

ON THE CODING AND REPORTING OF RACE AND ETHNICITY IN NEW HAMPSHIRE FOR PURPOSES OF CANCER REPORTING

Background: The objective was to investigate how data on race and ethnicity are collected by hospitals reporting to the New Hampshire State Cancer Registry (NHSCR).

Method: NHSCR surveyed hospitals asking how information on race and ethnicity were collected. A review of relevant legal mandates and national guidelines was undertaken.

Results: Many hospitals lack policies on collection, computer systems fail to support national guidelines, and staff rely on visual inspection.

Conclusions: Hospital staffs are not now culturally equipped to collect race and ethnicity in a meaningful way. The numerator in cancer incidence rates is most likely not accurate and for some smaller populations very biased. A new framework is needed that takes into account the needs of the democracy. (*Ethn Dis.* 2005;115:324–331)

Key Words: Ethnicity, Race, Registry

Bruce L. Riddle, PhD

INTRODUCTION

In the annual calculations of population incidence rates on cancer, two numbers are required. The denominator, the base population, comes from the US Census Bureau and the 2000 Census.¹ The numerator for a specific geographic region, usually a state, is drawn from reports submitted to a central cancer registry mostly by hospitals. When these reports are broken down by race or ethnicity,² the denominator, using 2000 Census data, was collected by using a process of self-identification under guidance of the Office of Management and Budget (OMB) Statistical Directive 15.³ The numerator is most often collected under a different set of guidelines and circumstances.^{4,5}

This paper examines the disconnect between the national standards on collecting race and ethnicity and what occurs in New Hampshire at the hospital level, at the admitting desk. A hospital cancer registrar in preparing information to submit to a central registry can only abstract what is in the medical record. This paper looks at how race and ethnicity enter the medical record, the mandates for collection, and what the implications of the data collection methods might be. Hospital staffs are not culturally equipped to collect race and ethnicity in any meaningful way. Hospitals are drawn to the minimum federal standards promulgated by the Office of Management and Budget⁶ while hospital cancer registrars are asked to use a more comprehensive set of standards put forth by the American College of Surgeons, Commission on Cancer,⁷ and the North American Association of Central Cancer Registries (NAACCR).⁸ What is required is a new framework if useful information is to be collected and

standards of privacy and confidentiality are to be met.

MANDATES

Most central cancer registries are not in a position to require medical providers to collect information on race and ethnicity. Rather, the legislative language requires identified providers to report a selected set of information to the central registry. The means or methods for collecting the information are left to the discretion of the medical provider. Therefore, a starting question becomes, "What is the legal and institutional mandate to collect information on race and ethnicity?" Put another way, "What mandates inclusion of information on race and ethnicity being recorded in the medical record?"

For example, the law establishing the national program of cancer registries⁹ sets up a funding stream "... to support the operation of population-based, statewide cancer registries in order to collect. ... (1) demographic information about each case of cancer." No explicit mention is made of race and ethnicity in the legislation. Perot and Youdelman¹⁰ provide a major review of the legal basis for data collection of race, ethnicity, and primary language in the healthcare system. They begin with the Civil Rights Act of 1964^{10(p3)}:

Title VI of the Civil Rights Act of 1964, which prohibits intentional discrimination on the basis of race or national origin in the provision of any services that are at all supported with federal funds, is considered the broadest mandate the federal government has to require collection and/or reporting of data on race, ethnicity, and primary language. ... Title VI provides a legal foundation for the collection of racial and ethnic data by and from recipients of federal financial assistance.

From Norris Cotton Cancer Center, Dartmouth Medical School, Hanover, New Hampshire.

Address correspondence and reprint requests to Bruce L. Riddle; NHSCR; P.O. Box 186; Hanover, NH 03755; 603-653-1036; 603-653-1050 (fax); bruce.l.riddle@dartmouth.edu

This paper examines the disconnect between the national standards on collecting race and ethnicity and what occurs in New Hampshire at the hospital level, at the admitting desk.

In addition to the Civil Rights Act, "For agencies that conduct clinical research or monitor disease, five statutes exist that address the collecting and reporting of racial and ethnic data."^{10(p4)} Four of the statutes relate to the activities of the Centers for Disease Control and Prevention. Some federal regulations require racial and ethnic data collection by some programs such as Medicaid.

