Longitudinal Data Collection for Sickle Cell Disease in California: History, Goals and Challenges
Developed by
Susan Paulukonis,¹ Faith Raider¹ and Mary Hulihan²

The Sickle Cell Disease Longitudinal Data Collection System project in California is made possible through funding support from Pfizer, Inc. and Biogen to the CDC Foundation.

December 2015

¹California Rare Disease Surveillance Program/California Environmental Health Tracking Program, Public Health Institute, Richmond, CA
²Centers for Disease Control and Prevention, National Center for Birth Defects and Developmental Disabilities, Division of Blood Disorders, Atlanta, GA
Executive Summary

Sickle cell disease (SCD) is the most common severe genetic disease. It impacts approximately 7,000 people in California and nearly 90,000 in the U.S. And yet remarkably little is known about the population impacted by SCD as a whole. Only recently has this disease become part of the nation’s public health agenda. Estimates of the prevalence of disease are not based on patient counts but rather on extrapolations of newborn screening data using data on life expectancy (itself extrapolated from other data) or based on patient counts from hospital discharge databases. While much is known of the pathophysiology of the disease, little is understood about the impact of the disease on people, particularly adults, living with SCD.

Estimates of the incidence of disease are not based on patient counts but rather on extrapolations of newborn screening data with life expectancy.

The Sickle Cell Disease Longitudinal Data Collection project in California extends over five years of successful SCD surveillance in the state as part of the Center for Disease Control and Prevention's (CDC) and National Heart, Lung and Blood Institute’s Registry and Surveillance System in Hemoglobinopathies (RuSH) and CDC’s Public Health, Research, Epidemiology and Surveillance in Hemoglobinopathies (PHRESH) cooperative agreements. The goals of this new effort are to continue using and to improve upon developed methods and data sources for understanding SCD at the population level in the state as well as to analyze the data collected and disseminate them to audiences that will drive policy and health care changes, leading to improvements in quality of life, life expectancy and health among those living with SCD.

Data sources for this surveillance system include administrative data (hospital discharge data, emergency department data, and Medicaid claims), newborn screening case reports, vital records and eventually clinical case reports. These data will be linked, de-duplicated and condensed to form a profile of the health conditions, health care and outcomes of this population.

After meetings with over 20 stakeholder groups in California and at the federal level, the suggested topics for analysis and dissemination using these data are as follows:

- Geography of patient population
- Transition from pediatric to adult care
- Hispanic SCD cases
- The aging SCD population
- High health care utilization patterns among SCD patients

This report describes in detail these and other potential topics and provides a plan for accomplishing these goals.
# Contents

- Executive Summary 2
- Introduction 4
- Sickle Cell Disease (SCD) Surveillance in California 5
- Key Surveillance Findings from the California RuSH and PHRESH Projects 6
- Other Current SCD Data Collection Efforts 7
- California’s SCD Longitudinal Data Collection System 7
- Goals and Areas of Focus 8
  - Geography of Patient Population 8
  - Transition from Pediatric to Adult Care 10
  - Hispanic SCD Cases 12
  - The Aging SCD Population 13
  - High Health Care Utilization Patterns Among SCD Patients 14
- Other Suggested Surveillance Topics of Interest 15
- Timeline 19
- References 22
- Appendix A: SCD LDC 2015 Stakeholder Meetings and Attendees 27
Introduction

Sickle cell disease comprises multiple genotypes that manifest in significant disease severity, such as hemoglobin (Hb) S/S and Hb S/β0, generally more severe, and Hb S/C, Hb S/β+ and other Hb S sub-types that typically manifest as less severe forms of the disease. It is estimated to affect 90,000 in the US1. SCD was once thought of as a childhood disease with the majority of cases dying before adulthood, but the screening of newborns and subsequent use of prophylactic penicillin in identified cases and implementation of comprehensive care models in the late 20th century dramatically changed the life expectancy and disease course for those with SCD2,3. Recent estimates suggest that mortality rates among young children with SCD are not significantly different than the general population, and over 95% of SCD cases born today will live into adulthood2,4.

This change in age distribution for the population living with SCD is welcome, but comes at a price. SCD is now a chronic, debilitating, complex, life-threatening disease with its primary impact on adolescents and adults5. Rates of emergency room and hospital utilization are high among adults with this disease6. For adults, use of the excellent comprehensive care models developed for the pediatric population is low due to a variety of challenges including lack of availability of care, lack of insurance, and the distance or lack of transportation to care. In California, as in the rest of the nation, there are few hematologists trained and willing to care for adult SCD patients; much of the care of these patients takes place in emergency rooms or other non-specialty settings7,8.

Among children, primary severe complications of SCD include anemia, septicemia or other severe infection (limited by use of penicillin), stroke, splenomegaly or splenic infarction, acute chest syndrome and debilitating acute pain9,10. Among adults, pain (both severe acute pain episodes called ‘crises’ and chronic pain) is the hallmark of the disease. Avascular necrosis in the large joints (hip, shoulder), stroke, and organ damage particularly in the kidneys and heart are among the primary causes of healthcare utilization by adults6,9,10. A recent study of mortality among the SCD population in California and Georgia reported a median age of death of 43 years, significantly lower than that of the general populations of those states4.

Beyond prophylactic penicillin for children up to age five, treatments for this disease are limited. Bone marrow transplant carries significant risks of severe complications or death, but has become lower risk as technology has improved. However, candidates must meet eligibility criteria and have a suitable match. Some patients with severe disease require regular blood transfusion to prevent stroke and other complications; this procedure also carries significant risks (e.g., fluid overload, iron overload, disease exposure, and antibody formation) and is time-intensive for the patient11. Hydroxyurea is a drug developed as a chemotherapeutic agent and approved for use in the adult Hb S/S and S/β0 populations in 1998. It has been shown to raise the body’s production of Hb F (fetal hemoglobin), which reduces the proportion of Hb S in the blood stream and lowers the rate of complications such as pain and acute chest syndrome as well as the need for blood transfusion among approximately 75% of study patients. Uptake of hydroxyurea is low for a number of reasons, including patient and provider reluctance to chronically use a drug with a 'black box' warning12-16.

