Using Bayesian Imputation to Assess Racial and Ethnic Disparities in Pediatric Performance Measures

By David P. Brown, Caprice Knapp, Meggen Kaufmann, and Kimberly Baker

Abstract

Objective. To analyze health care disparities in pediatric quality of care measures and determine the impact of data imputation.

Data Sources. Five HEDIS measures are calculated based on 2012 administrative data for 145,652 children in two public insurance programs in Florida.

Methods. The Bayesian Improved Surname and Geocoding (BISG) imputation method is used to impute missing race and ethnicity data for 42% of the sample (61,954 children). Models are estimated with and without the imputed race and ethnicity data.

Principal Findings. Dropping individuals with missing race and ethnicity data biases quality of care measures for minorities downward relative to non-minority children for several measures.

Conclusions. These results provide further support for the importance of appropriately accounting for missing race and ethnicity data through imputation methods.

Key Words. HEDIS, Disparities, Bayesian Imputation, Racial/Ethnic differences, Medicaid, Children

Introduction

In recent years there has been an increased emphasis on the importance on assessing quality of health care in pediatrics. In 2009 the Children's Health Insurance Program Reauthorization Act (CHIPRA) mandated the creation and refinement of a core set of measures, the creation of a Pediatric Quality Measures Program, and commissioned an annual report to Congress of the quality of care for children with Medicaid and CHIP coverage (Centers for Medicare and Medicaid Services 2009). Ultimately, the measures were designed to estimate quality at both the state and national level, conduct comparative analyses, and identify disparities (Centers for Medicare and Medicaid Services 2011).

A health care disparity is defined as "a particular type of health difference that is closely linked with social, economic, and/or environmental disadvantage" (U.S. Department of Health and Human Services 2008). While there are standards by which overall healthcare quality can be measured, it can often be difficult to easily and accurately identify health care disparities. This is in part due to the self-reported nature of race and ethnicity data resulting in missing data (Agency for Healthcare Research and Quality 2013). There have been attempts to improve the collection of self-reported race and ethnicity data nationwide but they have been met with limited success (Boston Public Health Commission 2006; National Health Plan Collaborative 2006; Bilheimer and Sisk 2008). Other attempts have been made to impute the missing data using geocoding or surname analysis (Fremont et al. 2005; Fiscella and Fremont 2006). A hybrid approach uses Bayesian methods to combine information on both an individual's surname and location using geocoding (Elliott et al. 2008; Elliott et al. 2009). These Bayesian methods have been shown to outperform the prior imputation methodologies (Elliott et al. 2009; Adjaye-Gbewonyo et al. 2014). However, this imputation methodology has not been widely used to analyze pediatric health care disparities.

Several studies have sought to investigate racial and ethnic disparities in the pediatric population using nationally recognized measures and large scale databases. Dougherty and colleagues showed that for more than 50 measures child and adolescent disparities have virtually remained unchanged over time (Dougherty et al. 2014). Shi and colleagues investigated racial disparities using data from the

Community Tracking Study (Shi, Green, and Kazakova 2004). Results from the study suggested that higher quality primary care was associated with fewer or decreased disparities. Although these studies are insightful, they did not comment on the accuracy of the race and ethnicity data used.

Our study makes two novel contributions to the literature: 1) we investigate plan-level disparities over an array of quality measures and 2) we use improved methods to impute missing data and determine what effect the missing data has on the measures. Results from our study demonstrate how improved accuracy can lead to better informed policies and interventions.

Methods

Study Sample

Data for this study came from the 2012 Healthcare Effectiveness Data and Information Set (HEDIS) measures that were reported for children in the Florida Healthy Kids (FHK) program and Children's Medical Services Network (CMSN). FHK serves healthy children ages 5 to 18 who are eligible for the Children's Health Insurance Program (CHIP). CMSN is Florida's Title V program for children with special health care needs. At the time of the study, CMSN was in the process of procuring a statewide Third Party Administrator to collect and process claims, encounter, and enrollment data for its Title XXI population. Only CMSN children on Medicaid were included.

