OFFICE OF THE SECRETARY
PATIENT-CENTERED OUTCOMES RESEARCH TRUST FUND

PROJECT REPORT

FINAL REPORT

Exploring Data Infrastructure Availability and Expansion Opportunities for Health Outcomes Research on Sickle Cell Disease

Prepared for
The Office of the Assistant Secretary for Planning and Evaluation (ASPE)
at the U.S. Department of Health and Human Services

by
NORC at the University of Chicago

November 2025

Office Of The Assistant Secretary For Planning And Evaluation

The Assistant Secretary for Planning and Evaluation (ASPE) advises the Secretary of the U.S. Department of Health and Human Services (HHS) on policy development in health, disability, human services, data, and science; and provides advice and analysis on economic policy. ASPE leads special initiatives; coordinates the Department's evaluation, research, and demonstration activities; and manages cross-Department planning activities such as strategic planning, legislative planning, and review of regulations. Integral to this role, ASPE conducts research and evaluation studies; develops policy analyses; and estimates the cost and benefits of policy alternatives under consideration by the Department or Congress.

Office Of The Secretary – Patient-Centered Outcomes Research Trust Fund

The Office of the Secretary Patient-Centered Outcomes Research Trust Fund (OS-PCORTF) was established as part of the 2010 Patient Protection and Affordable Care Act and is charged to build data capacity for patient-centered outcomes research. Coordinated by ASPE on behalf of the Department, OS-PCORTF has funded a rich portfolio of projects to meet emerging U.S. Department of Health and Human Services policy priorities and fill gaps in data infrastructure to enhance capabilities to collect, link, and analyze data for patient-centered outcomes research. For more information, visit https://aspe.hhs.gov/collaborations-committees-advisory-groups/os-pcortf.

This report was funded by the Office of the Secretary Patient-Centered Outcomes Research Trust Fund (OS-PCORTF) under Contract Number HHSP233201500048II of the HHS Office of the Assistant Secretary for Planning and Evaluation (ASPE). The work was carried out by NORC at the University of Chicago and ASPE. The authors are solely responsible for this document's contents, findings, and conclusions, which do not necessarily represent the views of HHS, ASPE, or NORC. Readers should not interpret any statement in this product as an official position of ASPE or of HHS.

Suggested Citation: Srinivasan M, Shah A, Peterson C, Aronoff A, Eloso J, and Dullabh P. Exploring Data Infrastructure Availability and Expansion Opportunities for Health Outcomes Research on Sickle Cell Disease. Washington, D.C.: Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services. November 2025.

Contributing Authors

Mithuna Srinivasan, PhD, Principal Research Scientist, NORC
Avantika Shah, MPH, Research Scientist, NORC
Caroline Peterson, MPH, Senior Research Associate II, NORC
Abigail Aronoff, MPH, Research Associate II, NORC
Jessica Eloso, PharmD, BCPS, Oak Ridge Institute for Science and Education Fellow, ORISE
Prashila Dullabh, MD, Vice President and Senior Fellow, NORC

Project Officers and Project Leadership

Jessica Eloso, PharmD, ORISE
Jesse Anderson, MPH, ASPE
Marcos Trevino, MPH, ASPE
Prashila Dullabh, MD, NORC
Mithuna Srinivasan, PhD, MA, NORC

Acknowledgements

The authors would like to acknowledge ASPE staff who provided feedback on previous versions of this report, including Adrienne Brown, Kang Shen, Jennifer Wiltz, Vetisha McClair, Scott R. Smith, and Marcos Trevino, as well as the project consultants and key informants listed below.

Consultants

Amanda Brandow, DO, MS, Medical College of Wisconsin
Lakshmanan Krishnamurti, MD, Yale School of Medicine

Key Informants

Sherif Badawy, MD, MS, MBA, Ann & Robert H. Lurie Children's Hospital of Chicago
Mandy S. David, PA-C, MPH, Sickle Cell Public Health Analyst
Titilope Fasipe, MD, PhD, Texas Children's Hospital
Julie Kanter, MD, University of Alabama at Birmingham
Sophie Lanzkron, MD, Thomas Jefferson University
Julie Panepinto, MD, MSPH, National Heart, Lung, and Blood Institute
Mariah Scott, MS, MPH, Rutgers University
David Wong, MD, Office of the Assistant Secretary for Health

This project was supported in part by an appointment to the Research Participation Program at the Department of Health and Human Services administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the U.S. Department of Energy and the Department of Health and Human Services.

Table of Contents

Table of Exhibits

Executive Summary

Background. More than 100,000 people in the United States live with sickle cell disease (SCD), a genetic hemoglobin disorder associated with severe acute and chronic complications, including pain, cardiovascular disease, and organ damage. Life expectancy for individuals with SCD is 25 years shorter than the national average—52.6 years vs. 77.5 years. While existing and newly developed treatment options offer potential benefits, data on long-term healthcare utilization and outcomes remain limited. There are significant gaps in data scope and quality of SCD data infrastructure for research, including underutilization of patient-reported outcomes (PROs), data missingness, inconsistent coding, and lack of interoperability across state-based registries and administrative data sources.

Objective. The objective of this report is to serve as a data source guide that helps policymakers and researchers understand the current data infrastructure landscape for SCD-focused health outcomes research. Specifically, the goals are to identify existing SCD data sources and data dashboards as well as key considerations to enhance data collection and use. This report is intended to inform federal, academic, and independent researchers on relevant resources for SCD-focused health outcomes research. Furthermore, considerations in this report can guide future SCD data infrastructure investments.

Methods. We first conducted an environmental scan of peer-reviewed and gray literature to identify existing federal and nonfederal SCD data sources. The scan also informed the development of a visual framework to illustrate the SCD data infrastructure and a preliminary list of gaps and opportunities to improve the infrastructure. We then conducted eight virtual key informant interviews with federal stakeholders and SCD researchers to gather feedback on the environmental scan findings and visual framework, identify additional data sources, and learn about considerations for strengthening the SCD data infrastructure.

