Improving an Administrative Case Definition for Longitudinal Surveillance of Sickle Cell Disease

Angela B. Snyder, PhD, MPH1; Mei Zhou, MS1; Rodney Theodore, MPH2; Maa-Ohui Quarmyne, MD3; James Eckman, MD4; and Peter A. Lane, MD3

Abstract

Objective: Several states are building infrastructure and data collection methods for longitudinal, population-based surveillance systems for selected hemoglobinopathies. The objective of our study was to improve an administrative case definition for sickle cell disease (SCD) to aid in longitudinal surveillance.

Methods: We collected data from 3 administrative data sets (2004-2008) on 1998 patients aged 0-21 in Georgia who had ≥1 encounter in which an SCD International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) code was recorded, and we compared these data with data from a laboratory and medical record review. We assessed performance (sensitivity, specificity, positive predictive value [PPV], and negative predictive value [NPV]) of case definitions that differed by number and type of SCD-coded encounters; addition of SCD-associated treatments, procedures, and complications; and length of surveillance (1 vs 5 years). We identified correct diagnoses for patients who were incorrectly coded as having SCD.

Results: The SCD case definition of ≥3 SCD-coded encounters in 5 years simplified and substantially improved the sensitivity (96.0% vs 85.8%) and NPV (68.2% vs 38.2%) of the original administrative case definition developed for 5-year, state-based surveillance (2 encounters in 5 years and 1 encounter for an SCD-related treatment, procedure, or complication), while maintaining a similar PPV (97.4% vs 97.4%) and specificity (76.5% vs 79.0%).

Conclusions: This study supports an administrative case definition that specifies ≥3 ICD-9-CM–coded encounters to identify SCD with a high degree of accuracy in pediatric patients. This case definition can be used to help establish longitudinal SCD surveillance systems.

Keywords

surveillance, sickle cell disease, administrative data

Sickle cell disease (SCD) describes a group of genetic blood disorders characterized by chronic hemolytic anemia and various acute and chronic complications, including episodes of pain, strokes, and premature death. SCD severity and treatment vary, depending in part on whether a person has homozygous SCD (HbSS), compound heterozygous forms of SCD (eg, HbSC), or various forms of S β-thalassemia. SCD is a rare disease: according to state-based newborn screening data, SCD primarily affects persons of African ancestry, or approximately 1 in 360 black or African American births in the United States.1-4 However, the prevalence of SCD in the United States is not known because many persons affected were born outside of the United States or before implementation of universal newborn screening in 2006.5 Prevalence is likely increasing because of increased life expectancy among persons living with SCD.6,7 Since 2008, multiple stakeholders have identified the need for longitudinal, population-based surveillance of SCD to establish the true prevalence of the disease and define how current management and novel therapies improve patient outcomes and survival.8-10 From 2010 through 2012, the

1Georgia Health Policy Center, Andrew Young School of Policy Studies, Georgia State University, Atlanta, GA, USA
2Sickle Cell Disease Program, Children’s Healthcare of Atlanta, Atlanta, GA, USA
3Emory University School of Medicine & Sickle Cell Disease Program, Children’s Healthcare of Atlanta, Atlanta, GA, USA
4Departments of Hematology & Oncology, Emory University School of Medicine, Atlanta, GA, USA

Corresponding Author:
Angela B. Snyder, PhD, MPH, Georgia State University, Andrew Young School of Policy Studies, Department of Public Management and Policy & Georgia Health Policy Center, 55 Park Pl NE, 8th Fl, Atlanta, GA 30303, USA.
Email: angiesnyder@gsu.edu

ª2019, Association of Schools and Programs of Public Health
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/0033354919839072
journals.sagepub.com/home/phr
Centers for Disease Control and Prevention (CDC) and the National Heart, Lung, and Blood Institute funded 7 states (California, Florida, Georgia, Michigan, New York, North Carolina, and Pennsylvania) to participate in 5 years of SCD surveillance through the Registry and Surveillance System for Hemoglobinopathies (RuSH) project. Since 2016, California and Georgia have leveraged RuSH data and methods to expand the SCD surveillance period from 5 years of data (2004 through 2008) to 15 years of data (2004 through 2018) through the Sickle Cell Data Collection program. Determining the best case definition to use for expanded surveillance and informing surveillance systems developing in other states, such as Michigan and North Carolina, may improve surveillance.

