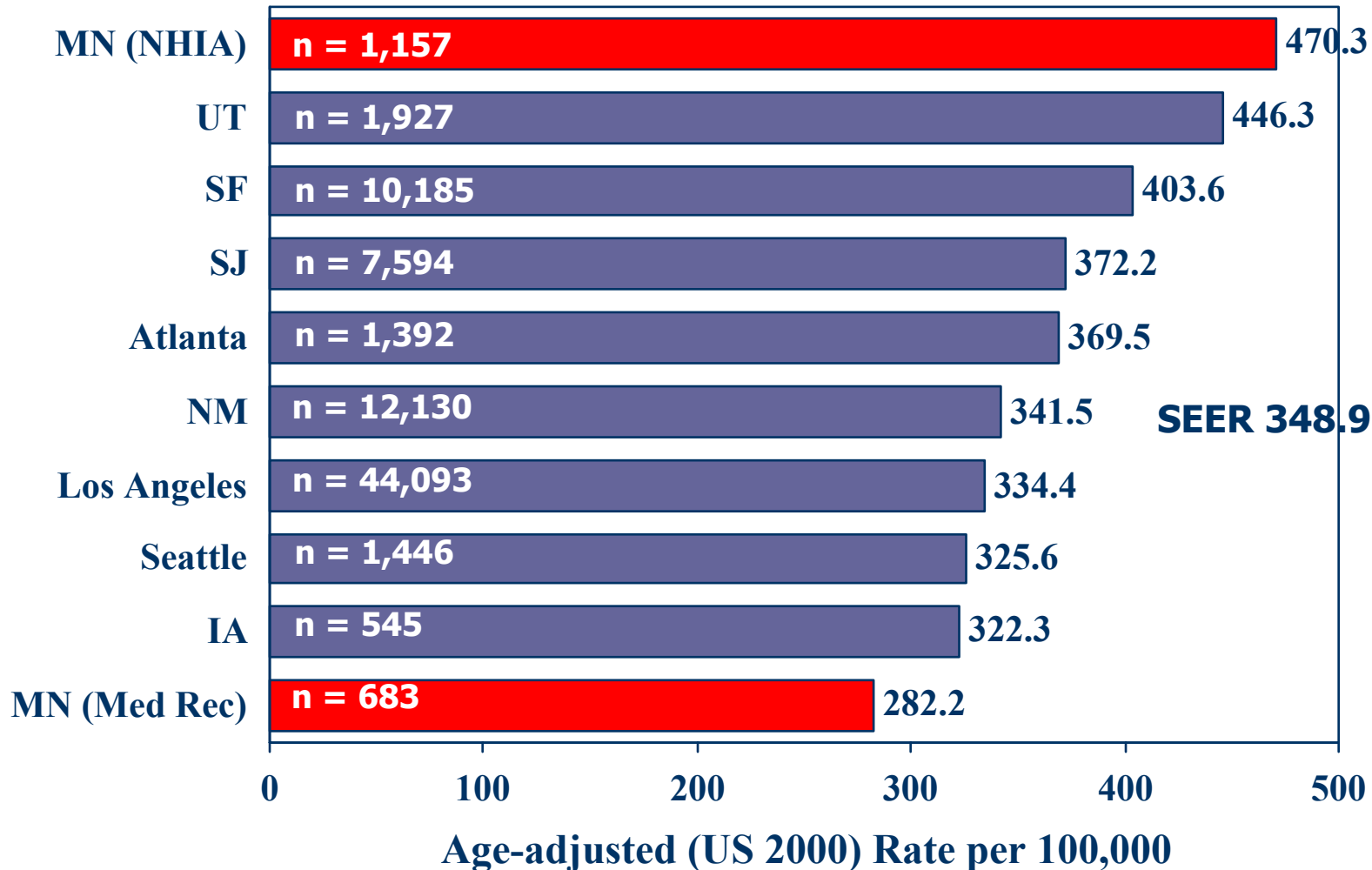
- ◆ Suspected that information on medical record did not identify many Hispanic individuals as such; therefore MCSS did not present Hispanic incidence.
- ◆ Increasing interest in burden of cancer among Hispanic Minnesotans due to development of comprehensive cancer control plan for the state
- ◆ MCSS began evaluating NHIA in spring of 2004.