Federal programs follow guidelines established in Executive Office of the President, Office of Management and Budget, Directive 15: Race and Ethnic Standards for Federal Statistics and Administrative Reporting.¹¹ The first directive was issued in May 1977 at the direction of Congress and described four races (American Indian or Alaskan Native, Asian or Pacific Islander, Black, and White) and two ethnic backgrounds (of Hispanic origin and not of Hispanic origin). The 1977 OMB directive reflected work begun in 1974 by an inter-agency committee of the federal government.

In October 1997, after an extensive period of task force activity, comment, and debate, the Office of Management and Budget issued "Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity."⁶ The revised standards delineated five minimum categories for data on race (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, or

White) and two categories for data on ethnicity ("Hispanic or Latino" and "Not Hispanic or Latino"). The OMB noted that "The categories represent a social-political construct designed for collecting data on the race and ethnicity of broad population groups in this country, and are not anthropologically or scientifically based."⁶ The OMB made other decisions⁶:

- i) retain the concept that the standards provide a minimum set of categories for data on race and ethnicity;*
- ii) permit the collection of more detailed information on population groups provided that any additional categories can be aggregated into the minimum standard set of categories;*
- iii) underscore that self-identification is the preferred means of obtaining information about an individual's race and ethnicity, except in instances where observer identification is more practical (eg, completing a death certificate);*
- iv) do not identify or designate certain population groups as "minority groups";*
- v) do not establish criteria or qualifications (such as blood quantum levels) that are to be used in determining a particular individual's racial or ethnic classification.*

The OMB definitions have been the subject of considerable criticism^{5(p 36)}:

- *reinforcement of the concept of "race" as reflecting genetic or biological differences between population groups;*
- *failure to reflect the fluid and dynamic nature of sociopolitical identity; and*
- *failure to reflect the way many Americans choose to define themselves.*

The new standards were to take effect on January 1, 2003. "It is important to note that OMB does not mandate the collection of racial, ethnic, and primary language data, but rather sets the standards by which the data are to be collected and presented."^{10(p6)} In December 2000, OMB issued "Provisional Guidance on the Implementation of the 1997 Standards for Federal Data on Race and Ethnicity."¹² The document again emphasized "that the Federal racial and ethnic data categories are social-political constructs and that they should

not be interpreted as being genetic, biological, or anthropological in nature."^{12(p9)} It is also emphasized that "the standard was not designed to capture the full complexity of race and ethnicity in the United States."^{12(p10)}

The "Provisional Guidance" document also discusses the importance of "self-reporting or self-identification as the preferred method for collecting data on race and ethnicity."^{12(p10)} Following the OMB decisions outlined above, the guidance document notes that "self-identification for race and Hispanic or Latino origin means that the responses are based on self-perception and therefore are subjective, but by definition, the responses are accurate."^{12(p10)} And it goes on to argue, because the standard allows individuals to report on one or more races, "... it is generally difficult for observers to report an individual's multiple racial heritages."^{12(p10)} Then the document describes one consequence of this self-reporting: "The use of self-identification coupled with the social nature of race also results in situations where an individual's response to questions on race may change over time as a result of the maturation process, the particular situation, and the changing environment."^{12(p10)}

The 2000 Census was the first nationwide implementation of the new federal standards for classification on race and ethnicity. The major change in the 2000 Census was to allow respondents the option of selecting one or more racial designations. In addition to the five categories in OMB Directive 15, "for respondents unable to identify with any of these five categories, the OMB approved including a sixth category—'Some other race'—on the Census 2000 questionnaire."^{3(p3)} The Census Bureau, in presenting the results of Census 2000, reports in Summary File 2 information on 249 race and ethnic groups.

In New Hampshire, the collection of race and ethnicity information rests on federal law and guidelines and at least

two state administrative rules. The administrative rules on reporting hospital discharge data to the state require the inclusion of information on race.¹³ The rules governing reporting to the state cancer registry specifically mention NAACCR Item Number 160, Race, in Version 6.¹⁴ The New Hampshire Hospital Association notes in a news article "it's neither illegal nor discriminatory to ask a patient for their race when admitting them. . ."¹⁵ The article goes on to note that "Extensive research was done with New Hampshire's Attorney General's Office before mandating the collection of this data element for the inpatient database."

