HHS Public Access

Author manuscript

Med Care. Author manuscript; available in PMC 2016 August 01.

Published in final edited form as:

Med Care. 2015 August ; 53(8): 666–672. doi:10.1097/MLR.0000000000000392.

Standardized Data Collection Practices and the Racial/Ethnic Distribution of Hospitalized Patients

Rosette J. Chakkalakal, MD, MHS[1], **Jeremy C. Green, PhD**[2], **Harlan M. Krumholz, MD**[1,3], and **Brahmajee K. Nallamothu, MD**[4]

[1]Robert Wood Johnson Foundation Clinical Scholars Program, Department of Internal Medicine, Yale University School of Medicine, 333 Cedar Street, New Haven, CT 06520

[2]Department of Health Management and Policy, Saint Louis University, 3545 Lafayette Avenue, St. Louis, MO 63104

[3]Center for Outcomes Research and Evaluation, Yale-New Haven Hospital, 333 Cedar Street, New Haven, CT 06510

[4]Division of Cardiovascular Medicine and Center for Health Outcomes and Policy, University of Michigan Medical School, 2800 Plymouth Rd. Ann Arbor, MI 48109

Abstract

Background—Although frequently used to track healthcare disparities, patient race/ethnicity data collected by hospitals can be unreliable, particularly for smaller minority groups. We sought to determine if the racial/ethnic distribution of hospitalized patients shifted after implementation of a statewide initiative to standardize data collection practices.

Methods—We conducted a difference-in-differences analysis of the State Inpatient Databases to estimate changes in the proportion of patients identified as non-Hispanic white, non-Hispanic black, Hispanic, Asian/Pacific Islander, and "other," before (2005–2006) and after (2008–2009) standardized practices were implemented in New Jersey (NJ) relative to New York (NY), a state with similar demographics but no changes to data collection.

Results—Among 12,552,702 hospital discharges, modest relative changes were noted in the proportion of patients identified as non-Hispanic white (+1.1%; 95% CI +0.9 to +1.2) and non-Hispanic black (+1.6%; 95% CI +1.1 to +2.1) in NJ that were attributed to its use of standardized data collection practices as compared with NY. Larger relative changes were noted in the proportion of patients identified as Hispanic (−7.1%; 95% CI −7.8 to −6.4), Asian/Pacific Islander (+26.5%; 95% CI +25.1 to +27.9) and "other" (−24.6%; 95% CI −26.4 to −22.8). This pattern was largely consistent in analyses stratified by gender, age, and Major Diagnostic Category.

Conclusions—Measurement of healthcare disparities fundamentally depends on the racial/ethnic categorization of individuals. By redistributing substantial proportions of patients across smaller minority groups, standardized data collection could lead to shifts in estimates of healthcare disparities for these rapidly growing populations.

Correspondence to: Rosette J. Chakkalakal, MD, MHS, 1215 21st Ave South, Suite 6000-North Tower, Nashville, TN 37232, Telephone: 615-936-1010, Fax: 615-936-1269, rosette.j.chakkalakal@vanderbilt.edu.

INTRODUCTION

In the landmark report *Unequal Treatment*, the Institute of Medicine (IOM) highlighted the need to track and compare healthcare utilization, delivery, and outcomes across racial and ethnic groups to ensure equitable care.[1] Although information on the race/ethnicity of patients collected by hospitals is frequently used to track such measures, several studies suggest these data are unreliable due to faulty data collection practices.[2–5] Hospitals frequently report patient race and/or ethnicity based on the observations of admitting clerks[6] rather than self-identification which is considered the gold standard.[7] Determinations of race/ethnicity by hospital staff are especially problematic for smaller minority groups.[3–5,8,9] Additionally, in 2011, only 16 of 43 states reported using current Federal standard categories for race and ethnicity established by the Office of Management and Budget (OMB) in hospital claims, with most discrepancies occurring in the classification of smaller minority groups such as Hispanic and Asian.[10] The absence of race and ethnicity fields on the national standard for hospital claims prior to 2007 and variable uptake of updates to the OMB standard in 1997 likely perpetuated these inconsistencies.[10]

To address these concerns, the IOM recommends that hospitals enable patients to self-identify their race and ethnicity using a uniform set of categories that expand on the current (1997) OMB standards.[11] The New Jersey Hospital Association (NJHA) is one of a few healthcare systems that have recently implemented these recommendations. In 2007, the Health Research and Education Trust (HRET) of NJHA launched an innovative statewide strategy to standardize practices for collecting more specific race and Hispanic origin data from patients in order to better understand the healthcare needs of the state.[12] Understanding if and how these changes to data collection practices affect assessments of the racial/ethnic distribution of hospitalized patients is unknown but of critical importance given the ubiquitous use of race/ethnicity data collected by hospitals to track healthcare disparities.