The RuSH case definition for identification of SCD required ≥2 encounters in which a sickle cell diagnosis code was recorded in an administrative data set and ≥1 code from a predetermined list of SCD-associated treatments, procedures, and complications (Box). In 2009, a national workgroup of clinicians with expertise in SCD developed this list because no algorithms had been validated. Since completion of the RuSH project, at least 2 SCD validation studies have reported several high-performing case definitions. Using Medicaid claims data for children aged ≤18 with a diagnosis of SCD, Reeves et al developed and compared 37 alternative case definitions for the disease and found that the definition of ≥3 paid SCD claims within a 1-year period, regardless of the type of service, was a straightforward definition to implement and identify SCD in children with the most accuracy (sensitivity, 89.7%; positive predictive value [PPV], 95.8%). Michalik et al did not compare algorithms but found that another high-performing algorithm in the study by Reeves et al (≥1 inpatient SCD claim or 2 SCD outpatient visits at least 30 days apart) performed well (sensitivity, 99.4%; PPV, 99.4%) when electronic health record data from 2 hospital systems were used. However, neither study assessed the usefulness of adding SCD-associated treatments, procedures, and complications to the case definition or the effect of varying periods of surveillance on performance.

Our study further examined and compared the performance of the RuSH administrative case definition with elements from these newly reported, high-performing case definitions across varying periods of surveillance to recommend the optimal administrative case definition for future longitudinal SCD surveillance.

**Methods**

**Data Sources**

Georgia’s RuSH surveillance system collected claims data from 3 administrative data sets (Medicaid, the Children’s Health Insurance Program, and the State Health Benefit Plan) and Georgia hospital discharge data for patients who had ≥1 encounter in which an SCD *International Classification of Diseases, Ninth Revision, Clinical Modification* (ICD-9-CM) code (282.6, 282.60-282.64, 282.68-282.69, 282.41-282.42) was recorded in any position from January 1, 2004, through December 31, 2008. Each claim could contain as many as 10 diagnosis codes. Each data set had previously been de-duplicated and matched in the RuSH surveillance system to state newborn screening data and clinical data from 3 large SCD treatment centers in Georgia.

The RuSH definition for identifying confirmed cases included in the RuSH probable case definition based on a consensus panel of SCD experts, Atlanta, Georgia, 2009

**Box. Sickle cell disease (SCD)–associated treatments, procedures, and complications**

| Treatments                  |
|-----------------------------|
| Hydroxyurea                 |
| Parenteral analgesics (morphine, meperidine, hydromorphone, ketorolac, butorphanol) |
| Iron chelators (deferasirox, deferoxamine) |
| Erythropoietin              |
| Folic acid                  |

| Procedures                 |
|-----------------------------|
| Red cell transfusion        |
| Red cell exchange           |
| Splenectomy                 |
| Cholecystectomy             |
| Transcranial Doppler        |

| Complications               |
|-----------------------------|
| Chronic renal failure/proteinuria |
| Pneumonia, acute chest syndrome |
| Pulmonary hypertension       |
| Stroke (ischemic or hemorrhagic), transient ischemic attack, seizures |
| Intracranial bleeding        |
| Priapism                     |
| Iron overload                |
| Gallstones/cholelithiasis, cholecystitis |
| Avascular necrosis           |
| Retinal disease              |
| Splenomegaly, splenic sequestration, hypersplenism |
| Leg ulcers                   |
| Dactylitis                   |
| Osteomyelitis                |