The landscape of care and treatment for those with SCD is poised to change dramatically, however. Pharmaceutical companies are currently testing at least 39 compounds in clinical trials for SCD or for conditions that impact those with SCD. The SCD community has high expectations that these treatments will lead to improved life expectancy, lower healthcare costs and higher quality of life for this population17,18. The compounds have diverse roles and a variety of expected endpoints. They include those that increase production of Hb F, target oxidative injuries and inflammation, or reduce cell adhesion among sickled cells. As these treatments
move into Phase III and IV clinical trials and wider usage, a challenge will be determining effectiveness vs. efficacy, uptake at the population level, the impact on health care utilization, and changes in outcomes due to the compounds. While life expectancy and descriptions of complications and health care utilization among those with SCD have been documented in small clinical populations, there have been only limited studies published using population-level data for SCD.

The SCD Longitudinal Data Collection System (SCD LDC) proposes to collect, synthesize and disseminate multi-source, population-based data collected over time that will establish a baseline health profile of the SCD population prior to the introduction of new therapies. It will then track changes in population outcomes over time. A longitudinal data system will ensure that the SCD community has credible, scientifically sound information to inform standards of care regarding these novel treatments.

The landscape of care and treatment for those with SCD is poised to change dramatically.

Sickle Cell Disease Surveillance in California

California began universal newborn screening for all forms of SCD in 1990. Approximately 133 cases of SCD (including all genotypes) are identified per year among approximately 500,000 births. The California Department of Public Health (CDPH) follows up on each case after birth to assure that the child has been seen by a pediatric hematologist. These providers offer parent education and, most importantly, prescribe prophylactic penicillin to be used for the prevention of severe infections. In 2011, the state implemented a ‘long-term follow up’ program that further tracks these children annually over the first five years of life, ensuring that they are in regular care, continuing on penicillin, and tracking any complications or other health events. The state does not track children with SCD after age five.

There are numerous hospital-based pediatric clinics for SCD and other genetic childhood disorders in California. Most of these have some form of internal clinical database for tracking patient care and outcomes over time. Some of these clinics also see adult patients, especially during young adulthood/transition. Typically, patient care that occurs outside of their system is not recorded in the SCD clinics’ electronic health records systems, including care that may not be directly related to SCD, such as for injuries or malignancies, or for illness that may be secondary to SCD such as renal disease and heart disease. Data collected by clinics in local databases or electronic health records belong to the clinic (and patients), and clinicians may or may not disseminate analyses of these data. It is also clear that in California, a substantial portion of adult patients are not seen in clinical settings nor are they included in data from these sites.

Researchers in California and other states have conducted state level or sub-state level surveillance of SCD using administrative data, such as hospital discharge data or large commercial claims databases. These efforts are closer to population-based surveillance and less biased than data from SCD clinics, as they do not rely on a clinical population and include health care utilization over the life course, but they have significant biases. Research suggests that analyses such as these may dramatically overestimate the number of cases (by including patients who have coding errors or ‘rule out’ diagnostic codes) and underestimate the utilization by true cases (by not including utilization that does not specifically include SCD diagnostic codes). Preliminary results with California and Georgia data (described below) show similar findings for death records, which suggests that these data alone are not a reliable source for determining SCD prevalence or life expectancy.

California was one of seven states that developed and systematically tested a multi-source surveillance system by participating in a cooperative agreement with the Centers for Disease Control and Prevention (CDC) and the National Heart, Lung and Blood Institute (part of the National Institutes of Health) called the Registry and Surveillance System for Hemoglobinopathies (RuSH, 2010-2012). All states that participated in RuSH agreed to collect data on the same populations, outcomes, complications,
and treatments at the state level; the methodology and data sources differed across states. California used data from the following sources:

- Newborn screening (SCD case identification 2000-2008)
- Hospital discharge data (all data 2004-2008, since expanded through 2013)
- Emergency room data (all data 2005-2008, 2004 was not available, data since expanded through 2013)
- Medicaid claims data for all patients with one or more SCD ICD diagnostic codes (2004-2008)
- Vital records (all deaths 2004-2008; births linked to SCD cases 2000-2008)
- Clinical case reports from two large SCD treatment centers (2004-2008)

California RuSH staff cleaned and standardized these data and developed a linking algorithm to create a profile by case across multiple data sources. For example, a child identified in newborn screening and linked to a birth certificate might also be seen among the cases reported by one of the care centers, appear in the hospital discharge data as having one or more emergency room visits or inpatient stays, and might have Medicaid utilization data. This child would be counted as an incident case only once, but her profile would include information from all of these sources. As a final step, a RuSH case definition was applied to describe the certainty of the evidence that each case had SCD: confirmed, probable or possible.

California was one of two states (along with Georgia) that performed validation and improvement work on the RuSH data in a subsequent cooperative agreement with the CDC called Public Health, Research, Epidemiology and Surveillance in Hemoglobinopathies (PHRESH, 2012-2014). During the PHRESH validation project, California refined its case definition and improved its linking and matching methods using the RuSH dataset and additional confirmed case reports from six comprehensive hemoglobinopathy centers in the state.

Key Surveillance Findings from the California RuSH and PHRESH Projects

The data from these projects led to surprising conclusions about the SCD population in the state. Among the published and in-development analyses is evidence of the importance of a broad-based, population-level surveillance effort for this disease.

- Clinical data sources are inadequate for tracking adult SCD cases; among the five largest hemoglobinopathy centers in California that see some adult SCD patients, a total of 492 adult patients were reported. However, estimates from the PHRESH project suggest there should be approximately 3,000 adult SCD patients in the state.
- 43% of confirmed and probable SCD cases who died during the 2004-2008 period and were linked to a death certificate did not have SCD or a condition linked to SCD listed among their causes of death.
- Preliminary results (not yet published) suggest the following:
  ◊ Patterns of high utilization are not consistent within a patient over time; patients are high utilizers (of emergency and inpatient services) for a period of some months, then return to expected levels of health care use.
  ◊ Older patients (over 40 years) have different health care utilization patterns than younger adult patients, with more visits coded for other health care problems related to organ failure (such as end stage renal disease) and other morbidities of older adults (diabetes, cancer, heart disease), rather than SCD.
  ◊ 51% of the state’s confirmed and probable adult SCD cases live in Los Angeles County, which no longer has an adult SCD clinic.
Other Current SCD Data Collection Efforts

The importance of data collection for the understanding of treatments, outcomes and access to care among those with SCD is well known. In addition to efforts described above (newborn screening, clinical databases), there are other projects proposing to collect data on people living with SCD. The similarities and differences among these programs in states participating in the SCD LDC should be clearly understood.