Accordingly, we sought to determine how patients may be shifted across different race/ethnicity categories in association with implementation of standardized data collection practices in New Jersey (NJ).

METHODS

Data Sources and the NJ HRET Program

We analyzed the State Inpatient Databases (SIDs) to compare the racial/ethnic distribution of hospitalized patients age 18 and older before (2005–2006) versus after (2008–2009) implementation of standardized data collection practices in NJ (intervention group) relative to New York (NY, comparison group). The SID, which is part of the Agency for Healthcare Research and Quality's Healthcare Cost and Utilization Project (HCUP), contains the universe of inpatient discharge abstracts provided by state data partners (the New Jersey Department of Health and Senior Services and the New York State Department of Health).[13] Training and implementation of the new system in NJ occurred throughout 2007; therefore, data from 2007 was excluded. Individuals with missing age, gender, or race data (less than 2% of discharges) were excluded from the analysis. The study was deemed exempt from review by the Yale Institutional Review Board.

The NJ HRET campaign included standardized guidelines and uniform protocols for the collection of self-identified patient race/ethnicity data which were disseminated statewide by: (1) conducting training programs for hospital intake workers, access managers, supervisors and registrars; (2) adapting information technology systems to be compatible with the standardized guidelines; and (3) designing and distributing educational tools, resources, and reference toolkits to hospital staff to ensure sustainability of the campaign.[12] As part of this effort, HRET also collaborated with the New Jersey Department of Health and Senior Services to mandate adoption of revised standardized categories for reporting race and ethnicity in January 2007 which expanded on the OMB standards to include more specific race and ethnicity categories consistent with US Census Bureau classifications (see Appendix Table 1).[12]

Study Outcomes

The outcome variable of interest was patient race/ethnicity, as reported according to categories used by HCUP. Because of known variation in coding race/ethnicity in claims data across states, HCUP uniformly reports the information received from all participating states using 6 mutually exclusive categories (see Appendix Table 1) equivalent to the 1977 OMB standards: non-Hispanic white, non-Hispanic black, Hispanic, Asian/Pacific Islander, Native American and “other”.[10] If patients were identified as Hispanic for the NJ or NY SID ethnicity variable, these individuals were reported to be Hispanic for the HCUP race variable regardless of their racial classification in the NJ or NY SID. We included Native Americans in the “other” category because of their relatively small populations in both states. The categories used to collect race/ethnicity data in NY did not change between the study periods (see Appendix Table 1), making it an ideal comparator for NJ in this analysis.

Statistical Analysis

Changes in the proportion of patients identified in each racial/ethnic category that were associated with the new data collection practices in NJ were estimated using a nonlinear difference-in-differences model.[14] We predicted the racial/ethnic distributions of hospitalized patients in NJ and NY as a function of the interaction of state and time (before versus after implementation of standardized data collection practices in NJ), adjusting for the two main effects of state and time, three age categories (18–44, 45–64, ≥65), and gender. For ease of reporting, we used multinomial logistic regression models[15] to generate estimates and then expressed these as the proportion of individuals identified within each race/ethnicity category per 10,000 hospitalized patients. We clustered the models on survey year and used an unconditional variance estimator to carry the clustering through predictions. The statistical significance of the difference-in-differences was assessed with 95% confidence intervals constructed from delta-method standard errors and corrected for multiple testing within each race/ethnicity category.[16,17] In order to assess the relative impact of standardized practices on the identification of patients within each racial/ethnic category, we divided the absolute difference-in-differences by the 2005–2006 proportions in NJ to report the percent change in the proportion of patients identified in each racial/ethnic category in NJ relative to NY (see Figure 1). This assessment of percent change was especially important given the smaller sizes of minority groups.

Inherent in our use of the difference-in-differences approach is the identifying assumption that the racial/ethnic distributions of the underlying populations of NJ and NY would have trended in parallel over time in the absence of the NJ program. To test the robustness of this assumption, we replicated the analysis using the Public Use Microdata Sample of the American Community Survey (ACS). The ACS provides annual estimates on the race/ethnicity of the general population for each state.[18] ACS respondents self-identify their race and ethnicity; additionally, since 2000, they have been able to identify multiple races if indicated.[19] For this analysis, we included multiracial respondents to the ACS in the "other" category. Importantly, the process for collecting race/ethnicity data on the ACS did not change between 2005 and 2009 in either state. Therefore, we were able to use the ACS data to replicate the primary analysis of the SID to test the identifying assumption that any observed shifts in the racial/ethnic distribution of hospitalized patients were independent of demographic shifts in the general population of either state.