The combined dataset contained 31,343 children in CMSN and 114,309 in FHK (total =145,652).

Quality Measures

The HEDIS measures used in this study are all part of the CHIPRA core measure set (Centers for Medicare and Medicaid Services 2013). The current analysis focuses on: (1) Child and Adolescent Access to Primary Care Practitioners (CAP); (2) Appropriate Testing of Children with Pharyngitis (CWP); (3) Well-Child Visits in the 3^{rd}- 6^{th} Years of Life (W34); (4) Adolescent Well-Care Visit (AWC); and (5) Follow-Up Care for Children Prescribed ADHD Medication (ADD). These measures were chosen because they represent a variety of health care scenarios including availability of care and preventive, acute, and chronic care. The HEDIS measures take on values of zero and one; a one indicates that a child received the care as specified in the HEDIS criterion.

The HEDIS measures used in this study required 12 months of continuous enrollment. Although HEDIS allows for supplementation of the administrative data with data from medical records for W34 and AWC, medical record data was not used.

Predictor Variables

The multivariate analyses used in the current study adjust for child's age (1 to 21), gender, race and ethnicity, geographic location (8 regions), health plan (9 plans; CMSN plus 8 FHK plans), and Clinical Risk Groups (CRG) (Neff et al. 2004). CRG is a categorical variable that classifies an individual's burden of illness using a combination of diagnoses and procedure codes into one of six categories: Non-Chronic Non-Acute, Significant Acute, Minor Chronic, Moderate Chronic, or Major Chronic. A residual category labeled as Unassigned is given to children who did not have adequate diagnoses to be categorized.

Imputation of Race and Ethnicity

Parents seeking to enroll their children in these programs are asked to fill out an application. The choices available for ethnicity are Hispanic or Not Hispanic and Latino or Not Latino. Choices for race are American Indian/Alaskan Native, Asian or Pacific Islander, Non-Hispanic Black, Non-Hispanic White, Hispanic, or Other. These fields are voluntary. If a parent reports the child's race and ethnicity, we define this as self-reported.

Once the race and ethnicity data are collected through the application process, the data is compiled. The enrollment file has five mutually exclusive categories: White, Black, Hispanic, Asian/Pacific Islander (API), and Other.

For individuals who did not report race or ethnicity, the BISG imputation methodology was used to estimate the probability that an individual is one of five race and ethnicity categories. The BISG imputation algorithm (Elliott et al. 2009) uses a Bayesian method to combine information on both the last name (surname) and geographical location using geocoding. Geocoding links a member's address to census data that contains the neighborhood's race and ethnicity makeup (Fremont et al. 2005). The surname approach infers race and ethnicity from last names (Abrahamse et al. 1994).

Analyses

Descriptive and multivariate analyses were performed. In the multivariate analyses, HEDIS performance was the outcome measure. In the models the reference groups for race and ethnicity, health plan, and CRG status were White, CMSN, and Non-Acute Non-Chronic.

Two sets of multivariate analyses were conducted. First, models were estimated using the BISG imputed race and ethnicity data. The predicted probabilities generated by the BISG algorithm are then used in place of the missing data. This approach has been shown to outperform other classification methods which classify probabilistic information about race and ethnicity into dichotomous variables (McCaffrey and Elliott 2007). Self-reported race and ethnicities were used when available. Second, the same analyses were performed, but individuals with missing data were dropped.

To interpret the impact that dropping the data with missing race and ethnicity has on the coefficient estimates, the percentage change in the coefficient estimates was calculated by:

$$\frac{\beta_{ij}^{DROPPED} - \beta_{ij}^{BISG}}{\beta_{ij}^{BISG}}$$

for each $i \in \{Black, Hispanic, Asian/\ API, Other\}$ and $j \in \{ADD, AWC, CAP, CWP, W34\}$.

Results

Summary Statistics

Table 1 provides the summary statistics. The p-values revealed that the children with and without reported race and ethnicity substantially differed. For example, children who self-reported had lower performance on the ADD measure, were older, less likely to be in CMSN, more likely to be located in a northern region, and more likely to be characterized in a more severe CRG.