*The RuSH definition of a probable case required ≥2 SCD *International Classification of Diseases, Ninth Revision, Clinical Modification* (ICD-9-CM) codes on ≥2 separate health care encounters plus ≥1 hemoglobinopathy-associated complication, treatment, or procedure from this list identified through *ICD*-9-CM codes, *Current Procedural Terminology* codes, or *National Drug Codes* but no laboratory-based disease confirmation available.
SCD or clinical data provided by the SCD treatment centers. The RuSH definition for identifying probable SCD cases in administrative data required \( \geq 2 \) encounters in which an SCD ICD-9-CM code was recorded plus \( \geq 1 \) encounter in which an ICD-9-CM code, Current Procedural Terminology codes, or National Drug Codes from a predetermined list of SCD-associated treatments, procedures, and complications was recorded (Box). In total, the Georgia RuSH project identified 4288 confirmed cases of SCD and 2721 probable cases of SCD from 2004 through 2008.

**Validation Sample**

We studied a cohort of 1998 patients aged 0-21 who (1) had \( \geq 1 \) health care encounter in which an SCD ICD-9-CM code was recorded in any of the 3 RuSH administrative data sets, (2) had \( \geq 1 \) encounter at Children’s Healthcare of Atlanta from 2004 through 2008, and (3) had a medical record available for review. Of these 1998 patients, 1560 were previously confirmed SCD cases, according to linkage with newborn screening confirmatory testing or clinical data from an SCD treatment center. For the 438 patients who were not confirmed cases, 1 of 2 pediatric hematologists (P.A.L. and M.O.Q.) with expertise in SCD reviewed Children’s Healthcare of Atlanta medical records to confirm case status. The pediatric hematologists confirmed the SCD diagnosis through a review of clinical and laboratory data, complete blood count, reticulocytes, and hemoglobin separation by electrophoresis, isoelectric focusing, and high-performance liquid chromatography. Criteria used to exclude the diagnosis of SCD were the absence of any documentation of SCD in provider notes and laboratory results inconsistent with SCD. Cases in which the data in medical records were inadequate to confirm or exclude SCD were categorized as indeterminate (Figure).

**Testing of Alternative Case Definitions**

We chose alternative case definitions based on the results of other published SCD validation studies that found that
simple counts of all SCD-related claims outperformed other algorithms. In our study, we counted the number of SCD-related encounters by using all medical claims data that contained an SCD-related ICD-9-CM code; we included inpatient and outpatient (emergency department, laboratory, and radiology) visits and excluded pharmacy visits. We counted consecutive inpatient claims as 1 encounter and similarly counted any outpatient claims that occurred during an inpatient episode as an encounter. We de-duplicated all encounters by patient identification, service date, and provider identification. Furthermore, we examined the additional treatment, procedure, and complication-related ICD-9-CM codes (Box) and length of the surveillance period (1 vs 5 years) to refine the RuSH case definition.

We examined alternative case definitions in 3 phases. First, we examined the effect of adding ≥1 SCD-associated treatment, procedure, or complication (Box) to the number of SCD-related medical encounters (1 to ≥6) during the 5-year surveillance period. We tabulated descriptive statistics, including the number of patients in which SCD was confirmed, excluded, or indeterminate, by number of encounters and stratified identified patients by the presence of ≥1 SCD-associated treatment, procedure, or complication defined by RuSH. We used the Mantel-Haenszel $\chi^2$ test for trend, with $P < .05$ considered significant. We reported true diagnoses for persons miscoded as SCD (excluded cases). We removed indeterminate cases from further analysis.

