Validity of Infant Race/Ethnicity from Birth Certificates in the Context of U.S. Demographic Change

Lisa Reyes Mason, Yunju Nam, and Youngmi Kim

Objective. To compare infant race/ethnicity based on birth certificates with parent report of infant race/ethnicity in a survey.

Data Sources. The 2007 Oklahoma birth certificates and SEED for Oklahoma Kids baseline survey.

Study Design. Using sensitivity scores and positive predictive values, we examined consistency of infant race/ethnicity across two data sources ($N = 2,663$).

Data Collection/Extraction Methods. We compared conventional measures of infant race/ethnicity from birth certificate and survey data. We also tested alternative measures that allow biracial classification, determined from parental information on the infant's birth certificate or parental survey report.

Principal Findings. Sensitivity of conventional measures is highest for whites and African Americans and lowest for Hispanics; positive predictive value is highest for Hispanics and African Americans and lowest for American Indians. Alternative measures improve values among whites but yield mostly low values among minority and biracial groups.

Conclusions. Health disparities research should consider the source and validity of infant race/ethnicity data when creating sampling frames or designing studies that target infants by race/ethnicity. The common practice of assigning the maternal race/ethnicity as infant race/ethnicity should continue to be challenged.

Key Words. Vital statistics, racial/ethnic differences in health and health care, infant health, survey research and questionnaire design

Birth certificates provide a valuable sampling frame for infant health and disparity studies, as they include information on almost everyone in the target population and on birth parents (e.g., race, ethnicity, and education), and allow researchers to construct subsamples of underrepresented groups (Schoendorf and Branum 2006). Since 1989, the National Center for Health Statistics (NCHS) has tabulated infant vital statistics primarily by assigning

the birth mother's race/ethnicity to the infant (NCHS 2005). Infant outcomes research often follows this convention, yet the practice is subject to critique (Ma 2008).

First, there is growing concern about the validity of birth certificate data on race and ethnicity (i.e., Hispanic origin). Comparisons of data from birth certificates with those from other sources suggest that racial/ethnic identification is often consistent for white, African American, and Hispanic birth mothers (Piper et al. 1993; Reichman and Hade 2001; Vinikoor et al. 2010) but inconsistent for American Indians and rarely studied for Asians (Hahn and Stroup 1994; Baumeister et al. 2000). Second, by using maternal race/ethnicity as the measure of infant race/ethnicity, researchers ignore the potential role of paternal race/ethnicity in infant outcomes. Recent studies find that paternal characteristics, including an indicator for missing racial/ethnic data, can have significant associations with infant outcomes (Getahun, Ananth, and Vintzileos 2006; Ma 2008). Third, the practice of using maternal race/ethnicity for the infant may not reflect U.S. demographics. Studies that use the NCHS convention will categorize non-Hispanic infants as white if the birth certificate reports that they have a white mother and African American father. However, parents can identify the same infant as white, African American, biracial, or some other race/ethnicity in surveys and administrative records. The prevalence of mismatch between conventional measures and parental perception is likely to grow with the increasing percentage of infants born from interracial/interethnic relationships in the United States (Atkinson, MacDorman, and Parker 2001; Johnson and Lichter 2010).

Existing studies suggest that the race/ethnicity variable produced by the conventional approach may be inconsistent with parents' view of their children. Using data from the 1990 Census, Qian (2004) focuses on children (under age 5) born to couples in whom one parent is white and the other is not. Qian finds that African American–white couples are most likely to identify their child with a racial/ethnic minority label; Asian American–white couples are least likely. Similar findings come from studies that focus on Asian–white couples' reporting of child race/ethnicity (Saenz et al. 1995; Xie and Goyette 1997). Using data from a nationally representative sample of kindergarteners

Address correspondence to Lisa Reyes Mason, Ph.D., College of Social Work, University of Tennessee, 1618 Cumberland Avenue, 321 Henson Hall, Knoxville, TN 37996-3333; e-mail: lmason@utk.edu. Yunju Nam, Ph.D., is with the School of Social Work, University at Buffalo, State University of New York, Buffalo, NY. Youngmi Kim, Ph.D., is with the School of Social Work, Virginia Commonwealth University, Richmond, VA.

and their parents, Brunσμα (2005) finds that 10.4 percent of the children are “multiracial by biological parentage,” but that parents report only 2.6 percent to be “more than one race” (2005, 1137). Both Brunσμα (2005) and Qian (2004) find that, among interracial couples, parents may tend to report the same race/ethnicity for children as for fathers.

Smith et al. (2010) compare child race/ethnicity based on birth certificates with that in health plan records. Using birth certificates as the gold standard (or criterion source), they report that 23.1–33.6 percent of health plan records misclassify child race/ethnicity, primarily due to missing data. Misclassification is low among Hispanic children but high among children in other racial/ethnic minority groups. Although one of the first validity studies of child race/ethnicity, the work by Smith et al. assumes that birth certificate data are accurate, exclusively uses the mother’s race for the child’s, and does not analyze data for infants.

This study contributes to the literature by providing the first comparison of infant race/ethnicity from birth certificates with parent report of infant race/ethnicity in a survey. We examine the NCHS convention to classify infant race/ethnicity based on the maternal race/ethnicity listed in infant birth certificates, and we test whether this practice captures the dynamic racial/ethnic distribution in the United States.