HEDIS Analysis

There were 147,568 children in the sample with 63,870 observations with missing race and ethnicity data. The BISG imputation methodology estimated the race and ethnicity probabilities for over 97% (61,954 observations) of those children. The remaining 3% (1,916 observations) were dropped from the analysis. Table 2 provides the linear regression results for the HEDIS analysis when the BISG imputed race and ethnicities were included. For brevity, the logistic regressions are not included in the paper but are available upon request. The conclusions in the logistic analysis are analogous.

The effect of an individual's age, gender, region, health status, and race and ethnicity varied across each HEDIS measure. However, several notable trends arose. For AWC and W34, being White had a negative and statistically significant lower probability of receiving the specified care as compared to being Hispanic, Black, or API. The effect ranges from 2.8% to 5.2% for AWC and 2.2% to 9.1% for W34. For CWP, being Hispanic or Black had a negative and statistically significant lower probability of receiving the specified care compared to being White. Being Hispanic or Black results in a 5.2% to 5.9% lower likelihood of receiving the specified care with CWP compared to being White.

For each HEDIS measure, the analysis was performed again using only those individuals who self-reported their race and ethnicity. Table 3 presents these results.

Dropping the individuals who did not self-report race had several impacts on the coefficient estimates. Table 4 summarizes the percentage change in the race and ethnicity coefficients. While the effects vary by HEDIS measure, Table 4 reveals that dropping individuals who did not self-report race and ethnicity has the potential to substantially impact these estimates. For example, for W34 the positive and statistically significant BISG race and ethnicity coefficient estimates on Black, Hispanic, and Asian were reduced downward towards zero. The magnitude of this effect ranged from a 2.7% to 25.6% reduction. Alternatively, for CWP the BISG coefficient estimates on Black, Hispanic, and Other were negative and statistically significant. Dropping those who did not self-report race and ethnicity amplifies these negative coefficient estimates leading one to potentially overestimate the negative impact of being non-White on the probability of receiving the specified care for the CWP measure. The magnitude of this effect ranged from 1.7% to 36.1%. In addition, the BISG race and ethnicity coefficients on the ADD and CAP measures for Black, Hispanic, and other are biased downwards when missing race and ethnicity data are dropped. However, the race and ethnicity coefficients are sometimes small and statistically insignificant.

Finally, dropping those who did not self-report race and ethnicity also affected the statistical significance of several coefficients. In particular, the coefficient estimates became either less statistically significant or lost their statistical significance completely for all but the ADD measure.

Discussion

Identifying racial and ethnic disparities in children's health care quality is important. Without an understanding of where these disparities occur in the health care delivery process, it is difficult to make decisions on how to allocate funds and prioritize initiatives. Inaccurate information could lead to wasted time and funds.

Our findings corroborate others that have identified disparities in pediatric health care quality measures such as the previously mentioned Dougherty and Shi studies. A study by Berdahl and colleagues similarly found racial disparities in pediatric quality measures as assessed by two large datasets: the Medical Expenditure Panel Survey and the Healthcare Cost and Utilization Project (Berdahl et al. 2010), particularly for Hispanics. To our knowledge, none of these studies assessed the completeness of the race or ethnicity data nor imputed missing data. Ignoring the missing race and ethnicity data has an effect on the identification and magnitude of disparities.

Our results reveal for all measures, the race and ethnicity coefficient estimates are often biased downwards either towards zero or more negative when data with missing race and ethnicity is dropped. Further, all but one measure (ADD) saw a decrease or total loss of significance in at least one race and ethnicity coefficient estimate. The percentage change in the coefficient estimates when the missing data is dropped ranged from about 2 to 200 percent for Asian, 22 to 108 percent for Black, 7 to 112 for Other, and 2 to 84 for Hispanic. These findings imply that failure to impute missing race and ethnicity data might lead researchers and policy makers to under-identify the presence of health care disparities. Alternatively, when missing data is dropped, the negative race and ethnicity coefficient estimates may be artificially inflated. These two confounding results make it difficult to disentangle when health care disparities are present and their underlying magnitude, highlighting the importance of imputation.