Second, we compared differences in the surveillance period (1 vs 5 years) and retained only 1 definition that included the additional SCD-related treatments, procedures, and complications—the RuSH case definition—because that addition did little to improve the identification of confirmed cases during phase 1. We compared the RuSH case definition with 11 alternative case definitions that included ≥2 to ≥6 encounters within the 5-year surveillance period and within any 1 year of the 5-year surveillance period. We defined the 1-year period on a rolling basis rather than as a calendar year. That is, for definitions requiring ≥2 SCD diagnosis codes within 1 year, the year was determined by the gap of time between ≥2 SCD encounters. We calculated sensitivity, specificity, PPV, and negative predictive value (NPV) for each case definition tested, using newborn screening, clinical, and medical record data as the reference standard. We calculated 95% confidence intervals (CIs) for these statistics and assumed significance for metrics that had nonoverlapping CIs.

Finally, to test whether including a single SCD ICD-9-CM code from an inpatient claim would improve performance of the preferred SCD case definition from phase 2 (≥3 encounters within 5 years), we subdivided the claims for patients with only 1 or 2 encounters within the 5-year surveillance period by whether or not ≥1 of the encounters was an inpatient admission. We conducted all analyses by using SAS version 9.4.22 We received institutional review board approval under a public health practice exemption from participating institutions (Georgia State University, Emory University, and Augusta University).

Results

Of 1998 cases initially identified, we confirmed SCD in 1763 (88.2%) cases, excluded SCD in 197 (9.9%) cases, and declared SCD indeterminate in 37 (1.9%) cases (Figure). Most indeterminate cases were patients who had few encounters in Children’s Healthcare of Atlanta medical records, with no mention or suspicion of SCD, and insufficient laboratory results to exclude the possibility of SCD. Of 197 patients miscoded for SCD by ICD-9-CM coding, 65 (33.0%) were carriers (usually HbAS or HbAC), 11 (5.6%) had non-sickle hemoglobin disorder (eg, HbCC), 9 (4.6%) had thalassemias, 18 (9.1%) had other nonmalignant hematologic disorder, 19 (9.6%) had malignancy, and 75 (38.1%) had other or unknown diagnosis.

The likelihood of having SCD significantly increased with an increasing number of SCD encounters ($P < .001$) (Table 1). Fifty of 105 (47.6%) patients with ≤2 SCD encounters during the 5-year period did not have SCD, whereas three-quarters of patients with ≥3 encounters had SCD. For patients with ≥3 encounters, the addition of ≥1 SCD-associated treatment, procedure, or complication did little to improve the accuracy of case identification, adding <50% of true cases, but did serve to exclude more than 60% of non-SCD cases. Of 1998 patients, 1607 (80.4%) had ≥6 encounters during the 5-year period.

The performance of case definitions applied within the 1-year rolling period was similar to the same encounter counts applied within the 5-year period (Table 2). Small differences favored the 5-year period, but these differences were significant only in the case definitions that included ≥4 encounters. For the 5-year period, the PPV of ≥2 encounters (94.8%; 95% CI, 93.6%-95.7%) increased to 97.4% (95% CI, 96.4%-98.1%) with the addition of ≥1 encounter with an SCD-associated treatment, procedure, or complication (RuSH case definition), but the number of missed cases increased from 28 (1.5%) to 251 (16.1%) SCD cases identified. The RuSH case definition substantially underperformed on both sensitivity (85.8%) and NPV (38.2%) compared with case definitions that included ≥3 SCD encounters (96.0% sensitivity, 68.2% NPV) and ≥4 SCD encounters (93.4% sensitivity, 58.7% NPV).

Upon further review of inpatient admissions for the 128 patients with 1 encounter, 9 had an SCD diagnosis recorded during an inpatient admission, 4 of whom had confirmed SCD. For the 92 patients with 2 encounters, 7 had an SCD diagnosis recorded during an inpatient admission, 4 of whom had confirmed SCD.

Discussion

Given the expansion of SCD surveillance to include 15 years or more of data, we sought to determine the best administrative case definition to use for longitudinal SCD surveillance. By taking elements from recent validation studies and examining different lengths of surveillance, we found an
administrative case definition that improves the accuracy of case identification and is straightforward to implement.