Health Resources and Services Administration (HRSA)/Sickle Cell Disease Association of America (SCDAA) GetConnected Project

SCDAA Press Release:
http://sicklecelldisease.org/index.cfm?page=news&id=85

One component of this HRSA grant awarded to SCDAA is a voluntary patient registry that will require patient consent to join, will focus on collecting data on access and quality of care among those patients engaged with SCD community-based organizations, and will enable communication to patients about educational and clinical opportunities and information

National Heart, Lung and Blood Institute Sickle Cell Disease Implementation Consortium (SCDIC) RFA-HL-16-010

NIH Funding Opportunity Announcement:

This upcoming award will include an extensive data collection effort among awardee patient populations. This project will collect data on patients 15 to 45 years of age who are identified, recruited and consented at clinical sites.

California’s SCD Longitudinal Data Collection System

With new funding, California plans to continue and expand its well-developed and broad based SCD surveillance system established through participation in the RuSH and PHRESH projects. The new project is now located in the California Environmental Health Tracking Program at CDPH, in the California Rare Disease Surveillance group (CRDS). This program brings with it new areas of expertise in database development and linkage, geographic/spatial analysis and mapping, data analysis and statistical methodology, and outreach and dissemination. There is staff continuity in project management and data acquisition for this project, and we have excellent documentation for new staff joining us for data linkage and database development. All state agencies and clinical partners remain supportive of this work and are planning to continue to contribute data as resources are available. As informed by the RuSH and PHRESH projects, the new system will do the following:

- Begin with the same administrative and statewide data sources as used previously, with data collected through 2013 (additional years as soon as available)
- Request case reports from our clinical partners for new cases seen since 2008
- Use the revised, validated case definition for SCD developed by California during the PHRESH project
- Revise the structure used for linking, matching and analyzing data
  ◊ Use a relational database format rather than flat file
  ◊ Develop a revised (based on lessons learned) matching and linking process
  ◊ Review ‘by hand’ cases that are unusual or ‘borderline’ for case definitions, and all newborn screening cases (which are difficult to link due to limited consistent personal identifiers)
  ◊ Create systems for more agile and flexible data analyses, rather than sorting/matching all data for each analysis
Goals and Areas of Focus

Over a two-month period, June-July 2015, California’s SCD LDC project management staff held over twenty small group meetings with community-based organizations, clinicians, state agency staff, and patients and their families throughout the state, as well as with federal agency partners, relevant national association partners, and stakeholders in the Washington, DC and Baltimore areas. A complete list of the participants in these meetings is in Appendix A. The goals of these meetings were the following:

- Inform stakeholders about the project and answer questions
- Enhance or develop collaborative relationships
- Assure continued availability of surveillance data
- Recognize gaps and overlaps in SCD surveillance among different programs and plan to mitigate
- Determine the most valuable products that could come from this effort, taking into account time, available resources and complexity

Based on the experience of RuSH and PHRESH, we learned that the data collected are highly valuable for answering certain types of question (e.g., disease prevalence, health care utilization, clinical outcomes) but not helpful for others (quality of life, education and employment status, some clinical markers such as units of blood transfused, or compliance with oral medications). We also learned that some clinical questions are so complex that considerably more data and analysis would be needed to answer them compared to useful but simpler questions that have never before been addressed in the literature using population-based data. Finally, we know that some areas of investigation are of great interest, but hold low probability of influencing change. We placed a high priority on those questions that had a clear path to changes in policy or health care practice for patients. With these parameters in mind, we encouraged the attendees of the meetings to share their highest priority questions of the data.

We propose here five broad areas of focus based on these conversations and highlight the reasons for the choices, the specific questions we may answer, the literature and background on the topics, and improvements in patient outcomes or quality of care that may result from this work.

Geography of Patient Population

With nearly 38 million residents and 164,000 square miles, California is larger than many nations; it is the most populous state and the third largest in land mass. The state’s population is highly diverse in its racial and ethnic makeup, country of birth, languages spoken, and socioeconomic status, but sub populations for all of these demographic variables are clustered. Our preliminary analyses found that SCD cases are clustered in different parts of the state, which has not been previously described in the literature. Our data offer us a unique opportunity to analyze and present the demography and geography of SCD in California. A simple representation of this power is shown in Figure 1, which identifies counts of newborn screening identified cases of SCD by county. With the proposed surveillance methods, we can also look at sub-county data, including city or, in the case of large cities such as Los Angeles, neighborhood. These data can also demonstrate the geographic challenges in gaining access to care.
Our data offer us a unique opportunity to analyze and present the demography and geography of SCD in California.

In addition, spatial disease data modeling can be a useful analytical tool, if the levels of aggregation (e.g., census block, city, county) are thoughtfully selected to best answer the research question\(^{31}\). Although early attempts at mapping to explain and intervene in disease patterns met significant challenges, recent research has led to improvements in methodology and standardization of best practices from which the SCD LDC project can benefit\(^{32}\).

SCD researchers in the US have explored sociodemographic factors affecting the disease using simple spatial techniques. Clinical studies outside of California have found that sociodemographic status based on patient zip codes was not associated with longer hospital stays or outcomes among SCD patients\(^{33,34}\). Other researchers have examined access to care via public transit, and the effect of such access on outcomes\(^{35}\). A recent California-based study assessed access to care by determining the approximate distance from patient’s home (the distance from the center of the patient’s zip code area) to the site of care for patients with SCD ICD 9 codes in emergency department data, as well as other socio-demographic factors associated with SCD hospital utilization and outcomes. It showed that geographic distance to care and patient insurance status are significant predictors of ED utilization\(^{36}\). No other relevant publications describing SCD distribution or care as a function of geography in California have been identified.