Given this data is at the health plan level, it is important to discuss what a health plan could infer from the findings. If health plans wish, or are mandated, to address health care disparities they could focusing on improving the data by increasing the number of enrollees who self-report and using imputation. A 2013 national report provides instructions, case studies, and resources on how to collect race, ethnicity, and language data so that it can be used in a meaningful way (Health Research & Educational Trust 2013). Regardless of whether or not health plans desire to allocate resources towards increasing self-report, they should consider the BISG imputation. This would allow for a better understand of enrollees including where they get care, improved resource planning, more accurate targeting of subgroups for improvement, and more targeted marketing.

Our study has implications at the policy and programmatic level. We demonstrated that a significant number of children were missing race and ethnicity data. Our findings are consistent with Medicaid health plans in other states (Michigan Department of Community Health 2012). Yet, without clear direction and incentives from policymakers this problem will persist. The National Committee for Quality Assurance does have a measure called Race/Ethnicity Diversity of Membership (NCQA 2011). Specifications call for plans to report on the percentage of their members by racial and ethnic category. Having health plans report on this measure is a first step, but there should be recommendations to address missing data. Health plans could be incentivized for improving the accuracy of their race and ethnicity data, the diversity of their membership, and the ability to use evidence based strategies to reduce disparities.

There are several limitations in this study. First, we focused on six HEDIS measures when many other pediatric quality measures exist. Second, we made no attempt to validate the race and ethnic data against birth vital statistics or medical records. In practice this validation is often not feasible and self-reported data are considered a gold standard. Third, there are other factors that could contribute to quality of care that were omitted. Fourth, the 12-month inclusion criterion may introduce bias if some children churn in and out of these programs and their performance is not captured. If this churning interrupts the continuity of care, it may have a negative effect on quality. Not including these children may bias the HEDIS estimates upwards. Fifth, there is some evidence which suggests that using BISG in regional samples may not take full advantage of the macro geographical variation (Elliot 2009). Finally, the health care disparities results may not be generalizable to other states. Although, we conjecture that the importance of appropriately accounting for missing race and ethnicity data will generalize.

Despite these limitations our study used a novel method, and we demonstrated how failure to address missing data could overestimate disparities.

Table 1. Summary Statistics and Comparisons by Availability of Race and Ethnicity Information

Variables	All (n=145,652)	Self-Reported (n=83,698)	Missing (n= 61,954)	p-value
HEDIS Measures				
ADD	37.4	39.9	32.8	<0.01
AWC	54.5	54.4	54.8	0.204
CAP	92.0	92.1	91.9	0.203
CWP	65.8	65.9	65.6	0.771
W34	71.0	70.7	71.5	0.319
Clinical Risk Groups				
Non-Chronic, Non-Acute	62.3	63.4	60.8	<0.01
Sign. Acute	7.6	7.8	7.3	<0.01
Minor Chronic	6.6	9.0	8.1	<0.01
Moderate Chronic	13.4	13.3	13.5	0.259
Major Chronic	7.1	5.3	9.5	<0.01
Unassigned	1.0	1.2	0.8	<0.01
Health Plan				
CMSN	21.4	21.9	20.7	<0.01
Plan A	28.3	27.9	28.8	0.01
Plan B	2.3	2.6	2.0	<0.01
Plan C	9.4	8.6	10.5	<0.01
Plan D	1.9	1.7	2.1	0.01
Plan E	0.4	0.5	0.3	<0.01
Plan F	14.2	15.0	13.1	<0.01
Plan G	22.1	21.8	22.4	0.01
Demographics				
Age	11.1	11.0	11.4	<0.01
Male	51.8	51.3	52.5	<0.01
Northwest	4.1	4.5	3.6	<0.01
Big Bend	1.8	1.9	1.5	<0.01
North Central	15.1	16.0	13.8	<0.01
Tampa Bay	15.1	15.4	14.8	<0.01
Central	11.9	12.0	11.8	0.136
Southwest	9.9	10.3	9.3	<0.01
Southeast	23.6	22.7	24.7	0.01
South	18.6	17.2	20.6	<0.01

Note: For all variables except age, the numbers represent percentages because they are categorical variables. Age is continuous and varies between 1 and 21.

