2009-2010 VIReC Database and Methods Cyber Seminar Series

2009-2010 VIReC Database and Methods Cyber Seminar Series

Assessing Race and Ethnicity

Session 10
February 1, 2010

Presented by:
Kevin Stroupe

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

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**

Audience Poll

- **Have you ever used race data?**
  - Yes
  - No

Audience Poll

- **How would you rate your overall knowledge of race data?**
  - 1 (Never Used)
  - 2
  - 3
  - 4
  - 5 (Used Frequently, Very familiar)

Introduction/Overview

- **Importance of race and ethnicity information for research**
  - Research on racial/ethnic disparities in health care
  - Research on clinical factors associated with race
- **Problems identified with race and ethnicity data in the VA**
  - Completeness
  - Consistency over time

Introduction / Overview

- **Well-documented and persistent healthcare disparities in the U.S.** (IOM, 2003; AHRQ, 2005)
  - Quantity and quality of health care is related to race and ethnicity (particularly African Americans and Hispanics)
- **Root causes and solutions not clear** (IOM; AHRQ)
- **Healthcare disparities are also found in VA where financial barriers to receiving care are minimized** (Saha, 2008)
- **Consequently, additional research on healthcare disparities is important**

Introduction/Overview

- **Obtaining race may also be important for addressing clinical questions in research**
- **For example:**
  - Race used when studying chronic kidney disease because African American race (along with age, gender, and serum creatinine) is in the calculation of the estimated glomerular filtration rate (eGFR), which is used to determine patients stage of chronic kidney disease
  - In studies of medications, it may be important to know race because appropriate medication management may differ by race in some cases (e.g., hypertension)

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

Race and Ethnicity in the VA
Race and Ethnicity Categories (2003-present)

- **Ethnicity**
  - Hispanic or Latino
- **Race**
  - American Indian or Alaskan Native
  - Asian
  - Black or African American
  - Native Hawaiian or Pacific Islander
  - White
- **Current method of VA race and ethnicity reporting:**
  - 2 question format: ethnicity, race
  - Self-reported
  - Multiple races

Race and Ethnicity in the VA

Directives Regarding Standards

- **OMB Directive No. 15 revision**
  - In 1997 the OMB revised Directive 15 and established the latest standards for maintaining, collecting, and presenting federal data on race and ethnicity (OMB, 1997)
- **VHA implementation**
  - The VHA, through Directive 2003-027, implemented the mandated revisions in 2003 (DVA, 2003)

Race and Ethnicity in the VA

The Who, What, Where, When, & How of race/ethnicity information acquisition in VHA

WHO:	WHAT:	WHEN:	WHERE:	HOW:
Information Source: • Patient (self-report) • Proxy	• VA Form 10-10EZ, Application for Health Benefits (online, paper, or by interview)	• Enrollment • Hospital admission • Outpatient visit or pre-registration	• Online • Telephone call from local VHA facility • In-person visit to local VHA facility	• VHA Facility Enrollment Coordinator or designee (e.g., Admission Interview Clerk, Enrollment Specialist) or • Outpatient clinic personnel • Collects the information and enters into VistA

Race and Ethnicity in the VA

Race data collection: Enrollment

- Race and ethnicity information for VA patients is obtained through completion of VA Form 10-10EZ (*Application for Health Benefits*) at enrollment
- The information is then entered into the Veterans Health Information Systems and Technology Architecture (VistA) and stored in the Race Information and Patient Information Sub-files
- Race and other demographic information is transmitted with patient healthcare encounter data to the Austin Information Technology Center (AITC) where it is stored in the National Patient Care Database (NPCD)
- National Data Systems directs the construction of extracts of these data, known as the Medical SAS (MedSAS) datasets

Race and Ethnicity in the VA

- **Race data collection: Clinical setting**
  - Recent requirement to ask ethnicity and race questions during pre-registration, if missing
  - Race and ethnicity may also be obtained by personnel if missing in VistA
  - This information is gathered directly from the patient (or proxy) at hospital admission or clinic registration
  - Race and ethnicity are then entered into VistA and stored in the Race Information and Patient Information Sub-files
  - Separate VistA field for method of data collection

Race and Ethnicity in the VA
Historical Variables

- **Historical method of race and ethnicity reporting in the VA**
  - Method of ascertainment uncertain, primarily observer-reported
  - No multiple race reporting option
  - Single question captured both race and ethnicity:
    - Hispanic, White
    - Hispanic, Black
    - American Indian
    - Black
    - White

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

Race and Ethnicity in the VA

- **Sources of race data in the VA: Other data sources**
  - 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
- VA Race Data Quality
- Recommendations
- Where to Go for More Help

Medicare Race/Ethnicity

- **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 (SSA) administrative records**
  - 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) is limited to 4 original categories
  - SSN application form – single question format and no multiple race reporting
- **Initiatives to improve race and ethnicity completeness and quality**
  - Periodic updates on American Indians and Alaskan Natives from Indian Health Service
  - 1997 mailed survey to 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
- VA Race Data Quality
- Recommendations
- Where to Go for More Help