Finally, we conducted a series of sensitivity analyses to test the robustness of our results by replicating the primary analysis on different subsets of the SID. For these stratified analyses we analyzed the SID according to (1) Major Diagnostic Categories (MDCs) and (2) age categories (18–44, 45–64, and >65) separately for men and women.

All analyses were conducted in Stata 12 (College Station, TX).[20]

RESULTS

The final study population from the SIDs of NJ and NY included 12,552,702 discharges. Table 1 lists the racial/ethnic distribution, age, gender, and MDCs for hospitalized patients stratified by state prior to the implementation of the new data collection practices in NJ. The distribution of patients across these categories was similar between both states with the exception of NJ having a slightly higher percentage of patients classified as non-Hispanic white and NY having a greater percentage of patients classified as "other".

Overall, we found significant changes in the racial/ethnic distribution of hospitalized patients associated with the new standardized data collection practices in NJ when compared with NY; however, the magnitude of these changes varied substantially across racial/ethnic categories (see Table 2 and Figure 2). For example, the proportion of patients identified as non-Hispanic white and non-Hispanic black changed only modestly between the study periods in NJ relative to NY. In contrast, we observed much larger changes among smaller racial/ethnic categories related to the new data collection practices in NJ. The proportion of patients identified as Hispanic decreased in NJ but increased in NY leading to a relative decrease in the proportion of patients identified as Hispanic in NJ (−7.1%). The proportion of patients identified as Asian/Pacific Islander increased to a greater extent in NJ than in NY, leading to a relative increase in the proportion of patients identified as Asian/Pacific Islander in NJ (+26.5%). Finally, the proportion of patients identified as "other" increased to a lesser extent in NJ than in NY, leading to a relative decrease in the proportion of patients identified as "other" in NJ (−24.6%).

Table 3 reports results from our analysis of changes in the racial/ethnic distribution of the general populations of NJ and NY for the same time periods using ACS data. As expected, assessments of the distribution of race/ethnicity in both populations from these states remained relatively stable over time, indicating our results were unlikely to be explained by underlying demographic shifts. In sensitivity analyses, the percent changes in the proportions for each racial/ethnic category were largely consistent across MDCs as well as across age categories among men and women (see Supplemental Digital Content Figure 1 and Supplemental Digital Content Figure 2).

DISCUSSION

We found substantial alterations in the racial/ethnic distribution of patients hospitalized in NJ following implementation of a statewide effort to standardize the processes and categories used to collect race/ethnicity data directly from patients in 2007. The most notable changes occurred among smaller minority groups with the proportion of patients identifying as Hispanic decreasing by 7%, “other” decreasing by 25%, and Asian/Pacific Islander increasing by 27%. These shifts may have occurred through several mechanisms. For example, the increase in the proportion of patients identified as Asian/Pacific Islander and decrease in the proportion of patients identified as Hispanic and “other” may have been due to misclassification of Asians/Pacific Islanders by hospital staff into the latter two categories in 2005–2006. Self-identification may have also allowed patients whose race/ethnicity was previously unassigned to identify themselves as Asians/Pacific Islanders. The decreases in the proportions of patients identifying as Hispanic and “other” were likely to have been additionally influenced by systemic issues with the conceptualization of race/ethnicity among these groups. A recent study found that individuals identifying as Hispanic, “other”, or multiracial (included in the “other” category in the NJ SID) were more likely to change their self-identified race and/or ethnicity than single-race non-Hispanic white, non-Hispanic black and Asian individuals when comparing responses on the 2000 and 2010 Census.[21] Hispanic individuals often define both their race and ethnicity as Hispanic even though Hispanic is considered an ethnic classification by the federal government; thus, responses to questions of race often vary between white and “other” for Hispanics.[21]

NJ hospitals are among a growing number of healthcare institutions seeking to improve race/ethnicity data collection systems in order to better monitor healthcare disparities. A few studies have evaluated the effect of such interventions on the racial/ethnic identification of specific patients followed before and after data collection changes were made. One study found that as much as 70% of the race/ethnicity data collected for American Indian/Alaskan Native and Asian/Pacific Islander patients by the Department of Veterans Affairs may have been inaccurate prior to 2003 when they mandated the use of self-report for racial/ethnic identification.[22] More recently, the “Expecting Success” program found that among a subset of patients seen both before and after changes were made to the collection of race/ethnicity data in participating hospitals there was no significant effect on the aggregate distribution of reported race and ethnicity for the OMB race categories.[23] However, very few patients from smaller minority groups were included in these studies and neither study assessed the effects of data collection changes on the racial/ethnic distribution of patients at a population level.

