

# 2012 VIREC Database and Methods Cyber Seminar Series



2012 VIREC Database and Methods Cyber Seminar Series

# Assessing Race and Ethnicity

May 7, 2012

Presented by: Maria Mor, PhD

Center for Health Equity Research and Promotion  
VA Pittsburgh Healthcare System



# Session Objectives

- Introduction/Overview
- Race and Ethnicity in the VA Data
- Race and Ethnicity in Medicare Data
- Quality of VA Race/Ethnicity Data
- Recommendations
- Where to Go for More Help



# Session Objectives

- Introduction/Overview
- Race and Ethnicity in the VA Data
- Race and Ethnicity in Medicare Data
- Quality of VA Race/Ethnicity Data
- Recommendations
- Where to Go for More Help



# Introduction/Overview

- Racial/ethnic disparities in health and health care are well-documented and persistent in the US
  - Root causes and solutions are not well understood
  - While quality is improving (in general), access and disparities are not improving (AHRQ 2010)
- Racial/ethnic disparities also exist in VHA, where financial barriers to receiving care are minimized
  - Potential contributors include patient factors, provider decision-making, and VAMC characteristics (Saha 2008)
- More research to detect, understand, and address disparities in health and health care is needed



# Introduction/Overview

- Accurate race/ethnicity data are essential to disparities research and research on clinical factors associated with race
- Problems with race/ethnicity data in the VA
  - Completeness
  - Accuracy
  - Consistency over time



# Session Objectives

- Introduction/Overview
- Race and Ethnicity in the VA Data
- Race and Ethnicity in Medicare Data
- VA Race Data Quality
- Recommendations
- Where to Go for More Help



# VA Race and Ethnicity Categories

## VHA Handbook 1601A.01 (2009)

### ■ Ethnicity

- Spanish, Hispanic, or Latino

### ■ Race (>1 may be selected)

- American Indian or Alaska Native
- Asian
- Black or African American
- Native Hawaiian or Other Pacific Islander
- White
- Unknown by Patient

### ■ Current reporting method

- 2 question format: ethnicity, race
- Self-reported



# Race and Ethnicity Data Collection Standards

- OMB Directive Revision No. 15 (1997)
  - Standards for maintaining, collecting, and presenting federal data on race and ethnicity
  - Implemented in DVA in 2003
- Joint Commission (2010)
  - Collection of patient demographic data, e.g., race and ethnicity key element of performance
- Affordable Care Act currently under review
  - New data collection standards for race, ethnicity, primary language, and sex

# VHA Health Equality Workgroup

- Chartered by Dr. Jesse in 2011
  - Advise leadership: how to coordinate VHA components to achieve equity in Veteran health care
- VHA accepts recommendations aligned with dimensions of National Partnership for Action to Eliminate Health Disparities:
  - leadership
  - awareness
  - health and health care outcomes
  - diversity and cultural competence of the workforce
  - data, research, and evaluation



# Workgroup Recommendations Related to Data, Research, and Evaluation

- Universal standard collection of self-reported data on race/ ethnicity, primary language, rural/urban residence, and sex
- Identify unrecognized and emerging vulnerable patient populations
- Assess/report differences stratified by vulnerable groups:
  - Quality of care
  - Patient satisfaction
  - Access to care
- Harmonize standards for vulnerable populations with HHS and Affordable Care Act data standards



# Acquisition of Race/Ethnicity Data in VHA

WHO:	WHAT:	WHEN:	WHERE:	HOW:
<p>Information Source:</p> <ul style="list-style-type: none"> <li>• Patient (self-report)</li> <li>• Proxy</li> </ul>	<ul style="list-style-type: none"> <li>• VA Form 10-10EZ, Application for Health Benefits (online, paper, or by interview)</li> </ul>	<ul style="list-style-type: none"> <li>• Enrollment</li> <li>• Hospital admission</li> <li>• Outpatient visit or pre-registration</li> </ul>	<ul style="list-style-type: none"> <li>• Online</li> <li>• Telephone call from local VHA facility</li> <li>• In-person visit to local VHA facility</li> </ul>	<ul style="list-style-type: none"> <li>• VHA Facility Enrollment Coordinator or designee (e.g., Admission Interview Clerk, Enrollment Specialist) or</li> <li>• Outpatient clinic personnel</li> <li>• Collects the information and enters into VistA</li> </ul>

# Race/Ethnicity Data Collection: Historical

- Method of ascertainment uncertain, primarily observer-reported
- No option for multiple race reporting
- Single question captured both race and ethnicity:
  - Hispanic, White
  - Hispanic, Black
  - American Indian
  - Black
  - Asian
  - White

# Race/Ethnicity Data Collection: Enrollment

- VA Form 10-10EZ (Application for Health Benefits)
- Data entered into the Veterans Health Information Systems and Technology Architecture (VistA)
  - Race Information and Patient Information Sub-files
- Demographic data (e.g., race) transmitted with encounter data to the Austin Information Technology Center
  - National Patient Care Database
- Medical SAS (MedSAS) datasets
  - Extracts of NPCD data



# Race/Ethnicity Data Collection: Clinical Setting

- Race and ethnicity obtained during pre-registration, if missing
- Race and ethnicity gathered directly from the patient (or proxy)
  - Hospital admission
  - Clinic registration
- Data entered into VistA
  - Race Information and Patient Information Sub-files
- Separate VistA field for method of data collection



# Sources of Race Data in the VA

## Medical SAS Datasets

Variable Name	MedSAS Dataset	Description
RACE	Inpatient (PTF Main File)	FY1970 - present
	Outpatient (Visit File)	FY1997 - present
	Outpatient (Event File)	FY1998 - present
RACE1-RACE6	Inpatient (PTF Main)	FY2003 - present
RACE1-RACE7	Outpatient (Visit, Event)	FY2004 - present
ETHNIC	Inpatient (PTF Main)	FY2003 - present
	Outpatient (Visit, Event)	FY2004 - present

# Other Sources of Race Data in the VA

## ■ VA Vital Status File

- Death dates from multiple VA and non-VA data sources (e.g., PTF, BIRLS, SSA, Medicare)
- Contains race data from Medicare
- VSF file structure
  - Master File contains one record for each SSN-date of birth (DOB)-gender combination found in VA data.
  - Some SSNs have more than one record.
  - Mini File contains one record for each SSN. An algorithm is used to select the “best” DOB, gender, and DOD for each SSN.
- Race is in Master File only

## ■ Decision Support System National Data Extracts



# Session Objectives

- Introduction/Overview
- Race and Ethnicity in the VA Data
- Race and Ethnicity in Medicare Data
- Quality of VA Race/Ethnicity Data
- Recommendations
- Where to Go for More Help



# Medicare Race/Ethnicity Data

- Potentially useful source for Veterans in Medicare
  - Age 65 and older (>95% of VA elderly)
  - Disabled (~20% of VA patients <65 years)
  - End stage renal disease
- Derived primarily from Social Security Administration
  - Obtained at the time of application for SSN and/or replacement card
  - Reporting sources: Usually self- or family
- Distinctions from current VA race/ethnicity data
  - 'Hispanic' is a race category
  - No multiple race

# Medicare Race Data from SSA

- Until 1980, 4 categories only
  - White
  - Black
  - Other
  - Unknown
- In 1980, 'Other' replaced by
  - Asian, Asian American or Pacific Islander
  - Hispanic
  - American Indian or Alaskan Native

# Medicare Race Data from SSA

- Medicare race data quality issues
  - Information on most enrollees (those who obtained SSN prior to 1980) limited to original 4 categories
  - SSN application form – single question format and no multiple race reporting
- Initiatives to improve quality of race/ethnicity data
  - Periodic updates on American Indians and Alaskan Natives from Indian Health Service
  - 1997 survey of enrollees classified as 'Other', 'Unknown', or with Spanish surname, requesting race/ethnicity self-report