VA Race Data Quality

- **Consistency over time**
  - Due to implementation of revised OMB directive there is discontinuity over time in race categories
  - The VA's implementation of the revised OMB directive directs that
    - Race and ethnicity be collected and reported separately
    - Race be collected by self-report or proxy
    - Race categories comply with the federal standard
  - New race/ethnicity data standards implemented:
    - MedSAS inpatient files 2003
    - MedSAS outpatient files 2004

Consistency Over Time

- **Sohn, et al. (2006)**
  Compared all unique users of VA healthcare services with self-reported race/ethnicity data in 2004 with their prior race/ethnicity data from a pre-VHA directive period, 1997–2002 (N = 988,277)
- **Results from comparison of self-report with pre-VHA directive values:**
  - Non-Hispanic White patients: 98% agreement
  - NH African American patients: 94% agreement
  - For non-African American minorities agreement was much less: American Indian/Alaska Native: 32%
    Asian/Pacific Islander: 35%

VA Race Data Quality

- **Consistency over time**
  - Agreement improved when categories were combined

Categories	Agreement	Kappa
White, African American, Other	95.3	.88
White, Other	95.7	.89
African American, Other	97.8	.93

VA Race Data Quality

- **Completeness:**

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

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

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

VIReC VA Race Data Quality Study
Background

- **Options for filling in missing race in VA databases**
  - Filling in missing race values using data from other years of VA data
  - Filling in missing race using non-VA data sources

- **For patients enrolled in Medicare**
  - Medicare databases serve as a source of non-VA race information
  - To evaluate the practice of using Medicare data to fill in *missing* VA race data, it is important to assess the agreement between the *non-missing* VA race data and Medicare race data

VIReC VA Race Data Quality Study

Objectives

1. **To estimate the extent missing "usable" race in MedSAS data could be reduced using multiple years of MedSAS data**

2. **To estimate the extent missing usable race in MedSAS data could be reduced after merging VA and Medicare data**

3. **To evaluate the agreement between Medicare and VA self-reported race data in MedSAS files**
   - Using data from 2004 to 2005 after VA implemented OMB Directive
   - Examining sensitivities & specificities
   - Examining Kappa statistic to evaluate agreement

VIReC VA Race Data Quality Study

Methods

- **Patient cohort**
  - 10% representative sample of VHA patients who obtained services during FY1997 – FY 2005
  - A unique patient was defined as a unique SSN, birth date, and gender combination
- **Race data for patient cohort from 1997 to 2005 in**
  - MedSAS Outpatient Visit files
  - MedSAS Inpatient Acute Main files
  - N = 939,021
- **We obtained Medicare race data from Medicare Vital Status file**

VIReC VA Race Data Quality Study

Results: Characteristics of Patients with & without Usable Race Data VA MedSAS

	Usable Race Value (n=416,611)	No Usable Race Value (n=522,410)
Age, mean (SD)	64.5 (15.6)	59.1(18.9)
Male, %	94	82
Number of Inpatient Stays:		
Mean (SD)	2.9 (3.4)	1.5 (1.1)
Median	2	1
Number of Outpatient Visits:		
Mean (SD)	61.6 (97.3)	12.2 (24.3)
Median	33	5
Years Followed in Study, Mean (SD)	4.9 (2.8)	2.6 (2.1)[31]