The standards supported by NAACCR and the American College of Surgeons Committee on Cancer separate race and ethnicity. A single variable on "Spanish/Hispanic Origin" has 7 categories that attempt to capture Hispanic or Non-Hispanic status but also country of origin (Mexico, Puerto Rico, Cuba, or South America).^{8(p304)} Five variables allow for multiple races, and the number of categories is limited^{8(p204-210)} and do not match the most frequently chosen categories of the 2000 Census.

THE LITERATURE

Much literature exists on issues of race and ethnicity in public health including many excellent reviews and discussions. In the main, five broad conclusions can be taken from more recently published articles:

1. Racial and ethnic groups (or ancestral groups, continental populations, or geographic origins) are social constructs and have no basis in biology.^{12,16-20}
2. Racial and ethnic differences on access to and use of health services by certain groups persist even when differences in income and health insurance are held constant.^{5,21-22}
3. A body of research consistently documents the misclassification of indi-

viduals into racial and ethnic groups.²³⁻³⁰ When visual inspection or secondary measures are used, misclassification occurs frequently for the specific populations. A summary of these findings is presented in Table 1.

4. Public health data systems use inconsistent methods of classifying people by race and ethnicity.³¹
5. Currently, insufficient data exist to determine the precise relationships that exist between socially defined populations and noteworthy genetic features.³²

SURVEY

In Fall 2002, New Hampshire State Cancer Registry (NHSCR) sent to its 27 hospital registrars a simple questionnaire asking how race and ethnicity were coded at their hospitals. Eleven of the 27 hospitals have fewer than 105 reportable cancer cases per year, serve small communities and rural areas, and do not have a cancer registrar. Sixteen are larger hospitals; 9 participate in the American College of Surgeons Commission on Cancer Community Hospital Cancer Program, one is a Veterans Administration Hospital, one is an academic medical center. The remaining five hospitals have large enough patient load to support a cancer registrar. Responses came in from 22 hospitals. The five non-responding hospitals include the Veterans Administration Hospital, two larger hospitals, and two smaller hospitals with fewer than 105 reportable cases per year.

In Spring 2001, the New Hampshire Hospital Association administered a survey on the coding of patient race at acute care facilities. The conclusions one can draw from the two independent surveys are similar; the results are summarized in Table 2 and Table 3.

The two surveys show that the collection of race and ethnicity is poorly done and strays from many of the fed-

eral guidelines. Most hospitals use a coding scheme supplied by the New Hampshire Hospital Association³³:

- 1—White, non-Hispanic
- 2—Black, non-Hispanic
- 3—Asian/Pacific Islander
- 4—American Indian/Alaskan Native
- 5—Hispanic
- 6—Other/Multi-Racial
- X—Unknown or Refused to Answer

Other hospitals reported that their admitting software vendor provided the following choices:

- 1—African American
- 2—Asian/Pacific Islander
- 3—Caucasian
- 4—Hispanic
- 5—Multi-Racial
- 6—Native American/Alaskan
- 7—Other
- 8—Unknown

Both classifications fail to separate race and ethnicity as suggested by OMB Directive 15 for self-reporting. Both classifications allow only one choice or code to be completed so multiple races cannot be checked by a respondent. This scheme is allowed under OMB guidelines but without 'Other' or 'Unknown' categories.^{10(p6)} Most of the hospitals' admitting staff, as the survey shows, do not ask and just visually check. While the focus here is on hospitals, similar issues arise for physician offices, clinics, and pathology labs.

New Hampshire's population grew 11.4% from 1990 to 2000.³⁴ Whites are still the majority racial group, constituting 96% of the total population. Blacks or African Americans increased 25.5%; Asians increased 74.7%; Hispanics or Latinos increased 80.9%. At least two New Hampshire cities, Manchester and Nashua, have minority populations of approximately 10%. The changing composition of the New Hampshire population is so far not reflected in NHSCR cancer statistics. This finding may be because minority populations are young so the overall incidence of cancer is low. Alternatively, the recoding

Table 1. Selected summary of findings on the misclassification of individuals into racial and ethnic groups