# Session Objectives

- Introduction/Overview
- Race and Ethnicity in the VA Data
- Race and Ethnicity in Medicare Data
- Quality of VA Race/Ethnicity Data
- Recommendations
- Where to Go for More Help



# Quality of VA Race/Ethnicity Data Completeness

- A substantial portion of patients do not have a “usable” race value\* in the VA Medical SAS Inpatient and Outpatient Datasets

FYear	Usable Race, %	FYear	Usable Race, %
1997	57	2005	67
1998	57	2006	72
1999	58	2007	75
2000	58	2008	76
2001	56	2009	78
2002	55	2010	80
2003	49	2011	83
2004	62		

\* A usable race value is any value that is not ‘missing’ or ‘unknown’ or ‘declined’

# Use of Medicare and DoD Data for Improving VA Race Data Quality

Stroupe, et al. (2010) *Journal of Rehabilitation Research & Development*

## Aims

1. To estimate the extent missing “usable” race data in VA MedSAS files can be reduced by using non-VA data sources
  - a. Medicare
  - b. DoD
2. To evaluate the agreement between VA self-reported race data in MedSAS files and
  - a. Medicare race data
  - b. DoD race data



# Use of Medicare and DoD Data for Improving VA Race Data Quality

Stroupe, et al. (2010) *Journal of Rehabilitation Research & Development*

- Patient cohort
  - 10% representative sample of VHA patients who obtained services during FY2004-2005
  - N = 570,018
- Medicare race data were obtained from Medicare Vital Status file
- DoD race data were obtained from the VA/DoD Identity Repository (VADIR) database for individuals <65 years
  - VA/DoD data-sharing agreement
  - Self-reported race/ethnicity obtained from servicemembers