Aim 1:
Improvement in Single Year Data with Additional Years of MedSAS Data

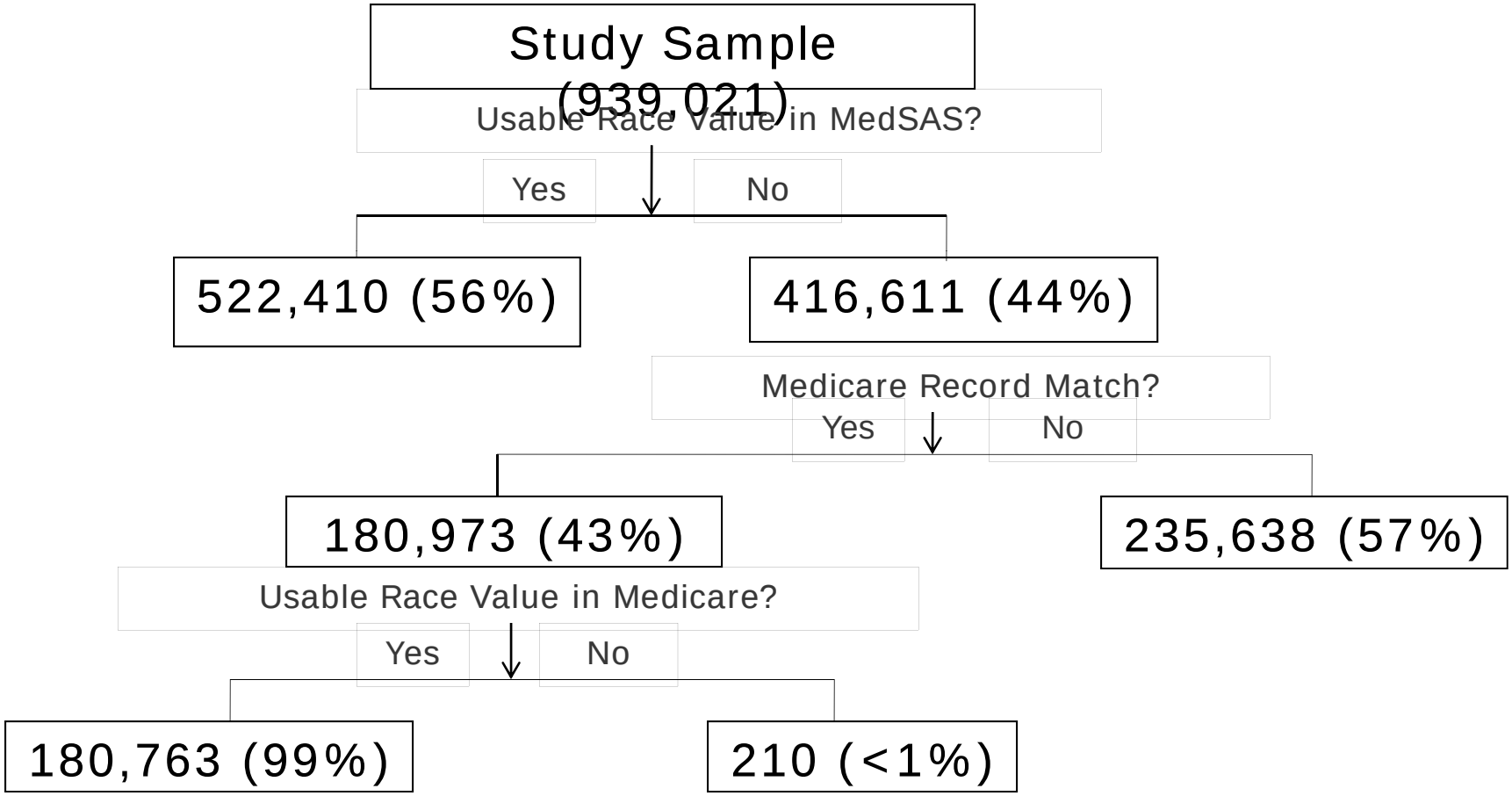
	N	Usable Race Single Yr Data, %	Usable Race All Yrs Data, %
1997	300,272	57	74
1998	328,050	57	73
1999	342,525	58	74
2000	367,004	58	73
2001	407,993	56	71
2002	446,416	55	69
2003	471,794	49	67
2004	488,947	62	68
2005	502,205	67	69