Results. We identified 43 data sources and five data dashboards that could be leveraged for SCD health outcomes research. Within these data sources, we also catalogued the availability of health outcomes data, including PROs. Notably, only 16 sources (37%) captured any PROs. Additionally, only 10–20% of sources capture psychological factors, environmental risk factors, medication adherence, reproductive health, and emerging therapies.

Conclusion. The environmental scan and key informant interviews identified several considerations for enhancing the available data infrastructure for SCD-focused health outcomes research. In the short term, stakeholders should prioritize the identification and use of core data elements and promote the consistent application of standardized definitions to improve the utility of existing datasets. Over the longer term, establishing systematic approaches to link data across the lifespan will be essential to enable longitudinal analyses and reduce fragmentation. Additional priorities include streamlining data collection efforts to minimize burden on patients and providers, as well as improving the capture of key patient-reported outcomes. An emerging opportunity lies in the use of remote data collection tools and methods to generate real-time, patient-centered data. These technologies may offer a scalable solution to long-standing gaps in SCD research. Central to all of these efforts is the need for collaborative, sustained engagement among multiple partners to ultimately ensure quality and relevance of data infrastructure that is responsive to the needs of the community.

1. Introduction

More than 100,000 people in the United States live with sickle cell disease (SCD), an inherited hemoglobin disorder caused by mutations in the β-globin.[1,2] Individuals with SCD experience significant sickness related to severe acute and chronic pain, cardiovascular disease, and organ damage, contributing to lower life expectancy—approximately 25 years less than the national average (52.6 years vs. 77.5 years).[3,4] SCD management includes pharmacological therapy, such as hydroxurea for pain management, and non-pharmacological therapies, such as cognitive behavioral therapy for the psychosocial management of chronic pain.[5,6] Other SCD treatments include blood transfusions, transplants, and cell and gene therapies (CGTs). While novel CGTs and other specialized treatments can have transformative health outcomes for individuals with SCD, there is little existing health outcomes research data on long-term treatment utilization and impacts.[7,8] Recognizing this, the American Society of Hematology (ASH) identified priority topic areas, updated in 2024, to advance SCD research in the next five years to improve health outcomes.[9]

To advance health outcomes research, numerous federal and nonfederal efforts have focused on improving data collection and strengthening the overall SCD data infrastructure, including individual registries and surveillance networks. Nevertheless, there is no national SCD surveillance effort, which contributes to significant knowledge gaps on care transitions, health care utilization, and patient-reported outcomes (PROs) for the SCD population.[10,11] The lack of a national surveillance system also poses challenges for SCD health outcomes researchers conducting large-scale population-based studies. Often, they must rely primarily on administrative data, which can cause vailidity issues related to timeliness, loss to follow up, data missingness, and misuse of diagnostic codes.[12] These concerns with data quality further impact data linkages, which multistate registries rely on. Data quality issues are further exacerbated by state-level variability in reporting administrative data, electronic health record (EHR) data, and birth and death data.[13]

This report qualitatively assesses the existing data infrastructure for conducting health outcomes research for the SCD population and identifies gaps and opportunities for enhancing or expanding the SCD data infrastructure. Through this report, researchers can gain a clearer understanding of what data and tools are available for SCD-focused research, and where limitations exist. This report also provides important considerations for strengthening the data systems that underpin SCD research and care, to enable improved research and evidence-based decision making by healthcare professionals and policymakers.

1.1 Objectives and Structure of the Report

This report describes the findings of a qualitative assessment of the SCD data infrastructure, as well as considerations for expansion and enhancement. This project is guided by the following research questions:

1. What existing data sources can researchers use to study SCD health outcomes?
2. What are the gaps in the SCD data infrastructure and opportunities for improvement?

The findings of this report will result in a data source guide intended to support federal, academic, and independent researchers and policymakers interested in SCD-focused health outcomes research. The subsequent sections of the report are organized as follows: the ***Methods*** section provides an overview of the data collection methods, including the implementation of an environmental scan and engagement with subject matter experts to identify SCD data sources; ***Research Domains and Data Elements Derived from ASH Research Priorities*** describes how we used the ASH priorities (2024 update)

as our guiding framework for identifying important data elements within identified data sources; the ***Findings on Data Sources*** section identifies and describes a) data sources that contain individual-level data at the person or encounter level available for research and b) data dashboards and/or tools that include compiled metrics or descriptive statistics in a visual or tabular form available for public consumption); the ***Discussion*** section describes gaps and opportunities to enhance the data infrastructure for SCD health outcomes research; and the report ends with a ***Conclusion*** section that summarizes findings and impacts for the SCD data infrastructure. The report also contains supplementary appendices described in the following sections.

2. Methods

Our approach to identifying existing data sources with individual-level data and dashboards and considerations for enhancing the SCD data infrastructure included two key activities. First, we identified data sources through an environmental scan. Next, we sought feedback on the environmental scan findings from subject matter experts by conducting virtual, semi-structured key informant interviews with federal stakeholders and SCD researchers. Both activities are detailed below.