Table 2. Linear Regression Results with BISG Race and Ethnicity Data

	ADD Performance	AWC Performance	CAP Performance	CWP Performance	W34 Performance
Variables	(1)	(2)	(3)	(4)	(5)
Demographics					
Age	-0.0032 (0.0788)	0.0016 (0.0116)	0.0098 *** (0.0007)	0.0352 *** (0.0069)	0.1314 *** (0.0301)
Age Squared	-0.0010 (0.0042)	-0.0012 *** (0.0004)	-0.0005 *** (0.0003)	-0.0019 *** (0.0003)	-0.0226 *** (0.0042)
Male	-0.0687 *** (0.0243)	-0.0206 *** (0.0033)	-0.0086 *** (0.0016)	-0.0057 (0.0097)	0.0061 (0.0074)
Black	-0.0190 (0.0414)	0.0292 *** (0.0056)	-0.0015 (0.0026)	-0.0516 *** (0.0183)	0.0223 ** (0.0109)
Hispanic	0.0206 (0.0205)	0.0279 *** (0.0042)	0.0082 *** (0.0021)	-0.0589 *** (0.0118)	0.0337 *** (0.0096)
Asian/API	0.0331 (0.1491)	0.0519 *** (0.0149)	0.0042 (0.0073)	-0.0154 (0.0449)	0.0911 *** (0.0355)
Other	0.1188 (0.1491)	0.0440 (0.0330)	-0.0092 (0.0198)	-0.2655 *** (0.0889)	-0.0353 (0.0831)
Clinical Risk Groups					
Sign. Acute	-0.0648 (0.0724)	0.1010 *** (0.0062)	0.0736 *** (0.0024)	0.0218 (0.0184)	0.0762 *** (0.0139)
Minor Chronic	0.0858 *** (0.0279)	0.0834 *** (0.0059)	0.0736 *** (0.0002)	0.0086 (0.0178)	0.0883 *** (0.0147)
Moderate Chronic	0.1525 *** (0.0376	0.0886 *** (0.0054)	0.0755 *** (0.0019)	0.0089 (0.0154)	0.0853 *** (0.0101)
Major Chronic	0.2324 *** (0.0674)	0.0719 *** (0.0082)	0.0835 *** (0.0019)	0.0097 (0.0231)	0.0579 *** (0.0119)
Unassigned	0.0351 (0.0338)	-0.0267 (0.0458)	-0.0543 (0.0425)	-0.0307 * (0.0166)	-0.3756 ** (0.1642)
Health Plan					
Plan A	-0.1169 ** (0.0536)	-0.0145 ** (0.0062)	-0.0469 *** (0.0025)	0.0608 *** (0.0161)	-0.0412 *** (0.0148)
Plan B	-0.1084 (0.0824)	-0.1240 *** (0.0121)	-0.0482 *** (0.0063)	-0.1056 *** (0.0290)	-0.2175 *** (0.0376)
Plan C	-0.0938 (0.0649)	-0.0095 (0.0076)	-0.0323 *** (0.0034)	0.0929 *** (0.0199)	-0.0543*** (0.0419)
Plan D	-0.1481 (0.1038)	-0.0102 (0.0134)	-0.0158 ** (0.0063)	0.0883 ** (0.0435)	-0.0506 (0.0419)
Plan E	-0.3876 *** (0.0608)	-0.0202 (0.0256)	-0.0377 ** (0.0145)	0.0983 (0.0737)	0.0106 (0.0687)
Plan F	-0.0630 (0.0548)	-0.0191 *** (0.0068)	-0.0175 *** (0.0028)	0.0183 (0.0205)	-0.0349 ** (0.0163)
Plan G	-0.1137 ** (0.0522)	-0.0955 *** (0.0064)	-0.0618 *** (0.0027)	0.0392 ** (0.0166)	-0.0961 *** (0.0159)
Regional Dummies	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
No. of Observations	1933	87,833	111,641	9,411	14,705

Note: Stars indicate statistical significance where * if p≤0.10, ** if p≤0.05, and *** if p≤0.01. Values in the parentheses are robust standard errors. All models also include location (Region) covariates.