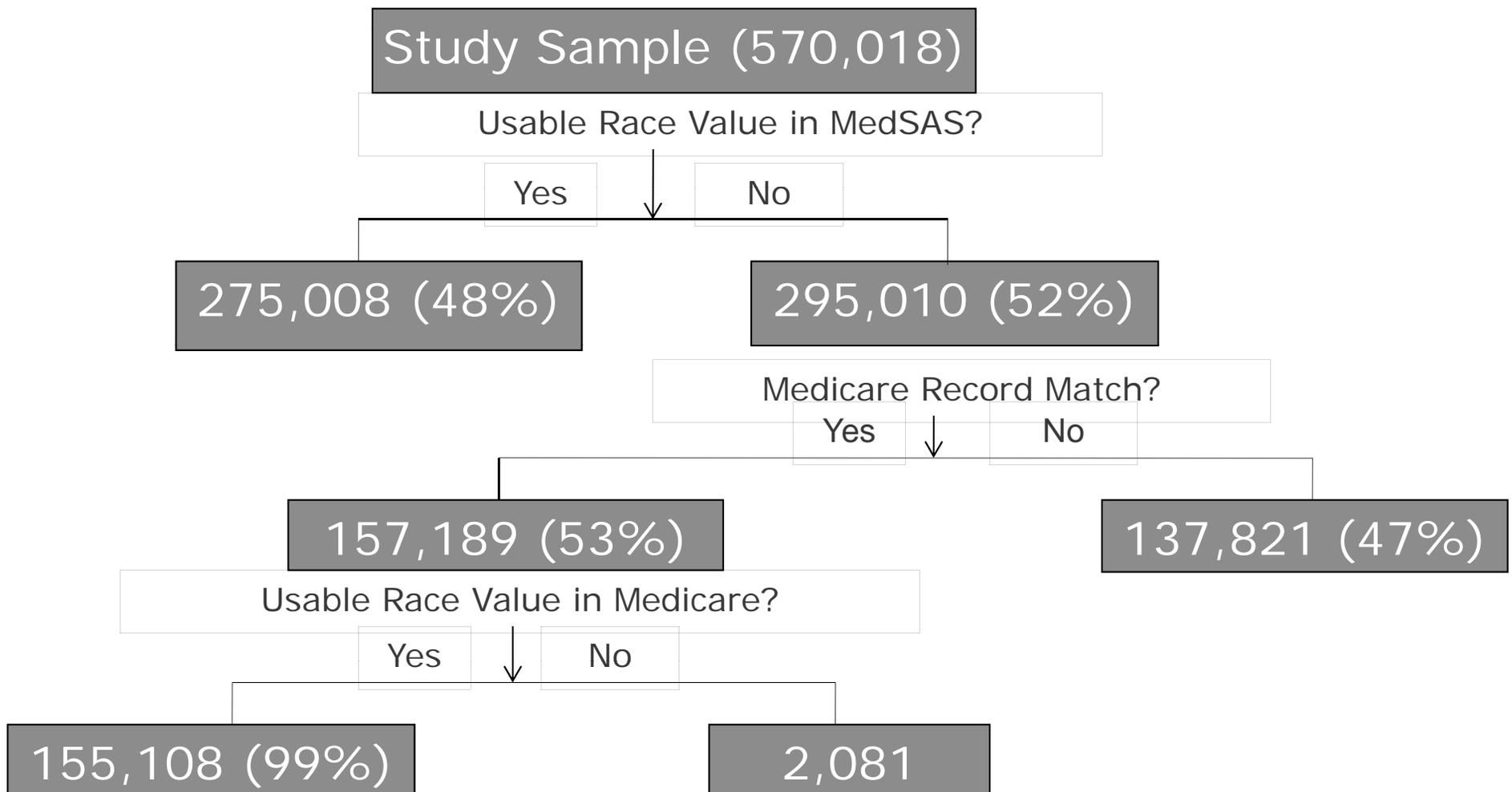


# Use of Medicare and DoD Data for Improving VA Race Data Quality

Stroupe, et al. (2010) *Journal of Rehabilitation Research & Development*

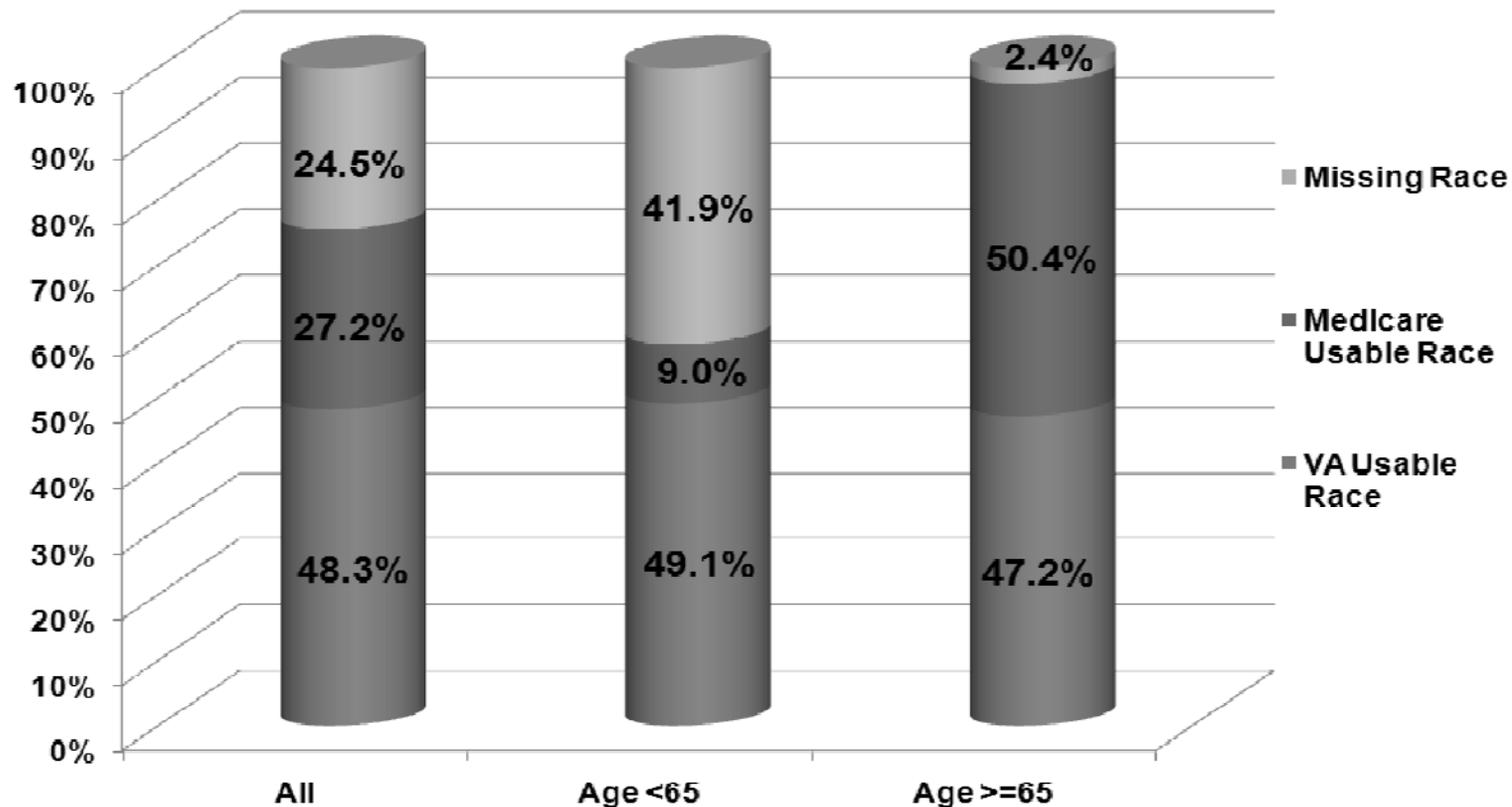
Patients with & without Usable Race Data VA MedSAS	Usable Race Value (n=275,008)	No Usable Race Value (n=295,010)
Age, $\geq 65$ years, %	43	45
Male, %	94	89
Married, %	56	56
Geographic Region, %		
Northeast	16	17
South	44	37
Midwest	23	21
West	17	25

# Aim 1a: Improvement in Race Completeness with Addition of Medicare Data from 2004-2005



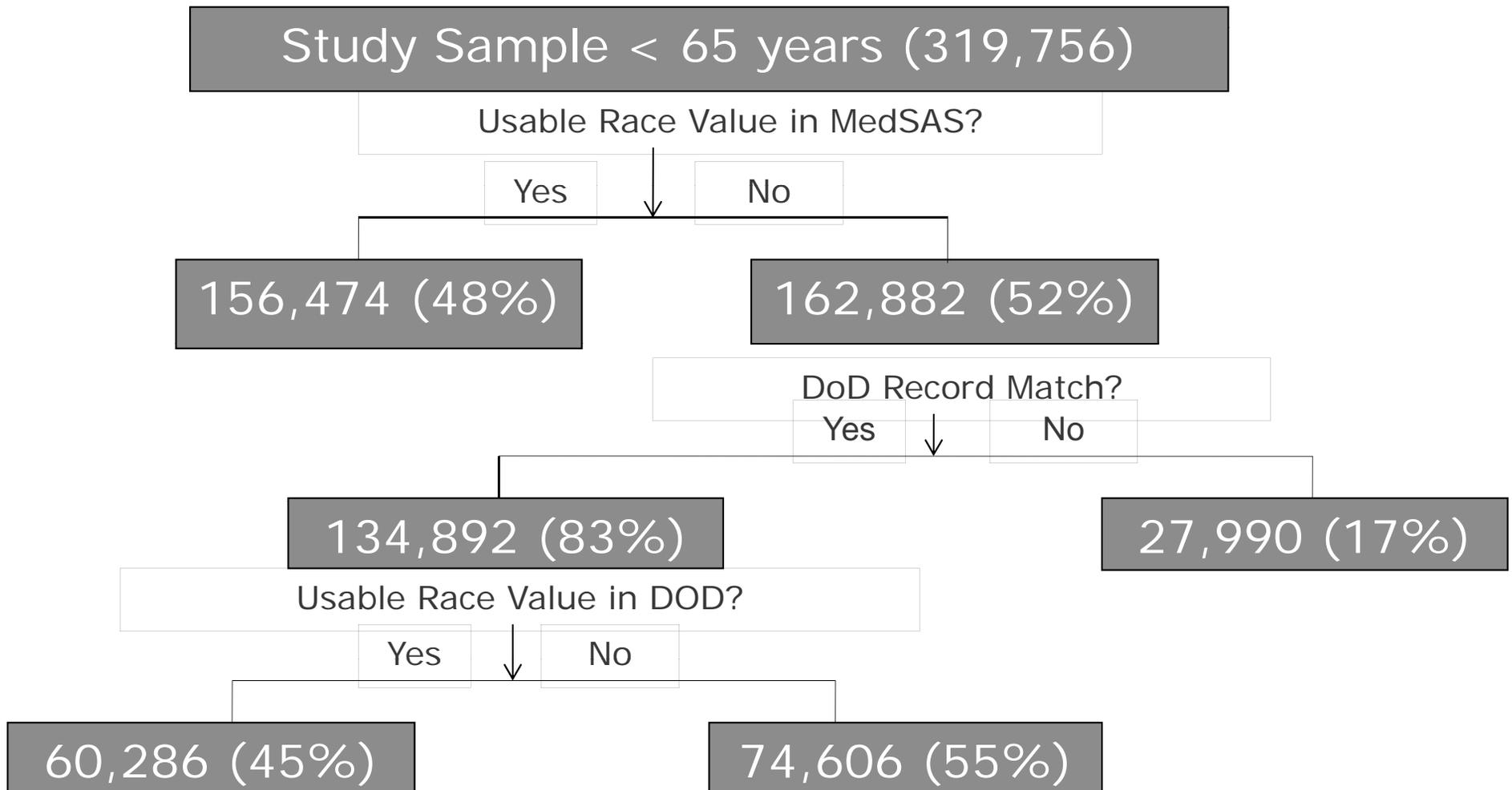
# Aim 1a: Improvement in Race Completeness with Addition of Medicare Data from 2004-2005