To our knowledge, our study is the first to assess the impact of standardized data collection practices at the population, rather than individual, level. Use of the difference-in-differences approach, in particular, enabled us to better attribute specific changes in the racial/ethnic distribution of patients to implementation of standardized data collection practices in NJ. Additionally, our study overcame prior limitations by evaluating an initiative that collected self-reported data and used standardized processes and categories in a state with large minority populations. In doing so, our findings highlight the particular benefit of using standardized data collection practices to obtain self-identified race/ethnicity data for smaller minority groups within hospital discharge data. Sample size concerns for smaller minority groups often limit researchers' ability to generate statistically reliable estimates for these groups in assessments of healthcare disparities.[24] For example, despite the fact that data were collected to identify Asians or Asians/Pacific Islanders in all of the measures included in the 2003 National Healthcare Disparities Report, reliable estimates for Asian Americans could be generated for only 47% of utilization measures because of inadequate sample sizes of this group.[24] Several recent articles have highlighted the need for better data on healthcare utilization and outcomes among Asian Americans.[25–29] By increasing the identification of Asians/Pacific Islanders in healthcare data, the newly implemented data collection system in NJ could significantly improve our ability to recognize disparities that may be affecting these groups.

Our study should be interpreted in the context of its limitations. Since we used cross-sectional data for our analysis, it is impossible to know exactly how standardized practices redistributed patients across racial/ethnic categories. Also, our study does not account for additional factors that could have contributed to the observed shift in the racial/ethnic distribution of hospitalized patients during the study period. For example, variation in the extent to which standardized practices were implemented by hospitals might have influenced our results. Demographic shifts in the general population of NJ and NY would have also significantly influenced the racial/ethnic distribution of hospitalized patients between the study periods. However, our analysis of ACS data indicates our findings were unlikely to be due to demographic shifts in the general population.

Despite these limitations, our results have significant implications for policymakers and providers. Measurement of healthcare disparities fundamentally depends on the accurate and reliable racial/ethnic categorization of individuals. By redistributing substantial proportions of patients across smaller minority groups, standardized data collection practices could lead to significant shifts in estimates of healthcare disparities for these rapidly growing populations. Additionally, the Department of Health and Human Services recently released new standards for the collection of data on race/ethnicity in national population health surveys that closely resemble the NJ program.[7,30] Our study demonstrates the significant effect these data collection changes could have if applied to data collected by hospitals as well. As healthcare institutions develop innovative methods of collecting high quality race/ethnicity data, we should continue to evaluate these strategies in order to identify "best practices" for monitoring healthcare disparities across all groups.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding: Dr. Chakkalakal was supported by the Robert Wood Johnson Foundation Clinical Scholars Program at Yale University (award ID 70033) at the time this work was completed. This research was also supported by the Center for Cardiovascular Outcomes Research at Yale University (U01 HL105270-04) which is funded by the National Heart, Lung, and Blood Institute.

REFERENCES

1. Institute of Medicine. Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care (full printed version). The National Academies Press; 2003.
2. Moscou S, Anderson MR, Kaplan JB, Valencia L. Validity of Racial/Ethnic Classifications in Medical Records Data: An Exploratory Study. American Journal of Public Health. 2003 Jul 1; 93(7):1084–1086. 2003. [PubMed: 12835189]
3. Blustein J. The reliability of racial classifications in hospital discharge abstract data. American Journal of Public Health. 1994 Jun; 84(6):1018–1021. [PubMed: 8203669]
4. Boehmer U, Kressin NR, Berlowitz DR, Christiansen CL, Kazis LE, Jones JA. Self-Reported vs Administrative Race/Ethnicity Data and Study Results. American Journal of Public Health. 2002 Sep 1; 92(9):1471–1472. 2002. [PubMed: 12197976]
5. Kressin NR, Chang BH, Hendricks A, Kazis LE. Agreement Between Administrative Data and Patients' Self-Reports of Race/Ethnicity. American Journal of Public Health. 2003 Oct; 93(10): 1734–1739. [PubMed: 14534230]
6. Hasnain-Wynia, RPD.; Pittman, M. Who, When, and How: The Current State of Race, Ethnicity, and Primary Language Data Collection in Hospitals. New York: Commonwealth Fund; 2004 May 18. 2004
7. Weissman JS, Hasnain-Wynia R. Advancing Health Care Equity through Improved Data Collection. New England Journal of Medicine. 2011; 364(24):2276–2277. [PubMed: 21675885]
8. Swallen KC, Glaser SL, Stewart SL, West DW, Jenkins CN, McPhee SJ. Accuracy of racial classification of Vietnamese patients in a population-based cancer registry. Ethn Dis. 1998; 8(2): 218–227. [PubMed: 9681287]
9. Pan CX, Glynn RJ, Mogun H. Definition of race and ethnicity in older people in Medicare and Medicaid. Journal of the American Geriatrics Society. 1999 Jun; 47(6):730–733. [PubMed: 10366176]
10. Andrews RM. Race and ethnicity reporting in statewide hospital data: progress and future challenges in a key resource for local and state monitoring of health disparities. J Public Health Manag Pract. 2011 Mar-Apr;17(2):167–173. [PubMed: 21297413]
11. Institute of Medicine. Race, Ethnicity, and Language Data: Standardization for Health Care Quality Improvement. Washington, DC: The National Academies Press; 2009.
12. Vali, F. Patient Race & Ethnicity: Improving Hospital Data Collection & Reporting. Princeton, NJ: 2004.
13. Agency for Healthcare Research and Quality. State Inpatient Databases (SID), Healthcare Cost and Utilization Project (HCUP), Agency for Healthcare Research and Quality, 2005–2009. 2009. www.hcup-us.ahrq.gov/sidoverview.jsp.
14. Karaca-Mandic P, Norton EC, Dowd B. Interaction Terms in Nonlinear Models. Health Services Research. 2011; 47(1pt1):255–274. [PubMed: 22091735]
15. Hoffman SD, Duncan GJ. Multinomial and conditional logit discrete-choice models in demography. Demography. 1988 Aug; 25(3):415–427. [PubMed: 3234576]
16. Oehlert G. A Note on the Delta Method. The American Statistician. 1992; 46(1):27–29.
17. Bland JM, Altman DG. Multiple significance tests: the Bonferroni method. BMJ. 1995 Jan 21.310(6973):170. [PubMed: 7833759]

