HCFA's Racial and Ethnic Data: Current Accuracy and Recent Improvements

Susan L. Arday, M.H.S, C.H.E.S., David R. Arday, M.D., M.P.H., Stephanie Monroe, Ph.D., and Jianyi Zhang, M.D., M.H.S., Ph.D.

Using linked data from the Medicare Current Beneficiary Survey (MCBS), the authors assessed the accuracy of racial/ethnic classifications in HCFA's enrollment data base (EDB) before and after the 1997 effort to update the EDB. After the update, the sensitivity of the EDB was 97 percent for white persons and 95 percent for black persons, but less than 60 percent for all other categories. The positive predictive value was above 96 percent for white, black, and Hispanic persons, but below 80 percent for all others. There was some improvement in accuracy for white persons and black persons from 1991-1997, and larger improvements for the non-black minorities from 1996-1997.

INTRODUCTION

Although the collection and use of race and ethnicity data in public health research has recently become more controversial (Fullilove, 1998; Kaufman, 1999), reliable information on race and ethnicity are still useful to health researchers, physicians, hospitals, and health administrators who strive to provide quality health care to increasingly diverse populations (Watson, 1997; Centers for Disease Control and Prevention, 1993; U.S. Department of Health and Human Services, 1999). In fact,

a number of recent studies using Medicare data have unfortunately demonstrated striking differences in access to care and treatment received between persons of different races or ethnicities (Gornick et al., 1996; McBean and Gornick, 1994; Siddique et al., 1998; Mark and Paramore, 1996).

Eggers and Greenberg (2000) provide a first look at recent racial and ethnic differences in hospitalizations and certain hospital-associated types of care among aged Medicare beneficiaries, using the recently expanded racial and ethnic codes in HCFA's administrative claims data. And by comparing 1998 U.S. Census Bureau population estimates with HCFA's EDB counts, they also document the continued large under-representation of minorities--other than black--in Medicare enrollment data.

RACIAL/ETHNIC CODING IN HCFA ADMINISTRATIVE DATA

The primary resource for identifying Medicare beneficiary race/ethnicity in HCFA's administrative claims data is the EDB. This information is released annually in the public-use Medicare denominator file, and the racial variable is also attached to or included in certain other files.[1] Therefore, we will only present a brief summary and add updating information here.

It should be noted that HCFA EDB records are not created by the agency itself, but are populated by information received from the Social Security

Susan L. Arday and David R. Arday are with the Office of Clinical Standards and Quality, Health Care Financing Administration (HCFA). Stephanie Monroe is with the Office of Strategic Planning, HCFA. Jianyi Zhang is with the Department of Epidemiology and Preventive Medicine, School of Medicine, University of Maryland. The views expressed in this article are those of the authors and do not necessarily reflect those of HCFA or the University of Maryland.

[1] Lauderdale and Goldberg, (1996) present an excellent history of race coding using HCFA data.

Administration (SSA) or the Railroad Retirement Board (RRB). The EDB receives the vast majority of its beneficiary demographic information, including age, sex, and race, from SSA's master beneficiary record (MBR) file. The race variable in the MBR includes only four codes–White, Black, Other, and Unknown–while the RRB files contain no racial information at all.

The original source of the MBR's race variable is the SSA's SS-5 Form, which is completed either when an individual initially applies for a Social Security number or for a replacement Social Security card. Racial and ethnic information on the SS-5 Form are collected on a voluntary basis. Prior to November 1980, the SS-5 form contained only three racial categories: White, Black, and Other, plus Unknown. Thereafter, in response to Office of Management and Budget (OMB) Statistical Directive 15 (*Federal Register*, 1978) SSA expanded the SS-5 to include the categories of: White; Black; Hispanic; Asian, Asian-American, or Pacific Islander; and North American Indian or Alaskan Native. The category of Other was eliminated, while Unknown was retained for persons not responding to the race/ethnicity question. The MBR record file was not expanded, but instead, the SS-5 data were manipulated to populate the four categories that were retained in SSA's MBR file.