# Overall Cancer Incidence among Hispanics (all races), 1995-2001, prior to applying NHIA to MCSS



\*SEER 11 + Alaska Regions for which both incidence and mortality rates presented for Hispanics – Seattle, SF, SJ, LA, UT, NM, IA, Atlanta.

# Overall Cancer Incidence among Hispanics (all races), 1995-2001, before & after applying NHIA to MCSS



\*SEER 11 + Alaska Regions for which both incidence and mortality rates presented for Hispanics – Seattle, SF, SJ, LA, UT, NM, IA, Atlanta.

# Factors Evaluated

- ◆ Accuracy of 2000 population estimates for Hispanics
- ◆ Potential role of migrant labor in inflating numerator but not denominator
- ◆ Incidence/mortality rate ratios
- ◆ Lack of maiden name
- ◆ Age-specific curves before and after NHIA
- ◆ Site-specific increases in incidence

## Sensitivity and Specificity of Heavily Hispanic Surnames based on the 1990 Census Spanish Origin Research file

Self-reported Ethnicity Hispanic		Heavily Hispanic Surname		Total
		+	-	
	+	135,131	25,040	160,171
	-	7,670	879,638	887,308
	Total	142,801	904,678	1,047,479

Sensitivity = proportion of self-reported Hispanics with heavily Hispanic surname  
 $= 135,131 / 160,171 = 84.37\%$

Specificity = proportion of non-Hispanics who don't have heavily Hispanic surnames  
 $= 879,638 / 887,308 = 99.14\%$

Predictive value positive = proportion of heavily Hispanic surnames who are Hispanic  
 $= 135,131 / 142,801 = 94.6\%$

Ratio of surname Hispanics to self-reported =  $142,801 / 160,171 = 0.8915$

Ratio of surname + self-reported Hispanics to self-reported =  $167,841 / 160,171 = 1.048$ .

## How PVP Changes with Proportion of Hispanics in the Population to which Surname List is Applied

Population is 5% Hispanic				
Self-reported Ethnicity Hispanic		Heavily Hispanic Surname		Total
		+	-	
	+	422	78	500
	-	82	9,418	9,500
	Total	504	9,496	10,000

Sensitivity = 84.37%

Specificity = 99.14%

Predictive value positive = proportion of heavily Hispanic surnames who are Hispanic  
 $= 422 / 504 = 83.7\%$

Ratio of surname Hispanics to self-reported =  $504 / 500 = 1.01$

Ratio of surname + self-reported Hispanics to self-reported =  $582 / 500 = 1.16$ .



# How PVP Changes with Proportion of Hispanics in the Population to which Surname List is Applied

Population is 2.9% Hispanic (Minnesota)				
Self-reported Ethnicity Hispanic		Heavily Hispanic Surname		Total
		+	-	
	+	245	45	290
	-	84	9,626	9,710
	Total	329	9,671	10,000