18. U.S. Census Bureau. A Compass for Understanding and Using American Community Survey Data: What PUMS Data Users Need to Know. Washington, D.C: U.S. Government Printing Office; 2009.

19. Grieco, EM.; Cassidy, RC. Overview of Race and Hispanic Origin: Census 2000 Brief. Washington, D.C: US Census Bureau; 2000.

20. StataCorp. Version Release 12. Stata Statistical Software; 2011. [computer program].

21. Liebler, CA.; Rastogi, S.; Fernandez, LE.; Noon, JM.; Ennis, SR. America's Churning Races: Race and Ethnic Response Changes between Census 2000 and the 2010 Census. Washington, DC: Center for Administrative Records Research and Applications; 2014.

22. Sohn MW, Zhang H, Arnold N, et al. Transition to the new race/ethnicity data collection standards in the Department of Veterans Affairs. Popul Health Metr. 2006; 4:7. [PubMed: 16824220]

23. Billings, J.; Berry, C.; Kaplan, S.; Mayer, A. Evaluation of Expecting Success: Excellence in Cardiac Care. The New York University Evaluation Team; 2010.

24. Moy E, Arispe IE, Holmes JS, Andrews RM. Preparing the national healthcare disparities report: gaps in data for assessing racial, ethnic, and socioeconomic disparities in health care. Med Care. 2005 Mar; 43(3 Suppl):I9–I16. [PubMed: 15746596]

25. Ghosh C. Healthy People 2010 and Asian Americans/Pacific Islanders: Defining a Baseline of Information. American Journal of Public Health. 2003 Dec 1; 93(12):2093–2098. 2003. [PubMed: 14652340]

26. Ghosh C. A National Health Agenda for Asian Americans and Pacific Islanders. JAMA: The Journal of the American Medical Association. 2010 Sep 22/29; 304(12):1381–1382. 2010. [PubMed: 20858884]

27. Islam NS, Khan S, Kwon S, Jang D, Ro M, Trinh-Shevrin C. Methodological issues in the collection, analysis, and reporting of granular data in Asian American populations: historical challenges and potential solutions. J Health Care Poor Underserved. 2010 Nov; 21(4):1354–1381. [PubMed: 21099084]

28. Narayan KMV, Aviles-Santa L, Oza-Frank R, et al. Report of a National Heart, Lung, and Blood Institute Workshop: Heterogeneity in Cardiometabolic Risk in Asian Americans in the U.S.: Opportunities for Research. Journal of the American College of Cardiology. 2010; 55(10):966–973. [PubMed: 20202512]

29. Palaniappan LP, Araneta MRG, Assimes TL, et al. Call to Action: Cardiovascular Disease in Asian Americans. Circulation. 2010 Sep 21; 122(12):1242–1252. 2010. [PubMed: 20733105]

30. Dorsey R, Graham G. New HHS Data Standards for Race, Ethnicity, Sex, Primary Language, and Disability Status. JAMA: The Journal of the American Medical Association. 2011 Dec 7; 306(21): 2378–2379. 2011. [PubMed: 22147383]