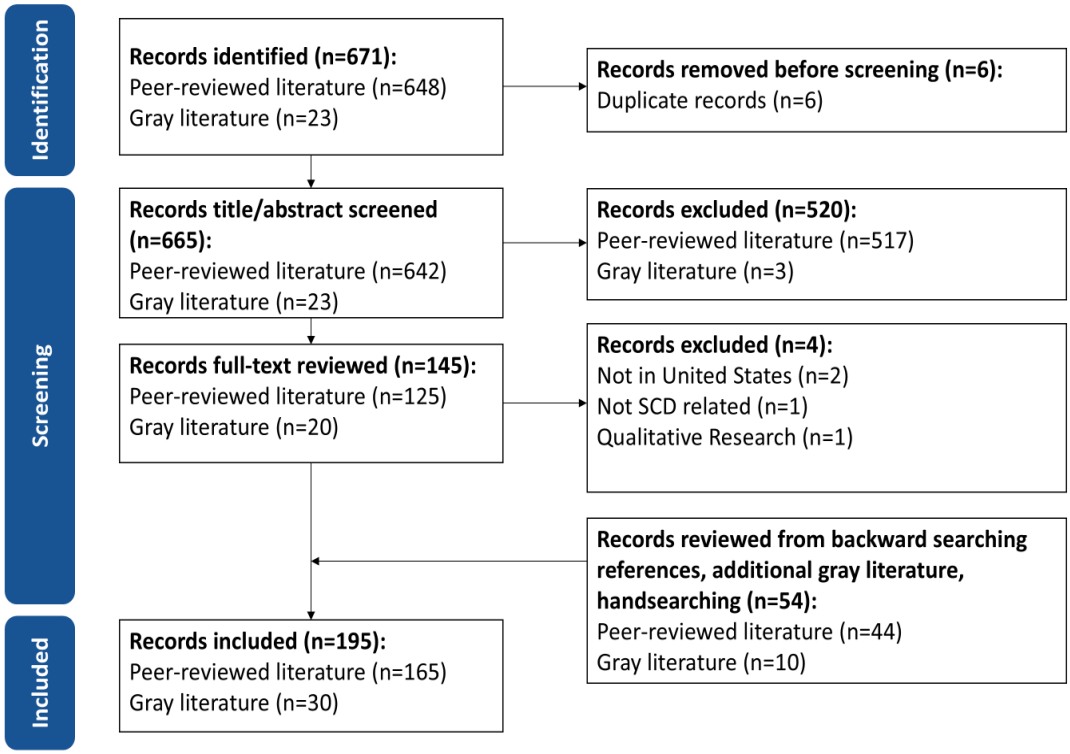
2.1 Environmental Scan

An environmental scan was conducted to systematically identify and describe existing data sources (administrative data, registries, and surveys) that can be leveraged to study health outcomes for individuals with SCD. Additionally, the scan identified gaps in the current data infrastructure and opportunities for improvement.

We conducted a web-based search for existing SCD-related data sources, supplemented by a targeted search of peer-reviewed and gray literature to identify additional relevant datasets. In conducting the web scan for SCD data sources/tools, we prioritized HHS agency websites, national professional associations, and nonprofit organizations focused on SCD. A list of prioritized websites is in **Appendix A, Exhibit A1**. After completing the web scan, targeted searches were conducted in PubMed, Google Scholar, and ClinicalTrials.gov to identify peer-reviewed literature reviewing or using SCD-relevant data sources. **Exhibit 1** illustrates the article selection process. Articles identified through the literature review underwent a two-step screening process: an initial title and abstract review, followed by a full-text review for eligibility. We applied search terms and inclusion/exclusion criteria as detailed in **Appendix A, Exhibits A2 and A3**, respectively.

Information was abstracted into an inventory (**Appendix A, Exhibit A4**) that was accessible to all team members to ensure transparency and collaborative input. A designated team member (who did not perform the original abstraction) validated the accuracy and consistency of the abstractions.

Exhibit 1. Article Selection Process



2.2 Engagement with Subject Matter Experts

We conducted eight virtual, semi-structured key informant interviews during March 2025 to solicit feedback on the environmental scan findings. Key Informants included individuals with SCD expertise from the federal government and university-based researchers, including researchers with lived experience. Interviewees reviewed the inventory of data sources and dashboards, suggested additional sources, and commented on data infrastructure gaps and opportunities for improvement. The interview questions are included in **Appendix B**. We performed a thematic analysis of the interview data, developing preliminary codes a priori based on the interview protocols and anticipated themes, and inductively and deductively coded each interview for these prespecified themes. Emerging themes were identified during the coding process when new patterns relevant to the research objectives appeared that were not previously identified.

3. Research Domains and Data Elements Derived from the ASH Research Priorities

The ASH SCD Research Priorities – 2024 Update[9] served as a guiding framework for identifying key research topics relevant to the population with SCD. Drawing from the seven broad priority areas outlined in the update—spanning basic, translational, clinical, health services, and implementation research—we identified the priorities most applicable to health outcomes research. We distilled relevant data elements from each of these priorities to identify within inventoried data sources (shown in **Appendix C**). These data elements, organized into three domains, are shown in **Exhibit 2**.

Exhibit 2. Identified Research Domains and Data Elements based on selected ASH SCD Research Priorities

Research Domain	Domain Description	Relevant Data Elements
Assessment of Disease-Related Risk Factors and Biomarkers	Evaluates risk factors that influence disease severity, progression, and patient outcomes	• Genotype • Sociodemographic characteristics • Comorbidities • Laboratory test results • Imaging results • Psychological risk factors • Environmental risk factors
Assessment of Health Outcomes	Measures and analyzes health outcomes, including those that are patient-reported for individuals with SCD	• Mortality • Health care costs • PROs • Comorbidity progression • Organ function • Accessibility • Reproductive health and fertility outcomes • Medication adherence • Pregnancy outcomes • Pain
Assessment of SCD Treatments and Therapies	Evaluates the use, access, and efficacy of various pharmacological and nonpharmacological treatments and therapies	• Pharmacological therapy • Blood transfusions • Transplant regimens • Gene therapy • Nonpharmacological therapies

4. Findings on Data Sources

Through the environmental scan and key informant interviews, we identified 43 data sources with individual-level data and five data dashboards and/or tools that can be used to study health outcomes for the population with SCD. The complete list of data sources are in a separate inventory accompanying the report on https://aspe.hhs.gov/. The inventory provides information on several characteristics of each included data source. These characteristics are listed in **Exhibit 3** below.