The preferred SCD case definition suggested by our analysis (≥3 encounters within a 5-year period) substantially improved the sensitivity (96.0% vs 85.8%) and NPV (68.2% vs 38.2%) for identifying SCD, while maintaining a PPV (97.4% vs 97.4%) and a specificity (76.5% vs 79.0%) that were similar to those resulting from use of the RuSH case definition. The higher specificity and lower sensitivity of administrative case definitions used to identify rare diseases, such as SCD. Because of a trade-off between sensitivity and specificity in choosing an administrative case definition for any condition, the optimal definition depends in part on the purpose for which the data are used. To define the burden of SCD for the purposes of public health resource allocation and planning, the number of cases identified should be maximized by choosing a definition that optimizes sensitivity. Conversely, for quality and outcome studies of SCD interventions, it is important to limit the number of non-SCD patients in the study sample by maximizing specificity. Thus, the performance measures we reported should be useful to health services planners and researchers who need to identify the best case definitions for their needs.

Findings from our study confirm and extend the results of recent studies that examined the validity of using administrative data to identify cases of SCD. Reeves et al. identified ≥3 SCD encounters within a 1-year period as the highest-performing and straightforward definition after testing 37 algorithms. The Reeves et al study reported a sensitivity of 89.7% and a specificity of 92.4%, compared with 96.0% sensitivity and 76.5% specificity in our study. Our lower specificity can be partially explained by including only patients with ≥1 SCD claim.

### Addition of SCD-Associated Codes

Like Reeves et al., who explored a wide array of services, pharmacy claims, and coding combinations in the tested algorithms, we examined the addition of SCD-associated treatments, procedures, and complications defined by the RuSH project to improve specificity. When we added treatments, procedures, and complications to the case definition of ≥2 SCD encounters, specificity increased to 79.0% but sensitivity decreased to 85.8%, because a substantial number of confirmed SCD cases were missed. Furthermore, 6 of the tested case definitions that relied only on the number of encounters outperformed the RuSH case definition on all 4 measures of performance, including specificity.

### Contribution of Inpatient Codes

Recent studies also found high-performing SCD case definitions that give higher weight to SCD diagnosis codes from inpatient admissions than from outpatient encounters, for example, requiring only 1 inpatient SCD diagnosis code to identify a case. Based on our results, giving a higher weight to SCD diagnosis codes from inpatient admissions rather than from other health care encounters did little to improve the performance of the case definition, because an equal number of cases and noncases would be identified for patients with few health care encounters.

### Length of Surveillance Periods

Our study also compared the performance of case definitions applied during various surveillance periods. Results suggest
Table 2. Performance measures for sickle cell disease (SCD) case definitions, including RuSH probable case definition, within 1 year and 5 years of administrative data, Atlanta, Georgia, 2004-2008