Author(s)	Year	Ref	Study Description	Summary
Hahn RA, et al	1996	#26	National survey and proxies	Self-classification of ancestry at initial interview was consistent with proxy reports at follow-up for 55% of subjects for whom proxy information was available. (p. 75)
Swallen KC, et al	1998	#27	Population based cancer registry	Among persons classified as Vietnamese by registry race, 26% identified themselves as being non-Vietnamese. In addition, registry race missed 10% of the self-identified Vietnamese. (p. 225)
Hahn RA, et al	1992	#25	Linked birth/infant death computer tapes 1983–1985	The coding of race and ethnicity at birth and death is remarkably inconsistent. Inconsistency in the coding of race is low for Whites (1.2%), greater for Blacks (4.3%), and greatest for races other than White or Black (43.2%). (p. 259)
Boehmer U, et al	2002	#29	Department of Veterans Affairs dental out patients	Between 76.4% and 77.1% of self-reported Whites, between 68.4% and 68.9% of self-reported African Americans, between 57.1% and 61% of self-reported Hispanics, between 33.3% and 54% of Asians, and between 1.4% and 4.6% of self reported American Indians were classified as such in the observer based administrative data. (p. 1471)
Kressin N, et al	2003	#30	Veterans Affairs administrative data and VA 1999 Large Health survey race/ethnicity data	Our results indicated fairly poor overall agreement between administrative data and self-reported race/ethnicity; the best rates of agreement (for Whites and African Americans) were approximately only 60% and rates of agreement were even lower for other racial groups. (p. 1737)
Arday S, et al	2000	#23	Medicare Current Beneficiary Survey	Using linked data from the Medicare Current Beneficiary Survey (MCBS), the authors assessed the accuracy of racial/ethnic classifications in the 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% for White persons and 95% for Black persons, but less than 60% for all other categories. (p. 107)
Sorlie P, et al	1992	#24	Validity of demographic characteristics on death certificates, 1979–1985	In a sample of United States population from the Census Bureau's Current Population Survey, we compared demographics characteristic with those recorded on death certificates for 43,000 decedents followed from 1979–1985. Overall percentage agreements were: race 99.4%, Hispanic origin 98.7%. Relatively fewer American Indians and Asian/Pacific Islanders had death certificates that agreed with baseline race (73.6% and 82.4%, respectively). (p. 181)
Poe GS, et al	1993	#28	Comparability of death certificate and 1986 national mortality followback survey	Overall, there was a high level of agreement (97.9%) on race between the death certificate and the questionnaire. However, for those reported to be American Indian on the death certificate, the level of agreement was lower. (92.9%) (p. 6)

Table 2. Summary of survey results, New Hampshire State Cancer Survey, Fall 2002/Winter 2003

Race and ethnicity is collected by the admitting staff	21 out of 22 responses (95%)
Admitting staff visually check	21 out of 22 responses (95%)
The categories come from the NH Hospital Association or software vendor	12 out of 22 responses (54%)
The responses are stored electronically	6 out of 22 responses
Race appears in the medical record	21 out of 22 responses
Just race is collected	15 out of 22 responses (68%)
Race and ethnicity are collected	14 out of 22 responses (64%)
Patients are asked to self-identify	3 out of 22 responses
Race and/or ethnicity is abstracted from a required element	8 out of 22 responses (36%)
	18 out of 22 responses (82%)

Source: NHSCR Survey of New Hampshire Hospitals.

and coding of race and ethnicity are not done consistently.

The New Hampshire Department of Health and Human Services (DHHS) has recognized, at a minimum, that no uniformity in race and ethnicity existed across many public health surveillance data sets—hospital discharge data, vital records, cancer registry, etc.—making comparisons and analysis very difficult.³⁵ The dramatic rise in New Hampshire's minority population in the 10 years from 1990–2000 has encouraged public health official and health advo-

Table 3. Spring 2001 survey of hospitals by hospital association; race coding at acute care hospitals

1 hospital said they consistently asked their patient his or her race when registering them.
13 hospitals said they recorded race by sight.
6 hospitals reported a combination by sight and by asking.
"Inconsistent coding across hospitals and even across registrars at the same facility, when it came to coding patients that looked to be one race but specified another."
"Responses were also inconsistent about how to code a newborn's race if the parent's race differed."
"43% of non-White admissions were coded as other/multi-racial."

Source: June 2001 DataLink from the Foundation for Healthy Communities. Available at: http://www.nhha.org/fhc/healthcare_data/newsletter/June2001.php.

cacy groups to look for data that include information on race and ethnicity. Hospitals, in turn, take their lead from the DHHS.