Exhibit 3. Data Source Characteristics Included in the Inventory

Data Source Characteristics	
• Data Source Name • Data Source Type • Timespan of Data • Data Steward/Funder • Does the Source Include International Data? • Geographic Coverage of the Data in the U.S. • Is the Data Representative at Its Unit of Geographic Scope?	• Was Source Developed Specifically for SCD? • Data Accessibility • Periodicity of Data Collection • Example Variables Pertinent to SCD-Focused Health Outcomes Research • Inclusion of PROs? (Y/N) • Description of PROs • Publications Related to SCD Health Outcomes Research

These data sources are summarized in the visual summary shown in **Exhibit 4**, which also illustrates the main funders of these sources spanning federal and state agencies, foundations, professional societies, and other organizations. (Acronyms in the framework are defined in **Appendix D**.) The framework also illustrates the extent to which the data elements shown in **Exhibit 2** are available in the identified data sources, with high-coverage variables (that is, variables that are represented in 50-100% of data sources) shaded in dark green, medium-coverage variables (in 30-49% of data sources) in a medium shade of green, and low-coverage variables (in 0-29% of data sources) in light green. In what follows, we provide details about the identified data sources and dashboard.

Exhibit 4. Visual Summary of the SCD Data Infrastructure

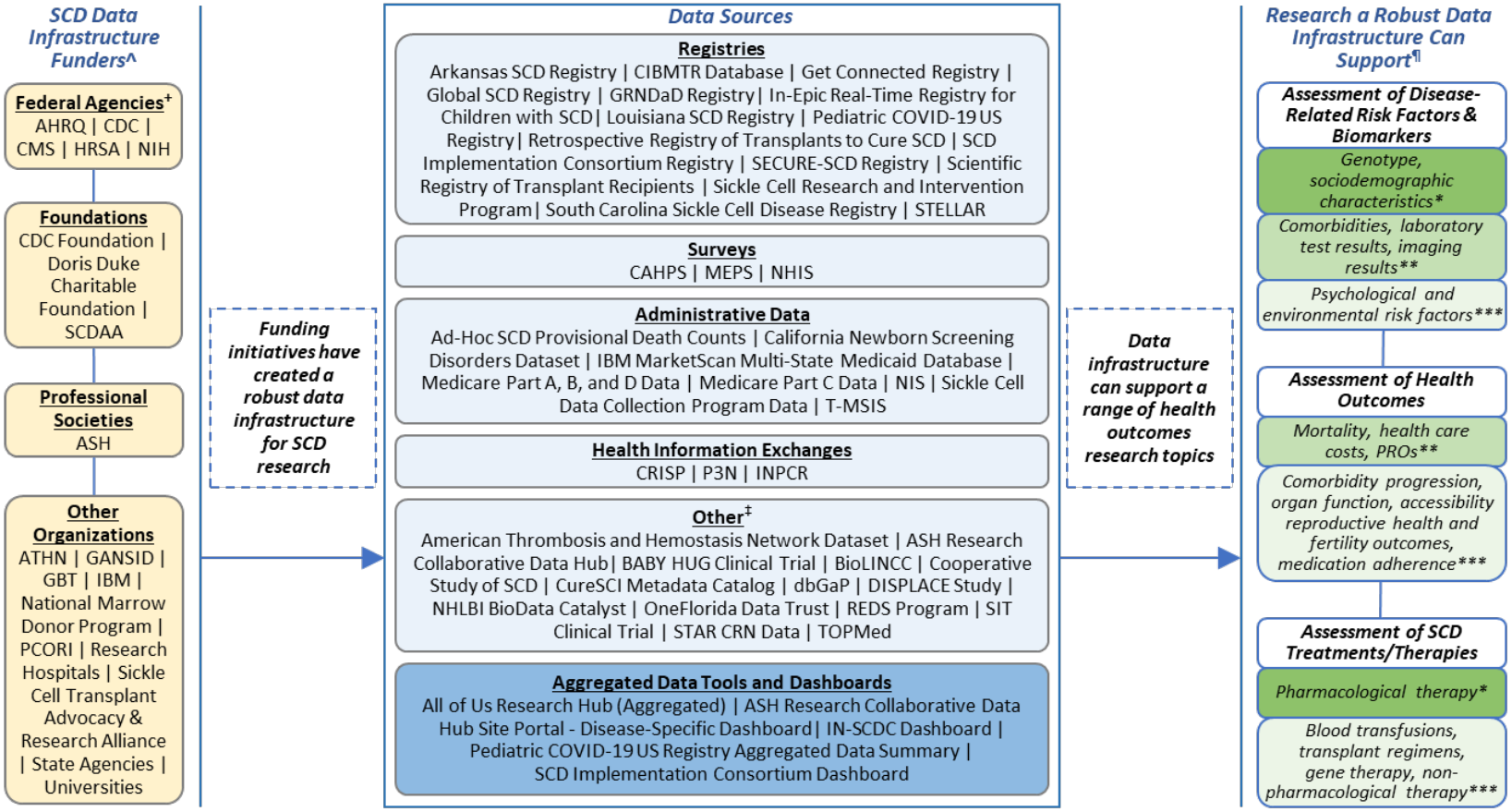
Access the full Data Sources Inventory here: https://aspe.hhs.gov/

Key Color

*High coverage (found in 50-100% of data sources)

**Moderate coverage (found in 30-49% of data sources)

***Low coverage (found in 0-29% of data sources)



^ Funders who fund and lead data infrastructure and SCD research.

+ The Office of the Secretary Patient-Centered Outcomes Research Trust Fund provides funding to a range of federal agencies that conduct SCD-focused research, including AHRQ, CDC, CMS, and NIH.

‡ The "Other" category constitutes data sources such as research studies, networks, and repositories.