According to internal HCFA records, the expanded race/ethnicity categories were fully integrated into the SSA's electronic SS-5 file–otherwise known as the numerical identification file (NUMIDENT)—by 1982. As new enrollees matriculated into Medicare, however, HCFA continued to populate its EDB only with the more limited racial data from SSA's MBR file. There was little reason for HCFA to immediately expand their existing race/ethnicity data files in the early 1980s, as very few beneficiaries would have had any new informa-tion available to either change what was already in the EDB or to populate an expanded record upon matriculation. Even today, the great majority of the elderly Medicare population provide their racial information using the earlier version of the SS-5, and relatively few have had reason to update their information over the years using the newer form with more categories. Furthermore, beneficiary spouses who never applied for and received a Social Security number of their own have their EDB racial variable information populated by the racial information of their wage earner spouse, which may or may not accurately reflect their own race/ethnicity. As many as 18 percent of Medicare beneficiaries (and thus their EDB records) fall into this latter category.

The EDB is updated daily with information on all newly enrolling or dying Medicare beneficiaries. In 1994, HCFA expanded the racial codes for the single race variable in the EDB from three categories (White, Black, and Other—plus Unknown) to six (White, Black, Hispanic, Asian American or Pacific Islander, North American Indian, and Other—plus Unknown). Over the next several months, HCFA then populated the expanded racial categories for every Medicare beneficiary in the EDB known to be alive, using original source data from SSA's NUMIDENT file. After this single effort, newly entitled Medicare beneficiaries continued to have their EDB demographic information populated from the MBR file, using only the three racial categories that file contains (plus Unknown).

In May 1997, HCFA again undertook a NUMIDENT update, targeting all beneficiaries added to the EDB since the 1994 update. Also in May 1997, HCFA updated EDB race information on 858,000 people who responded to a direct-mail survey. The survey had been sent to nearly 2.2 million

beneficiaries who had an EDB race of unknown or other, who had a Hispanic country of birth (as defined by SSA), or who had a surname identified as Hispanic (using U.S. Census Bureau name information). Newly matriculating beneficiaries continued to be identified in the EDB only as white, black, other, or unknown following the second update.

NEW INITIATIVES AND DATA QUALITY IMPLICATIONS

In February 1998, President Clinton committed the Nation to eliminate the disparities in six areas of health status experienced by racial and ethnic minority populations: (1) infant mortality, (2) cancer screening and management, (3) cardiovascular disease, (4) diabetes, (5) human immunodeficiency virus infection/ acquired immunodeficiency syndrome, and (6) immunizations (U.S. Department of Health and Human Services, 1999). As a first step in improving baseline data about the effectiveness of Department of Health and Human Services (HHS) programs in reaching minority populations, the department has adopted a policy that will require all HHS-sponsored data collection and reporting systems to include standard racial and ethnic categories (U.S. Department of Health and Human Services, 1999). These categories must be compliant with OMB Statistical Directive 15, which after much debate during the mid 1990s (Evinger, 1996), underwent three relatively minor revisions in October 1997. These revisions were: (1) the separation of the Asian or Pacific Islander racial category into two separate categories, Asian and Native Hawaiian or Other Pacific Islander, (2) the changing of the ethnic term Hispanic to Hispanic or Latino, and (3) the changing of the instructions to allow the marking of one or more racial categories by the respondent (*Federal Register*, 1997).

Because the existing racial variable in HCFA's administrative claims data may be used to segment the data or to identify certain subpopulations for intervention targeting in support of the Federal initiative to eliminate disparities, it is important to know how useful the data are for those purposes. Misclassifications of race/ethnicity in source data may lead to erroneous conclusions, if those misclassifications are overlooked or ignored (Kelley et al., 1996; Lauderdale and Goldberg, 1996). We undertook this study both to assess the improvement in the EDB race/ethnicity classification data from 1991-1997, as a result of HCFA's efforts, and to quantify the accuracy of these data.

METHODS

The MCBS is a continuous, multipurpose survey of a nationally representative sample of the Medicare population, the design of which has previously been described in detail (Adler, 1994). Respondent information on race/ethnicity is collected during an initial face-to-face interview, either with the respondent, or in the case of a Medicare beneficiary who is unable to answer, a designated proxy. Once these data are collected and processed, they are linked with respondent Medicare claims data and other administrative data, including the EDB data.