**Figure 3. Adding Medicare Data Improves Race Data Completeness <sup>a</sup>**



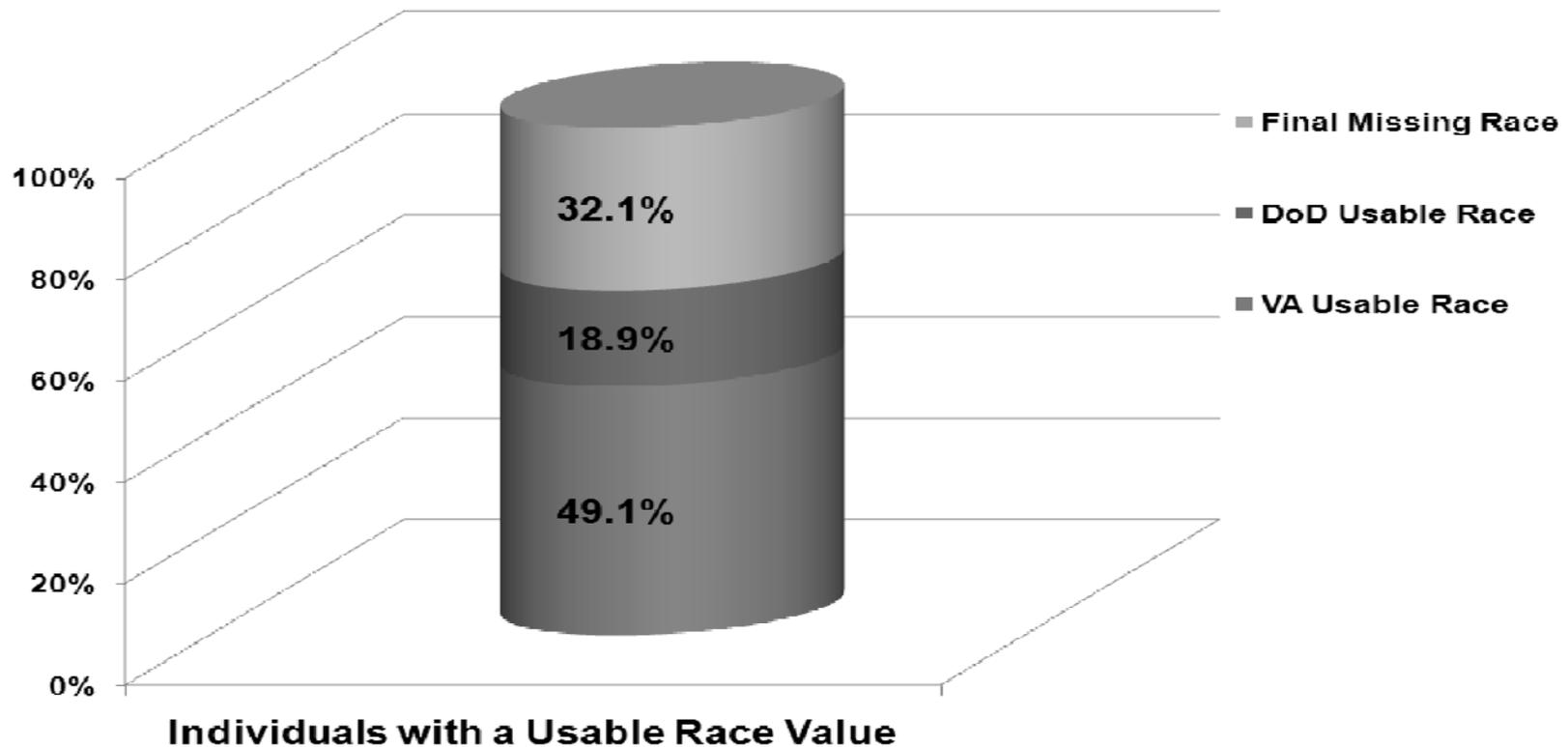
<sup>a</sup> Sample sizes: All 570,018, Age <65 319,756, Age >=65 250,262

# Aim 1b: Improvement in Race Completeness with Addition of DoD Data from 2004-2005



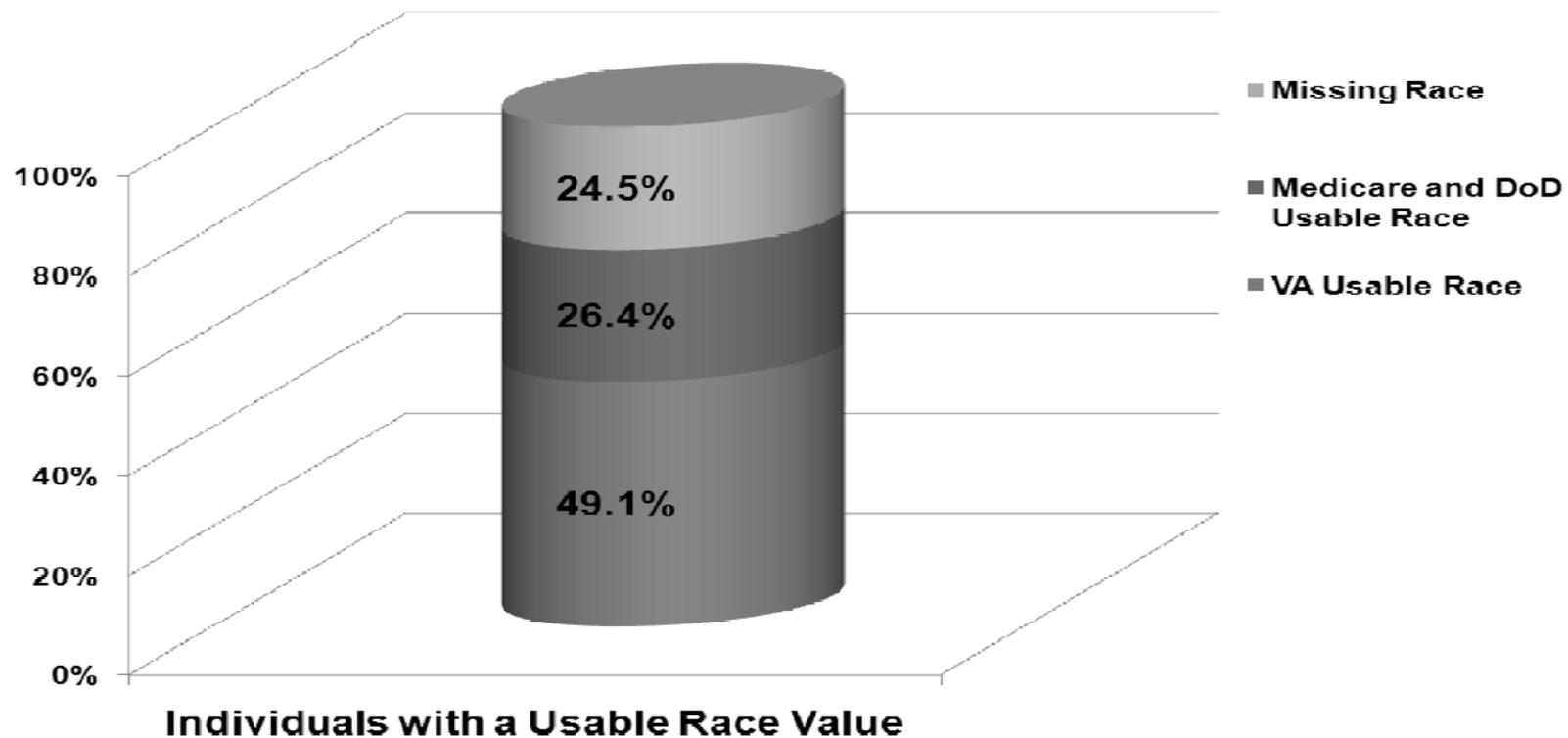
# Aim 1b: Improvement in Race Completeness with Addition of DoD Data from 2004-2005