Sensitivity = 84.37%

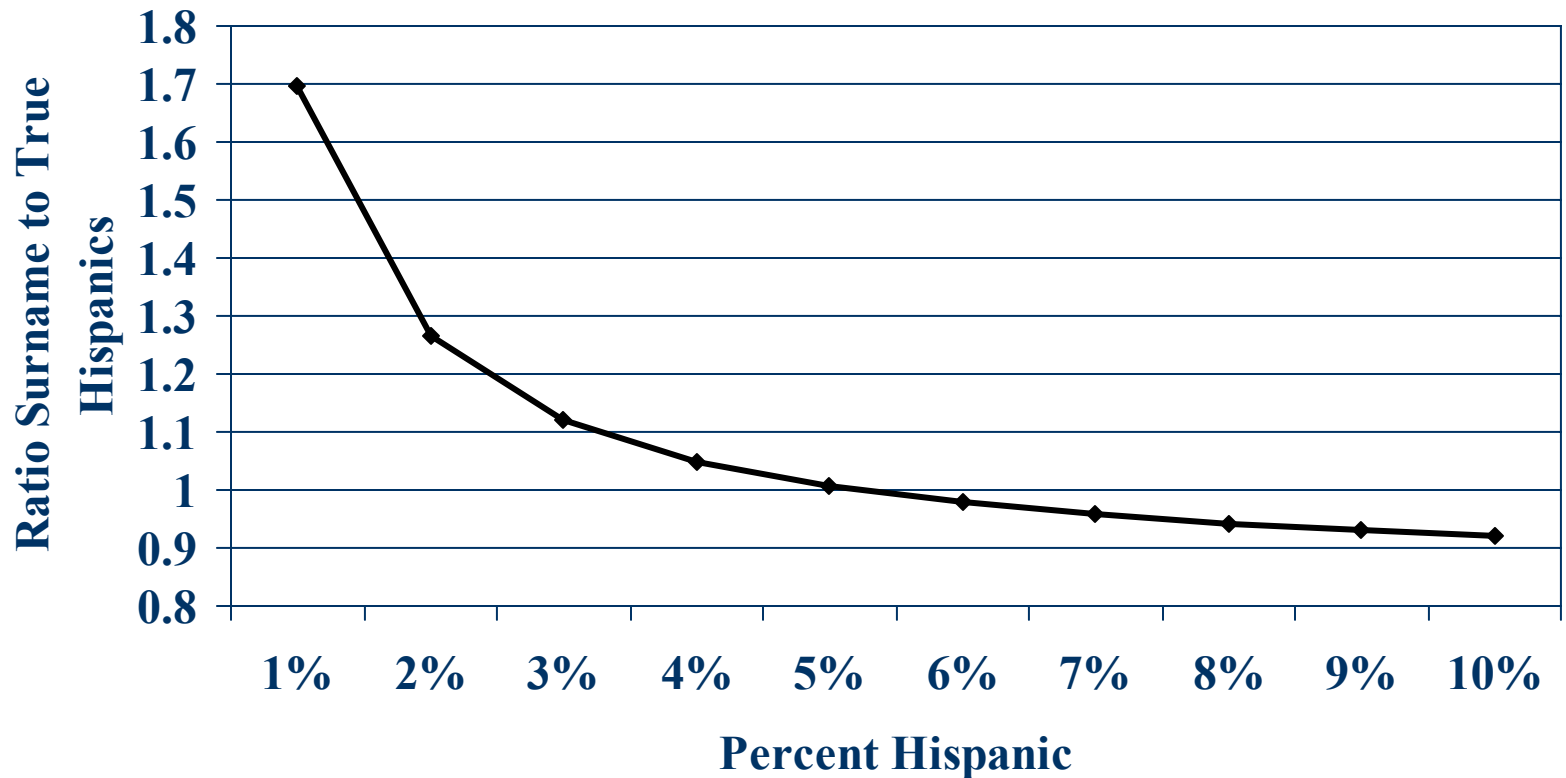
Specificity = 99.14%

Predictive value positive = proportion of heavily Hispanic surnames who are Hispanic  
 =  $245 / 329 = 74.4\%$

Ratio of surname Hispanics to self-reported =  $329 / 290 = 1.13$

Ratio of surname + self-reported Hispanics to self-reported =  $374 / 290 = 1.29$

# Inflation of Hispanic cases using 1990 Census heavily Hispanic surnames to identify Hispanics as a function of the proportion of Hispanics in the population