In accordance with the OMB Statistical Directive 15 (*Federal Register*, 1978), the MCBS collects race and ethnicity information using two questions as follows. The first question assigns race using the following categories: American Indian, Asian or Pacific Islander, Black/African American, White, or Other (specify). The

Table 1

Definitions of Measures Used to Assess Accuracy and Improvement in Agreement for HCFA's Enrollment Data Base (EDB) Race/Ethnicity Classifications

Sensitivity	The chance that the EDB will correctly identify a person of the given race/ethnicity. The sensitivity is given by the formula: (true positives)/(true positives+false negatives).
Specificity	The chance that the EDB will correctly exclude a person not of the given race/ethnicity. The specificity is given by the formula: (true negatives)/(true negatives+false positives).
Positive Predictive Value (PPV)	The chance that a person identified as a given race/ethnicity in EDB will actually be that race/ethnicity. The PPV is given by: (true positives)/(true positives + false positives).
Negative Predictive Value (NPV)	The chance that a person identified as not having a given race/ethnicity in the EDB will actually not be of that race/ethnicity. The NPV is given by: (true negatives)/(true negatives+false negatives).
Kappa	A statistical measure of overall agreement which corrects for random agreement. Kappa values greater than 0.75 are considered excellent agreement beyond what could be expected by chance alone. The general formula for kappa is: $(I_o - I_e)/(1 - I_e)$, where I_o is the observed value of the index and I_e is the expected value of agreement. Kappa will go to +1 as sensitivity and specificity together go to 100 percent.

NOTE: HCFA is Health Care Financing Administration.

SOURCES: (Fletcher, Fletcher, and Wagner, 1988.) A Dictionary of Epidemiology, 3rd Edition. Oxford University Press, May 1995.

second question asks if the respondent "is of Hispanic origin?" Additional values are assigned to either variable if the respective question is refused or the respondent answers "don't know."

We used data collected in rounds 1, 16, and 19 of the MCBS. Data collection for these rounds took place during the months of September-December: in 1991 for round 1, 1996 for round 16, and 1997 for round 19. We excluded data obtained from proxy respondents (roughly 11 percent of community interviews), so all MCBS racial/ethnic data used was self-provided. We also excluded a small number of MCBS respondents who refused, answered "don't know," or whose racial category was otherwise unascertained (e.g., a total of 16 respondents in 1997), leaving 10,429 self-respondents from round 1, 14,700 from round 16, and 15,168 from round 19 who provided answers to the question on race.

As previously noted, at the time of MCBS round 1, the linked EDB racial variable only contained the classifications White, Black, or Other for each of the respondents. For rounds 16 and 19, the EDB data included the previously mentioned three categories plus Hispanic,

Asian/Pacific Islander, and American Indian. Linkage and preparation of the final analytic files took place in March of the year following the MCBS collection year (e.g., March 1998 for round 19). Thus, the 1996 and 1997 linkages occurred immediately before and about 9 months after the 1997 EDB update.

For the latter two MCBS rounds, individual racial/ethnic classifications in the EDB race variable were compared with those in the MCBS collected race and ethnicity variables. The sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of the EDB were calculated for each category (Fletcher, Fletcher, and Wagner, 1988), taking the self-reported values in the MCBS as the reference. (Table 1 provides definitions of these terms.)

Kappa statistics (Shoukri and Edge, 1996) were calculated as a statistical measure of the agreement between the EDB and MCBS for each classification, for each year. This was done to better measure change in agreement over time. Although traditionally used to evaluate interrater agreement or reliability, the designation of one measure as the reference standard

Accuracy of HCFA's Enrollment Data Base (EDB) Racial/Ethnic Classifications Using the Medicare Current Beneficiary Survey (MCBS) for Comparison: 1996

Racial/Ethnic Classification	MCBS			Accuracy and Agreement Measures for EDB				
	EDB	Yes	No	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value	Kappa
				Percent				
White	Yes	12,275	265					
	No[1]	425	1,735	96.6	86.8	97.9	80.3	0.81
Black	Yes	1,426	70					
	No[1]	67	13,137	95.5	99.5	95.3	99.5	0.95
Hispanic	Yes	199	4					
	No[1]	824	13,673	19.4	100.0	98.0	94.3	0.31
Asian/Pacific Islander	Yes	31	16					
	No[1]	125	14,528	19.9	99.9	66.0	99.2	0.30
American Indian	Yes	1	6					
	No[1]	179	14,514	0.6	100.0	14.3	98.8	0.19
Other	Yes	25	317					
	No[1]	188	14,170	11.7	97.8	7.3	98.7	0.07

[1] No represents any person not positively identified in the respective category in EDB, including unknowns.