Table 3. Linear Regression Results when Missing Race and Ethnicity Data is Dropped

	ADD Performance	AWC Performance	CAP Performance	CWP Performance	W34 Performance
Variables	(1)	(2)	(3)	(4)	(5)
Demographics					
Age	0.0486 (0.0997)	0.0152 (0.0154)	0.0111 *** (0.0009)	0.0478 *** (0.0087)	0.1459 *** (0.0398)
Age Squared	-0.0039 (0.0053)	-0.0018 *** (0.0005)	-0.0006 *** (0.0001)	-0.0026 *** (0.0004)	-0.0236 *** (0.0056)
Male	-0.0768 ** (0.0306)	-0.0288 *** (0.0044)	-0.0089 *** (0.0022)	-0.0047 (0.0122)	-0.0029 (0.0095)
Black	-0.0395 (0.0462)	0.0357 * (0.0186)	-0.0019 (0.0030)	-0.0702 * (0.0420)	0.0166 (0.0130)
Hispanic	0.0033 (0.0313)	0.0384 *** (0.0053)	0.0061 ** (0.0026)	-0.0599 *** (0.0139)	0.0237 *** (0.0119)
Asian/API	0.0129 (0.1566)	0.0648 *** (0.0168)	0.0039 (0.0081)	-0.0478 (0.0500)	0.0886 ** (0.0381)
Other	0.1005 (0.3767)	0.0384 (0.0346)	-0.0187 (0.0213)	-0.2851 *** (0.0925)	-0.0748 (0.0871)
Clinical Risk Groups					
Sign. Acute	-0.2112 ** (0.0837)	0.1008 *** (0.0084)	0.0738 *** (0.0030)	0.0394 * (0.0228)	0.0637 *** (0.0169)
Minor Chronic	0.0598 * (0.0361)	0.0912 *** (0.0078)	0.0725 *** (0.0029)	0.0291 (0.0221)	0.0984 *** (0.0177)
Moderate Chronic	0.1038 ** (0.0473)	0.0922 *** (0.0072)	0.0728 *** (0.0025)	0.0212 (0.0189)	0.0876 *** (0.0124)
Major Chronic	0.1586 * (0.0851)	0.0662 *** (0.0110)	0.0793 *** (0.0025)	0.0090 (0.0320)	0.0589 *** (0.0177)
Unassigned	0.0248 (0.0425)	-0.0511 (0.0582)	-0.0713 (0.0628)	-0.0355 * (0.0204)	-0.1132 (0.3313)
Health Plan					
Plan A	-0.1411 ** (0.0697)	-0.0398 *** (0.0078)	-0.0582 *** (0.0032)	0.0576 *** (0.0195)	-0.0784 *** (0.0186)
Plan B	-0.1598 (0.0999)	-0.1629 *** (0.0154)	-0.0559 *** (0.0081)	-0.0992 *** (0.0347)	-0.2288 *** (0.0451)
Plan C	-0.1309 (0.0846)	-0.0420 *** (0.0101)	-0.0372 *** (0.0045)	0.0879 *** (0.0249)	-0.0875*** (0.0261)
Plan D	-0.2812 ** (0.1283)	-0.0295 (0.0183)	-0.0298 *** (0.0093)	0.1463 ** (0.0579)	-0.1193 ** (0.0556)
Plan E	-0.4448 *** (0.0791)	-0.0418 (0.0320)	-0.0529 *** (0.0181)	0.0934 (0.0933)	-0.0008 (0.0763)
Plan F	-0.1061 (0.0709)	-0.0420 *** (0.0085)	-0.0268 *** (0.0036)	0.0165 (0.0248)	-0.0599 *** (0.0198)
Plan G	-0.1191 * (0.0677)	-0.1092 *** (0.0080)	-0.0660 *** (0.0035)	0.0639 *** (0.0199)	-0.1297 *** (0.0197)
Regional Dummies	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
No. of Observations	1257	48,885	61,046	5,922	8,874

Note: Stars indicate statistical significance where * if p≤0.10, ** if p≤0.05, and *** if p≤0.01. Values in the parentheses are robust standard errors. All models also include location (Region) covariates.