DATA AND SAMPLE

This study uses data from 2007 Oklahoma birth certificates and the SEED for Oklahoma Kids (SEED OK) baseline survey. Both datasets come from SEED OK, a longitudinal, statewide policy experiment (Zager et al. 2010). The SEED OK experiment is unique in that the sampling frame comprises birth certificates of all infants born in Oklahoma during April–June 2007 and August–October 2007. Using stratified random sampling, SEED OK selected a sample of 7,297 infants, oversampling non-Hispanic African American, non-Hispanic American Indian, and Hispanic infants to permit separate analyses by race/ethnicity (Marks, Rhodes, and Scheffler 2008). Of these, 182 were ineligible (e.g., maternal/infant death). Among the 7,115 remaining, 2,704 had a birth mother or main caretaker (if the birth mother did not live with the infant) join the study and complete the baseline survey in August–December 2007 or January–April 2008 (Zager et al. 2010). The response rate was 38 percent; study participation rates do not differ in a statistically significant way among whites (38.27 percent), African Americans (39.81 percent),

American Indians (38.30 percent), and Hispanics (35.45 percent; Nam et al. 2013). This study obtained an analytic sample of 2,663 cases by excluding 41 infants whose main caretakers were not the birth mother (e.g., a father, grandparents, or siblings).

ANALYSES AND RESULTS

Conventional Measures

We generated conventional measures of infant race/ethnicity from 2007 Oklahoma birth certificate data and SEED OK survey data. In 2007, Oklahoma and at least 22 other states used the 1989 U.S. Certificate of Live Birth (Martin et al. 2011). The form allowed more than one race/ethnicity for each parent, but the Oklahoma State Department of Health entered only one race in its 2007 vital statistics dataset following NCHS protocol at the time. Although the department recommended that birth mothers complete the birth certificate worksheet, it is possible that some hospitals used other data collection procedures (e.g., observation). Northam, Polancich, and Restrepo (2003) document this in other states.

We collapse race and ethnicity into one construct, creating a conventional variable to capture infant race/ethnicity from birth certificate data (hereafter, conventional birth certificate measure, or CBC):

1. If the birth certificate reported that the mother was Hispanic (Mexican, Puerto Rican, Cuban, Central or South American, or other and unknown Hispanic), we coded the infant as Hispanic.
2. If the certificate did not record the mother's Hispanic origin but indicated that the father was Hispanic, we would code the infant as Hispanic. No certificate, however, met this criterion.
3. We used the certificate's information on the mother's race to categorize non-Hispanic infants as non-Hispanic white (white), non-Hispanic African American (African American), non-Hispanic American Indian (American Indian), and non-Hispanic Asian (Asian).

The CBC measures race/ethnicity in five groups: white, African American, American Indian, Asian, and Hispanic. As no case in the study sample is identified as Hispanic because of father's Hispanic origin, CBC corresponds to the mother's race/ethnicity.

It is important to distinguish between our approach—which treats race/ethnicity as a single construct—and current Office of Management and Budget

(OMB) standards for collection of data on race and ethnicity. To some extent, the above categories can be considered analogous to the OMB's minimum categories for race and ethnicity: white, black or African American, American Indian or Alaska Native, Asian, Native Hawaiian or other Pacific Islander, and Hispanic or Latino (an exception, the Asian category used here, combines Asians, Native Hawaiians, and other Pacific Islanders into one group; OMB 1997). An important difference, however, is that we treat Hispanic or Latino origin (ethnic identity in OMB standards), and white, African American, American Indian, and Asian (racial identities in OMB standards) as categorical values of a single construct. Although OMB standards (1997) indicate that collection and reporting of ethnic and racial identity as two separate constructs is preferred, we decide to treat race/ethnicity as a single construct in light of recent and growing evidence that many people of Hispanic or Latino origin do not distinguish their Hispanic identity from their race (Buescher, Gizlice, and Jones-Vessey 2005; IOM [Institute of Medicine] 2009; Perez and Hirschman 2009). Both the OMB (1997) and the IOM (2009) recommend additional research to examine the possible combination of Hispanic origin and race into a single-item format in national surveys. (For reviews of race and ethnicity as sociopolitical constructs, see Hahn and Stroup [1994]; Mays et al. [2003]; Ford and Kelly [2005]; and IOM [2009]).

We created an equivalent conventional variable to capture infant race/ethnicity from the survey data (hereafter, conventional survey measure, or CS). Congruent with the OMB recommendation for self-reported data, the SEED OK survey posed two questions about infant race/ethnicity, asking mothers about their infant's identity and permitting more than one response in their description of the infant's race. Because Oklahoma's 2007 vital statistics dataset includes only one value from each parent's racial information on the infant birth certificate and because CBC categorizes racial/ethnic identity in five groups, we generated CS using the first answer to the SEED OK race question; we assume that mothers mentioned their infant's primary racial identity first. We generated CS as follows:

1. If the mother identified her infant as Hispanic, we coded the infant as Hispanic.
2. If the mother identified her infant as non-Hispanic or did not answer the infant's Hispanic origin question, and if the mother identified her infant as white, African American, American Indian/Alaska Native, or Asian/Pacific Islander/Hawaiian Native, we assigned the infant to the corresponding non-Hispanic group.