NOTES: HCFA is Health Care Financing Administration. $n = 14,700$.

SOURCE: Author's calculations based on HCFA's EDB and MCBS, 1996.

does not violate any of the required assumptions for the use of kappa (Shoukri and Edge, 1996). We also did stratified analyses by sex, age, and education level on the 1996 data. Finally, we used the weighted MCBS samples to estimate the changes in the distributions of the EDB racial and ethnic classifications among community-dwelling Medicare beneficiaries from 1991-1997.

RESULTS

Table 2 presents the unweighted round 16 overall comparisons, including the sensitivity, specificity, PPV, NPV, and kappa statistics for each of the five specific racial/ethnic categories in HCFA's EDB. These comparisons show accuracy just prior to the second NUMIDENT update. As of 1996, the EDB's sensitivity was roughly 95 percent or better in its identification of both white and black Medicare beneficiaries, but was less than 20 percent for all other classifications. Specificity was over 99 percent for most non-white beneficiaries, but was only 87 percent for white beneficiaries. The PPV was over 95 percent

for white, black, and Hispanic beneficiaries, but considerable lower for all remaining classifications. The NPV was above 94 percent for all minority classifications, but lower for white beneficiaries. Furthermore, the kappa statistics indicated agreement was only slight for minorities other than black beneficiaries.

Table 3 presents the same comparisons for round 19. After the second NUMIDENT update, there were significant improvements in several categories. For the Hispanic classification in the EDB, the sensitivity doubled, from 19 percent to 39 percent, and kappa increased from 0.31 (95 percent confidence limits [CL] 0.28, 0.34) to 0.54 (0.51, 0.58). For the Asian category, sensitivity nearly tripled, from 20 percent to 58 percent, PPV rose from 66 percent to 79 percent, and kappa increased from 0.30 (95 percent CL 0.22, 0.38) to 0.66 (0.59, 0.73). For the American Indian category, sensitivity jumped from 0.6 percent to 11 percent, PPV rose from 14 percent to 78 percent, and kappa increased from 0.01 (95 percent CL -0.01, 0.03) to 0.19 (0.10, 0.27). Specificity for the White category also improved slightly, from 87 percent to 89

Table 3

Accuracy of HCFA's Enrollment Data Base (EDB) Racial/Ethnic Classifications Using the Medicare Current Beneficiary Survey (MCBS) for Comparison: 1997

Racial/Ethnic Classification	MCBS EDB	Yes	No	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value	Kappa
						Percent		
White	Yes	12,798	209					
	No [1]	423	1,738	96.8	89.3	98.4	80.4	0.82
Black	Yes	1,435	59					
	No [1]	79	13,595	94.8	99.6	96.1	99.4	0.95
Hispanic [2]	Yes	343	8					
	No [1]	532	14,271	39.2	99.9	97.7	96.4	0.54
Asian/Pacific Islander	Yes	80	21					
	No [1]	59	15,008	57.6	99.9	79.2	99.6	0.66
American Indian	Yes	14	4					
	No [1]	115	15,035	10.9	100.0	77.8	99.2	0.19
Other	Yes	14	131					
	No [1]	151	14,872	8.5	99.1	9.7	98.9	0.08

[1] No represents any person not positively identified in the respective category in the EDB, including unknowns.

[2] Only 15,154 MCBS respondents answered the question about Hispanic origin with a yes or a no.

NOTES: HCFA is Health Care Financing Administration. n =15,168.

SOURCE: Author's calculations based on HCFA's EDB and MCBS, 1997.

percent. The remaining non-kappa values changed by 2 percentage points or less, while the kappa for the Other classification increased by only 0.01.

While many of the 1997 improvements were dramatic, the EDB's sensitivity remained less than 60 percent for all classifications other than Black and White. Similarly, the PPV was less than 80 percent for the Asian/Pacific Islander, American Indian, and Other categories. The pattern of lower specificities and NPVs for the White category also remained. While improved, the kappa statistics indicated the agreement was still little better than random chance for both the American Indian and Other categories, and only fair for the Hispanic and Asian/Pacific Islander categories.