Specific questions regarding patient location include the following:

- Where in the state (by city or zip code) are patients located?
- What facilities are seeing SCD patients, and in what setting (e.g., county hospitals vs. private hospitals, emergency department vs. outpatient clinics)?
- Are individual patients being seen in multiple care facilities (e.g., more than one ED)?
- Are there nearby hemoglobinopathy treatment centers that may be better able to serve these patients?
- Based on these numbers, where are target areas for outreach?
- What are the differences in access to care by distance (proximity to comprehensive care centers and high quality emergency care) across

---

**Figure 1: California Newborn Screening Identified SCD Births, 2004-2008**

![Map showing the distribution of SCD births in California from 2004 to 2008.]
SCD is a special case; with widespread use of prophylactic penicillin and monitoring for stroke risk in the pediatric population, the period of transition coincides with the onset of the most severe symptoms and high health care utilization for patients in regions with high quality pediatric care.\(^6,10,37\).

Teen and young adult patients suffer from more frequent SCD-related complications than younger pediatric patients.\(^10,37\). In particular, pain crises, chronic pain, and avascular necrosis are frequently seen beginning in the later teen years and into adulthood and the frequency of diagnoses for SCD complications increases markedly after age 16.\(^16\). The increase in complications, accompanied by a decline in transfusion frequency, is likely due in part to a decrease in close medical follow-up and preventive care as patients transition to adulthood, and may be in part due to psychosocial factors and access to care issues.\(^8,38-43\).

Additional studies that analyzed Medicaid and other administrative data have documented the increase in the frequency of ED and inpatient and outpatient visits post-transition to adulthood.\(^8,25,35\). While Hemker et al. examined data from different types of providers, the study primarily focused on whether lack of outpatient care during transition leads to increased use of the ED rather than on provider usage patterns.\(^8\). The authors found that increased ED utilization after transition suggests lack of access to primary care providers for SCD patients. Andemariam et al. found in a clinical trial that longer travel distance to an adult SCD center is a risk factor for an unsuccessful transition.\(^35\).

Additionally, at least one cohort study has found that shortly following the transition to adult medical care, young adults are at high risk for death.\(^4\). Hamideh and Alvarez examined death certificates to find that young adults with SCD, 20–24 years of age, were at over double the risk for mortality (1.4/100,000 patients) than 15- to 19-year olds (0.6/100,000 patients).
Few studies have examined conditions in childhood that predict outcomes after transition. Boyd et al. found in a clinic-based study that asthma is associated with higher mortality rates in SCD patients, but did not explore whether childhood asthma specifically predicts poor adult outcomes. Platt examined a variety of acute and chronic conditions in adults with SCD that were risk factors for early death in patients 20 and older, including renal failure, acute stroke and chest syndrome, and pain episodes.

Our stakeholders noted the importance of quantitative data on complications, outcomes and utilization during the transition period. Specifically, areas of interest that mesh well with the planned data include:

- Where are patients seen during the different phases of transition (e.g., 16-18 years, 19-21, 22-24)?
  - Pediatric comprehensive care clinic?
  - Primary care physician?
  - Adult hematologist?
  - Emergency room?
  - Primary, secondary or tertiary hospitals?
- What is the age of onset of complications such as acute pain crises, avascular necrosis, and other identified complications?
- Can we determine to what degree increased health care utilization is a function of disease vs. lack of care options?
- What conditions in childhood predict poor young adult outcomes?
  - Asthma
  - Acute chest syndrome
  - Stroke
  - Infection
  - Others?
- What proportion of transition-aged patients die, and what are the causes of death?
  - Are there events or diagnoses that predict mortality in this age group?

Answers to most of these areas of study will require collecting multiple years of data (e.g., predicting adult outcomes based on childhood complications, mortality) and tracking patients over time. These analyses may be best conducted once a significant number of years of data have been linked. The SCD community has expressed strong interest in the results of these analyses. We believe publication of results in peer reviewed journals will ensure a wide audience. Implications and consequences of such disseminated analyses may include informing workforce development efforts to recruit and train more qualified providers for the transitioning and adult population, changes in the structure of pediatric hemoglobinopathy
clinics to allow for internal transition to in-house adult programs (as has been done in some California clinics), increased screening for complications at specific ages, and better understanding of health care and social needs of all of California’s young adults with SCD.

**Hispanic SCD Cases**

People with SCD born in California in 1990 or later were diagnosed at birth due to universal newborn screening for SCD, but those born prior to 1990 or outside the state (particularly outside the US) may have been diagnosed later or may never have been diagnosed at all47. Brousseau et al. used SCD newborn screening prevalence rates for Black/African Americans and Hispanic/Latinos in conjunction with more recent census data to estimate that about 10% of SCD patients in the United States are Hispanic1, which may be an underestimate in California due to high immigration of Latino populations. Prevalence rates for Hispanic children of non-Mexican heritage were almost 30 times higher than for Hispanic children of Mexican ancestry. This may be due to the introduction of the sickle cell trait gene from Africa into the Americas through the slave trade, which disproportionately affected the eastern coasts of the US and Brazil, as well as the Caribbean47.

California has a large Latin American population, but Mexican immigrants predominate. The RuSH project found that 8% of all SCD patients in California were White, Hispanic30. Black Hispanic cases are included under the administrative data coding category of ‘Mixed’ and are therefore difficult to identify as such. The proportion of the Hispanic population living with SCD remains low in California, but the increasing proportion of the state’s population that is Hispanic and a decreasing Black/African American population (from 6% in 2010 to <5% in 2040) means that a higher proportion of SCD cases are non-Black and/or mixed race48.

Anecdotal reports and one publication suggest that there are challenges to diagnosis and treatment among Hispanic and particularly immigrant cases, with physicians not screening non-Black patients despite presentation with relevant symptoms 49. With the exception of reports on rates of Hispanic newborn cases of SCD, there is no published literature on the changing race/ethnic makeup of the disease in California to date50.

In a clinical study, researchers found no difference in baseline and hydroxyurea-induced percent fetal hemoglobin, a determinant of SCD disease severity, in African American and Hispanic patients aged 4 to 21 years51. No other relevant literature addresses whether conditions, symptoms, complications and outcomes differ in SCD patients based on race/ethnicity.