Table 4. Percent Change in Coefficient Estimates Dropping Those with Missing Race and Ethnicity Data

Variables	ADD Performance	AWC Performance	CAP Performance	CWP Performance	W34 Performance
Black	107.9 ↓	22.3 †	26.7 ↓	36.1 ↓ †	-25.6 ↓ ††
Hispanic	-84.0 ↓	37.6	-25.6 ↓ †	1.7 ↓	-29.7 ↓
Asian/API	-61.0 ↓	24.9	2.4	210.4 ↓	-2.7 ↓ †
Other	-15.4 ↓	-12.7 ↓	103.3 ↓	7.4 ↓	111.9 ↓

Note: The numbers reflect the percent change in the coefficient estimates. ↓ reflects a reduction in the coefficient estimates either towards zero (if coefficient is positive) or more negative (if coefficient is negative). † reflects a reduction in the degree of statistical significance. †† reflects a loss of statistical significance.

References

Abrahamse, A., P. Morrison, and N. Bolton. 1994. "Surname analysis for estimating locate concentration of Hispanics and Asians." *Population Research and Policy Review* 13(4): 383-398.

Adjaye-Gbewonyo, D., R. Bednarczyk, R. Davis, and S. Omer. 2014. "Using the Bayesian Improved Surname Geocoding (BISG) to Create Working Classifications of Race and Ethnicity in a Diverse Managed Care Population: A Validation Study." *Health Services Research* 49(1): 268-283.

Agency for Healthcare Research and Quality. 2013. "2012 National Healthcare Disparities Report" [accessed on April 21, 2014]. Available at: http://www.ahrq.gov/research/findings/nhqrdr/nhdr12/index.html

Berdahl, T., P. Owens, D. Dougherty, M. McCormick, Y. Pylypchuk, and L. Simpson. 2010. "Annual Report on Health Care for Children and Youth in the United States: Racial/Ethnic and Socioeconomic Disparities in Chilren's Health Care Quality." *Academic Pediatrics* 10(2): 95-118.

Bilheimer, L. and J. Sisk. 2008. "Collecting Adequate Data on Racial and Ethnic Disparities in Health: The Challenges Continue." *Health Affairs* 27(2): 383-391.

Boston Public Health Commission. 2006. *Data Collection Regulation*. Boston Public Health Commission.

Centers for Medicare and Medicaid Services. 2009. "CHIPRA Initial Core Set of Children's Health Care Quality Measures" [accessed on April 21, 2014]. Available at: http://www.medicaid.gov/Medicaid-CHIP-

Program-Information/By-Topics/Quality-of-Care/CHIPRA-Initial-Core-Set-of-Childrens-Health-Care-Quality-Measures.html

Centers for Medicare and Medicaid Services. 2011. “Initial Core Set of Children’s Health Care Quality Measures: Technical Specifications and Resource Manual for Federal Fiscal Year 2011 Reporting” [accessed on April 21, 2014]. Available at: http://www.medicaid.gov/mltss/docs/CHIPRACoreSetTechManual-1.pdf

Centers for Medicare and Medicaid Services. 2013. “Core Set of Children’s Health Care Quality Measures for Medicaid and CHIP (Child Core Set): Technical Specifications and Resource Manual for Federal Fiscal Year 2013” [accessed on April 21, 2014]. Available at: http://www.medicaid.gov/Medicaid-CHIP-Program-Information/By-Topics/Quality-of-Care/Downloads/Medicaid-and-CHIP-Child-Core-Set-Manual.pdf

Dougherty, D., X. Chen, D. Gray, and A. Simon. 2014. “Child and Adolescent Health Care Quality and Disparities: Are We Making Progress?” *Academic Pediatrics* 14(2): 137- 148.