Aim 1:
Improvement in Single Year Data with Additional Years of MedSAS Data

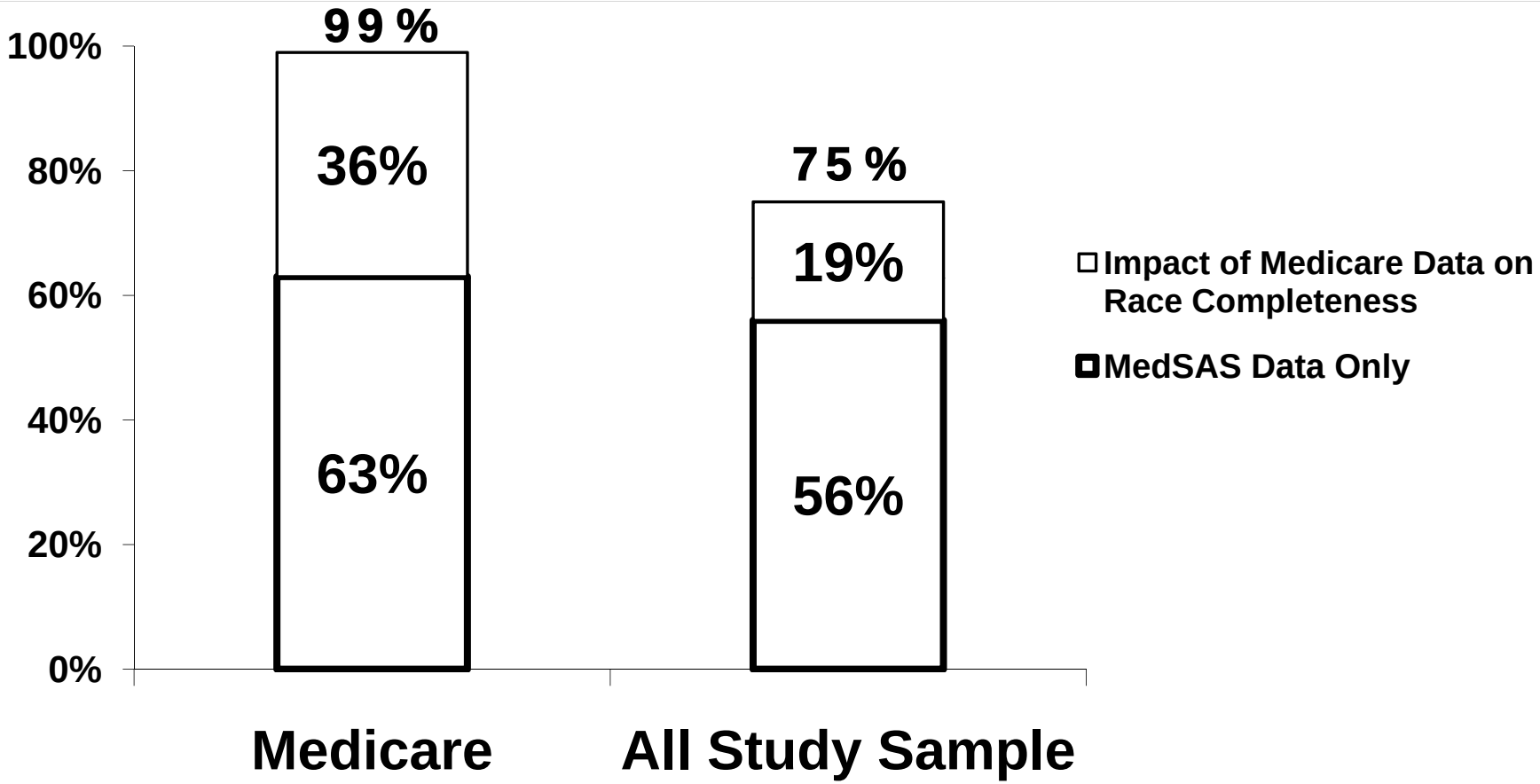
- **This approach resulted in a 3% (in 2005) to 29% (in 1997) improvement in race data completeness**

- **The largest increases were found in patient samples from earlier years**
  - Reflecting the relative utility of searching later rather than earlier years of data

Aim 2:
Improvement in Race Completeness with Addition of Medicare Data from 1997-2005



Aim 2:
Improvement in Race Completeness with Addition of Medicare Data from 1997-2005



Aim 3:
Comparison of Medicare with VA Data, 2004-2005

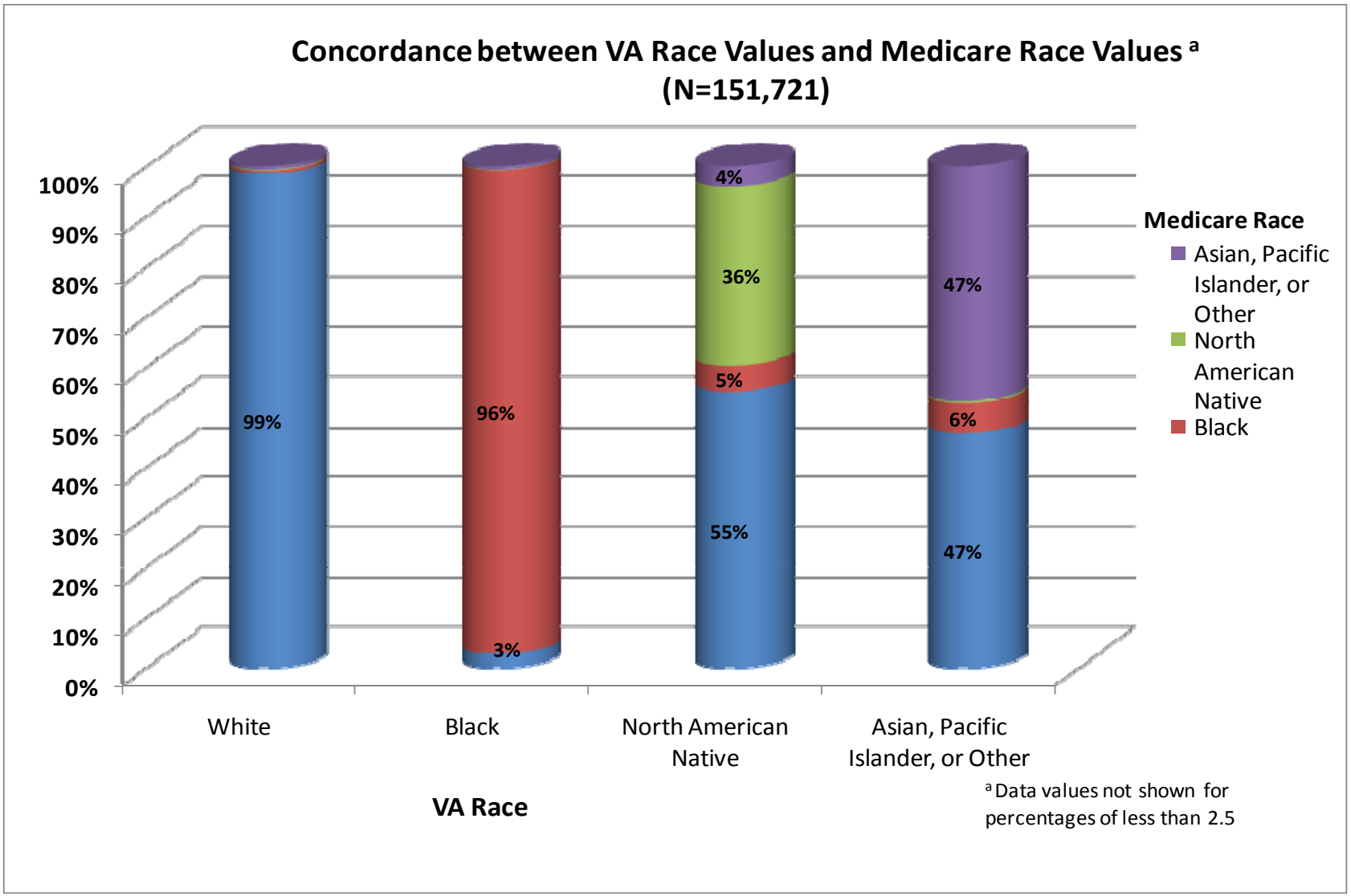
- **Analysis restricted to FY 2004 – 2005 data in order to use VA self-reported data as the "gold standard"**