IMPEDIMENTS TO COLLECTION

The New Hampshire Hospital Association sponsored a meeting on data quality to review how race and ethnicity were collected in New Hampshire hospitals in winter 2003.³⁵ Part of the discussion focused on why admitting staff just "visually check." The list included:

- a) lack of administrative buy-in to support data collection
- b) admitting staff feel patients are defensive
- c) patient paranoia/patient anger
- d) patients refusing to answer
- e) language barriers, literacy/residence issues
- f) lack of training of admitting staff coupled with high rates of turnover
- g) discrimination issues
- h) phone admissions—communicating a list of choices hard to accomplish
- i) information technology problems—long lead times to change software
- k) privacy issues

As one participant put it, "Staff just visually check because they are afraid of being yelled at."³⁶ Hospital admitting clerks start a medical record and enter the most basic demographic information into the computer systems that

drive most hospital administrative systems. In a few moments, they are asked to collect identifying information, demographic information, information on survivorship, and details on billing and insurance information. These hospital staff members are not culturally equipped to quiz an individual, who may be distraught or anxious, on the country of birth or origin, primary language, and race and ethnicity. If staffs do ask, they may face bewilderment, confusion, hostility and even anger. The next person in the chain, usually a nurse, has the primary task of soliciting, as clearly as possible, medical information that will affect triage decisions or placement. Again, they are not culturally equipped to ask questions on country of birth, primary language, and race and ethnicity.

If collecting accurate information on ethnicity and race is important, then creating a cultural environment in the hospital setting that allows that information to be collected is a necessary condition. The above list of impediments provides a long agenda if the collection of race and ethnicity is to meet the standards outlined by federal rules and national standards bodies. In a state (and region) with small minority populations, creating the cultural environment where staff members feel comfortable to ask, and where patients are willing to answer, will require strong leadership from many sectors, inside and outside the healthcare system.

GENETIC INFORMATION

In May 2003, Baton Rouge Louisiana police used a DNA sample to identify a suspect's race.³⁷ This seminal event raises a number of troubling issues that greatly affect the debate on race and ethnicity in public health. On one hand, this event confirms some of the worst fears of those who oppose collection of race and ethnicity at all, since it might legitimize and impute a scientific imprint on categorization.^{19,31} Put another way, race and genetics would be coupled. On the other hand, for others it shows the power of DNA tools in the search for the genetic basis of diseases. Now, at least one measure of race and ethnicity can be performed away from the patient but presumably with their knowledge.

But what was measured? The report "typed the crime scene sample as being 85% African ancestry and 15% American Indian." The developer of the test used the term "major continental races." So, presumably, the test reveals phenotypic traits such as skin color and general features or race as biological inheritance. The test might answer one set of questions—phenotype—but will not be adequate to answer others such as how the patient views their own race and ethnicity.

PREREQUISITES FOR ACCURATELY COLLECTING INFORMATION

As noted above, the primary mandate for collecting information on race, ethnicity, and primary language is in support of Title VI of the Civil Rights Act of 1964¹⁰ and by extension, the continuing examination of unequal treatment of certain groups in receiving medical care. What information should legitimately be part of a medical record? What information should legitimately be part of a public health surveillance record? What are the needs of medical researchers, lawyers and administrators

working on civil rights issues, and public health officials doing surveillance? Who should collect this information and when? The genetic tests make these questions much more cogent. The language and method used, and the use to which the information is put, may indicate how individuals will view themselves and their place in society. The United States is on the verge of creating the first multi-racial/ethnic democracy in the world. How individuals are treated by public health institutions is a critical element if the creation of that democracy is to be successful.

Therefore, the routine collection of information on race, ethnicity, and primary language must be done in a framework that respects patient confidentiality and privacy and that acknowledges the fluidity of identification. In sum, if we conclude that we need this information, we need to collect it correctly. The question should be, "What is the best practice that should be used by public health institutions that supports the democracy?" and not, "What can we get away with?" The data collection moment at a hospital admitting desk or outpatient center is fraught with difficulties of time and privacy so the tendency is to take the easy way out: visual inspection or just skipping it. A data collection scheme not based on visual inspection will require development of a framework that addresses many issues of public health surveillance. For example:

1. Successfully collecting data on race and ethnicity in New Hampshire will require establishing a new culture that clearly defines the importance, methods, rules, and norms for collecting race and ethnicity in a medical setting. To say "it should be collected" is not enough, but a clear, credible rationale for the data needs to be presented to both staff and patients. Many New Hampshire hospitals do not have any written policies on the collection of race and ethnicity. Hospital staff members must feel comfortable asking an individual patient for the information.
2. Race and ethnicity should not be on the face sheet of a medical record. For many people, race and ethnicity are a privacy issue. Since race and ethnicity should not, *prima facie*, affect any treatment given, they should not appear next to sex or age. Patients should have the right not to disclose race or ethnicity. Visual inspection should not be allowed. A reasonable case could be made for not including it at all in the viewable medical record but only storing it to provide data files submitted to public health authorities or researchers. Information regarding primary language, because it allows hospital staff to plan for translators or staff members who speak the language, should be allowed.
3. Written policies, procedures, and norms must be provided by the state and individual providers. A listing of acceptable codes is not sufficient. A policy manual should define the method for collection, procedures for storing and disclosing the data, and the norms that system is trying to achieve. See, for example, a policy document from Rhode Island that is becoming a model for other states.³⁸
4. Healthcare staff need to receive constant training on how to collect the information. The staff who collect the information suffer from high turnover and the need to collect more and more information. The need for constant training—just like for any employee in the private service economy—is obvious.
5. The language used in collection matters. Census 2000 established a pattern of collection with which respondents were comfortable. To gather data that are compatible with Census 2000, we should use the same format, at least as a starting point. It may be desirable, as some have noted, to broaden the language, collapse the terms race and ethnicity into one term, and add ancestry.^{31,39}

The framework must deal with the fact, as some researchers have noted, that an increasing number of people will choose not to respond.⁴⁰ This risk of not allowing visual inspection is that significant numbers of people will be classified as missing. For the purposes of the democracy and increasing confidence in the data, that risk might be necessary. By some estimates, 70% of US population growth will come through immigration.^{41,42} Others have noted that with intermarriage, each generation is more racially and ethnically mixed than the previous generation and that in some parts of the country, minorities have become the majority.⁴³ The methods and categories used in collection need to be broader and be a part of a larger framework.

A new framework will carry with it an administrative burden: developing policy manuals, training staff, new forms, new software, and new efforts at compliance. Larger bureaucratic organizations such as larger hospitals can absorb such a burden. Small hospitals, physician offices, and clinics, where more and more diagnosis and treatment of cancer is shifting, have fewer resources to draw on. Over time, without strong leadership and legal mandates, already incomplete data on race and ethnicity could become even more incomplete. For the framework to be successful, third parties such as insurance companies and HMOs, who also routinely collect data on race and ethnicity, will have to buy in to the new framework.

A key component of democracy is equal treatment. The ability of governments to monitor treatments received by certain groups depends on collecting accurate information on race and ethnicity. One consequence of people either choosing not to respond (or, worse, giving a false identification) will be to hinder government's ability to monitor equal treatment. Another consequence might be to make it difficult to identify and recruit adequate numbers of member of subpopulations for clinical trials.⁴⁴ (p1174)

Misclassifications of race and ethnicity at the hospital level carry forward into numerators used to calculate incidence rates at county, regional, state, and national levels.

SUMMARY

The tacit support of visual inspection and questionable categorization schemes used in hospital settings in New Hampshire (and elsewhere) together lead to enough misclassification as shown by studies cited earlier to question the usefulness of providing incidence rates by race and ethnicity. In regions with small populations, errors of misclassification are probably high. Misclassifications of race and ethnicity at the hospital level carry forward into numerators used to calculate incidence rates at county, regional, state, and national levels. Comparisons of incidence rates become questionable; incidence rates on specific sites (ie, lung) by a subpopulation for a geographic area may not be valid.

A small but vocal, and maybe growing, minority question the role of all public health surveillance, including cancer registries. Poor collection methods, lack of interest in defining and supporting a standard, and a systematic failure to maintain the confidentiality of the data all work to undermine the larger public health surveillance system of which cancer registries are a part. Cancer registries at all levels need to tighten definitions and collect only information that is robust and meets adequate standards of validity and reliability.