**Figure 5. Adding DoD Data Improves Race Data Completeness Among Non-Elderly (N=319,756)**



# Aim 1: Improvement in Race Completeness with *Addition of Medicare and DoD Data* from 2004-2005

**Figure 6. Adding Medicare and DoD Data  
Improves Race Data Completeness Among Non-Elderly  
N= (319,756)**

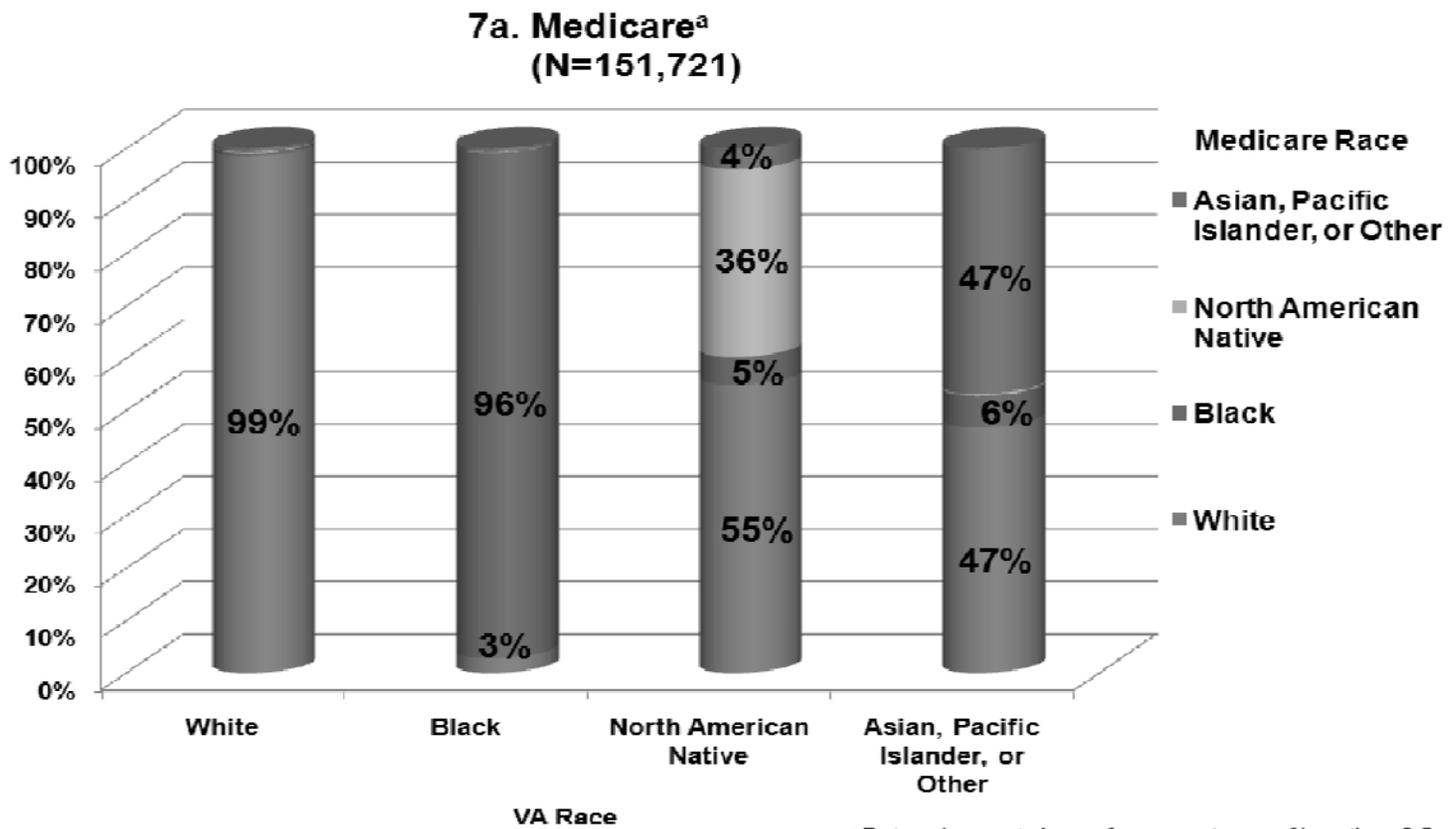


## Aim 2: Comparison of Medicare with VA and DoD Data, 2004-2005

<b>VA Race</b>	<b>Medicare Race</b>	<b>DoD Data</b>	<b>Classification Constructed for Consistency Analysis</b>
White	White	White	White
Black or African American	Black	Black	Black or African American
American Indian or Alaska Native	North American Native	American Indian or Alaska Native	North American Native
Asian	Asian	Asian or Pacific Islander	Asian, Pacific Islander, or Other
Native Hawaiian or Other Pacific Islander	Other	Other	

# Aim 2a: Comparison of VA with Medicare Data, 2004 2005

**Figure 7. Concordance Between VA Race Values and Race Values from External Sources**



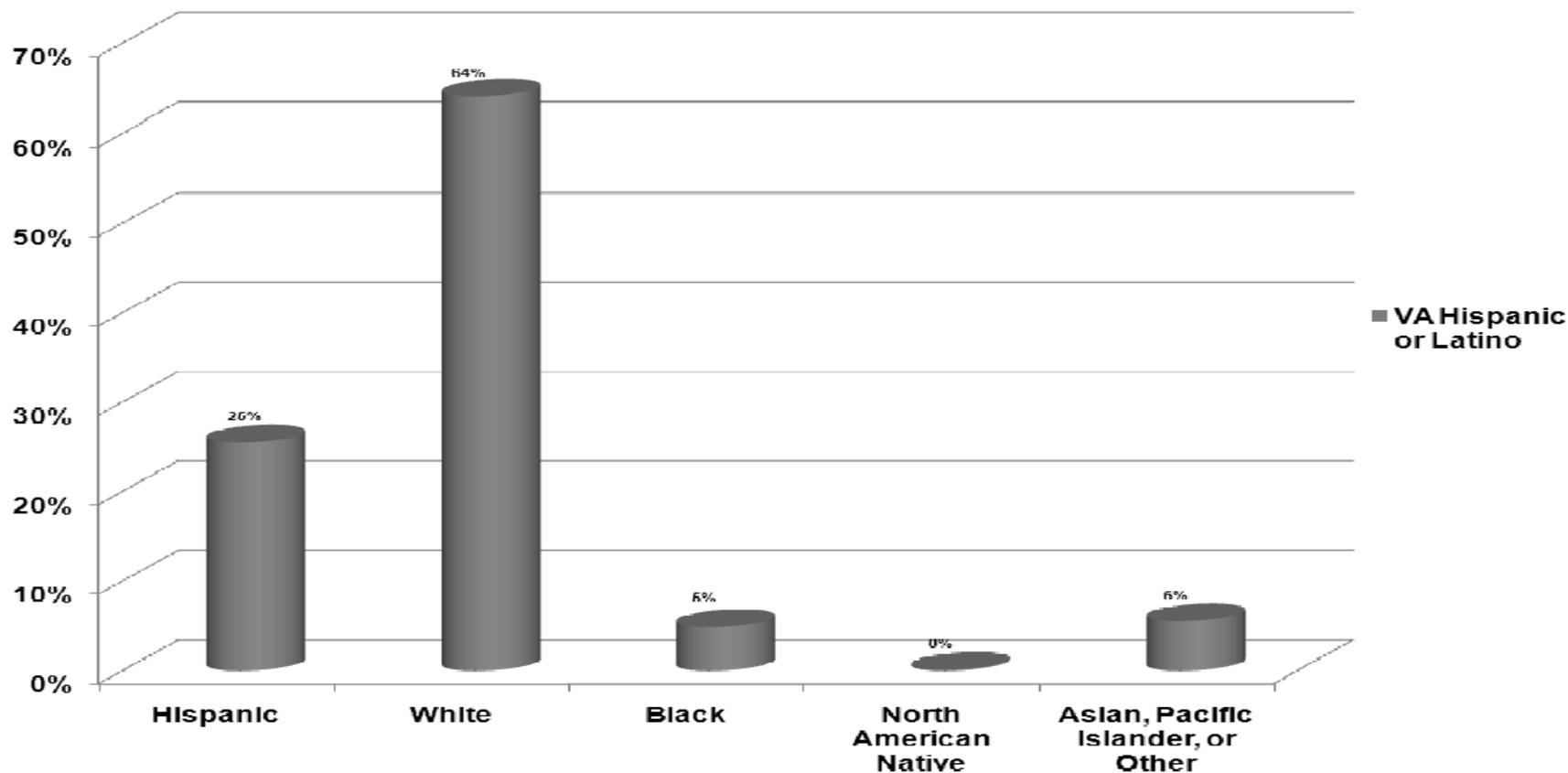
<sup>a</sup>Data values not shown for percentages of less than 2.5

# Aim 2a: Comparison of VA with Medicare Data, 2004-2005

Race	Sensitivity	Specificity	Kappa
White	98.5	91.3	0.89
Black or African American	96.4	99.3	0.95
North American Native	35.8	99.8	0.37
Asian, Pacific Islander or Other	46.8	99.3	0.47