Table 4 shows the estimated distributions of the EDB racial/ethnic classifications for 1991, 1996, and 1997, derived from the weighted samples of community dwelling MCBS respondents. From 1991-1997, there was a relatively large decrease in the percentage of unknowns (from 3.0 percent to 0.3 percent) and others (from

Table 4

Estimated Distribution of Racial/Ethnic Classifications in the Enrollment Data Base (EDB) Community-Dwelling Medicare Beneficiaries: 1991, 1996, and 1997[1]

Racial/Ethnic Classification	1991[2]	1996	1997
		Percent	
White	86.4	86.9	87.0
Black	8.2	8.6	8.4
Hispanic	—	1.2	2.3
Asian/Pacific Islander	—	0.2	0.8
American Indian/Alaska Native	—	0	0.1
Other	2.4	2.0	1.1
Unknown	3.0	1.1	0.3

[1] Estimates derived from weighted samples of community-dwelling Medicare Current Beneficiary Survey (MCBS) respondents.

[2] In 1991, the EDB contained only the four categories for which numbers are shown.

SOURCE: Author's calculations based on HCFA's EDB and MCBS, 1997.

2.4 percent to 1.1 percent), while the percentages for the White and Black categories changed very little.

From 1991-1996, however, there was measurable improvement in the agreement between the EDB and MCBS for the Black and White categories. The unweighted kappa statistic for the White category increased from 0.76 (95 percent CL 0.74, 0.78) to 0.81 (0.79, 0.82), and for the Black category it increased from 0.93 (95 percent CL 0.92, 0.94) to 0.95 (0.94, 0.96).

Stratified analyses of the 1996 data revealed that for Hispanics and Asian/Pacific Islanders, those under age 65 had somewhat better agreement between their race/ethnicity in the EDB and that in the MCBS (kappa values greater than 0.50) than did those age 65 or over (kappa values less than 0.40). For the Black and White categories, however, there was no age-related difference. Nor was there any variation in the accuracy of the EDB racial classifications by sex or educational level.

DISCUSSION

By using administrative data linked to the MCBS respondent data, we were able to quantify the accuracy of HCFA's EDB in terms of sensitivity, specificity, etc. We found that the additional categories added to the EDB in 1994, plus the 1994 and 1997 efforts to update the EDB with NUMI-DENT and other data, reduced the percentages of beneficiaries classified as other or unknown in the EDB, and increased its overall accuracy. Despite these improvements, our findings do not yet allow us to refute those of Lauderdale and Goldberg (1996), who concluded that HCFA's EDB race/ethnicity data remained incomplete and biased after the first update in 1994.

As of 1997, it appears that many minority beneficiaries remain misclassified in the EDB, as evidenced by the lower sensitivities for categories other than Black or White. Interpretation of the sensitivity and specificity patterns suggests that most are misclassified as either white or other. In addition, the SSA SS-5 Form's use of only a single variable to classify race and ethnicity means that a large number of white Hispanics in the EDB are identified only as Hispanic, without any additional racial categorization. This latter fact is highlighted by the EDB's lower specificity for the White category. But because of the low sensitivity of the Hispanic category, it would be incorrect to assume that the EDB classification of White really means "White, non-Hispanic." On the other hand, there appear to be few non-Hispanic or non-black persons misclassified as Hispanic or black respectively, as evidenced by the high PPV values for these two categories. The high specificities and NPVs for all the minority categories indicate that if the objective is to exclude members of a particular minority from a sample, the EDB data may be used effectively.

In a recent comparison of Medicare and Medicaid administrative data in a single State (Pan et al., 1999), only modest overall agreement between the race/ethnicity data from the two sources was found. Using either data source as the reference, the authors found a similar gradient to ours, in that agreement was better for the categories White and Black, but poorer for other categories. Although their approach did not allow them to determine the accuracy of the sources, they found that the race and ethnicity depictions in the two sources were often contradictory, and concluded that overall accuracy was approximate at best.