- **151,721 patients enrolled in Medicare from 2004-2005 with usable race value in both VA and Medicare data**

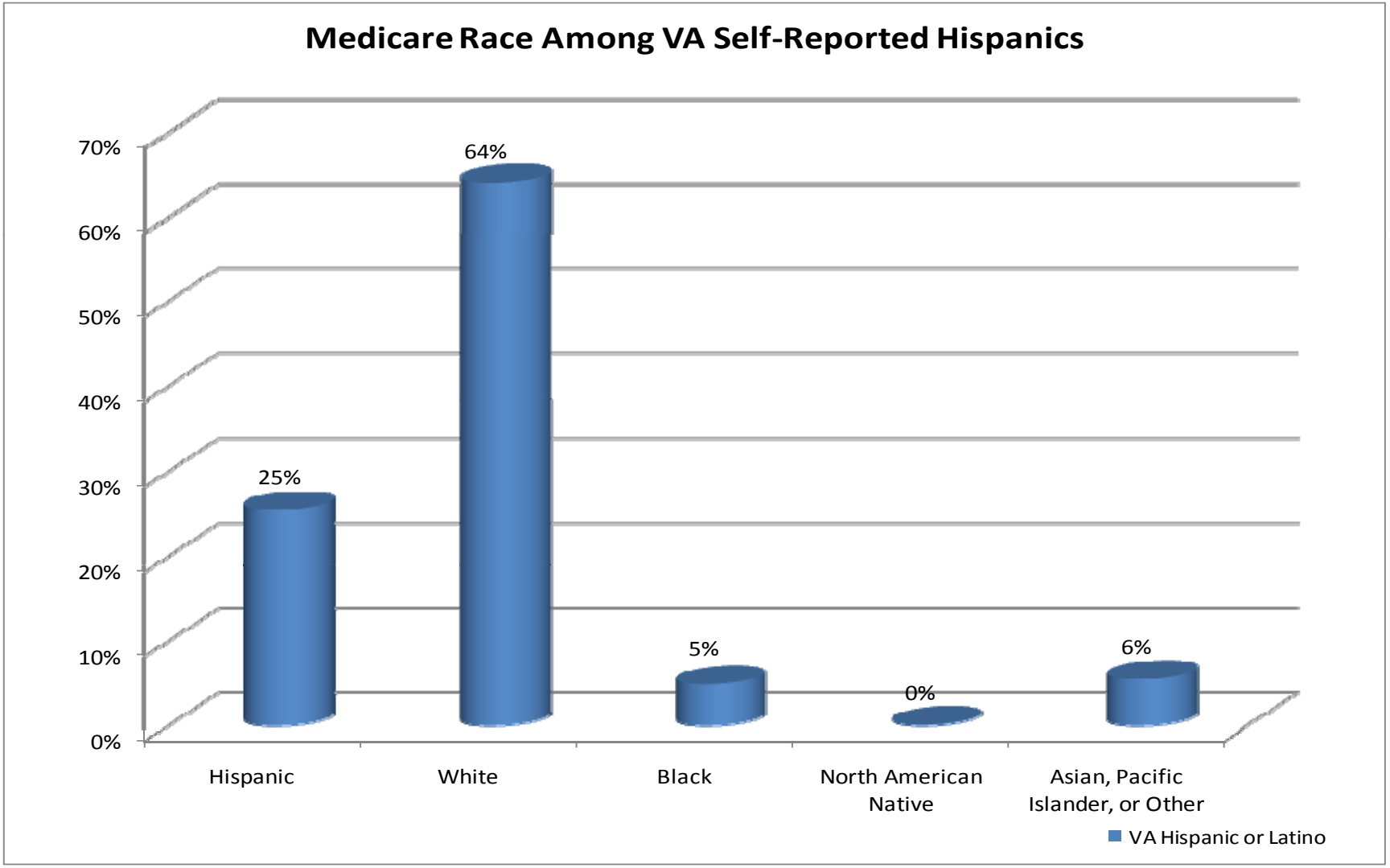
Aim 3:
Comparison of Medicare with VA Data, 2004-2005