¶ These research areas were identified from ASH's Sickle Cell Disease Research Priorities – 2024 Update, published September 25, 2024. Retrieved from: https://www.hematology.org/research/sickle-cell-disease-and-sickle-cell-trait

4.1 General Characteristics of Data Sources

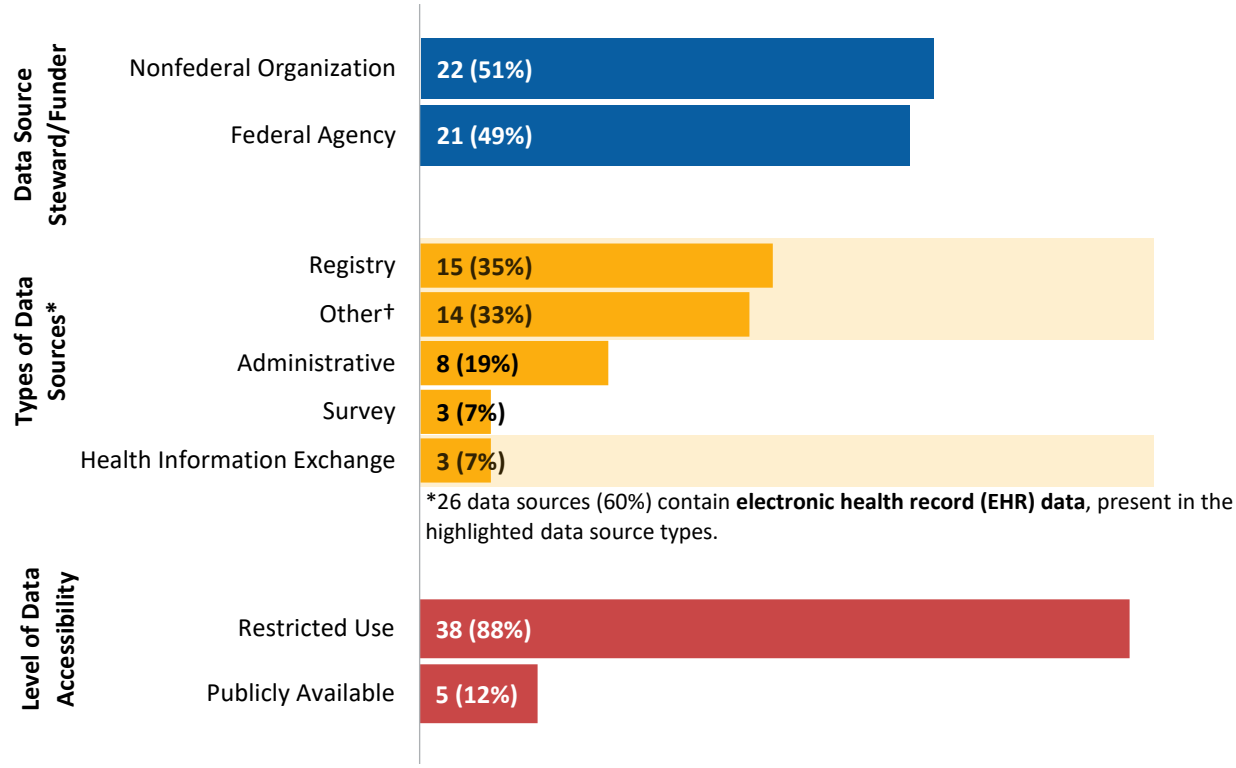
The data inventory categorized each of the 43 data sources that provide individual-level data by their general characteristics, such as data funder, type of data source, and level of data access, shown in **Exhibit 3**. A slightly higher proportion of data sources was funded by nonfederal organizations compared to federal agencies. The most common federal funding source was the National Institutes of Health (NIH), which funded 12 data sources—10 through the National Heart, Lung, and Blood Institute (NHLBI).[14] The most common **type of data source** identified was registries, which accounted for 15 data sources.

Identified registries (n=15) include:

- *10 multistate registries*
- *3 statewide registries*
- *1 site-specific registry*
- *1 nationally representative registry*

In terms of **data source accessibility, Exhibit 5** shows that most data sources are classified as restricted use (88%). These sources require users to either be affiliated with a participating site, contribute data, or establish a formal data use agreement with the data holder. However, all NIH-funded studies and data sources can be made accessible to the public through data use agreements and are housed in data repositories such as BioLINCC and BioData Catalyst. In contrast, five data sources were publicly available (that is, their data files are readily downloadable by researchers without review or permission).

Exhibit 5. General Characteristics of SCD Data Sources



*26 data sources (60%) contain **electronic health record (EHR) data**, present in the highlighted data source types.

† The "Other" category constitutes data sources such as research studies, networks, and repositories.

4.2 Identifying the Population with SCD in Data Sources

Newborn screening disorders data provide an opportunity to examine early life identification and outcomes for children born with SCD. All states are required to collect newborn screening data, but California is the only state to publicly report the data.

Of the 43 identified data sources, almost half (47%) were developed specifically for SCD. Examples of SCD-specific data sources with ongoing multistate data collection include the Globin Research Network for Data and Discovery (GRNDaD) registry,[15,16] ASH Research Collaborative Data Hub,[17,18] and Sickle Cell Disease Implementation Consortium (SCDIC) registry.[19] Data sources that were developed specifically for SCD identify the population with SCD by default. **Exhibit 6** shows how individuals with SCD can be identified across data sources not specifically developed for SCD. Identification of the population with SCD in these data sources relies most commonly on diagnosis codes (e.g., ICD-10 D57.x), though some surveys include self-reported SCD status. Other data sources like research studies and networks may support more nuanced phenotyping approaches.