For our analyses, we treated the EDB and MCBS as two independent sources of data, even though they largely originate with the same individual self-identifying his or her race and ethnicity. (Although there have been anecdotal reports of proxies completing the racial data on the SS-5 form, this is believed to occur very infrequently.) We considered the MCBS to be both independent of the EDB and to be the reference standard because the information was generally obtained more recently, was largely derived in face-to-face interviews, and because the MCBS used separate questions about race and Hispanic ethnicity, which the SS-5 data source for the EDB did not.

Because the MCBS does not significantly oversample minority beneficiaries, there were relatively few respondents who identified themselves as either Asian/Pacific Islander or American Indian in 1996 and 1997. Still, there were more than 100 MCBS respondents who self-identified in each of these categories in each year, which is sufficient to allow reasonable comparisons. It should be recognized, however, that our estimates for these categories may be somewhat unstable, due to the relatively fewer numbers of individuals identified in these categories in the EDB. Therefore, the large improvements noted between 1996 and 1997 may be partly due to sampling variations introduced by the use of the MCBS. The size and consistency of these improvements, though, argue that the 1997 EDB update effort had a positive effect. Although this sample size limitation would be especially true for our stratified analyses, the patterns we observed in these analyses were consistent with those previously reported (Lauderdale and Goldberg, 1996), which also supports our findings.

Our 1997 comparison occurred shortly after the most recent update to the EDB, and therefore our results represent something of a best-case scenario over the short run. As new enrollees continue to be added to the EDB, using only the limited data in SSA's MBR, the accuracy of the EDB will decay somewhat, and the number of others and unknowns increase. This will continue until another update from the full NUMIDENT file is performed, something which is next scheduled to take place late in calendar year 2000. In the short run, HCFA may wish to consider increasing the frequency of NUMIDENT updates to annually, so as to minimize decay of race/ethnicity data in the EDB and ensure at least some improvement each year. If HCFA wants to improve the quality of its administrative race/ethnicity data more rapidly, however, HCFA may ultimately need to consider additional sources of race/ethnicity data, such as in-house or external survey data, or perhaps U.S. Census Bureau data, which could be electronically linked and used to update the EDB data periodically.

For many analytic approaches, the sensitivity and PPV of the EDB race/ethnicity classifications are the most important values in determining the utility of these data. And we have shown that the EDB remains relatively insensitive in its identification of Hispanics, Asian/Pacific Islanders, and American Indians. As noted by Eggers and Greenberg (2000), our work does not show that using the expanded racial/ethnic categories will necessarily result in biased utilization estimates for hospital services. We can only say that attempts to stratify HCFA claims data to study racial and ethnic differences (other than between Black and White categories) may give biased results due to the observed misclassifications.

Unfortunately, it is hard to know a priori exactly what the introduced bias may be, because it will differ depending upon the question at hand, the study design, and the outcome measure's relationship to race/ethnicity. This problem can be completely overcome only with an independent source of race/ethnicity data to allow appropriate reclassification of the subjects of the study, something that is usually not feasible. At the present time, we feel it would be most ideal to undertake an external validation before relying on estimates for non-black minority populations derived from HCFA administrative data alone. Otherwise, one might use the accuracy measures presented here to estimate the selection probabilities for the various racial/ethnic categories included in a study based on HCFA's administrative claims data, and then further estimate the resulting direction of any bias using the methods

presented by Kleinbaum, Morgenstern, and Kupper (1981). Knowing at least the direction of the introduced bias may greatly reduce the chances of drawing an erroneous conclusion. For some analytic approaches, it may be possible to actually correct for the misclassification using further methods (Greenland and Kleinbaum, 1983). Finally, one might consider calculating synthetic estimates for comparison purposes, by applying models of racial/ethnic differences derived from other sources to HCFA administrative claims data.

One may certainly argue that race is a concept fraught with problems of meaning that go beyond those of measurement that we have documented here. If one accepts that racial classifications may be informative, it remains that one may reasonably use HCFA's administrative race/ethnicity data when comparing White with Black categories. We believe, however, that one cannot yet utilize all the other categories with equal confidence, despite the improvements we have documented. This does not mean one should automatically ignore the expanded EDB racial/ethnic categories. But pending further improvements, we would continue to recommend appropriate caution when using the EDB alone to stratify HCFA claims data for analyses, or to identify and select certain Medicare minority populations for intervention targeting.

ACKNOWLEDGMENTS

REFERENCES

Adler, G.S.: A Profile of the Medicare Current Beneficiary Survey. *Health Care Financing Review* 15(4):153-163, Summer 1994.