Questions raised by our stakeholder groups include the following:

- Number/proportion of California SCD cases by race/ethnicity and age
- Origin of immigrant cases (region, country or state of birth)
- Challenges to diagnosis of cases born outside the country or prior to newborn screening (while there is strong interest in this topic, this may be outside the scope of the project due to lack of available data)
- Differences in complications, co-morbidities, outcomes, mortality and utilization among SCD cases by race/ethnicity. Whether the following
Anecdotal reports and one publication suggest that there are challenges to diagnosis and treatment among Hispanic and particularly immigrant cases, with physicians not screening non-Black patients despite presentation with relevant symptoms.

The Aging SCD Population

Only recently have there been a significant number of SCD cases with severe genotypes to study into older age; the widespread use of preventative treatment among the pediatric population has allowed patients to live into adulthood. Consequently, little has been published about complications, co-morbidities and health care utilization among cases over the age of 45 years. A recent paper by Sandhu and Cohen addressed the most common complications and co-morbidities among a clinic-based cohort of adult patients in New Jersey. It found that the majority were not undergoing routine screenings for cancers and that a high number had important co-morbidities, such as hypertension and diabetes, and early onset complications, such as chronic renal disease, iron overload and cardiovascular disease.

Previous clinical research has shown that pain crises tend to lessen in number and possibly in severity in SCD patients over the age of 40, although not all studies agree on this point. Sanders et al. compared groups of younger adults with SCD (age 18-36) and older adults with SCD (age 37-62) on pain, complications and health care utilizations. The study found a significant difference only in the patterns of health care utilization; older adult patients had more outpatient clinic visits, while younger adults had more ED visits and hospital admissions. Serjeant et al. studied a Jamaican cohort of surviving patients aged 60 and older with Hb SS. Survival was associated with clinical features indicating a milder course of the disease.

Hamideh and Alvarez analyzed mortality statistics to compare the average life expectancy among African-Americans with and without SCD in 2009. Most of the deaths for both groups were attributed to cardiovascular causes; these cardiovascular deaths occurred much earlier in SCD patients than in the non-SCD African American population, with a peak in the late thirties to early fifties. They found that cancer and diabetes, the second and fourth largest causes of death among African-Americans without SCD, were not leading causes among African-American SCD patients.

Questions on this subject raised by our stakeholder groups include the following:

- What are common complications and co-morbidities over age 45 in the SCD population?
- Are common diseases of adults (such as cancers, heart disease, diabetes) routinely screened for in this population? Do these diseases occur at the same rate as in the general population? At the same mean/mean...
• How do health care utilization patterns change from one age group to another?
• Whom do older patients see for care? Do they see specialists for complications of SCD, such as nephrologists for kidney disease, cardiologists for heart disease?
• How does cause of death differ among older and younger adults?
• What is the rate of preventable causes of death in SCD compared to that of other, similar chronic diseases?
• What is the rate of renal disease compared to the general population and the Black/African American population? Is age of onset earlier?
  ◊ Rates of kidney transplant
  ◊ Rates of adverse outcomes
  ◊ Autoimmune hemolytic failure
• What is the rate of pain crises, iron overload, leg ulcers and avascular necrosis among this older population compared to younger groups?

As with results of analyses on the transition from pediatric to adult care, we believe analyses of these data on the aging SCD population will be informative to a national and international audience. We suggest that peer reviewed journals are the best venue for dissemination. These analyses will also require as many years of data as possible to accurately portray the population in the older ages, as well as capture events or complications that may be predictive of future outcomes, so we contemplate scheduling these analyses for later years of this work. Intended outcomes for this work would be increased understanding of the health care utilization, clinical outcomes, access to care and support needs of this population, outreach efforts to non-hematologist specialists who may be seeing these patients for complications, increased ability to identify and follow these older patients in administrative data, updated mortality and life expectancy analyses, and improved and expanded standards of care that include more information on complications among older SCD patients.

Little has been published about complications, co-morbidities and health care utilization among cases over the age of 45 years.

High Utilization Patterns Among SCD Patients

By contrast with the other topics, much has been written on high hospital and emergency department (ED) utilization by SCD patients. Most analyses are based on limited, clinic-based cohorts or on the use of single administrative (hospital discharge or ED) data sets. As noted above, these data sources are limited in their capacity to identify confirmed or highly probable cases and to look at all utilization, not just ED or inpatient, and not just encounters with SCD diagnostic codes included.

Brousseau et al. studied administrative data to estimate SCD population utilization, finding 29% had no ED visits or hospitalizations while 16.9% had 3 or more per year. These data sources are limited in their capacity to identify confirmed or highly probable cases and to look at all utilization, not just ED or inpatient, and not just encounters with SCD diagnostic codes included.

Brousseau et al. studied administrative data to estimate SCD population utilization, finding 29% had no ED visits or hospitalizations while 16.9% had 3 or more per year. That study also found age to be a factor in readmissions. Ezenwa et al.’s longitudinal comparative investigation found that a SCD patient-reported measurement of pain in outpatient visits is a predictor of acute health care utilization in the following year, along with age but not gender. Studies by Carroll have looked at high utilization patterns. Her 2009 and 2011 study of administrative data sets found that among high utilizers, a period of high utilization could be followed...
by moderate periods that were not likely to reverse\textsuperscript{66,67}. The 2011 study found that in California, high utilizers were associated with prior hospitalizations and previous diagnoses of bone death and renal disease. Carroll et al. found in a clinical study that age, disease severity, greater parental education, and psychiatric illness are significant factors in high care utilization\textsuperscript{62}. A retrospective cohort study found that care for adolescents differed between children's hospitals and general hospitals in terms of rate of intubation and length of stay\textsuperscript{68}.

The SCD LDC data are uniquely able to identify the utilization patterns among the majority of cases in California, including hospital, ED and outpatient, and to look for patterns of utilization over time. Additionally, we have the capacity to look at events or complications that precede periods of high utilization. Among areas of interest for our stakeholders were the following:

- Are the high utilizers of ED and inpatient services the same people over long periods of time?
- If not, what are the factors that predict high utilization starting and stopping?
- What demographic and other variables predict whether a patient becomes a high utilizer?
- Do the rates of ED to admission vary by facility type, region, condition or other factors?
- What are the patterns of outpatient care in between ED/hospital stays?
- What factors predict readmissions to inpatient or ED settings?
- Are outcomes predicted by type of facility?