| Surveillance Period | SCD Case Definition | No. of SCD Cases Identified | No. of SCD Cases Confirmed | No. of SCD Cases Missed | Sensitivity, % (95% CI) | Specificity, % (95% CI) | PPV, % (95% CI) | NPV, % (95% CI) |
|---------------------|---------------------|-----------------------------|---------------------------|-------------------------|-------------------------|-------------------------|----------------|----------------|
| 5 years             | 1 SCD ICD-9-CM code | 1959                        | 1763                      | 0                       | 98.4 (97.7-98.9)        | 51.0 (43.8-58.2)       | 94.8 (93.6-95.7) | 78.1 (70.0-85.0) |
| 5 years             | 2 SCD ICD-9-CM codes | 1831                      | 1735                      | 28                      | 100                     | 85.8 (84.0-87.4)       | 97.4 (96.4-98.1) | 38.2 (33.4-43.1) |
| 5 years             | 2 SCD ICD-9-CM codes plus ≥1 ICD-9-CM or CPT code a | 1558                      | 1512                      | 251                      | 85.8 (84.0-87.4)       | 97.4 (96.4-98.1)       | 38.2 (33.4-43.1) |                  |
| 5 years             | 3 SCD ICD-9-CM codes | 1739                      | 1693                      | 70                      | 96.0 (95.0-96.9)        | 76.5 (70.0-82.3)       | 97.4 (96.5-98.1) | 68.2 (61.7-74.3) |
| 5 years             | 4 SCD ICD-9-CM codes | 1678                      | 1647                      | 116                     | 93.4 (92.2-94.5)        | 84.2 (78.3-89.0)       | 98.2 (97.4-98.7) | 58.7 (52.7-64.5) |
| 5 years             | 5 SCD ICD-9-CM codes | 1632                      | 1612                      | 151                     | 91.4 (90.0-92.7)        | 89.8 (84.7-93.7)       | 98.8 (98.1-99.3) | 53.8 (48.3-59.3) |
| 5 years             | 6 SCD ICD-9-CM codes | 1602                      | 1586                      | 177                     | 90.0 (88.5-91.3)        | 91.8 (87.1-95.3)       | 99.0 (98.4-99.4) | 50.4 (45.1-55.7) |
| Within 1 year       | 1 SCD ICD-9-CM code | 1959                      | 1763                      | 0                       | 98.2 (97.5-98.8)        | 51.5 (44.3-58.7)       | 94.8 (93.7-95.8) | 67.5 (64.8-83.4) |
| Within 1 year       | 2 SCD ICD-9-CM codes | 1827                      | 1732                      | 31                      | 95.2 (94.1-96.1)        | 77.0 (70.5-82.7)       | 97.4 (96.5-98.1) | 64.0 (57.5-70.5) |
| Within 1 year       | 3 SCD ICD-9-CM codes | 1723                      | 1678                      | 85                      | 95.2 (94.1-96.1)        | 77.0 (70.5-82.7)       | 97.4 (96.5-98.1) | 64.0 (57.5-70.5) |
| Within 1 year       | 4 SCD ICD-9-CM codes | 1656                      | 1629                      | 134                     | 92.4 (91.1-93.6)        | 86.2 (80.6-90.7)       | 98.4 (97.6-98.9) | 55.8 (50.0-61.5) |
| Within 1 year       | 5 SCD ICD-9-CM codes | 1596                      | 1580                      | 183                     | 89.6 (88.1-91.0)        | 91.8 (87.1-95.3)       | 99.0 (98.4-99.4) | 49.6 (44.3-54.9) |
| Within 1 year       | 6 SCD ICD-9-CM codes | 1544                      | 1530                      | 233                     | 86.8 (85.1-88.3)        | 92.9 (88.3-96.0)       | 99.1 (98.5-99.5) | 43.9 (39.0-48.8) |

Abbreviations: CPT, Current Procedural Terminology; ICD-9-CM, International Classification of Diseases, Ninth Revision, Clinical Modification; RuSH, Registry and Surveillance System for Hemoglobinopathies.

The RuSH definition of a probable case required ≥2 SCD ICD-9-CM codes on ≥2 separate health care encounters plus ≥1 hemoglobinopathy-associated complication, treatment, or procedure identified through ICD-9-CM codes, Current Procedural Terminology codes, or National Drug Codes but no laboratory-based disease confirmation available.