- 3. If the mother identified her infant’s race/ethnicity as Other (19 cases), we reviewed her specified answer. We coded nine of these cases as Hispanic and six as white. We coded four nonclassifiable cases (e.g., “American”) as Unknown.
- 4. We coded two cases as Unknown because the mothers refused to answer or chose “Do Not Know” for race and also chose “Do Not Know” or reported “No” for Hispanic origin.

We use six categories in the CS: white, African American, American Indian, Asian, Hispanic, and Unknown.

As Table 1 shows, the percentages of white, African American, and Asian infants are comparable in the two datasets. The proportion identified as American Indian is higher in CBC than in CS; the pattern is reversed among Hispanic infants.

To examine the relationship between the two conventional measures, we ran a cross-table of CBC and CS, estimating sensitivity scores (or true positive rate) and positive predictive values (PPV; see Table 2; Piper et al. 1993; Baumeister et al. 2000; Reichman and Hade 2001; Reichman and Schwartz-Soicher 2007). The sensitivity score represents the percentage of infants in the criterion data source (SEED OK survey) whose racial/ethnic assignment

Table 1: Infant Race/Ethnicity from Birth Certificates and Survey Data: Conventional Approach (*N* = 2,663)

<i>Race/Ethnicity</i>	<i>CBC</i>		<i>CS</i>	
	<i>n</i>	<i>%</i>	<i>n</i>	<i>%</i>
White	1,215	45.63	1,179	44.27
African American	463	17.39	472	17.72
American Indian	512	19.23	401	15.06
Asian	26	0.98	23	0.86
Hispanic	447	16.79	582	21.86
Unknown	0	0.00	6	0.23
Total	2,663	100.00	2,663	100.00

Note. CBC data are based on responses to the birth certificate questions: (1) “OF HISPANIC ORIGIN? (Specify No or Yes—If yes, specify Cuban, Mexican, Puerto Rican, etc.)”; and (2) “RACE – American Indian, Black, White, etc. (*Specify below*).” Data for CS are based on responses to two survey questions: (1) “Is [*insert infant’s name*] of Hispanic or [Latina/Latino] origin?” and (2) “What is [*insert infant’s name*] race?” For the survey’s race question, SEED OK instructed interviewers to “Select all that apply. Prompt by reading categories if necessary.” Categories included white, African American, American Indian/Alaska Native, Asian/Pacific Islander/Hawaiian Native, and Other.

CBC, conventional birth certificate measure; CS, conventional survey measure.

Table 2: Comparing Two Conventional Measures of Infant Race/Ethnicity: CBC and CS ($N = 2,663$)

CS	CBC					
	White	African American	American Indian	Asian	Hispanic	Total
White	1,022	20	116	6	15	1,179
African American	21	424	23	1	3	472
American Indian	83	4	309	0	5	401
Asian	1	0	3	19	0	23
Hispanic	87	15	61	0	419	582
Unknown	1	0	0	0	5	6
Total	1,215	463	512	26	447	2,663
Sensitivity score	86.68	89.83	77.06	82.61	71.99	
PPV	84.12	91.58	60.35	73.08	93.74	

Note. CBC, conventional birth certificate measure; CS, conventional survey measure; PPV, positive predictive value.

matches their assignment in the birth certificate data, and separate scores are calculated for each group. For example, the sensitivity score for American Indian infants indicates the percentage of infants whose (a) mothers identified them as American Indian during the survey and (b) birth certificates also identify them as American Indian. This measure assumes that birth mothers' survey reports accurately identify infant race/ethnicity. The PPV represents the percentage of infants whose assignment to a given racial/ethnic group in the birth certificate data matches their assignment in the survey data. In other words, PPV captures the ability of birth certificate data to correctly predict the infant race/ethnicity identified in the survey.

In using single racial/ethnic classifications, we find that the two measures produce consistent estimates for the majority of cases (Table 2). Sensitivity scores are highest for African Americans (89.83 percent), followed by whites (86.68 percent), Asians (82.61 percent), American Indians (77.06 percent), and Hispanics (71.99 percent). Because sensitivity scores are estimated with the assumption that surveyed mothers correctly identified their infant's race/ethnicity, these results suggest that birth certificate data are less accurate at capturing infant racial/ethnic identity among American Indians and Hispanics than among other groups. The patterns from PPVs differ somewhat. Hispanics have the highest PPV (93.74 percent) but the lowest sensitivity score. These results, along with a higher proportion of Hispanic infants in the survey data (21.86 percent) than in the birth certificate data (16.79 percent), suggest that many mothers identify their infant as Hispanic but do not identify themselves

as such. We note that the conventional measures produce low sensitivity scores and PPVs for American Indian infants.