Exhibit 6. Identifying the SCD Population in Data Sources Not Developed Specifically for SCD

Data Source Type	Example Data Sources Not Specifically Developed for SCD	SCD Identification Method
Registries	Center for International Blood and Marrow Transplant Research Database Scientific Registry of Transplant Recipients	Diagnosis codes
Administrative	T-MSIS, IBM MarketScan, Medicare	Diagnosis codes
Surveys	Consumer Assessment of Healthcare Providers and Systems, Medical Expenditure Panel Survey	Self-reported status
Health Information Exchange Organizations	Chesapeake Regional Information System for Our Patients, PA Patient & Provider Network	Diagnosis codes
Other	OneFlorida Data Trust, Recipient Epidemiology and Donor Evaluation Study Program, American Thrombosis and Hemostasis Network Dataset	Diagnosis codes or self-reported status

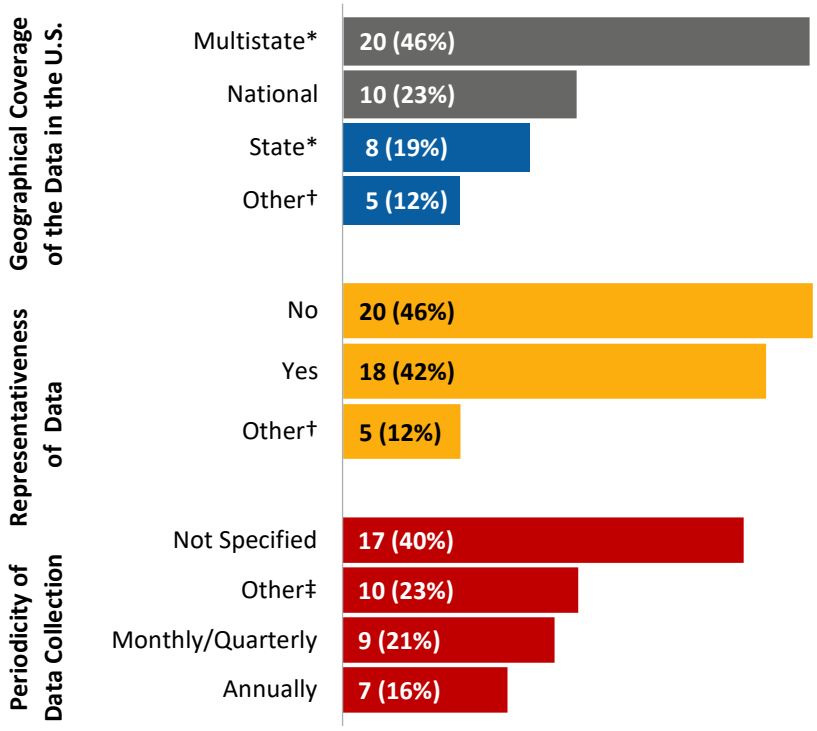
4.3 Geographic Scope, Representativeness, and Periodicity of Data Sources

The CDC's Sickle Cell Data Collection (SCDC) Program is the largest surveillance effort for SCD in the US, spanning 16 states. The program collects newborn screening and health care utilization data.

We categorized data sources by their geographic scope, representativeness, and periodicity of data collection, as shown in **Exhibit 7**. In terms of **geographic scope,** only 10 (or 23%) sources have national coverage. Most of these sources were not specific to SCD and consisted primarily of administrative data sources or large-scale surveys. The remaining data sources either contained data for a single state or a set of states (e.g., the CDC's Sickle Cell Data Collection [SCDC] Program). In terms of **representativeness**, a similar proportion of studies were representative of their unit of geographic coverage (20, or 46%) as those that were not (18, or 42%).

Information on the **periodicity of data collection** was available for only 26 data sources, or 60% (**Exhibit 7**). Of these, only seven (or 16%, and all surveys) collect data annually. There were nine data sources with quarterly or monthly collection, and they were mostly in the category of administrative or health information exchange organization data sources.

Exhibit 7. Geographic Scope and Periodicity of SCD Data Sources



* Some multistate/statewide data sources cover select cities, and certain multistate data sources include international regions.
† The "Other" category includes data repositories that comprise multiple studies with varied geographical coverage and data representativeness.
‡ The "Other" data sources include those that collect data on an ad hoc, weekly, or other basis not listed.

4.4 ASH Research Priorities-Based Data Elements Available in Data Sources

Exhibit 8 reports the number of data sources that include one or more data elements across the three research domains (see **Exhibit 2** for a description of these domains). Data element availability across research elements is color-coded by coverage across data sources: dark green for high coverage (50–100% of data sources), medium green for moderate coverage (30–49%), and light green for low coverage (0–29%) and summarized below:

- Among the 43 data sources reviewed, data elements related to **risk factors and biomarkers** were the most frequently represented. In this domain, at least half of the identified data sources included information on genotype and sociodemographic characteristics such as age, geographic location or insurance status. In contrast, psychological and environmental risk factors were far less common.

- In the **health outcomes** domain, mortality was the most commonly observed outcome (available in 18 sources, or 42%), followed by health care costs (present in 12 sources, or 28%). The least represented data elements were reproductive health and fertility outcomes (n=6), medication adherence (n=4), and pregnancy outcomes (not present in the data sources).
- For **treatments and therapies,** pharmacological therapy was the most commonly observed treatment (present in 23 sources, or 53%). The least commonly observed treatments were gene therapies (present in seven sources, or 16%) and nonpharmacological therapies (present in six sources, or 14%)

*Clinical- and utilization-related research elements are generally well represented across the identified SCD data infrastructure, while significant **gaps remain in psychological and environmental risk factors, medication adherence, reproductive health, and emerging therapies.***

Exhibit 8. Summary of Key Variables Relevant to SCD Research in Data Sources

Research Domain	Research Topic	Number (n=43)
Risk Factors and Biomarkers	Genotype*	26
	Sociodemographic characteristics*	23
	Comorbidities**	21
	Laboratory test results**	21
	Imaging results**	20
	Psychological risk factors***	9
	Environmental risk factors***	6
Health Outcomes	Mortality**	18
	Health care costs**	12
	PROs**	16
	Comorbidity progression***	10
	Organ function***	11
	Accessibility***	9
	Reproductive health and fertility outcomes***	6
	Medication adherence***	4
	Pregnancy outcomes	Unknown
Treatments and Therapies	Pharmacological therapy*	23
	Blood transfusions***	12
	Transplant regimens***	11
	Gene therapy***	7
	Nonpharmacological therapies***	6