VA Race	Medicare Race	Classification Constructed for Consistency Analysis
White	White	White
Black or African American	Black	Black or African American
American Indian or Alaska Native	North American Native	North American Native
Asian	Asian	Asian, Pacific Islander, or Other
Native Hawaiian or Other Pacific Islander	Other	

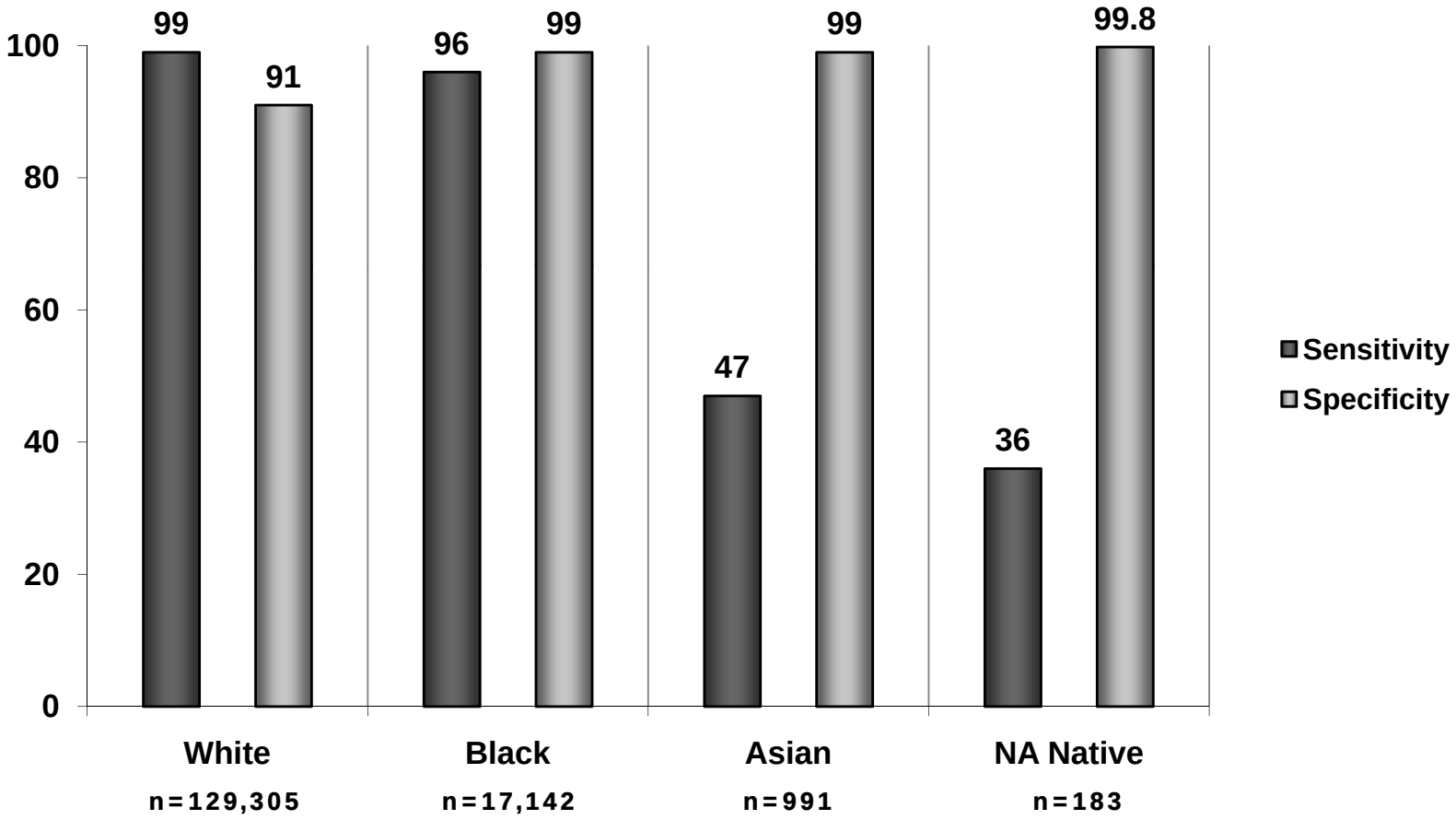
Aim 3:
Comparison of Medicare with VA Data, 2004-2005