ACKNOWLEDGMENT

Support for this research was provided by the New Hampshire State Cancer Registry. Funding for NHSCR comes from the New Hampshire Department of Health and Human Services and the Centers for Disease Control and Prevention. An early version of this paper was presented at the NAACCR 2003 Annual Meeting.

REFERENCES

1. The county population estimates used in the SEER*Stat software to calculate cancer incidence and mortality rates represent a modification of the annual time series of July 1 county population estimates by age, sex, race, and Hispanic origin produced by the US Census Bureau's Population Estimates Program, with support from the NCI through an interagency agreement. Available at: <http://seer.cancer.gov/popdata/>
2. Hotes JL, Wu XC, McLaughlin CC, et al, eds. *Cancer in North America, 1996–2000. Volume One: Incidence*. Springfield, Ill: North American Association of Central Cancer Registries; 2003.
3. Grieco E, Cassidy RC. *Census 2000 Brief: Overview of Race and Hispanic Origin*. Washington, DC: US Census Bureau; 2001. C2KBR/01–1.
4. Laws MB, Heckscher RA. Racial and ethnic identification practices in public health data systems in New England. *Public Health Rep*. 2002;117:50–61.
5. Smedley BD, Stith AY, Nelson AR, eds. *Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care*. Washington, DC: The National Academies Press; 2003.
6. Office of Management and Budget (OMB). *Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity. Statistical Policy Directive No. 15, Race and Ethnic Standards for Federal Statistics and Administrative Reporting*. Washington, DC: OMB; 1997.
7. American College of Surgeons Commission on Cancer. *Facility Oncology Registry Data Standards (FORDS)*. Standards of the Commission on Cancer. Vol 2. Chicago, Ill: American College of Surgeons Commission on Cancer; 2002.
8. Hultstrom D, ed. *Standards for Cancer Registries Volume II: Data Standards and Data Dictionary, Version 10.1*. 7th ed. Springfield, Ill: North American Association of Central Cancer Registries; 2003.
9. National Program of Cancer Registries Act §280e (a), 42 USC (1998).
10. Perot RT, Youdelman M. *Race, Ethnic, and Primary Language Data Collection in the Healthcare System: An Assessment of Federal Policies and Practices*. New York, NY: The Commonwealth Fund; 2001. Publication No. 492.
11. Wallman KK, Evinger S, Schechter S. Measuring our nation's diversity: developing a common language for data on race/ethnicity. *Am J Public Health*. 2002;90(11):1704–1708.
12. Office of Management and Budget. *Provisional Guidance on the Implementation of the 1997 Standards for Federal Data on Race and Ethnicity*. Washington, DC: Tabulation Working Group, Interagency Committee for Review of Standards for Data on Race and Ethnicity; 2000.
13. New Hampshire Code of Administrative Rules. Part He-C 1502 Acute Care Hospitals—Health Care data, He-C 1502.02 data format.
14. New Hampshire Code of Administrative Rules. Part He-P304 Cancer Registry Rules.
15. Foundation for Healthy Communities. Remember, It Doesn't Hurt to Ask! *DataLink*. June 2001.
16. Williams DR. Race and health: basic questions, emerging directions. *Ann Epidemiol*. 1997;7:322–333.
17. Williams DR. The concept of race in health services research: 1966–1990. *Health Serv Res*. 1994;29(3):261–274.
18. Kaplan JB, Bennett T. Use of race and ethnicity in biomedical publication. *JAMA*. 2003;289(20):2709–2716.
19. Cooper RS, Kaufman JS, Ward R. Race and Genomics. *N Engl J Med*. 2003;318(12):1166–1170.
20. American Medical Association. CSA Reports: Report 5 of the Council on Scientific Affairs (I-00) Full Text, Classification of Data on Race and Ethnicity. 2000. Available at: <http://www.ama-assn.org/ama/pub/print/article/2036-7205.html>. Accessed September 3, 2003.
21. De Lew N, Weinick RM. An overview: eliminating racial, ethnic, and SES disparities in health care. *Healthcare Financing Rev*. 2000; 21(4):1–7.
22. Haynes MA, Smedley BD, eds. *The Unequal Burden of Cancer: An Assessment of NIH Research and Programs for Ethnic Minorities and the Medically Underserved*. Washington, DC: National Academy Press; 1999.
23. Arday SL, Arday DR, Monroe S, Zhang J. HCFA's racial and ethnic data: current accuracy and recent improvements. *Healthcare Financing Rev*. 2000;21(4):107–116.
24. Sorlie PD, Rogot E, Johnson NJ. Validity of demographic characteristics on the death certificate. *Epidemiology*. 1992;3(2):181–184.
25. Hahn RA, Mulinare J, Teutsch SM. Inconsistencies in coding of race and ethnicity between birth and death in US infants: a new look at infant mortality, 1983–1985. *JAMA*. 1992;267(2):259–263.
26. Hahn RA, Truman BI, Barker ND. Identifying ancestry: the reliability of ancestral identification in the United States by self, proxy, interview, and funeral director. *Epidemiology*. 1996;7(1):75–80.
27. Swallen KC, Glaser SL, Stewart SL, West