Miscoding of SCD

Lastly, we sought to determine the correct diagnosis for cases that were incorrectly coded as SCD. Of 197 patients who were coded as SCD, 56% were reclassified as a different ICD-9-CM code. The reclassification was based on the presence or absence of clinical evidence of SCD, such as HbC or HbE trait, or other conditions associated with SCD, such as HbSC trait or another hemoglobinopathy carrier state (eg, HbAC, HbAD). A second review of these cases revealed that 33% of patients had been miscoded, with the majority being due to the presence of other hemoglobinopathies and hematological disorders. Our findings suggest that there is a significant degree of miscoding of SCD, which may have implications for surveillance and disease management. Additionally, the use of ICD-9-CM codes for disease surveillance in adult populations may have limitations, and the use of ICD-9-CM codes for SCD surveillance in children may be different. It is unclear how rates might differ if coding for other carrier states were also included. Furthermore, the findings suggest that there is a significant degree of miscoding of SCD, which may have implications for surveillance and disease management. Additionally, the use of ICD-9-CM codes for disease surveillance in adult populations may have limitations, and the use of ICD-9-CM codes for SCD surveillance in children may be different. It is unclear how rates might differ if coding for other carrier states were also included.
later will need to use ICD-10 coding, which has the potential to better distinguish SCD from thalassemias, thereby potentially improving case identification.

Conclusion
This study supports an administrative case definition that includes ≥3 ICD-9-CM codes to identify SCD with a high degree of accuracy in pediatric patients and suggests that this definition can be useful for establishing longitudinal SCD surveillance systems. Results can also be used by public health practitioners and epidemiologic researchers to estimate the prevalence of SCD and to determine overall rates of disease-specific complications, emergency department use, and inpatient hospital use.

Acknowledgments
The authors acknowledge the contributions of other valued members of the Sickle Cell Data Collection program in Georgia and our colleagues in California.

Declaration of Conflicting Interests
The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was funded by the Centers for Disease Control and Prevention (CDC) through a cooperative agreement with the Association of University Centers on Disabilities (5U38OT000140) and through the help and support of the CDC Foundation and Pfizer, Inc. Support was also provided by a grant from the Abraham J & Phyllis Katz Foundation.