Along with transition and outcomes among older adults, analyses of high utilization would be improved with more years of data, may be helpful to other states and countries, and are likely to be of high interest to researchers. We suggest that these analyses be informed by as many years of data as possible and that peer reviewed journal articles will be an important mode of dissemination. Among the changes in policy, practice and outcomes that may result from this work are a greater understanding of the triggers for periods of high utilization as well as prevention strategies. It could result in increased education of ED providers on the importance of timely and adequate pain relief in vaso-occlusive crises to prevent re-admissions and of post ED follow up, and to workforce development efforts such that there are enough knowledgeable and trained providers to ensure adequate access for SCD patients to outpatient and preventative services.

### Other Suggested Surveillance Topics of Interest

Our stakeholder groups raised the above focus areas often. We feel analysis and dissemination on these areas are achievable within the scope of the planned data collection and are likely to have high impact on policy, clinical practice or other factors that drive patient outcomes. Many other areas of interest were also raised in these conversations; some may be candidates for additional analyses, while others may require collection of different data to address or other resources that are outside the scope of this surveillance effort. Some of these areas of interest are discussed below.

#### Health Care Quality for SCD in California

Stakeholders familiar with care patterns among those with SCD were concerned about systemic problems with care at the state level. Differences in access to care based on payer were cited above all else. Although no publications describe the situation specifically in California, stakeholders in our meetings described that the situation is similar to other states: adults with SCD who are on Medicaid have few or no physicians who will see them due to low reimbursement rates, physician unwillingness to manage patients with a highly complex disease, or physician unwillingness to prescribe opioids/narcotics to meet standards of care\textsuperscript{66,69}. The issue of specific challenges with enrollment in the state’s Genetically Handicapped Persons Program (GHPP), which pays for services and treatments not covered by Medicaid or private insurers, was raised frequently (RuSH data suggested that fewer than 10% of those who qualify enroll). People living with SCD and clinical SCD experts alike pointed to the inherent and systemic discrimination and bias against those with SCD in the health care
system and wondered how to quantify it. Their primary questions were as follows:

- How do treatments and outcomes differ by payer?
- What is the rate of enrollments for GHPP (state-based payer)?
- How do outcomes for GHPP recipients compare to outcomes for those not receiving benefits?
- How do medical reimbursement rates for SCD compare with other complex, chronic diseases (diabetes, cancer)?
- Is there a way to quantify systemic bias/discrimination in the health care system?
- Can Consumer Assessment of Healthcare Providers and Systems (CAHPS, from the Agency for Healthcare Research and Quality) data inform our understanding of health care quality for SCD?
- What is average time spent in emergency department waiting room before being seen? How does this compare to other similar diseases? Does it differ by facility type or other factors?
- What is the ED ‘left without being seen’ rate for SCD, and how does it relate to wait time?
- How will systemic changes in health care and payment structures in the Affordable Care Act impact Californians with SCD? Will the shift of Medicaid enrollees to managed care plans improve health care for the population, or will it present new challenges for access to care?

Quantifying and addressing these problems is an enormous challenge. Some efforts have been made to do this69-73, and more are needed. However the SCD LDC has limited data on costs of care outside of Medicaid claims, no information on care quality for those in the general population or those with chronic illnesses similar to SCD, nor encounter-level variables such as emergency department wait time. We feel the answers to these questions are best addressed by other means.

Costs of Care

As with quality of health care, there was a great deal of interest among stakeholders in understanding potential cost savings to insurers if standards of care were followed by all providers treating SCD, especially those treatments and practices that help prevent complications. Some investigation into these questions has taken place74-76. Stakeholder questions included the following:

- What are the costs of preventative treatment and cost savings associated with them?
- Is there a cost savings in treating SCD patients in a day hospital setting (for scheduled transfusions as well as management of pain episodes)?

Such analyses are possible as a part of the SCD LDC using the Medicaid claims data for identified SCD cases (for Medicaid-covered claimants only); these data include outpatient treatments, screenings, and use of preventative drugs, as well as the outcomes of interest. These are complex analyses, however. Such work is best undertaken when a plan for what to do with the results of such analyses to best impact policy and practice change is in place. For this reason, this topic is beyond the scope of this project for now.

Narcotic Use and Pain Management

Appropriate and adequate use of narcotics to treat pain is an important aspect of care for adolescents and adults with SCD. This is one of the most stigmatized and misunderstood areas of disease treatment and management, however. Many ED providers believe that patients presenting to the ED with specific requests for pain management are ‘drug seeking’ and many other physicians not familiar with SCD standard protocols express an unwillingness to prescribe the appropriate amount of medication to treat acute pain crises by SCD standard of care guidelines69,77,78. Stakeholders expressed an urgent need to understand how, where, what, by whom and how often pain medications are prescribed for SCD patients, and what the outcomes are based on adherence to SCD standards of care and Ameri-
Stakeholders expressed an urgent need to understand how, where, what, by whom and how often pain medications are prescribed for SCD patients, and what the outcomes are based on adherence to SCD standards of care. It is also clear that in California, a substantial portion of adult patients are not seen in clinical settings nor are they included in data from these sites.

Can Pain Society protocols. Their questions included the following:

- Can we quantify 'normal' narcotic use in SCD?
- Is 'high' narcotic use correlated with high utilization at a population level?
- Are narcotics prescribed to meet standards of care for SCD?
- What are challenges to patients of obtaining narcotic treatment in different settings?

As critical as these questions are to addressing health, quality of life and outcomes among those living with SCD, we feel strongly that the resources California has been awarded are insufficient to address this issue at this time. This is one of the most complex research areas affecting those with SCD, and it would be well served by a multi-agency cooperative effort.