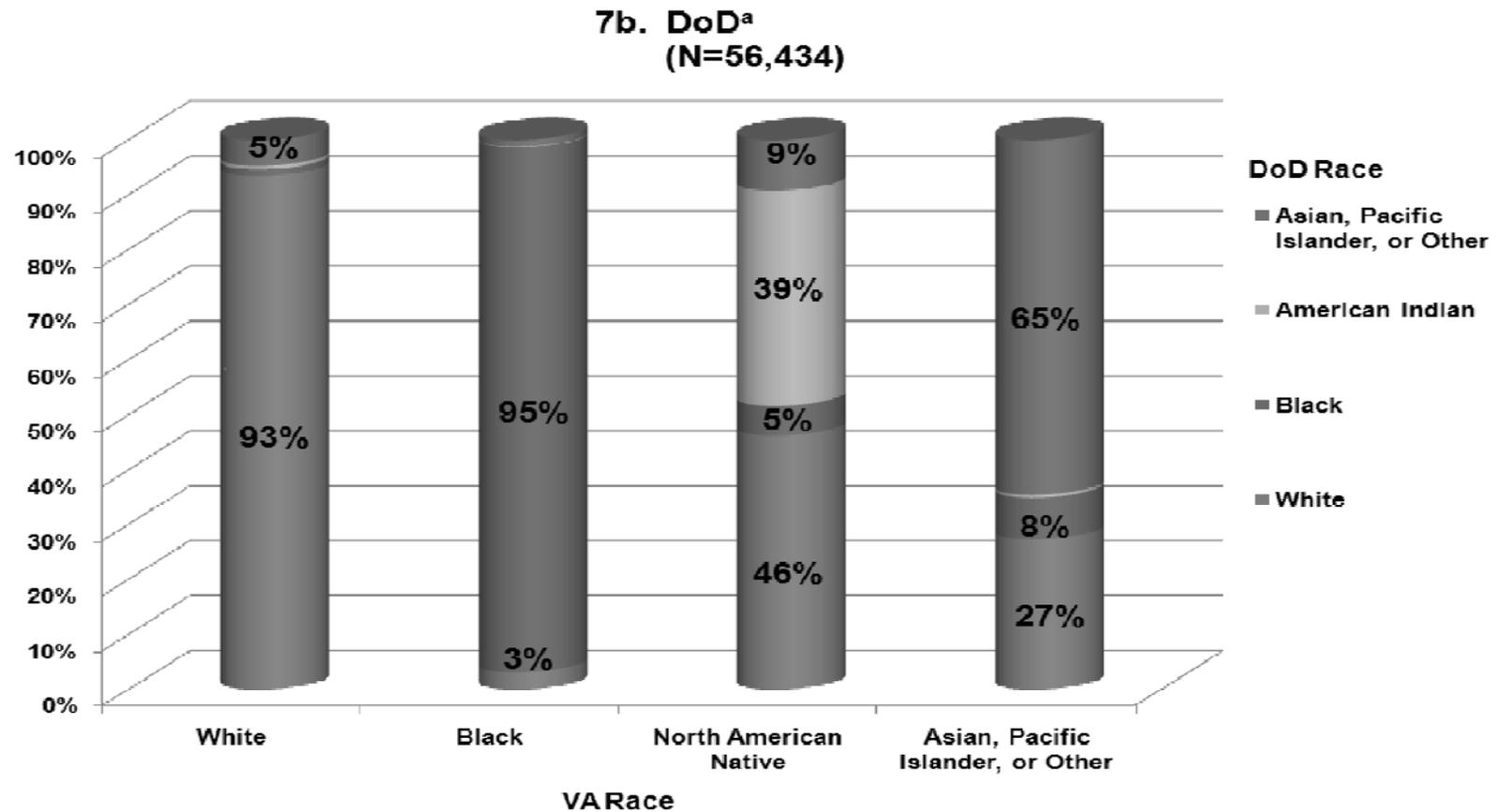
# Aim 2a: Comparison of VA with Medicare Data, 2004-2005

Figure 8. Medicare Race Among VA Self-Reported Hispanics



# Aim 2b: Comparison of VA with DoD Data, 2004 2005

**Figure 7. Concordance Between VA Race Values and Race Values from External Sources**



<sup>a</sup>Data values not shown for percentages less than 2.5

## Aim 2b: Comparison of VA with DoD Data, 2004-2005

Race	Sensitivity	Specificity	Kappa
White	93.4	94.2	0.86
Black or African American	95.4	98.5	0.95
North American Native	39.2	99.5	0.37
Asian, Pacific Islander or Other	64.5	96.3	0.39

# Use of Medicare and DoD Data for Improving VA Race Data Quality

Stroupe, et al. (2010) *Journal of Rehabilitation Research & Development*

## Conclusion

- Supplementing VA with Medicare and DoD data improves VA race data completeness substantially
- More study is needed to understand poor rates of agreement between VA and external sources in identifying non-African-American minority individuals