- DW, Jenkins CNH, McPhee SJ. Accuracy of racial classification of Vietnamese patients in a population-based cancer registry. *Ethn Dis.* 1998;8:218–227.
28. Poe GS, Powell-Griner E, McLaughlin JK, et al. Comparability of the death certificate and the 1986 National Mortality Followback Survey. National Center for Health Statistics. *Vital Health Stat.* 1993;2(118).
 29. Boehmer U, Kressin NR, Berlowitz DR, Christiansen CL, Kazis, LE, Jones JA. Self-reported vs administrative race/ethnicity data and study results. *Am J Public Health.* 2002; 92(9):1471–1473.
 30. Kressin NR, Chang BH, Hendricks A, Kazis LE. Agreement between administrative data and patients' self-reports of race/ethnicity. *Am J Public Health.* 2003;93(10):1734–1739.
 31. Laws MB, Heckscher RA. Racial and ethnic identification practices in public health data systems in New England. *Public Health Rep.* 2002;117:50–61.
 32. Foster MW, Sharp RR. Race, ethnicity, and genomics: social classifications as proxies of biological heterogeneity. *Genome Res.* 2002; 12:844–850.
 33. Foundation for Healthy Communities. *In Patient Data Specification Manual.* Concord, NH: Foundation for Healthy Communities; 1997.
 34. US Census Bureau. *Census 2000 Redistricting Data (P.L. 94–171) Summary File* [tables PL1, PL2, PL3 and PL4]. Washington, DC: US Census Bureau. 1990 Census of Population, General Population Characteristics (1990 CP-1).
 35. Augustine D. *Data Quality Committee Minutes.* Concord, NH: Foundation for Healthy Communities; February 7, 2003.
 36. Author's notes from Foundation for Healthy Communities meeting, February 7, 2003.
 37. Wade N. Unusual use of DNA aided in serial killer search. *New York Times.* June 3, 2003. Available at: <http://query.nytimes.com/search/>.
 38. Brown-Small V, Buechner JS. *Policy for Maintaining, Collecting, and Presenting Data on Race and Ethnicity.* Providence, RI: Rhode Island Dept of Health; 2000.
 39. Kaufman JS, Cooper RS. Commentary: considerations for use of racial/ethnic classification in etiologic research. *Am J Epidemiol.* 2001;154(4):291–298.
 40. Hirschman C. Race and ethnicity. Testimony before National Institute of Child Health and Human Development, Center for Population Research, Demographic and Behavioral Sciences Branch, Visions of the Future: DBSB Long-Range Planning. 2001. Available at: <http://www.nichd.nih.gov/cpr/dbs/sp/hirschman1.htm>. Accessed December 18, 2003.
 41. US immigration and population growth. SUSPS web pages. Available at: <http://www.susps.org/overview/immigration.html>. Accessed October 20, 2003.
 42. Hollman FW, Mulder TJ, Kallan JE. *Methodology and Assumptions for the Population Projections of the United States: 1999 to 2100.* Washington, DC: Populations Projections Branch, Population Division, US Census Bureau, Dept of Commerce; 2000.
 43. Rodriguez G. Mongrel America. *The Atlantic Monthly.* January/February 2003;95–97.
 44. Burchard EG, Ziv E, Coyle N, et al. The importance of race and ethnic background in biomedical research and clinical practice. *N Engl J Med.* 2003;318(12):1170–1175.

AUTHOR CONTRIBUTIONS

Design and concept of study: Riddle

Acquisition of data: Riddle

Data analysis and interpretation: Riddle

Manuscript draft: Riddle