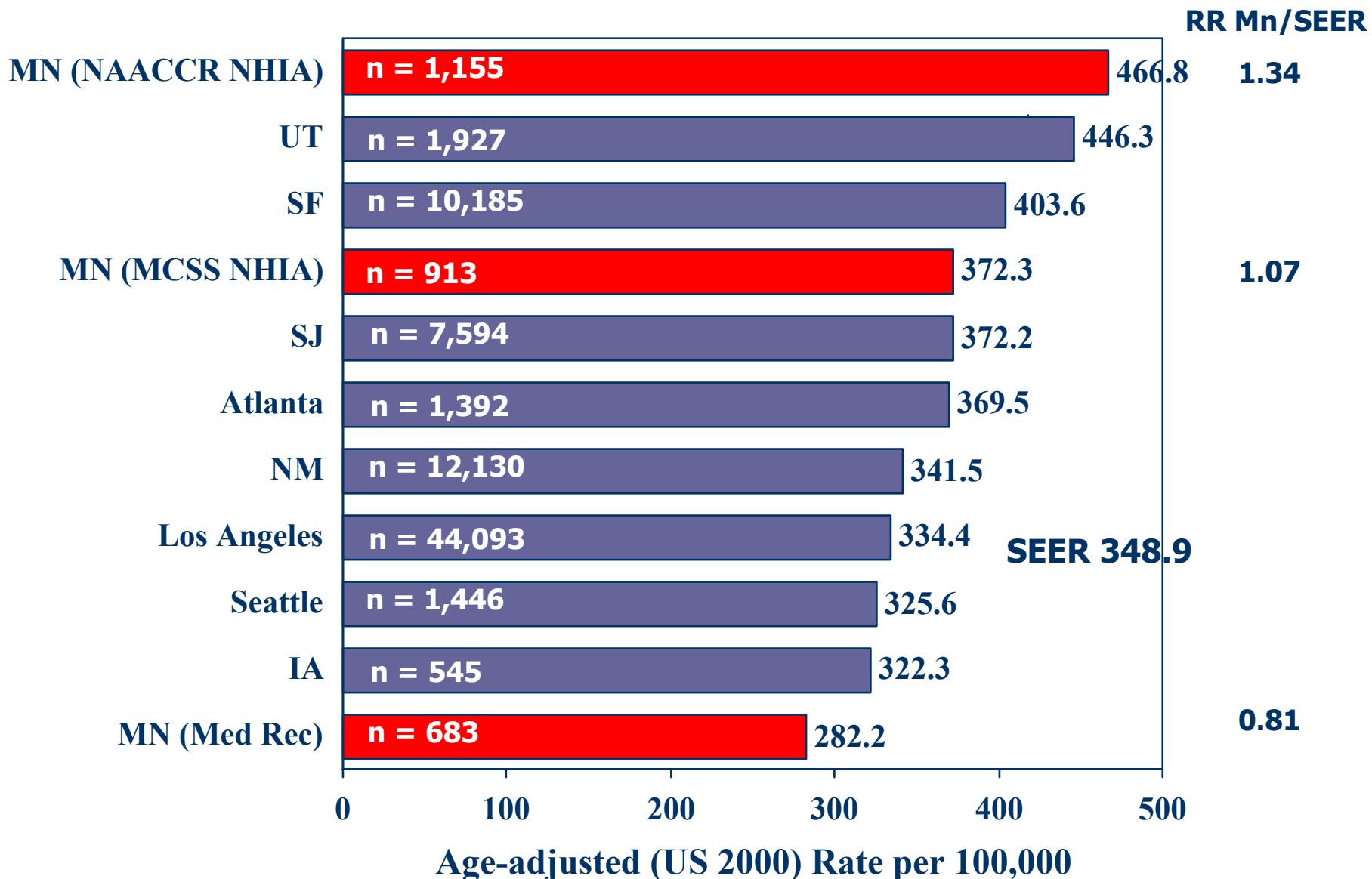


Sensitivity = 0.8437; specificity = 0.9914.

# What to do in Minnesota?

- ◆ Use race and birthplace fields per NHIA.
- ◆ But, apply surname list only to cases residing in the 11 Minnesota counties with 4.0% or more Hispanics in 2000 Census.
- ◆ Those 11 counties account for about 63% of Hispanics in the state.
- ◆ For other counties, recode “unknown” as “non-Hispanic.”