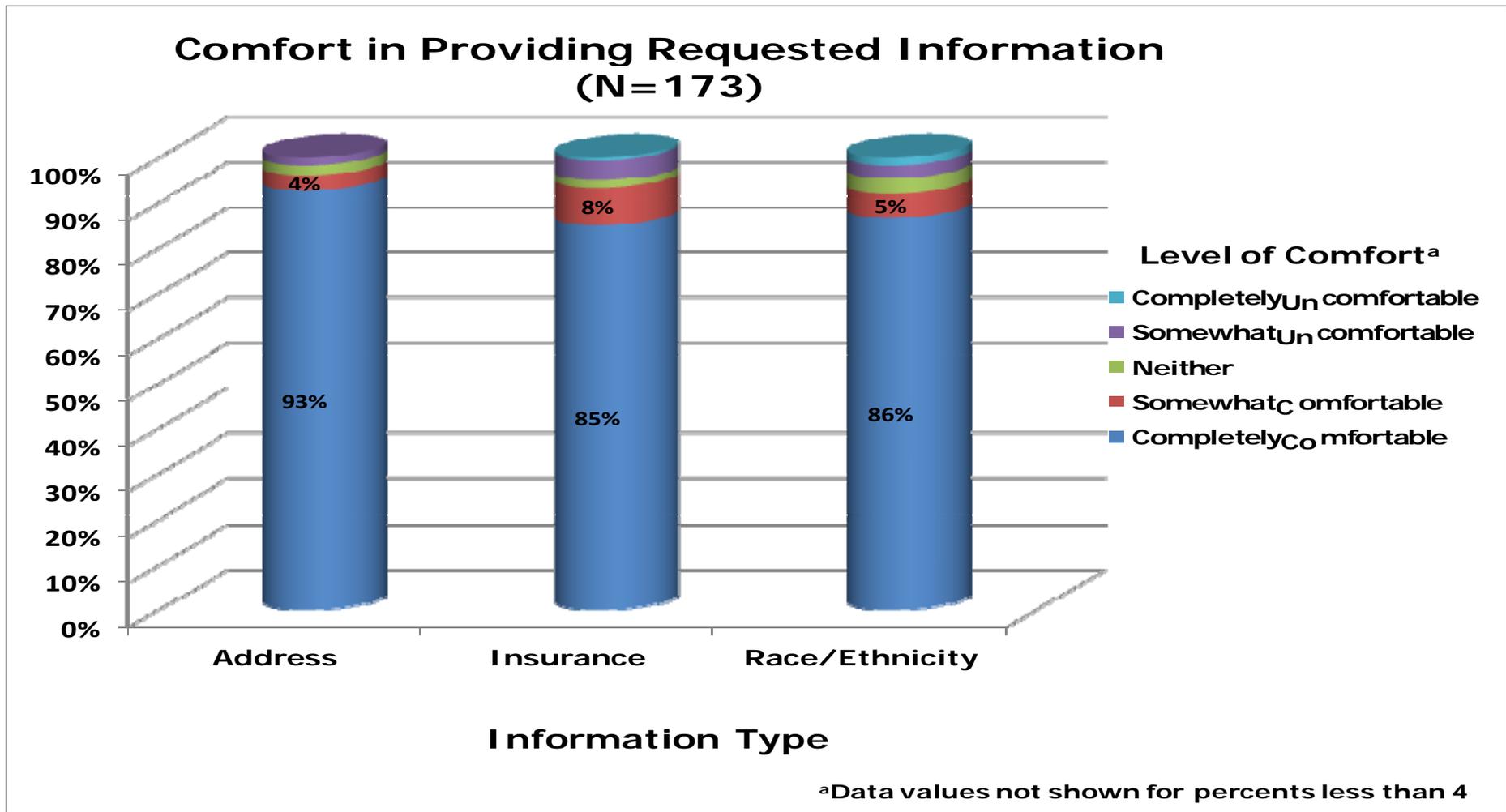


# Improving Patient Demographic Information in VA Databases

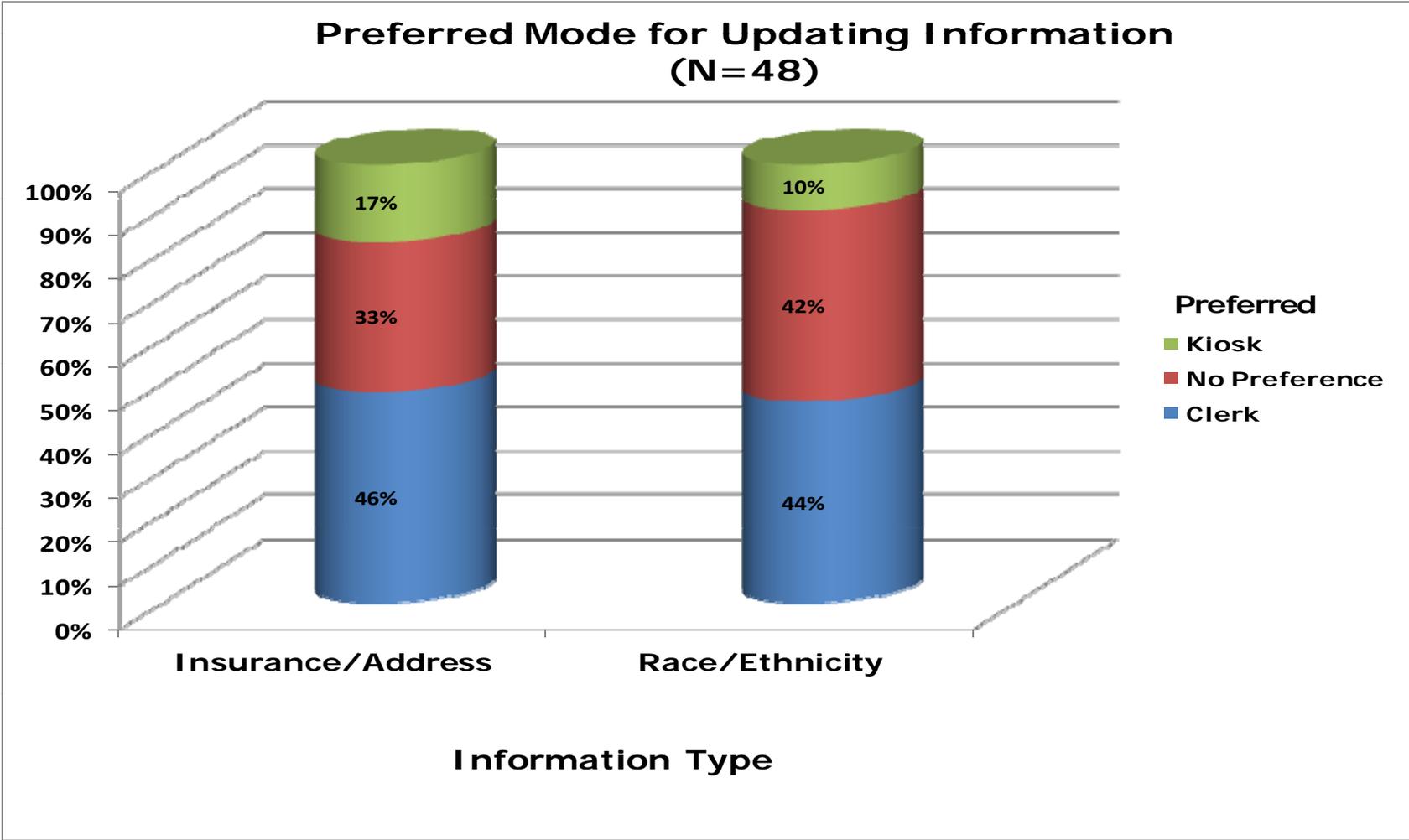
- T21-funded collaboration between Center for Health Equity Research and Promotion and Veterans Engineering Resource Center
- Patients (n=173) with missing or declined race at VA Pittsburgh Healthcare System were surveyed by telephone regarding their:
  - Comfort with being asked to provide race/ethnicity, address/telephone, and insurance when coming to the VA
  - Preference for providing that information to a clerk or computer kiosk



# Comfort in Providing Race/Ethnicity Data



# Preference for Collection of Race/Ethnicity Data



# Session Objectives

- Introduction/Overview
- Race and Ethnicity in the VA Data
- Race and Ethnicity in Medicare Data
- Quality of VA Race/Ethnicity Data
- Recommendations
- Where to Go for More Help



# Recommendations

- When using VA VSF, match on date of birth and gender, in addition to (scrambled) SSN
  - Researchers will be most likely to identify the right individuals in the VSF if they use all three elements when conducting their VSF-study cohort record match
- Use of Medicare race information will reduce the problem of missing race in VA studies using administrative data

# Recommendations

- Use of a dichotomous race classification of Black/African American and Other in VA studies using Medicare race information results in higher rates of accurate classification than other groupings
  - VA North American Natives and Hispanics frequently misclassified as White (and Non-Hispanic) in Medicare
  - Medicare White and African-American categories, both had high predictive values

# Recommendations

- Medicare data cannot be used to identify Hispanics with any degree of accuracy or completeness
- Consider other supplementary data sources
  - Department of Defense
  - Special surveys

# Session Objectives

- Introduction/Overview
- Race and Ethnicity in the VA Data
- Race and Ethnicity in Medicare Data
- Quality of VA Race/Ethnicity Data
- Recommendations
- Where to Go for More Help



# VIReC Help

- VIReC Webpage (<http://www.virec.research.va.gov>)
  - Information on VA data sources and how to access data
  - Documentation on some VA datasets, e.g., MedSAS datasets
- HSRData Listserv
  - Join at the VIReC Web site
  - Discussion among >500 data stewards, managers, and users
  - Past messages; in archive (on intranet)
- VIReC Help Desk
  - VIReC staff will answer your question and/or direct you to available resources on topics
  - [VIReC@va.gov](mailto:VIReC@va.gov) or (708) 202-2413