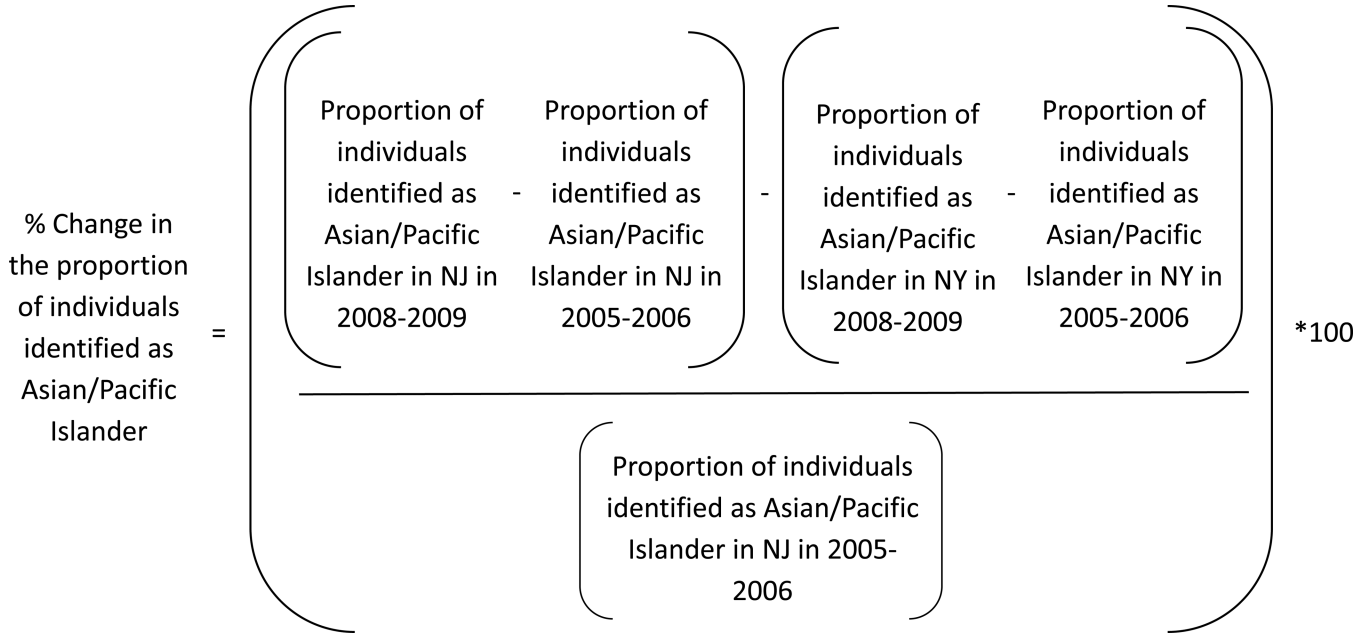

Figure 1.
Calculation of Percent Change in the Proportion of Patients Identified in Each Racial/Ethnic Category