Centers for Disease Control and Prevention: Use of Race and Ethnicity in Public Health Surveillance. Summary of the CDC/ATSDR Workshop. *Morbidity and Mortality Weekly Report 42* (Recommendations and Reports No.10), June 25, 1993.

Eggers, P.W., and Greenberg, L.G.: Racial and Ethnic Differences in Hospitalization Rates Among Aged Medicare Beneficiaries, 1998. *Health Care Financing Review 21* (4):81-95, Summer 2000.

Evinger, S.: How to Record Race. *American Demographics* 18(5):36-42, May 1996.

Federal Register: Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity. *Federal Register* Vol. 62, No. (210):58782-58790. National Archives and Records Administration. Washington, DC. U.S. Government Printing Office, October 30, 1997.

Fletcher, R.H., Fletcher, S.W., and Wagner, E.H.: *Clinical Epidemiology, 2nd Edition*. Baltimore. Williams and Wilkens, 1988.

Fullilove, M.T.: Comment: Abandoning "Race" as a Variable in Public Health Research–An Idea Whose Time Has Come. *American Journal of Public Health* 8 (9):1297-1298, September 1998.

Gornick, M.E., Eggers, P.W., Reilly, T.W., et al.: Effects of Race and Income on Mortality and Use of Services Among Medicare Beneficiaries. *New England Journal of Medicine* 335(11):791-799, 1996.

Greenland, S., and Kleinbaum, D.G.: Correcting for Misclassification in Two-Way Tables and Matched-Pair Studies. *International Journal of Epidemiology* 12 (1):93-97, March 1983.

Kaufman, J.S.: How Inconsistencies in Racial Classification Demystify the Race Construct in Public Health Statistics. *Epidemiology* 10(2):101-103, March 1999.

Kelley, J.J., Chu, S.Y., Diax, T., et al.: Race/Ethnicity Misclassification of Persons Reported With AIDS. *Ethnicity and Health* 1(1):87-94, 1996.

Kleinbaum, D.G., Morgenstern, H., and Kupper, L.L.: Selection Bias in Epidemiologic Studies. *American Journal of Epidemiology* 113(4):452-463, April 1981.

Lauderdale, D.S., and Goldberg, J.: The Expanded Racial and Ethnic Codes in the Medicare Data Files: Their Completeness of Coverage and Accuracy. *American Journal of Public Health* 86(5):712-716, May 1996.

Mark, T.L., and Paramore, L.C.: Pneumococcal Pneumonia and Influenza Vaccination: Access to and Use by U.S. Hispanic Medicare Beneficiaries. *American Journal of Public Health* 86(11):1545-1550, November 1996.

McBean, M., and Gornick, M.E.: Differences by Race in the Rates of Procedures Performed in Hospitals for Medicare Beneficiaries. *Health Care Financing Review* 15(4)77-90, Summer 1994.

Office of Management and Budget: Statistical Policy Directive No. 15, Race and Ethnic Standards for Federal Statistics and Administrative Reporting. In: *Statistical Policy Handbook*. Washington, DC. Office of Federal Statistical Policy and Standards, U.S. Department of Commerce, 1978.

Pan, C.X., Glynn, R.J., Mogun, H., Choodnovskiy, I., and Avorn, J.: Definition of Race and Ethnicity in Older People in Medicare and Medicaid. *Journal of the American Geriatric Society* 47(6):730-733, 1999.

Shoukri, M.M., and Edge, V.L.: *Statistical Methods for Health Sciences*. Boca Raton, FL. CRC Press, 1996.

Siddique, R.M., Amini, S.B., Connors, A.F., and Rimm, A.A.: Race and Sex Differences in Long-Term Survival Rates for Elderly Patients with Pulmonary Embolism. *American Journal of Public Health* 88(10):1476-1480, October 1998.

U.S. Department of Health and Human Services: The Initiative to Eliminate Racial and Ethnic Disparities in Health. Available from URL:http://raceandhealth.hhs.gov, February 21, 1999. (Cited August 18, 1999.)

Watson, S.D.: Race, Ethnicity and Hospital Care: The Need for Racial and Ethnic Data. *Journal of Health and Hospital Law* 30(2):125-132, 1997.