Transfusions

Stakeholders said it was important to know how many SCD patients receive transfusions, rates of chronic vs. intermittent transfusion, outcomes associated with transfusion, development of antibodies and infection due to transfusion, whether standards of care are being followed in determining whether to transfuse and how blood should be handled, and costs associated with the procedure compared to costs of the complications prevented. A high proportion of patients will be transfused over their lifetime; CA RuSH data suggests that over 70% of those with SCD will receive one or more transfusions during a five-year period. This treatment is critical for preventing many of the life-threatening complications of the disease. Specific questions raised by SCD stakeholders include the following:

- Are clinical settings adhering to standards of care for transfusion in SCD treatment?
- How many patients are chronically transfused? In what setting? How often?
- How many patients are intermittently transfused and under what circumstances?
- How many units do patients receive? How does this differ by circumstance (chronic, ED, surgery)?
- How do transfusions impact short- and long-term outcomes?
- What are patient genotype differences in need for transfusion?
- Is it feasible to create a database of antibodies that different blood banks and transfusing hospitals can use to avoid antibody-induced transfusion reactions?
- What are the rates and severity of transfusion reactions (e.g., TACO, TRALI, hives, exposure to infectious agents) and adverse outcomes?

As noted in the 2014 California PHRESH Validation Report, identifying and understanding transfusions in administrative and claims data proved to be more challenging than anticipated. Transfusions (both straight red blood cell and exchange transfusions) are billed differently in diverse settings (day hospital vs. inpatient, for instance), for different procedures/purposes, and across health care provider and insurance systems. None of the available data include number of units transfused during a procedure. Few data sources include information on non-life threatening transfusion reactions or acute/immediate adverse outcomes. In short, given the data we plan to collect, answering the questions stakeholders proposed would be challenging.

In 2014 the CDC awarded funds to a consortium of researchers and state agencies in Georgia to investigate many of these issues, with additional
funding awarded to sites in California and Florida in late 2015. The SCD LDC will work with the California site to support their collection and analyses of data related to transfusions in SCD; we hope to answer some of these questions during this effort.

Quality of Life, Burden of SCD

Chronic illness affects not just the body, but the whole person, family and community. Stakeholders hoped we could explore how SCD brings challenges to having a productive life. Their questions focused primarily on 'countable' variables such as the following:

- How do SCD patients’ employment and educational histories and status differ from the general public?
- How many days of work or school are missed annually (by age group) due to medical care and health related problems among those with SCD compared to the general population?
- What is the financial and quality of life burden of these challenges?

We believe that some data on school (K-12) attendance exists, and we are working to identify the data sets and determine whether they are linkable to LDC data. These data may enable understanding of the effect of SCD on school attendance; the other questions, however, are best addressed via qualitative methods such as questionnaires.

Population-Based Surveillance of Preventative Treatments/Screenings and Outcomes

Researchers typically conduct extensive clinical research before treatments and screenings become standard in clinical settings. The efficacy of treatments such as hydroxyurea, prophylactic penicillin and screenings such as transcranial Doppler use for stroke risk is well known, therefore. Effectiveness in the entire SCD population, however, has not been determined. The drop in infant and child mortality that followed universal new-born screening and subsequent wide adoption of penicillin for the first five years of life is widely attributed to that practice, but no such clear indications exist for outcomes stemming from the increased use of hydroxyurea, specialized vaccinations such as the pneumococcal vaccine, or other practices. Stakeholders were interested in the following related questions:

- How has use of hydroxyurea changed over time? At the population (effectiveness) level, what is the change in outcomes that can be attributed to its use?
- How has penicillin prescription and adherence changed over time? What proportion of the population remains on penicillin after age five, and how do outcomes differ compared to those who discontinue at age five?
- What proportion of eligible children has had regular transcranial Doppler screenings, and what is the impact on stroke rate in children?
- What proportion of children has had recommended vaccines, and what is the impact on health for these children vs. those with SCD who do not have all vaccines on time?

Monitoring of treatments and screenings and their relationship to outcomes can only take place when data are complete. Pharmacy data and outpatient screenings only appear in insurance claims data among our data sources, and we do not plan to have claims for sources beyond Medicaid (i.e., private insurer claims). This means there is a potential for gaps in the record that would distort conclusions based on these data – a child may be ineligible for Medicaid insurance for a period of several years, for instance, with no record of vaccinations or screenings. Did she have them while covered under private insurance, or did she fail to receive them? We feel that with the expected data sources, the completeness of these data is insufficient for answering these questions. We suggest that collection of data from a wider variety of payers would address this problem, but is beyond the scope of this project. However, data from the SCD LDC are ideal for monitoring the large scale changes in SCD population health over time as new treatments become available and as new treatment guidelines and health care policies create change. Such changes, including reduced
childhood morbidity and a greatly increased life expectancy, came about after the introduction of universal newborn screening for SCD and the widespread use of prophylactic penicillin for the prevention of childhood infection.

**Rates of Complications**

Some stakeholders wondered whether there was enough information about the prevalence and rates of complications associated with SCD, such as stroke, other neurologic complications, and acute chest syndrome. Their questions included the following:

- What are rates of pediatric and adult strokes, time to subsequent stroke by age and preventative measures taken, and long-term implications for health?
- Can we describe neurologic complications beyond stroke?
- Acute chest syndrome and pneumonia rates – have these changed over time with treatments/standards of care?

Using single source administrative data sets such as hospital discharge data or medical claims data, previous researchers have published extensively on rates of complications among SCD patients. Rates of pediatric stroke have been assessed by age, as have pulmonary complications of SCD, avascular necrosis, and other common conditions in this population. Because data on most of these conditions are found in inpatient hospitalization data and most states have complete data in these files, we do not believe we can add information substantially different from that already published in these areas.

**Fertility and Pregnancy**

Interest in fertility and pregnancy included questions about whether adult women with SCD have complications in getting pregnant and having healthy babies compared to the general population, whether there were threats to the health of mothers with SCD during pregnancy, interventions or settings that improve pregnancy outcomes in SCD, and rates of teenage pregnancy among those with SCD. Some of these questions have been addressed in clinic-based studies by other researchers. We may be able to add to the information on this topic. It is unclear, however, how this information would be used to change policy or outcomes for people living with SCD.

**Sickle Cell Trait**

There is a high level of interest among some policy makers and the general public about describing health problems that may be associated with sickle cell trait. Linking the approximately 80,000 sickle cell trait cases identified to date in California with health care records is outside the scope of this project, however.