Key Color

*High coverage (found in 50-100% of data sources)

**Moderate coverage (found in 30-49% of data sources)

***Low coverage (found in 0-29% of data sources)

4.5 Patient-Reported Outcomes Available in Data Sources

Data related to PROs are variably captured across the identified data sources. Only 16 data sources (37%) include one or more PROs. As shown in **Exhibit 9**, 11 out of 16 data sources include PROs related to quality of life (QoL). The surveys identified in the inventory, such as the Consumer Assessment of Healthcare Providers and Systems (CAHPS),[20,21] the National Health Interview Survey (NHIS),[22,23] and the Medical Expenditure Panel Survey (MEPS),[24,25] typically serve as useful sources of patient-reported data and include measures of functional status, mental health symptoms, and barriers to care. Registries, such as the GRNDaD registry[15,16] or the Sickle Cell Transplantation Evaluation of Long-Term and Late Effects Registry (STELLAR),[26] offer more detailed PRO assessments pertinent to SCD, including pain, fatigue, chronic graft-versus-host disease symptoms, economic impact of treatment, and sexual function. In general, most (27, or 63%) identified data sources either do not include PROs comprehensively or do not provide details on which specific PROs are collected. Registries and clinical trials could serve as sources for longitudinal PRO data. In contrast, national surveys such as CAHPS and NHIS primarily provide cross-sectional snapshots, limiting their utility for longitudinal analysis.

Exhibit 9. Examples of PROs Captured Across Individual-Level Data Sources

Outcome	Data Sources That Include the Outcome	Measures or Instruments Used for Data Collection (If Available)
Quality of Life	• GRNDaD • Arkansas Sickle Cell Disease Registry • Sickle Cell Clinical Research Intervention Program • STELLAR • Silent Cerebral Infarct Multi-Center Clinical Trial • MEPS • NHIS • Cure Sickle Cell Initiative (CureSCI) Metadata Catalog • The Database of Genotypes and Phenotypes (dBGaP) • NHLBI BioData Catalyst • Trans-Omics for Precision Medicine (TOPMed)	• Peds Quality of Life (QL) Inventory (generic, fatigue, SCD modules)[27,28] • Patient Reported Outcomes Measurement Information System (PROMIS) Pediatric Profile-25 survey[26,29] • Child Health Questionnaire Parent Form 50[30,31]
Pain	• GRNDaD • SCDIC • STELLAR • MEPS • CureSCI Metadata Catalog • dBGaP • NHLBI Biodata Catalyst • TOPMed	• PhenX Frequency of Sickle Cell Pain Episodes Per Year[32]
Mental Health Symptoms	• SCDIC • STAR CRN • MEPS • CureSCI Metadata Catalog	• ASCQ-ME Emotional Impact[33]
Fatigue	• GRNDaD • CureSCI Metadata Catalog	• PROMIS Fatigue[34]

Outcome	Data Sources That Include the Outcome	Measures or Instruments Used for Data Collection (If Available)
Functional Status	• SCDIC • MEPS • CureSCI Metadata catalog	• PROMIS Global Health[34]
Sexual Function	• STELLAR	• PROMIS Sexual Function SexFSv2.0 survey[34]
Economic Impact	• STELLAR • MEPS	• Dana Farber Cancer Institute Finances and Employment Scale
Missed Appointments or Barriers to Care	• SCDIC • STAR CRN • CAHPS • MEPS	• Standardized survey question
Medication Adherence	• SCDIC	• PROMIS Scale v1.0 – Medication Adherence[34]

4.6 SCD Data Dashboards

The environmental scan identified five dashboards that complement the previously described data sources by providing aggregate data and accessible insights through user-friendly visual formats. For example, the ASH Research Collaborative Data Hub Site Portal features a disease-specific dashboard with deidentified, aggregate clinical data from participating sites on individuals with SCD, including treatment patterns and health outcomes.[17,18] This data source is only accessible to participating research sites. The SCDIC Dashboard presents state-level aggregate data on health care utilization and complications among people with SCD using administrative and clinical sources.[35] The All of Us Research Hub includes some aggregate data on individuals with SCD, enabling exploration of demographic and health-related trends from a nationwide cohort.[36,37] **Exhibit 10** below summarizes characteristics of the five aggregate data sources identified and example visuals from these sources are in **Appendix E**.

Exhibit 10. Characteristics of SCD Data Dashboards

Data Source Name	Data Accessibility	Types of Aggregate Data Available	Format of Dashboard	Relevance to SCD Research
ASH Research Collaborative Data Hub: Disease-Specific Dashboard[17, 18]	Restricted use: data requests can be made to the ASH Research Collaborative	• Demographics • Risk factors and biomarkers, such as laboratory test results • Health outcomes, such as health care utilization • Treatments and therapies, such as pharmacological therapy	Contains charts and metrics across the following categories: • Patient characteristics • Acute care events • Cerebrovascular health • Other sequelae • Pharmacotherapy • Renal health • Transfusions	This dashboard allows users to monitor outcomes and evaluate care gaps for individuals with SCD by visually linking multiple types of data, including EHR data, clinical quality measures, and patient-reported data.