# Selected Recent References on Race/Ethnicity Data

- AHRQ (Agency for Healthcare Research and Quality) (2005). *National healthcare disparities report, 2005* (Rep. No. AHRQ Publication No. 06-0017). Rockville, MD: U.S. Department of Health and Human Services, Agency for Healthcare Research and Quality.
- Baker DW, Cameron KA, Feinglass J, Thompson JA, Geor gas P, Fos ter S, et al. (2006). A system for rapidly and accurately collecting patients' race and ethnicity. *Am J Public Health, 96*, 532-537.
- Bertolli J, LeeLisa M, Sullivan PS. (2007). Racial Misidentification of American Indians/Alaska Natives in the HIV/AIDS Reporting Systems of Five States and One Urban Health Jurisdiction, U.S., 1984–2002. *Public Health Reports, 122*, 382-392.
- Blustein J. (1994). The Reliability of Racial Classifications in Hospital Discharge Abstract Data. *American Journal of Public Health, 84*, 1018-1021.
- Boehmer U, Kressin NR, Berlowitz DR, Christiansen CL, Kazis LE, Jones JA. (2002). Self-reported vs administrative race/ethnicity data and study results. *Am J Public Health, 92*, 1471-1472.
- Brahan D, Bauchner H. (2005). Changes in reporting of race/ethnicity, socioeconomic status, gender, and age over 10 years. *Pediatrics, 115*, e163-e166.
- Clegg LX, Reichman ME, Hankey BF, Miller BA, Lin YD, Johnson NJ, et al. (2007). Quality of race, Hispanic ethnicity, and immigrant status in population-based cancer registry data: implications for health disparity studies. *Cancer Causes Control, 18*, 177-187.
- Eicheldinger C, Bonito A. (2008). More accurate racial and ethnic codes for Medicare administrative data. *Health Care Financ Rev, 29*, 27-42.
- Elliott MN, Fremont A, Morrison PA, Pantoja P, Lurie N. (2008). A new method for estimating race/ethnicity and associated disparities where administrative records lack self-reported race/ethnicity. *Health Serv Res.*
- Ford ME, Kelly PA. (2005). Conceptualizing and categorizing race and ethnicity in health services research. *Health Serv Res, 40*, 1658-1675.
- Friedman DJ, Cohen BB, Averbach AR, Norton JM. (2000). Race/ethnicity and OMB Directive 15: implications for state public health practice. *Am.J Public Health, 90*, 1714-1719.
- Gomez SL, Kelsey JL, Glaser SL, Lee MM, Sidney S. (2005). Inconsistencies between self-reported ethnicity and ethnicity recorded in a health maintenance organization. *Ann Epidemiol, 15*, 71-79.
- Gomez SL, Glaser SL. (2006). Misclassification of race/ethnicity in a population-based cancer registry (United States). *Cancer Causes Control, 17*, 771-781.
- Hahn RA. (1992). The state of federal health statistics on racial and ethnic groups. *JAMA, 267*, 268-271.
- Hahn RA, Stroup DF. (1994). Race and ethnicity in public health surveillance: criteria for the scientific use of social categories. *Public Health Rep, 109*, 7-15.
- Hamilton NS, Edelman D, Weinberger M, Jackson GL. (2009). Concordance between self-reported race/ethnicity and that recorded in a Veteran Affairs electronic medical record. *N C Med J, 70*, 296-300.
- Institute of Medicine. (2003). *Unequal treatment: Confronting racial and ethnic disparities in health care* Washington, DC: National Academies Press.
- Jones CP, Truman BI, Elam-Evans LD, Jones CA, Jones CY, Jiles R, et al. (2008). Using "socially assigned race" to probe white advantages in health status. *Ethn Dis, 18*, 496-504.
- Kashner TM. (1998). Agreement between administrative files and written medical records: a case of the Department of Veterans Affairs. *Med Care, 36*, 1324-1336.
- Kramer BJ, Wang M, Hoang T, Harker JO, Finke B, Saliba D. (2006). Identification of American Indian and Alaska Native veterans in administrative data of the Veterans Health Administration and the Indian Health.
- Laws MB, Heckscher RA. (2002). Racial and ethnic identification practices in public health data systems in New England. *Public Health Rep, 117*, 50-61.

# Selected Recent References on Race/Ethnicity Data

- Long JA, Bamba MI, Ling B, Shea JA. (2006). Missing race/ethnicity data in Veterans Health Administration based disparities research: a systematic review. *J Health Care Poor Underserved*. 17(1):128-40. Review.
- Mays VM, Ponce NA, Washington DL, Cochran S D. (2003). Classification of race and ethnicity: implications for public health. *Annu Rev Public Health*, 24, 83-110.
- McAlpine DD, Beebe TJ, Davern M, Call K T. (2007). Agreement between self-reported and administrative race and ethnicity data among Medicaid enrollees in Minnesota. *Health Serv Res*, 42, 2373-2388.
- McBean AM. (2006). Improving Medicare's Data on Race and Ethnicity. National Academy of Social Insurance. Medicare Brief, No. 15.  
Ref Type: Serial (Book, Monograph)
- Morgan RO, Wei II, Virnig BA. (2004). Improving identification of Hispanic males in Medicare: use of surname matching. *Med Care*, 42, 810-816.
- Office of Management and Budget *Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity, Notice of Decision* (Rep. No. 62).
- Pan CX, Glynn RJ, Mogun H, Choodnovskiy I, Avorn J. (1999). Definition of race and ethnicity in older people in Medicare and Medicaid. *J Am Geriatr Soc*, 47, 730-733.
- Polednak AP. (2001). Agreement in race-ethnicity coding between a hospital discharge database and another database. *Ethn Dis*, 11, 24-29.
- Rhoades D. (2005). Racial Misclassification and Disparities in Cardiovascular Disease Among American Indians and Alaska Natives. *Circulation*, 111, 1250-1256.
- Saha S, Freeman M, Toure J, Tippens KM, Weeks C, Ibrahim S. (2008). Racial and ethnic disparities in the VA Health Care System: A Systematic Review. *Journal of General Internal Medicine*, 23, 654-671.
- Sohn M, Zhang H, Arnold N, Stroupe K, Taylor B, Wilt T, et al. (2006). Transition to the new race/ethnicity data collection standards in the Department of Veterans Affairs. *Population Health Metrics*, 4.
- Sondik EJ, Lucas JW, Madans JH, Smith, SS. (2000). Race/ethnicity and the 2000 census: implications for public health. *Am.J Public Health*, 90, 1709-1713.
- Stehr-Green P, Bettles J, Robertson LD. (2002). Effect of racial/ethnic misclassification of American Indians and Alaska Natives on Washington State death certificates, 1989-1997. *American Journal of Public Health*, 92, 443-444.
- Stroupe KT, Tarlov E, Zhang Q, Haywood T, Owens A, Hynes DM. Use of Medicare and DoD data for improving VA race data quality. *Journal of Rehabilitation Research & Development*. 2010;47(8):781-795.
- Sugarman J, Soderberg R, Gordon J, Rivara, FP. (1993). Racial misclassification of American Indians: its effect on injury rates in Oregon, 1989 through 1990. *Am J Public Health*, 83, 681-684.
- Sugarman J, Holliday M, Oss, A, Astorina J, Hui Y. (1996). Improving American Indian cancer data in the Washington State Cancer Registry using linkages with the Indian Health Service and Tribal Records. *Cancer*, 78, 1564-1568.
- The Joint Commission: Advancing Effective Communication, Cultural Competence, and Patient- and Family-Centered Care: A Roadmap for Hospitals. Oakbrook Terrace, IL: The Joint Commission, 2010.
- Thoroughman DA, Frederickson D, Cameron D, Shelby L, Cheek, JE. (2002). Racial misclassification of American Indians in Oklahoma State Surveillance Data for Sexually Transmitted Diseases. *American Journal of Epidemiology*, 155, 1137-1141.
- US Department of Veterans Affairs (2003). *VHA Directive 2003-027, Capture of Race and Ethnicity Categories* Washington, DC: US Department of Veterans Affairs.
- US Department of Veterans Affairs (2009). *VHA Handbook 1601A.01, Intake Registration* Washington, DC: U.S. Department of Veterans Affairs.
- Veterans Health Administration Decision Support Office (2009). *National Data Extract Technical Guide* Bedford, MA: U.S. Department of Veterans Affairs.
- Wei II, Virnig BA, John DA, Morgan RO. (2006). Using a Spanish surname match to improve identification of Hispanic women in Medicare administrative data. *Health Serv Res*, 41, 1469-1481.



**Questions?**



# Upcoming Seminar

- June 4
- Using VA Corporate Data Warehouse (CDW) to Assess Vitals