The Role of Biracial Heritage

In searching for sources of inconsistency between birth certificate and survey data, we examine the possible role of biracial/biethnic heritage (hereafter, biracial heritage). We posit that a mother's perception of her infant's race/ethnicity may be less consistent with infant race/ethnicity derived from birth certificates, if each parent has a different racial/ethnic background than if they share a common background. To test this hypothesis, we used birth certificate data on both parents' race/ethnicity to generate an indicator of infant racial heritage: single heritage represents infants whose parents belong to the same racial/ethnic group, biracial heritage represents those whose parents belong to different groups, and Unknown represents 78 cases in which the father's information is missing from birth certificate data. Because we treat race/ethnicity as a single construct, we coded the heritage variable as biracial if the certificate identifies one parent as Hispanic and the other as a member of another racial/ethnic group (e.g., Hispanic mother and white father, regardless of mother's response to the specific race question). Hispanic identity studies inform this decision (Buescher, Gizlice, and Jones-Vessey 2005; IOM 2009; Perez and Hirschman 2009).

Table 3 demonstrates that racial heritage influences the consistency of identification between birth certificate data and survey data. For 92.69 percent of infants of single heritage, the race/ethnicity based on the birth certificate is the same as that reported in the survey. However, only 62.57 percent of infants of biracial heritage have the same information in both data sources. Among infants whose birth certificates lack information on the father's race/ethnicity (Unknown), consistency is 88.46 percent. The table's second panel presents consistency rates across heritage and racial/ethnic groups. Results indicate that the relationship between biracial heritage and consistent identification differs by race/ethnicity. Across all five groups, birth certificate and survey data are more consistent for infants with single heritage than for infants with biracial heritage, but the two datasets are much more consistent for African American and Hispanic infants with biracial heritage (over 80 percent) than for other infants with biracial heritage (51 percent or lower). Results suggest that some African American and Hispanic mothers may identify their infants' racial/ethnic identity as the same as theirs even when fathers belong to different racial/ethnic groups than mothers, or fathers' racial/ethnic identity is unknown.

Table 3: Consistency of Conventional Measures, for Infants with Single, Biracial, and Unknown Heritage (N = 2,663)

Sample (CBC)	Single Heritage	Biracial Heritage	Unknown*	Total
Whole sample	92.69 (91.70–93.68)	62.57 (60.73–64.41)	88.46 (87.25–89.67)	82.35 (80.90–83.80)
Infant race/ethnicity				
White	92.07 (90.55–93.59)	45.59 (42.79–48.39)	80.00 (77.75–82.25)	84.12 (82.06–86.18)
African American	97.50 (96.08–98.92)	81.97 (78.47–85.47)	100.00 (100.00–100.00)	91.58 (89.05–94.11)
American Indian	80.12 (76.66–83.58)	51.02 (46.69–55.35)	33.33 (29.25–37.41)	60.35 (56.11–64.59)
Asian	92.86 (82.96–102.76)	50.00 (30.78–69.22)	NA	73.08 (56.03–90.13)
Hispanic	98.50 (97.37–99.63)	87.58 (84.52–90.64)	80.00 (76.29–83.71)	93.74 (91.49–95.99)

Note. Results are percentages, with 95% confidence intervals (in parentheses).

*Father's information is missing.

CBC, conventional birth certificate measure.

Table 4 presents the numbers and percentages of infants by race/ethnicity and heritage. Combined with the information in Table 3, results in Table 4 suggest why sensitivity scores and PPVs in Table 2 are higher for some racial/ethnic groups than for others. Whites' scores and PPVs appear higher because the proportion of infants with biracial heritage is much lower among whites than among other groups (16.79 percent, Table 4). Meanwhile, African Americans' scores are high despite a high proportion of infants with biracial heritage (39.52 percent, Table 4) because the consistency rate is high even among biracial infants (81.97 percent, Table 3). Among American Indians, the rate of biracial heritage is very high (66.99 percent, Table 4) and consistency between the datasets is low, both for single and biracial heritage groups. These findings may partially explain the lower sensitivity score and PPV among American Indian infants. The percentage of biracial infants is lower among Hispanics than among all other groups except whites. In addition, consistency rates for infants of single and biracial heritage are higher among Hispanics than among other racial/ethnic groups. A low sensitivity score and high PPV, along with these findings, indicate that a high proportion of mothers may identify their infants as Hispanic even if only one biological parent is Hispanic. Additional analyses confirm this: infants are identified as Hispanic in the survey by 85 percent of Hispanic mothers whose infant's father is non-Hispanic (141 of 159) and by 91 percent of non-Hispanic mothers whose infant's father is Hispanic (83 of 91).

Seeking Alternative Measures

Because biracial heritage is a possible source of inconsistency between our initial measures (CBC and CS), we developed and tested two alternative

Table 4: Single, Biracial, and Unknown Heritage by Infant's Race/Ethnicity, Conventionally Measured ($N = 2,663$)

<i>Infant's Race/Ethnicity</i>	<i>Single Heritage (%)</i>	<i>Biracial Heritage (%)</i>	<i>Unknown (%)*</i>
White	996 (81.98)	204 (16.79)	15 (1.23)
African American	240 (51.84)	183 (39.52)	40 (8.64)
American Indian	166 (32.42)	343 (66.99)	3 (0.59)
Asian	14 (53.85)	12 (46.15)	0 (0.00)
Hispanic	266 (59.51)	161 (36.02)	20 (4.47)
Total	1,682 (63.16)	903 (33.91)	78 (2.93)