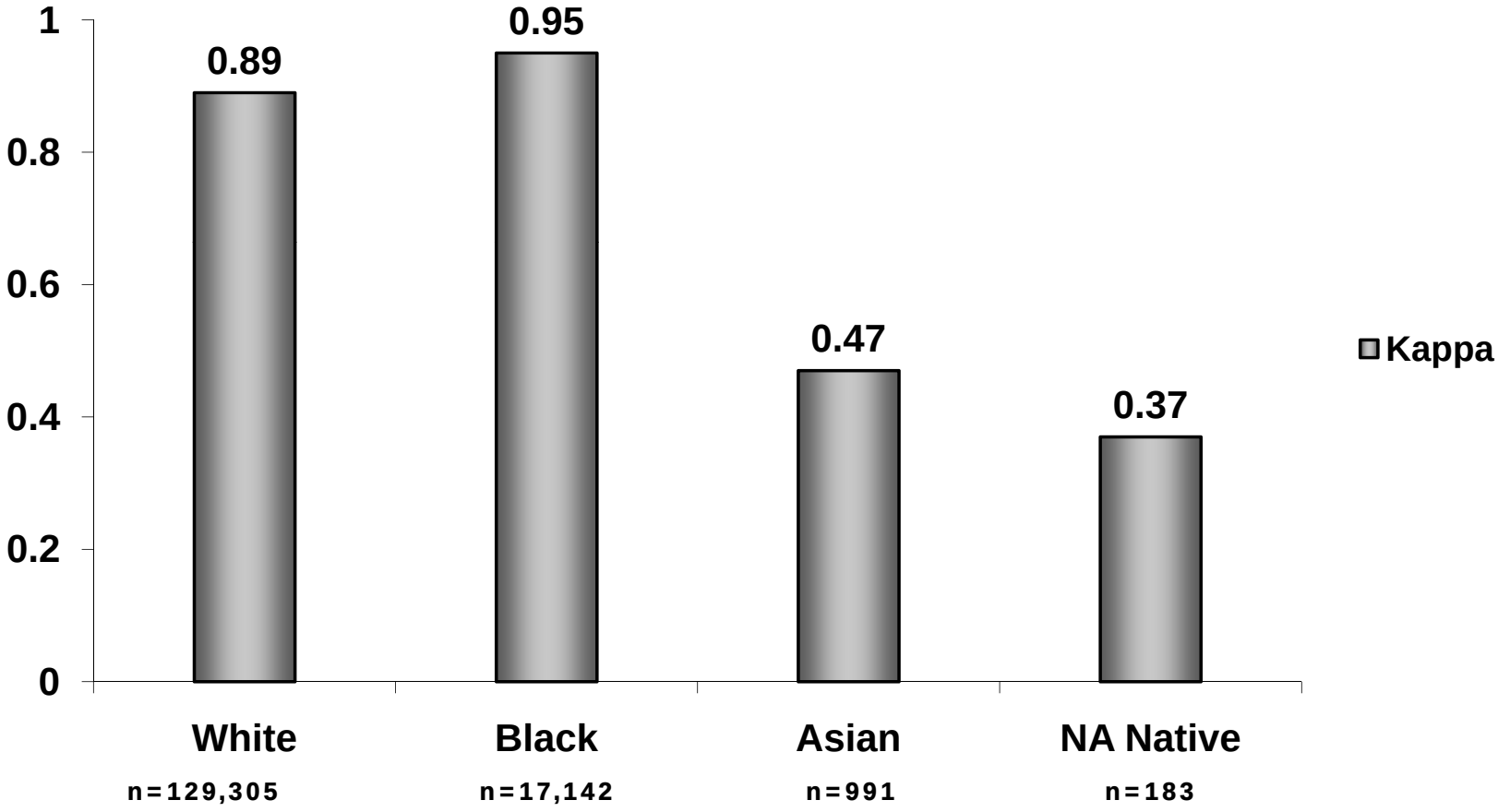
Aim 3:
Comparison of Medicare with VA Data, 2004-2005



Aim 3:
Comparison of Medicare with VA Data, 2004-2005



Aim 3:
Comparison of Medicare with VA Data, 2004-2005



Aim 3:
Comparison of Medicare with VA Data, 2004-2005

- **Among patients self-identified as White or Black in VA data**
  - 99% and 96%, respectively, had concordant information in Medicare records (Kappa 0.89 and 0.95)
- **Among the 1,174 non-Black minority patients**
  - 36% of North American Natives and 47% of Asians, Pacific Islanders, or Others were coded the same in Medicare
  - For these minority groups, Medicare data had low sensitivity and low overall rates of agreement
  - **VA minorities misclassified in Medicare**
    - Were most often classified as White