References
1. Feuchtbaum L, Carter J, Dowray S, Currier RJ, Lorey F. Birth prevalence of disorders detectable through newborn screening by race/ethnicity. Genet Med. 2012;14(11):937-945. doi:10.1038/gim.2012.76
2. Hassell KL. Population estimates of sickle cell disease in the U.S. Am J Prev Med. 2010;38(4 suppl):S512-S521. doi:10.1016/j.amepre.2009.12.022
3. Michlitsch J, Azimi M, Hoppe C, et al. Newborn screening for hemoglobinopathies in California. Pediatr Blood Cancer. 2009;52(4):486-490. doi:10.1002/pbc.21883
4. Wang Y, Kennedy J, Caggana M, et al. Sickle cell disease incidence among newborns in New York State by maternal race/ethnicity and nativity. Genet Med. 2013;15(3):222-228. doi:10.1038/gim.2012.128
5. Benson JM, Therrell BL Jr. History and current status of newborn screening for hemoglobinopathies. Semin Perinatol. 2010;34(2):134-144. doi:10.1053/j.semperi.2009.12.006
6. Quinn CT, Rogers ZR, McCavit TL, Buchanan GR. Improved survival of children and adolescents with sickle cell disease. Blood. 2010;115(17):3447-3452. doi:10.1182/blood-2009-07-233700
7. Yanni E, Grosse SD, Yang Q, Olney RS. Trends in pediatric sickle cell disease-related mortality in the United States, 1983-2002. J Pediatr. 2009;154(4):541-545. doi:10.1016/j.jpeds.2008.09.052
8. Brawley OW, Cornelius LJ, Edwards LR, et al. National Institutes of Health consensus development conference statement: hydroxyurea treatment for sickle cell disease. Ann Intern Med. 2008;148(12):932-938.
9. Hassell K, Pace B, Wang W, et al. Sickle cell disease summit: from clinical and research disparity to action. Am J Hematol. 2009;84(1):39-45. doi:10.1002/ajh.21315
10. Paulukonis ST, Eckman JR, Snyder AB, et al. Defining sickle cell disease mortality using a population-based surveillance system, 2004 through 2008. Public Health Rep. 2016;131(2):367-375. doi:10.1177/003335491613100221
11. Huilhian MM, Feuchtbaum L, Jordan L, et al. State-based surveillance for selected hemoglobinopathies. Genet Med. 2015;17(2):125-130. doi:10.1038/gim.2014.81
12. Paulukonis ST, Harris WT, Coates TD, et al. Population based surveillance in sickle cell disease: methods, findings and implications from the California Registry and Surveillance System in Hemoglobinopathies project (RuSH). Pediatr Blood Cancer. 2014;61(12):2271-2276. doi:10.1002/pbc.25208
13. Centers for Disease Control and Prevention. Sickle Cell Disease Data Collection (SCDC) Program. 2018. https://www.cdc.gov/ncbddd/hemoglobinopathies/scdc.html. Accessed January 31, 2019.
14. Grigorescu V, Kleyen MJ, Korzeniewski SJ, Young WI, Whitten-Shurney W. Newborn screening follow-up within the lifespan context: Michigan’s experience. Am J Prev Med. 2010;38(4 suppl):S522-S527. doi:10.1016/j.amepre.2010.01.002
15. Korzeniewski S, Grigorescu V, Copeland G, et al. Methodological innovations in data gathering: newborn screening linkage with live births records Michigan, 1/2007-3/2008. Matern Child Health J. 2010;14(3):360-364. doi:10.1007/s10995-009-0464-3
16. Centers for Disease Control and Prevention. International classification of diseases, ninth revision, clinical modification (ICD-9-CM). Reviewed November 2015. https://www.cdc.gov/nchs/icd/icd9cm.htm. Accessed February 4, 2019.
17. American Medical Association. Current Procedural Terminology. Chicago, IL: American Medical Association; 1998.
18. US Food and Drug Administration. National Drug Code Directory. 2016. http://www.fda.gov/cder/ndc/database/default.htm. Accessed February 4, 2019.
19. Grosse SD, Boulet SL, Amendah DD, Oyeku SO. Administrative data sets and health services research on hemoglobinopathies: a review of the literature. Am J Prev Med. 2010;38(4 suppl):S557-S567. doi:10.1016/j.amepre.2009.12.015
20. Reeves S, Garcia E, Kleyen M, et al. Identifying sickle cell disease cases using administrative claims. Acad Pediatr. 2014;14(5 suppl):S61-S67. doi:10.1016/j.acap.2014.02.008
21. Michalik DE, Taylor BW, Panepinto JA. Identification and validation of a sickle cell disease cohort within electronic health records. Acad Pediatr. 2017;17(3):283-287. doi:10.1016/j.acap.2016.12.005
22. SAS Institute Inc. *SAS Version 9.4*. Cary, NC: SAS Institute, Inc; 2015.

23. Tu K, Campbell NRC, Chen ZL, Cauch-Dudek KJ, McAlister FA. Accuracy of administrative databases in identifying patients with hypertension. *Open Med*. 2007;1(1):e18-e26.

24. Guttmann A, Nakhla M, Henderson M, et al. Validation of a health administrative data algorithm for assessing the epidemiology of diabetes in Canadian children. *Pediatr Diabetes*. 2010;11(2):122-128. doi:10.1111/j.1399-5448.2009.00539.x

25. Ellison AM, Bauchner H. Socioeconomic status and length of hospital stay in children with vaso-occlusive crises of sickle cell disease. *J Natl Med Assoc*. 2007;99(3):192-196.

26. Raphael JL, Mei M, Mueller BU, Giordano T. High resource hospitalizations among children with vaso-occlusive crises in sickle cell disease. *Pediatr Blood Cancer*. 2012;58(4):584-590. doi:10.1002/pbc.23181

27. Dampier C, Kanter J, Howard R, et al. Access to care for Medicaid and commercially-insured United States patients with sickle cell disease. *Blood*. 2017;130(suppl 1):4660.