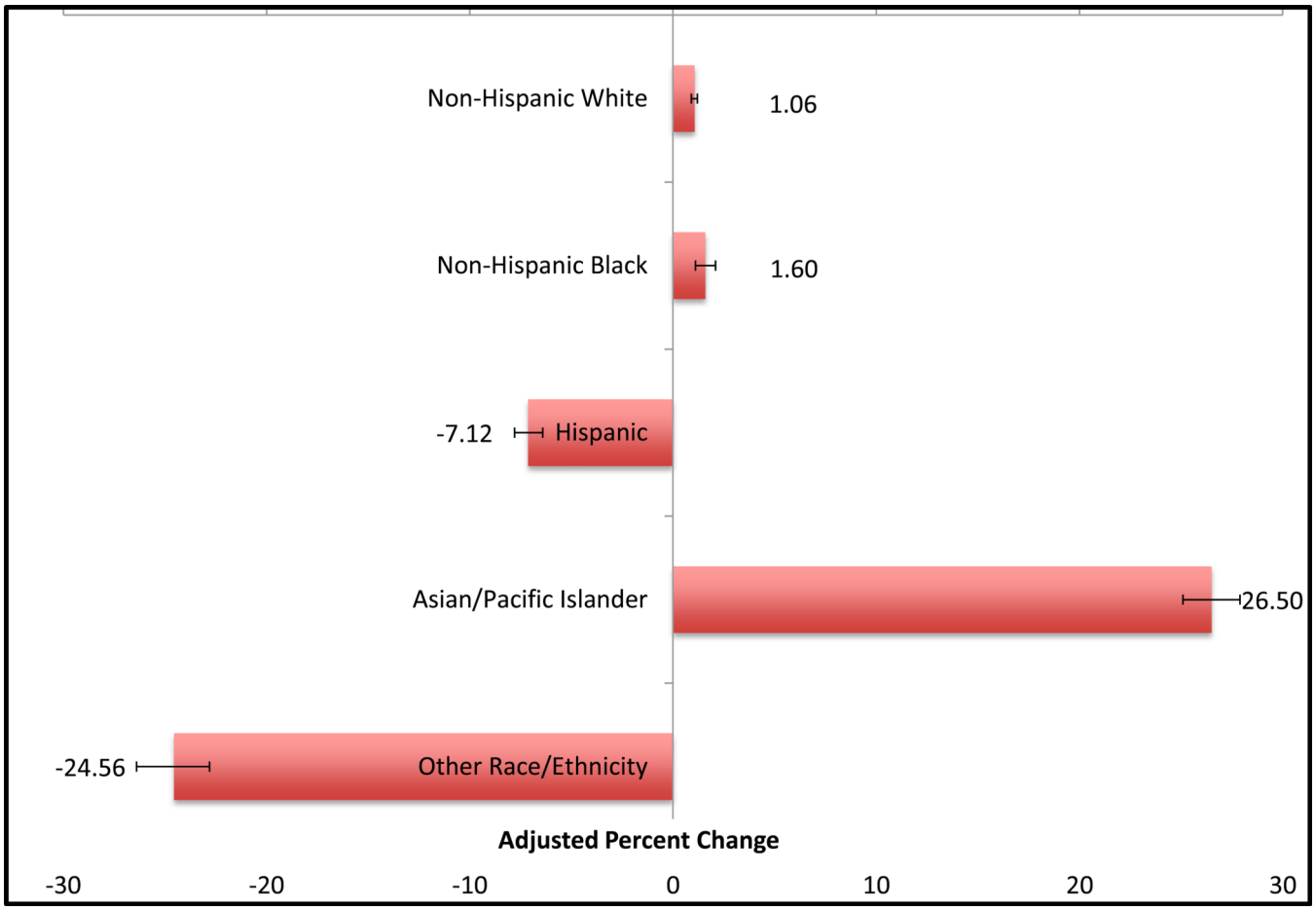

Figure 2.
Adjusted Percent Change for Each Racial/Ethnic Category

Table 1

Characteristics of 2005–2006 Hospital Discharges by State, Ages 18+

State	New Jersey	New York
Number of hospital discharges	1,833,110	4,313,490
Demographic Characteristics		
Non-Hispanic White	66.7%	61.0%
Non-Hispanic Black	16.3	18.4
Hispanic	11.6	11.9
Asian/Pacific Islander	2.3	2.5
Other Race/Ethnicity	3.1	6.2
Age 18–44	30.1	32.0
Age 45–64	26.5	28.0
Age 65 and above	43.4	40.0
Female	59.4	58.1
Male	40.6	41.9
Major Diagnostic Categories		
Circulatory	22.3	19.6
Pregnancy/Childbirth	12.7	12.1
Digestive	10.1	9.3
Respiratory	9.9	9.1
Musculoskeletal	7.3	7.8
Other diagnostic category	37.7	42.2

Note. Excludes hospital discharge abstracts missing data on gender, age, or race/ethnicity (less than 2% of the study sample).

Table 2

Adjusted Racial/Ethnic Distribution per 10,000 Hospitalizations in New Jersey and New York; State Inpatient Databases, Ages 18+ (N = 12,552,702)

	2005–2006	2008–2009	Difference	Difference-in-Differences	Percent change
Non-Hispanic White					
NJ	6,669 (6660, 6677)	6,520 (6512, 6529)	−148 (−159, −138)	+71 (59, 82)***	+1.1 (+0.9, +1.2)
NY	6,104 (6098, 6110)	5,885 (5879, 5891)	−219 (−226, −212)		
Non-Hispanic Black					
NJ	1,633 (1627, 1640)	1,663 (1656, 1669)	+29 (21, 38)	+26 (17, 35)***	+1.6 (+1.1, +2.1)
NY	1,838 (1833, 1842)	1,841 (1836, 1845)	+3 (−3, 9)		
Hispanic					
NJ	1,153 (1148, 1159)	1,114 (1109, 1120)	−39 (−46, −32)	−82 (−90, −74)***	−7.1 (−7.8, −6.4)
NY	1,185 (1182, 1189)	1,228 (1225, 1232)	+43 (38, 48)		
Asian/Pacific Islander					
NJ	233 (231, 236)	344 (340, 347)	+110 (107, 114)	+62 (58, 66)***	+26.5 (+25.1, +27.9)
NY	253 (251, 254)	301 (299, 303)	+49 (46, 51)		
Other Race/Ethnicity					
NJ	311 (308, 315)	359 (356, 363)	+48 (44, 52)	−76 (−81, −72)***	−24.6 (−26.4, −22.8)
NY	620 (617, 623)	744 (741, 747)	+124 (120, 128)		