Note. Estimates come from CBC data.
*Father's information is missing.
CBC, conventional birth certificate measure.

measures of infant race/ethnicity; both include a biracial category. We created the first, an alternative birth certificate measure (ABC), as follows:

1. If the certificate listed both parents as Hispanic, we coded the infant as Hispanic.
2. If the certificate indicated that a non-Hispanic infant's parents are both from the same racial group, we coded the infant as a member of that group. If the certificate lacked a parent's Hispanic origin information but included race, we coded the infant solely from the listing of parents' race.
3. If the certificate indicated that the parents belong to different racial groups, we coded the infant as biracial. We also coded the infant as biracial if one parent was listed as Hispanic and one non-Hispanic.
4. If the certificate lacked either parent's race information, we coded the infant as Unknown.
5. The ABC employs seven categories: white, African American, American Indian, Asian, Hispanic, Biracial, and Unknown.

We generated the second alternative measure (AS) from survey information on the infant's Hispanic origin and race. Like CS, AS captures responses to the SEED OK survey. However, CS considers only the first answer to the race question; AS uses multiple answers if mothers provided more than one racial identity for the infant and/or if they described the infant as Hispanic, with a non-Hispanic racial identity. We generated AS as follows:

1. If the mother identified the infant as non-Hispanic or Hispanic origin information was missing (13 cases), and if the mother reported single heritage for the infant (the mother gave only one answer for infant race), we coded the infant based on race.
2. If she identified the infant as Hispanic and identified the infant's race in a way that can only be classified as Hispanic (e.g., her response to infant race question was "Other," and we upcoded the response to Hispanic), we coded the infant as Hispanic.
3. If she identified the infant as Hispanic and chose one or more of five racial categories (white, African American, American Indian, Asian, or Other), and we did not upcode the response as Hispanic or Unknown, we coded the infant as biracial.
4. If she identified the infant as non-Hispanic or if Hispanic origin information was missing, and she reported that the infant has more than

one race (i.e., two or more answers for infant race), we coded the infant as biracial.

- 5. If she identified the infant as non-Hispanic or if Hispanic origin information was missing, and if she did not report the race, we coded the infant as Unknown.
- 6. Like ABC, AS employs seven categories: white, African American, American Indian, Asian, Hispanic, Biracial, and Unknown.

We compared results from the alternative measures to test whether ABC and AS are more consistent with each other than are CBC and CS. We also calculated alternative sensitivity scores and PPVs. We excluded 83 cases with missing values, reducing the sample for ABC and AS analyses to 2,580.

Results from the alternative measures show a complicated relationship between biological heritage and mothers' perception of infant race/ethnicity (Table 5). The majority of infants identified as single race/ethnicity in ABC have the same heritage in AS. The PPV is higher than 80 percent for all single heritage groups except American Indians. Almost 63 percent of infants identified as American Indian in the ABC are also identified as such in AS; AS identifies about 31 percent (51 of 166) as biracial. Over half of infants classified as biracial in ABC are classified as single heritage in AS; about 47 percent (425 of 901) are identified as biracial. These results suggest that many mothers may identify their infants as members of a single racial/ethnic group even if each parent's race/ethnicity differs from other. Sensitivity scores are low for all

Table 5: Comparing Two Alternative Measures of Infant Race/Ethnicity: ABC and AS (*N* = 2,580)

<i>AS</i>	<i>ABC</i>						<i>Total</i>
	<i>White</i>	<i>African American</i>	<i>American Indian</i>	<i>Asian</i>	<i>Hispanic</i>	<i>Biracial</i>	
White	857	0	7	1	2	85	952
African American	2	201	1	0	0	128	332
American Indian	15	0	104	0	0	161	280
Asian	1	0	2	13	0	3	19
Hispanic	6	0	1	0	228	99	334
Biracial	114	39	51	0	34	425	663
Total	995	240	166	14	264	901	2,580
Sensitivity score	90.02	60.54	37.14	68.42	68.26	64.10	
PPV	86.13	83.75	62.65	92.86	86.36	47.17	

Note. ABC, alternative birth certificate measure; AS, alternative survey measure; PPV, positive predictive value.

groups except whites; in other words, if the information in AS is correct, the probability is low that racial/ethnic identities are accurate in ABC.

Table 5 demonstrates that, in this study, ABC does not improve consistency of infant racial/ethnic identification, when AS is used for comparison. Sensitivity scores and PPVs from conventional measures (Table 2) are often higher than those from the alternative measures, but the differences are not sizable (e.g., conventional sensitivity score for whites is 86.68 percent and the alternative is 90.02 percent). In fact, alternative sensitivity scores and PPVs are lower than the respective conventional values for several groups (e.g., for African Americans, the alternative sensitivity score is 60.54 percent and the conventional one is 89.83).

DISCUSSION

Taking advantage of birth certificate and survey data from a unique policy experiment, this study adds to the critique of deriving infant race/ethnicity from maternal race/ethnicity in birth certificates (Ma 2008). It also contributes to discussions about the influence of changing demographics on the measurement of race/ethnicity in health disparities research (Mays et al. 2003; Ramirez et al. 2005; Zaslavsky, Ayanian, and Zaboriski 2012). Our results suggest that infant race/ethnicity cannot be measured reliably across all groups if the measure employs either single-race classification from maternal race/ethnicity listed on birth certificates or multiple race classification from both parents' race/ethnicity on certificates.