Elliott, M., A. Fremont, P. Morrison, P. Pantoja, and N. Lurie. 2008. “A New Method for Estimating Race and Ethnicity and Associated Disparities Where Administrative Records Lack Self-Reported Race and ethnicity.” *Health Services Research* 43(5): 1722-1736.

Elliott, M., P. Morrison, A. Fremont, D. McCaffrey, P. Pantoja, and N. Lurie. 2009. “Using the Census Bureau’s Surname List to Improve Estimates of Race and ethnicity and Associated Disparities.” *Health Services and Outcomes Research Methodology* 9(2): 69-83.

Elliot, M. “Using Medicare CAHPS Data to Validate Indirect Estimation of Racial/Ethnic Disparities” AcademyHealth Annual Meeting. 27 June 2009.

Fiscella K. and A. Fremont. 2006. “Use of Geocoding and Surname Analysis to Estimate Race and Ethnicity.” *Health Services Research* 41(1): 1482-1500.

Florida Department of Health. 2008. “Children’s Medical Services Program Overview” [accessed on April 21, 2014]. Available at: www.floridahealth.gov/alternatesites/cms-kids/home/resources/documents/cms_overview_08.pdf

Fremont, A., A. Bierman, S. Wickstrom, C. Bird, M. Shah, J. Escarce, T. Horstman, and T. Rector. 2005. "Use of Geocoding in Managed Care Settings to Identify Quality Disparities." *Health Affairs* 24(2): 516-526.

Health Research & Educational Trust. 2013. "Reducing Health Care Disparities: Collection and Use of Race, Ethnicity and Language Data" [accessed on May 5, 2014]. Chicago: Health Research & Educational Trust. Available at http://www.hpoe.org/resources/hpoehretaha-guides/1431

Jacobs, E. and D. Lauderdale. 2001. "Receipt of Cancer Screening Procedures Among Hispanic and non-Hispanic Health Maintenance Organization Members." *Cancer* 91(1): 257-261.

McCaffrey, D., and M. Elliott. 2007. "Power of Tests for a Dichotomous Independent Variable Measured with Error." *Health Services Research* 43(3): 1085-1101.

Michigan Department of Community Health. 2012. "Michigan Medicaid HEDIS 2012 Results Statewide Aggregate Report" [accessed on May 6, 2014]. Available at: http://www.michigan.gov/documents/mdch/16623_MI_Medicaid_HEDIS_2012_statewide_aggregate_report_422984_7.pdf

National Health Plan Collaborative. 2006. *Phase 1 Summary Report: Reducing Racial and Ethnic Disparities and Improving Quality of Health Care*. Hamilton, NJ: National Health Plan Collaborative.

National Committee for Quality Assurance (NCQA). 2011. "NCQA's Multicultural Health Care Distinction Program" [accessed on May 6, 2014]. Available at: http://www.ncqa.org/Portals/0/Public%20Policy/NCQA%20MHC%20fact%20sheet.pdf

Neff, J., V. Sharp, J. Muldoon, J. Graham, and K. Myers. 2004. "Profile of Medical Charges for Children by Health Status Group and severity Level in a Washington State Health Plan." *Health Services Research* 39(1): 73-90.

Shi, L., E.L. Green, and S. Kazakova. 2004. "Primary Care Experience and Racial Disparities in Self-Reported Health Status." *Journal of the American Board of Family Practice* 17(6): 443-52.

U.S. Department of Health and Human Services. 2008. "The Secretary's Advisory Committee on National Health Promotion and Disease Prevention Objectives for 2020. Phase I report:

Recommendations for the framework and format of Healthy People 2020. Section IV. Advisory Committee findings and recommendations" [accessed on April 21, 2014]. Available at: http://healthypeople.gov/2020/about/advisory/PhaseI.pdf