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

Recommendations

- **Use of Medicare race information will reduce the problem of missing race in VA studies using administrative data**
- **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

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**
  - Given the frequency with which VA North American Natives and Hispanics were classified as White (and Non-Hispanic) in Medicare and since Medicare White and African-American categories both had high predictive values

Recommendations

- **To supplement missing VA self-reported ethnicity data, it may be preferable to use pre-2004 VA data, rather than Medicare data** (Sohn, 2006)
  - 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
- VA Race Data Quality
- 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

  – (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, D. W., Cameron, K. A., Feinglass, J., Thompson, J. A., Georgas, P., Foster, 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, P. S. (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, N. R., Berlowitz, D. R., Christiansen, C. L., Kazis, L. E., & Jones, J. A. (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, L. X., Reichman, M. E., Hankey, B. F., Miller, B. A., Lin, Y. D., Johnson, N. J. 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, M. N., Fremont, A., Morrison, P. A., 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, M. E. & Kelly, P. A. (2005). Conceptualizing and categorizing race and ethnicity in health services research. *Health Serv Res, 40,* 1658-1675.

Friedman, D. J., Cohen, B. B., Averbach, A. R., & Norton, J. M. (2000). Race/ethnicity and OMB Directive 15: implications for state public health practice. *Am.J Public Health, 90,* 1714-1719.

Gomez, S. L., Kelsey, J. L., Glaser, S. L., Lee, M. M., & Sidney, S. (2005). Inconsistencies between self-reported ethnicity and ethnicity recorded in a health maintenance organization. *Ann Epidemiol, 15,* 71-79.

Gomez, S. L. & Glaser, S. L. (2006). Misclassification of race/ethnicity in a population-based cancer registry (United States). *Cancer Causes Control, 17,* 771-781.

Hahn, R. A. (1992). The state of federal health statistics on racial and ethnic groups. *JAMA, 267,* 268-271.

Hahn, R. A. & Stroup, D. F. (1994). Race and ethnicity in public health surveillance: criteria for the scientific use of social categories. *Public Health Rep, 109,* 7-15.

Hamilton, N. S., Edelman, D., Weinberger, M., & Jackson, G. L. (2009). Concordance between self-reported race/ethnicity and that recorded in a Veteran Affairs electronic medical record. *N C Med J, 70,* 296-300.

IOM (Institute of Medicine) (2003). *Unequal treatment: Confronting racial and ethnic disparities in health care* Washington, DC: National Academies Press.

Jones, C. P., Truman, B. I., Elam-Evans, L. D., Jones, C. A., Jones, C. Y., Jiles, R. et al. (2008). Using "socially assigned race" to probe white advantages in health status. *Ethn Dis, 18,* 496-504.

Kashner, T. M. (1998). Agreement between administrative files and written medical records: a case of the Department of Veterans Affairs. *Med Care, 36,* 1324-1336.

Kramer, B. J., Wang, M., Hoang, T., Harker, J. O., 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

Selected Recent References on Race/Ethnicity Data

Laws, M. B. & Heckscher, R. A. (2002). Racial and ethnic identification practices in public health data systems in New England. *Public Health Rep, 117,* 50-61.

Mays, V. M., Ponce, N. A., Washington, D. L., & Cochran, S. D. (2003). Classification of race and ethnicity: implications for public health. *Annu Rev Public Health, 24,* 83-110.

McAlpine, D. D., Beebe, T. J., 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, A. M. (2006). Improving Medicare's Data on Race and Ethnicity. National Academy of Social Insurance. Medicare Brief, No. 15. Ref Type: Serial (Book,Monograph)

Morgan, R. O., Wei, I. I., & Virnig, B. A. (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, C. X., Glynn, R. J., 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, A. P. (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, K. M., 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, E. J., Lucas, J. W., Madans, J. H., & Smith, S. S. (2000). Race/ethnicity and the 2000 census: implications for public health. *Am.J Public Health, 90,* 1709-1713.

Stehr-Green, P., Bettles, J., & Robertson, L. D. (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.

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.

Thoroughman, D. A., 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.

U.S.Department of Veterans Affairs (2003). *VHA Directive 2003-027, Capture of Race and Ethnicity Categories* Washington, DC: U.S.Department of Veterans Affairs.

U.S.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, I. I., Virnig, B. A., John, D. A., & Morgan, R. O. (2006). Using a Spanish surname match to improve identification of Hispanic women in Medicare administrative data. *Health Serv Res, 41,* 1469-1481.

Questions?

Next Seminar

- **March 1, 2010**
  - Using VA Corporate Data Warehouse to assess vitals
  - Ruth Perrin