Instead, research must seek a more nuanced understanding of racial/ethnic identity. We could move in this direction by revising birth certificates to collect parents' self-report of infant race/ethnicity as well as each parent's race/ethnicity. Collection of more detailed *granular ethnicity* may also be considered, as recommended by the IOM (2009, 33). States might collaborate with NCHS to test new data collection techniques. This suggestion is timely because the Patient Protection and Affordable Care Act of 2010 (Pub. L. 111-148, 124 Stat. 119) standardizes data collection requirements, mandating the use of OMB standards for measuring race and ethnicity in federally supported programs or surveys. It also requires that data on a minor be collected directly from a parent/legal guardian.

For African American infants in this study, single-race conventional measures more consistently identify race/ethnicity than alternative (biracial classification) measures. This finding is consistent with prior research on

African American identity. Those studies find that the so-called *one-drop rule* is persistent: having any African American ancestry increases the likelihood that one will self-identify as African American or be identified as such by others (Bratter 2007; Perez and Hirschman 2009). However, our results differ from previous findings that a parent's odds of identifying his/her child as white are higher if the parents are an African American mother and white father (Brunsma 2005).

We also find that conventional measures are more consistently accurate identifiers of American Indian identity than are the alternatives in this study, although conventional sensitivity scores and PPVs are generally low for this group. That neither approach seems to consistently identify American Indian infants may reflect the somewhat contradictory and evolving research on American Indian identity. Eschbach (1995) examines unions between American Indian and non-American Indian parents, finding that children born from those unions tend to be identified with the non-American Indian group. However, Eschbach (1995) and Liebler (2010) suggest that geography may matter. Because the American Indian population in Oklahoma is sizeable and more concentrated than in many other parts of the United States, one might expect our study findings of higher rates of American Indian infant identification and lower rates of multiracial or non-American Indian identification.

For Hispanic infants, racial/ethnic identification is complicated by separation of Hispanic origin from racial identity in vital statistics, a separation required by OMB standards. Studies of noninfant populations show that assigning Hispanics to the white racial category can artificially inflate or deflate health disparities outcome measures (Buescher, Gizlice, and Jones-Vessey 2005). This complication also seems present here. There is moderate-to-high consistency when Hispanic infants of any racial origin in birth certificates are classified solely as Hispanic (conventional measure), but less consistency when Hispanic infants identified with at least one other racial origin are classified as biracial (alternative measure). The alternative assignment of some Hispanic infants as biracial may not appropriately reflect that many Hispanics see their Hispanic identity as primary and more central than other racial origins they may have (Perez and Hirschman 2009). To address this, and to meet the Affordable Care Act's requirements concerning use of OMB minimum standards, new studies might examine the possibility of multiple items for the Hispanic origin question. For example, individuals might first indicate if they are of Hispanic origin. If yes, they might indicate if they are of full or partial Hispanic heritage, or if they consider their Hispanic identity as primary, secondary, or equal to any racial identities they may choose. Such studies would

inform further research into possible multiple response to the single-item Hispanic origin question, analogous to the multiple response now allowed to the question on race. The concept of multiple responses to Hispanic origin was first mentioned in the OMB's 1997 revised standards on race and ethnicity (OMB 1997).

Finally, we note a possible explanation of the low consistency with which the alternative measures identify biracial infants: people of different backgrounds may respond differently to the option of biracial classification. Thus, using parental race/ethnicity from birth certificates to identify biracial infants may not be an appropriate response to the problem of conventional reliance on maternal race/ethnicity data as the indicator of infant race/ethnicity. Although using birth certificate data on birth parent couplings (e.g., white/white, white/African American, African American/Asian; Ma 2008) provides useful information, our results suggest that researchers should use this approach cautiously: a substantial proportion of mothers in SEED OK identify their infant as a member of a single racial/ethnic group even when the infant has a parent of a different race/ethnicity. Future disparities research must consider and test alternate classifications before analyzing outcomes for racial/ethnic disparity studies.

Despite the unique contributions of this study, there are limitations. We chose the survey as the criterion source, but numerous factors influence birth parent report of infant race/ethnicity. Gender is an example (birth parents in the survey sample are female). Also, Oklahoma did not adopt the 2003 version of the U.S. Standard Certificate of Live Birth (which allows multiple listings for parent race) until after collection of the 2007 birth certificate data used here, and the data do not include multiple race information provided by some parents because this information was not entered into the dataset. The inclusion of multiple race data for parents might change our sensitivity scores and PPVs. Finally, the number of Asian cases is small because the Asian population in Oklahoma is small; readers should exercise caution in generalizing findings for this group.