# Overall Cancer Incidence among Hispanics (all races), 1995-2001

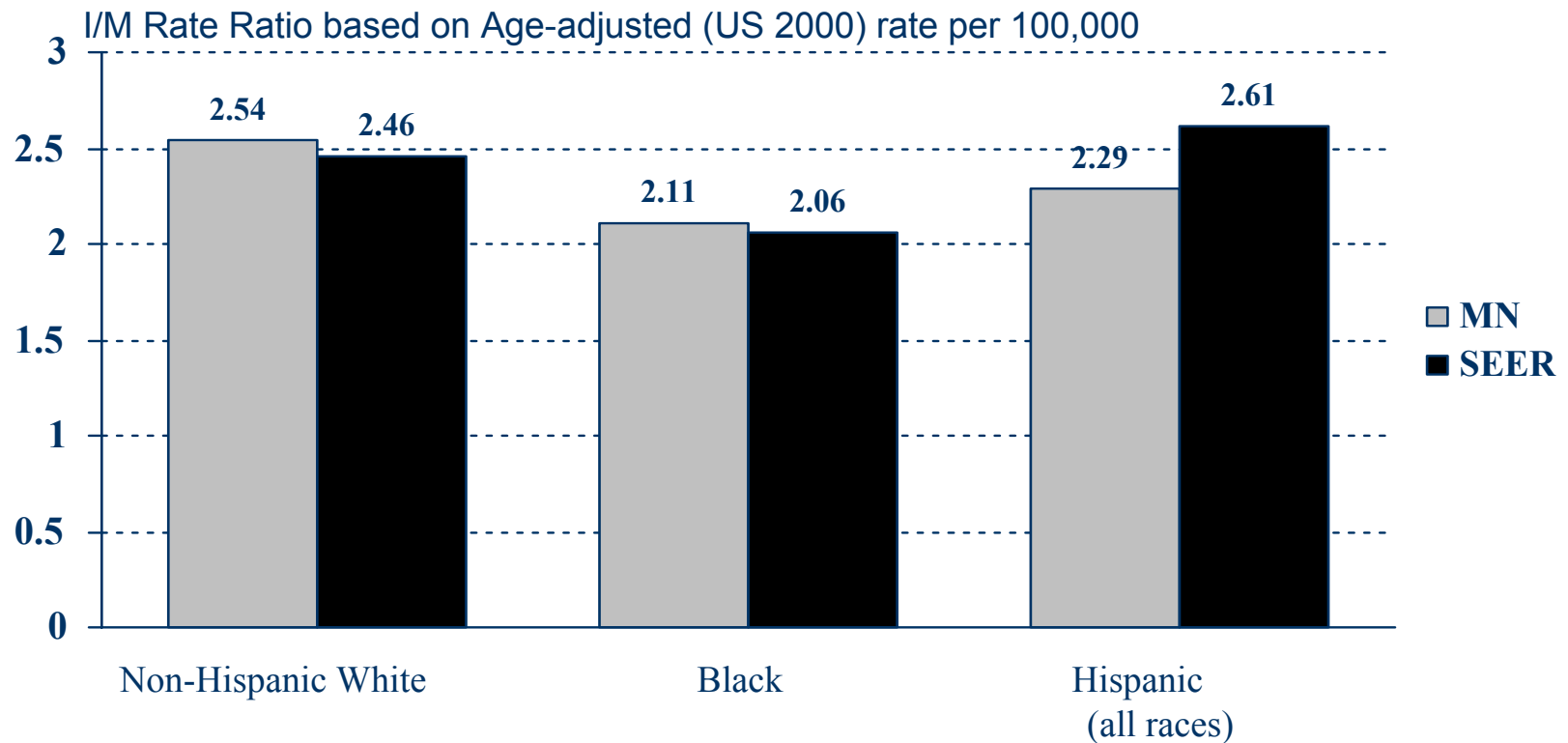


\*SEER 11 + Alaska Regions for which both incidence and mortality rates presented for Hispanics – Seattle, SF, SJ, LA, UT, NM, IA, Atlanta.

**Cancer Incidence among Hispanics (all races): As Reported on Medical Record, after NHIA applied to 11 counties (MCSS), and applying NHIA to other counties (NAACCR)**

	Medical Record	MCSS Added	Total MCSS	NAACCR Added	Total NAACCR
	Cases 1995-2001				
11w/ 4%+ Hispanics	405	210	615	0	615
Other Counties	278	20	299	228	540
Minnesota	683	230	913	228	1,155
	Age-adjusted Incidence Rates 1995-2001				
11w/ 4%+ Hispanics	255.0		384.2		384.2
Other Counties	332.0		353.0		613.6
Minnesota	282.2		372.3		466.8

# Race-specific I/M RR, Minnesota and SEER (1998-2002)



# Conclusions

- ◆ NHIA should not be applied unless there is reason to believe that Hispanics are not being reported as such through routine channels.
- ◆ Applying NHIA to geographic areas with a low proportion of Hispanics will overcount Hispanics, even if none are reported as Hispanic on the medical record.
- ◆ Public health efforts are best supported by accurate data.