**Timeline**

The proposed detailed timeline for the California SCD LDC Program follows. These plans are subject to the availability of planned sources of data, availability of resources and staffing, and changes in plan due to emerging and mutually agreed upon areas of interest in SCD surveillance in the state. Examples of such changes include the approval of new treatments, policy and practice changes by insurers or other changes in patient care.

**Year 1: Design of data collection system**

We intend to use the first year (April 2015-March 2016) of the project for information gathering, infrastructure development and planning for data collection. In the first four months of the project, we met with over 20 stakeholder groups to gather information about high priority goals and objectives for surveillance of SCD at the state level. We will continue to work with these stakeholders in group webinar con-
versations on a quarterly basis as well as in small conversations as needed for input.

We have developed the surveillance agenda described here for discussion. We have also begun work on a data collection protocol that will soon be submitted to the state’s Committee for the Protection of Human Subjects for clearance to begin data requests. We will follow this clearance by initiating a request for data from the state’s Medicaid office, which historically has been the most challenging and slow when seeking to obtain data. This will be followed by data requests from the state’s newborn screening program and vital records agency, as well as updates on emergency room and hospital discharge data through 2014 when available (CRDS already has possession of these data through 2013, pending permission to use for this project). Over time, we will request subsequent years of data and will also explore and request additional data sources that can fit well within the developed infrastructure and add to our knowledge about this population. We will further develop the case definitions for SCD validated during the PHRESH project.

Since the beginning of the funding period, we have hired a part-time health communications expert to help with external conversations and dissemination of data. Our database development has begun, using in-house staff and expertise. We have identified a data linkage consultant with over 10 years of experience linking data such as those we plan to collect, and plan to develop a sub-contract to work with him throughout Year 2 of the project. Security for the data is already in place through CDPH physical and procedural data protection systems. We anticipate that we will begin linking and reviewing some data by mid-winter, prior to the end of Year One.

**Years 2-4: Implementation of data collection system**

We anticipate that completion of data cleaning and standardization across all years of data (2004-2014), linking and de-duplication, application of the case definition and checking of anomalous cases, and development of the database and query systems will take much of Year Two. This process will be closely monitored by the CDC Division of Blood Disorders project officer, with interim milestones established. Once the data systems have been developed, new data will be incorporated annually (administrative data) or every six months (newborn screening, Medicaid claims) as available, through the remainder of the project.

With additional funding, we anticipate requesting data from clinical sites with which we already have established working relationships. These data will also be incorporated into the system, as they were in RuSH and PHRESH data.

Finally, we will remain in regular contact with stakeholders, continuing quarterly meetings and other forms of communication.

**Years 3-5: Data analysis and information dissemination**

As the data are processed to become health care and outcome profiles for each patient in the state, we will plan for analysis and dissemination. A meeting with stakeholders to discuss preliminary results of data analyses will be held in Year Three (along with continued quarterly web conference meetings). We plan to use the meeting to develop a list of fact sheets, presentations and publications to be completed.

Dissemination will proceed as planned and with on-going input from stakeholders. In Year Four we will create a description of how we developed and implemented the data collection and linking systems to inform those who wish to replicate this work. It will be refined and completed in Year Five.

We will hold a final meeting with stakeholders in Year Five to plan for next steps in SCD surveillance at the state level.

**Years 6 and Beyond: Continued growth and expansion of surveillance in sickle cell disease**

Community based organizations, national professional organizations, researchers, clinicians and members of the affected community continue to encourage Congress to devote dedicated funds to sickle cell disease surveillance at the national level. It is our belief that these efforts will be fruitful and that demonstration of the utility and power of a longitudinal data collection system for this disorder in California will inform these efforts.
Beyond California, it is vital that other states also engage in statewide, unbiased surveillance of their SCD affected populations. Each state has a unique demographic makeup, distinct policies and implementation of Medicaid and ACA, medical and research centers, and access to care. Our intention is to work closely with other states that have funding to do this work to inform all states’ efforts and compare and contrast differences in care and outcomes across states.

Until recent surveillance and health education efforts from the CDC, SCD had been seen primarily as an important clinical issue but had not been addressed as a public health concern. Although SCD impacts an estimated 7,000 Californians, little is known about the health status of those living with the disease, particularly adults. This longitudinal data collection project is intended to follow all patients over time, including those not seen in specialty care centers and not identified by newborn screening. The profile of healthcare utilization, complications, demographics and access to care created by the collected data over time will allow us to investigate and report on key areas of interest, including patient geography and proximity to care, healthcare outcomes associated with transition from pediatric to adult care, the emerging Hispanic/Latino SCD population, the health status of those who live into their 40’s and beyond with SCD, and triggers and outcomes for those patients who experience periods of high emergency room utilization.
References


27. Cluster S, Termuhlen A, Schrager S, et al. Pitfalls of Using Administrative Data Sets to Describe Clinical Care in Sickle Cell Disease. 7th Annual Sickle Cell Disease Research and Educational Symposium and 36th Annual National Sickle Cell Disease Scientific Meeting; April 14-17, 2013, Miami, FL.


Appendix A: SCD LDC June-July 2015 Stakeholder Meetings and Attendees

- American Society of Hematology
- Association of Public Health Laboratories
- Blood Centers of the Pacific/Blood Systems Research Institute
- California Healthcare Foundation
- Center for Inherited Blood Disorders
- Children's Hospital Los Angeles
- Children's Hospital Orange County
- Health and Human Services/Office of Minority Health
- Health Resources and Services Administration/Maternal and Child Health Bureau
- Kaiser Permanente Southern California
- The KIS Foundation
- Loma Linda University Medical Center
- National Institutes of Health/National Heart, Lung, and Blood Institute
- National Medical Association
- Northern California Sickle Cell Advisory Council
- Sickle Cell Disease Association of America
- Sickle Cell Disease Foundation of California
- UC Davis Hematology Clinic
- UC Irvine Medical Center
- UCSF Benioff Children's Hospital Oakland
- Parents of children living with sickle cell disease
- Adults living with sickle cell disease
- Expert on hemophilia Universal Data Collection system
- Los Angeles City Councilman Sponsored Meeting on SCD