The findings point to important recommendations for health services policy, practice, and research. First, policy makers and practitioners should evaluate the source of infant race/ethnicity data when targeting services to particular infant groups when attempting to address health disparities. Second, research studies should clearly identify how infant race/ethnicity is measured and conceived. Third, all remaining states should move toward adopting the 2003 U.S. Standard Certificate of Live Birth, which allows birth parents to identify themselves with multiple race categories. Fourth, research

should continue to challenge the convention of assigning birth mother's race/ethnicity to the infant. Future efforts should identify best practices for collecting infant race/ethnicity data. From in-depth interviews with mothers, fathers, and health care providers, qualitative research might provide information on how racial/ethnic constructs are defined and how these definitions may affect perceptions of infant racial/ethnic identity. New quantitative research might test parent self-report of infant race/ethnicity on birth certificates. In addition, new quantitative research should be carefully designed to include adequate subsamples of racial/ethnic minority groups, to improve general knowledge about racial/ethnic identity for these groups, and to improve the quality of data on which health disparity programs and policies may rely.

ACKNOWLEDGMENTS

Joint Acknowledgment/Disclosure Statement: Support for this research comes from the Ford Foundation and Charles Stewart Mott Foundation. The authors thank former Oklahoma State Treasurer Scott Meacham; Tim Allen, Deputy Treasurer for Policy and Administration; and James Wilbanks, former Director of Revenue and Fiscal Policy, who facilitated this research. We also appreciate contributions by staff at the Oklahoma State Department of Health, Kelly Baker, Derek Pate, and Sue Mallonee; and the survey research team at RTI International, Ellen Marks, Bryan Rhodes, and Jun Liu. At the Center for Social Development, we are grateful to Michael Sherraden and Margaret Clancy for designing and directing the SEED OK policy experiment, to Robert Zager for work with data management, and to Chris Leiker and Julia Stevens for editorial assistance.

Disclosures: None.

Disclaimer: None.

REFERENCES

- Atkinson, J. O., M. F. MacDorman, and J. D. Parker. 2001. "Trends in Births to Parents of Two Different Races in the United States: 1971–1995." *Ethnicity and Disease* 11 (2): 273–85.
- Baumeister, L., K. Marchi, M. Pearl, R. Williams, and P. Braveman. 2000. "The Validity of Information on 'Race' and 'Hispanic Ethnicity' in California Birth Certificate Data." *Health Services Research* 35 (4): 869–83.

- Bratter, J. 2007. "Will 'Multiracial' Survive to the Next Generation? The Racial Classification of Children of Multiracial Parents." *Social Forces* 86 (2): 821–49. doi:10.1093/sf/86.2.821.
- Brunsma, D. L. 2005. "Interracial Families and the Racial Identification of Mixed-Race Children: Evidence from the Early Childhood Longitudinal Study." *Social Forces* 84 (2): 1131–57. doi:10.1353/sof.2006.0007.
- Buescher, P. A., Z. Gizlice, and K. A. Jones-Vessey. 2005. "Discrepancies between Published Data on Racial Classification and Self-Reported Race: Evidence from the 2002 North Carolina Live Birth Records." *Public Health Reports* 120 (4): 393–8.
- Eschbach, K. 1995. "The Enduring and Vanishing American Indian: American Indian Population Growth and Intermarriage in 1990." *Ethnic and Racial Studies* 18 (1): 89–108. doi:10.1080/01419870.1995.9993855.
- Ford, M. E., and P. A. Kelly. 2005. "Conceptualizing and Categorizing Race and Ethnicity in Health Services Research." *Health Services Research* 40 (5 Pt II): 1658–75. doi:10.1111/j.1475-6773.2005.00449.x.
- Getahun, D., C. V. Ananth, and A. M. Vintzileos. 2006. "Uteroplacental Bleeding Disorders during Pregnancy: Do Missing Paternal Characteristics Influence Risk?" *BMC Pregnancy and Childbirth* 6 (1): 2. doi:10.1186/1471-2393-6-2.
- Hahn, R. A., and D. F. Stroup. 1994. "Race and Ethnicity in Public Health Surveillance: Criteria for the Scientific Use of Social Categories." *Public Health Reports* 109 (1): 7–15.
- IOM (Institute of Medicine). 2009. *Race, Ethnicity, and Language Data: Standardization for Health Care Quality Improvement*. Consensus Report. Washington, DC: National Academies Press.
- Johnson, K. M., and D. T. Lichter. 2010. "Growing Diversity among America's Children and Youth: Spatial and Temporal Dimensions." *Population and Development Review* 36 (1): 151–76. doi:10.1111/j.1728-4457.2010.00322.x.
- Liebler, C. A. 2010. "Homelands and Indigenous Identities in a Multiracial Era." *Social Science Research* 39 (4): 596–609. doi:10.1016/j.ssresearch.2010.02.003.
- Ma, S. 2008. "Paternal Race/Ethnicity and Birth Outcomes." *American Journal of Public Health* 98 (12): 2285–92. doi:10.2105/AJPH.2007.117127.
- Marks, E. L., B. B. Rhodes, and S. Scheffler. 2008. "*SEED for Oklahoma Kids: Baseline Analysis*." Report. Research Triangle Park, NC: RTI International.
- Martin, J. A., B. E. Hamilton, S. J. Ventura, M. J. K. Osterman, S. Kirmeyer, T. J. Mathews, and E. C. Wilson. 2011. "Births: Final Data for 2009." National Vital Statistics Reports Vol. 60, no. 1 [accessed January 2, 2013]. Available at http://www.cdc.gov/nchs/data/nvsr/nvsr60/nvsr60_01.pdf.
- Mays, V. M., N. A. Ponce, D. L. Washington, and S. D. Cochran. 2003. "Classification of Race and Ethnicity: Implications for Public Health." *Annual Review of Public Health* 24: 83–110. doi:10.1146/annurev.publhealth.24.100901.140927.
- Nam, Y., L. R. Mason, Y. Kim, M. Clancy, and M. Sherraden. 2013. "Survey Response in a Statewide Social Experiment: Differences in Being Located and Collaborating by Race and Hispanic Origin." *Social Work Research* 37 (1): 64–74.
- NCHS (National Center for Health Statistics). 2005. "Technical Appendix." In *Vital Statistics of the United States: 2003, vol. 1: Natality*. Hyattsville, MD: National