Notes. Adjustments are for gender and age categories (18/44, 45/64, 65+); Adjusted proportions are predictive margins (proportions), differences are contrasts of margins, and differences-in-differences are incremental effects of the interaction term in the model. Percent changes are DDs divided by 2005–2006 proportions for NJ, and multiplied by 100. 95% confidence intervals in parentheses were computed using the delta-method. CIs for each race are corrected for multiple testing ($\alpha/4$ for the proportions and $\alpha/2$ for the single differences). P-values are from chi-square tests of contrasts;

*** $p < 0.001$.

Table 3

Adjusted Racial/Ethnic Distribution per 10,000 Individuals in New Jersey and New York; American Community Survey, Ages 18+ (N = 838,461)

	2005–2006	2008–2009	Difference	Difference-in-Differences	Percent change
Non-Hispanic White					
NJ	6,489 (6486, 6492)	6,343 (6339, 6346)	−146 (−150, −142)	−65 (−69, −60)***	−1.0 (−1.1, −0.9)
NY	6,276 (6274, 6278)	6,194 (6192, 6197)	−82 (−85, −79)		
Non-Hispanic Black					
NJ	1,217 (1215, 1220)	1,261 (1259, 1263)	+44 (41, 47)	+10 (7, 13)***	+0.8 (+0.6, +1.1)
NY	1,375 (1373, 1376)	1,408 (1407, 1410)	+33 (31, 35)		
Hispanic					
NJ	1,438 (1435, 1440)	1,516 (1514, 1519)	+79 (76, 82)	+44 (40, 47)***	+3.0 (+2.8, +3.3)
NY	1,500 (1498, 1501)	1,535 (1533, 1536)	+35 (33, 37)		
Asian/Pacific Islander					
NJ	739 (737, 741)	754 (752, 755)	+15 (13, 17)	+3 (1, 6)**	+0.4 (+0.1, +0.8)
NY	692 (691, 693)	704 (702, 705)	+12 (10, 13)		
Other Race/Ethnicity					
NJ	117 (117, 118)	127 (126, 127)	+9 (8, 10)	+8 (7, 9)***	+6.5 (+5.6, +7.3)
NY	158 (157, 158)	159 (159, 160)	+2 (1, 2)		

Notes. Rates are estimated using the American Community Survey frequency weights. Adjustments are for gender and age categories (18/44, 45/64, 65+); Adjusted proportions are predictive margins (proportions), differences are contrasts of margins, and differences-in-differences are incremental effects of the interaction term in the model. Percent changes are DDs divided by 2005–2006 proportions for NJ, and multiplied by 100. 95% confidence intervals in parentheses were computed using the delta-method. CIs for each race are corrected for multiple testing ($\alpha/4$ for the proportions and $\alpha/2$ for the single differences). P-values are from chi-square tests of contrasts;

** $p < 0.01$,

*** $p < 0.001$.

Appendix Table 1

Race/Ethnicity Categories by Data Source

New Jersey SID		HCUP 2005–2009	New York SID 2005–2009
2005–2006	**2008–2009**		
Mexican	Mexican, Mexican American, Chicano	Hispanic	Spanish/Hispanic origin
Puerto Rican	Puerto Rican		
Cuban	Cuban		
Central or South American	Central or South American		
Other and Unknown Hispanic	Other Spanish/Hispanic/Latino		
White	White	White	White
Black	Black or African American	Black	African American (Black)
Asian Indian	Asian Indian	Asian or Pacific Islander	Asian
Chinese	Chinese		
Filipino	Filipino		
Japanese	Japanese		
Korean	Korean		
Vietnamese	Vietnamese		
Native Hawaiian	Native Hawaiian		Native Hawaiian or Other Pacific Islander
Guamian or Chamorro	Guamian or Chamorro		
Samoan	Samoan		
Other Asian or Pacific Islander	Other Asian		
	Other Pacific Islander		
American Indian or Alaska Native	American Indian or Alaskan Native	Native American	Native American (American Indian, Eskimo, Aleut)
Multiracial: White and Black or African American	Multiracial: White and Black or African American	Other	Other
Multiracial: White and American Indian or Alaskan Native	Multiracial: White and American Indian or Alaskan Native		
Multiracial: White and Asian	Multiracial: White and Asian		
Multiracial: Black or African American and American Indian or Alaskan Native	Multiracial: Black or African American and American Indian or Alaskan Native		
Other races	Other Race		
Unknown, Missing Declined to Answer Blank	Unknown/Unavailable Declined to answer Blank	Missing	Missing Blank