- Center for Health Statistics [accessed January 2, 2013]. Available at http://www.cdc.gov/nchs/data/TechApp03_1-09.pdf.
- Northam, S., S. Polancich, and E. Restrepo. 2003. "Birth Certificate Methods in Five Hospitals." *Public Health Nursing* 20 (4): 318–27. doi:10.1046/j.1525-1446.2003.20409.x.
- OMB (Office of Management and Budget). 1997. "Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity." *Federal Register* 62 (210 October 30): 58781–90.
- Perez, A. D., and C. Hirschman. 2009. "The Changing Racial and Ethnic Composition of the US Population: Emerging American Identities." *Population and Development Review* 35 (1): 1–51. doi:10.1111/j.1728-4457.2009.00260.x.
- Piper, J. M., E. F. Mitchel Jr, M. Snowden, C. Hall, M. Adams, and P. Taylor. 1993. "Validation of 1989 Tennessee Birth Certificates Using Maternal and Newborn Hospital Records." *American Journal of Epidemiology* 137 (7): 758–68.
- Qian, Z. 2004. "Options: Racial/Ethnic Identification of Children of Intermarried Couples." *Social Science Quarterly* 85 (3): 746–66. doi:10.1111/j.0038-4941.2004.00243.x.
- Ramirez, M., M. E. Ford, A. L. Stewart, and J. A. Teresi. 2005. "Measurement Issues in Health Disparities Research." *Health Services Research* 40 (5 Part II): 1640–57. doi:10.1111/j.1475-6773.2005.00450.x.
- Reichman, N. E., and E. M. Hade. 2001. "Validation of Birth Certificate Data: A Study of Women in New Jersey's HealthStart Program." *Annals of Epidemiology* 11 (3): 186–93. doi:10.1016/S1047-2797(00)00209-X.
- Reichman, N. E., and O. Schwartz-Soicher. 2007. "Accuracy of Birth Certificate Data by Risk Factors and Outcomes: Analysis of Data from New Jersey." *American Journal of Obstetrics and Gynecology* 197 (1): 32.e1–8. doi:10.1016/j.ajog.2007.02.026.
- Saenz, R., S. Hwang, B. E. Aguirre, and R. N. Anderson. 1995. "Persistence and Change in Asian Identity among Children of Intermarried Couples." *Sociological Perspectives* 38 (2): 175–94. doi:10.2307/1389289.
- Schoendorf, K. C., and A. M. Branum. 2006. "The Use of United States Vital Statistics in Perinatal and Obstetric Research." *American Journal of Obstetrics and Gynecology* 194 (4): 911–5. doi:10.1016/j.ajog.2005.11.020.
- Smith, N., R. L. Iyer, A. Langer-Gould, D. T. Getahun, D. Strickland, S. J. Jacobsen, W. Chen, S. F. Derose, and C. Koebnick. 2010. "Health Plan Administrative Records versus Birth Certificate Records: Quality of Race and Ethnicity Information in Children." *BMC Health Services Research* 10 (November): 316. doi:10.1186/1472-6963-10-316.
- Vinikoor, L. C., L. C. Messer, B. A. Laraia, and J. S. Kaufman. 2010. "Reliability of Variables on the North Carolina Birth Certificate: A Comparison with Directly Queried Values from a Cohort Study." *Paediatric and Perinatal Epidemiology* 24 (1): 102–12. doi:10.1111/j.1365-3016.2009.01087.x.
- Xie, Y., and K. Goyette. 1997. "The Racial Identification of Biracial Children with One Asian Parent: Evidence from the 1990 Census." *Social Forces* 76 (2): 547–70. doi:10.1093/sf/76.2.547.

- Zager, R., Y. Kim, Y. Nam, M. Clancy, and M. Sherraden. 2010. "The SEED for Oklahoma Kids Experiment: Initial Account Opening and Savings." CSD Research Report 10–14 [accessed on January 7, 2013]. Available at <http://csd.wustl.edu/Publications/Documents/RP10-14.pdf>.
- Zaslavsky, A. M., J. Z. Ayanian, and L. B. Zaborski. 2012. "The Validity of Race and Ethnicity in Enrollment Data for Medicare Beneficiaries." *Health Services Research* 47 (3 Part II): 1300–21. doi:10.1111/j.1475-6773.2012.01411.x.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

Appendix SA1: Author Matrix.