Data Source Name	Data Accessibility	Types of Aggregate Data Available	Format of Dashboard	Relevance to SCD Research
Pediatric COVID-19 US Registry Aggregated Data Summary[38]	Publicly accessible	• Demographics • Risk for COVID-19 • Underlying conditions • Symptoms • Imaging • Treatment • Outcomes	Contains column charts, numbers within a Microsoft Power BI tool that enables filtering	This data summary allows users to view the incidence and co-occurrence of individuals with SCD and COVID-19.
SCD Implementation Consortium Dashboard[35]	Publicly accessible	• Demographics • Sickle cell type • Care provider • Pain crises • Other health conditions • Depression • Blood transfusion • Hydroxyurea use	Azure web-based dashboard that includes column charts, pie charts, and counts	This dashboard presents relevant data from the SCDIC registry on health outcomes and treatments and therapies for individuals with SCD.
All of Us Research Hub[36,37]	Publicly accessible	• Demographics • Conditions • Drug exposures • Labs and measurements • Procedures • Genomics • Physicals and wearable data • Surveys include lifestyle, overall health, health care access and utilization, personal and family history	Web-based data browser that includes column charts and/or participants counts and percentages	This data hub allows users to view a wide range of biomedical data, including SCD-related diagnoses and laboratory test results.
Indiana Sickle Cell Data Collection Dashboard	Publicly accessible	• Demographics • Number of cases • Positive newborn screening • Hospitalization • Mortality • Prevalence by age and time	Web-based dashboard representing SCD data for Indiana that can be filtered by county	This dashboard allows users to view the incidence, geographic distribution, and healthcare utilization of individuals with SCD.

5. Discussion

Along with identifying data sources to support SCD-focused health outcomes research, the environmental scan and key informant interviews identified several considerations for improving or enhancing the available data. These include strengthening longitudinal data collection, streamlining ongoing data collection, improving the representation of PROs, and promoting standardization efforts.

Key informants emphasized that dedicated funding is a fundamental requirement for implementing these considerations, which often require committed staff and resources to be effectively sustained. The considerations are described in detail below.

5.1 Considerations to Strengthen the SCD Data Infrastructure

Improving collection of longitudinal data for the population with SCD. Multiple key informants noted that the limited availability of longitudinal data on people with SCD hinders health outcomes research. Since there is no longitudinal registry or designated surveillance program focused on SCD, researchers must rely on disparate, less comprehensive sources like administrative data that do not fully capture the scope of service use or long-term outcomes for people with SCD.[11,39] Several factors contribute to this reality: many young adults with SCD are lost to follow-up when they shift from pediatric to adult care; some data sources provide information on patients with Medicaid, and if Medicaid coverage ends, they are no longer represented in the data; and state-level data collection efforts face difficulties in tracking individuals who move out of state due to data access issues.[40] Key informants also described how researchers still rely on data from the Cooperative Study of Sickle Cell Disease, a large prospective cohort study. While detailed, the study ended data collection in the 1990s, meaning that it does not reflect the most recent advancements in SCD treatment.[41]

To address this challenge, SCD stakeholders should prioritize core data elements that can facilitate collection of longitudinal data. When addressing the lack of longitudinal SCD data, key informants emphasized that stakeholders should consider the inherent cost of data collection. To promote feasible, sustainable data collection, SCD stakeholders can prioritize a set of critical SCD data elements that should be a part of any large-scale data collection effort. These could include foundational population-level variables, such as diagnosis and mortality, as well as patient experience data that is important to SCD patients and caregivers, such as SCD symptoms, physical function, QoL, emotional and mental health, and pain events.[42,43]

> *"Every piece of data you collect has a cost, right? Whether it's resources, time, patient time, computer storage, whatever it is, so think about…core data elements that would be absolutely essential that you need."*
>
> ***– Key Informant***

Collecting and linking SCD data systematically across the lifespan presents longer-term opportunities to address this challenge. Several key informants identified strengthening longitudinal data collection as an overarching goal. While achieving this goal would require substantial coordination across stakeholders and sustained investment, it would provide researchers with the foundational information needed to assess a wide range of short- and long-term outcomes, as well as track unique patients across care settings and geographic locations.[43] Specifically, stakeholders have advocated for the creation of a linked surveillance system and a longitudinal clinical registry, which provide complementary information. A surveillance system would include "all individuals with SCD regardless of payer" and determine the "true SCD prevalence and geographic distribution of individuals with SCD."[44] A longitudinal clinical registry would include more detailed information—such as lab test results and PRO data—on a subset of patients with SCD who consent to sharing data.[40] It is important to note that while registries provide critical data for research, they inherently involve the collection of personally identifiable, sensitive information.[39] As a result, patients should be informed of their data's potential uses so they can provide fully informed consent.[43] Together, these data sources would provide researchers with robust data to support health outcomes research on SCD. Key informants also highlighted the CDC's leadership in conducting SCD surveillance through its SCDC program,[45] experience

that could make the CDC well positioned to serve as a key player in managing such a nationally linked surveillance system and registry.

Streamlining multiple data collection efforts to reduce participant burden. Several key informants noted that disparate data collection efforts and registries create burdens for researchers, due to the time-consuming process of submitting data to multiple state and national data efforts, each with a particular focus and reporting structure. Informants also recognized the potential for duplication and inconsistencies that emerge from unaligned data projects, which further complicates collecting and providing SCD data for research.

> *"What we were hearing from the sickle cell community was, people were being asked to join like four or five different registries, and it was very frustrating. It was very time-consuming. It didn't send a good message to the community that we were working together and working together well to really serve their needs, to really understand where the gaps in care and research were."*
>
> ***– Key Informant***

To address this challenge, SCD stakeholders should prioritize the linkage of existing data sources to support health outcomes research. Since many SCD-relevant data sources already exist, key informants emphasized that creating linkages among sources would be an effective strategy to reduce burden on researchers and strengthen the overall SCD data infrastructure. Key informants specifically highlighted creating linkages between key registries, such as GRNDaD,[15,16] and surveillance data. These linkages are necessary for developing "a systematic and comprehensive understanding of the natural history of [SCD], its outcomes, and trajectory under different care models."[40] A coordinated effort that has stakeholders submit data to a limited number of key data sources—which are then linked—would support researchers in conducting more comprehensive studies, ultimately improving outcomes for individuals with SCD.

> *"I hear so many more [SCD] warriors talking about quality of life, and it's very empowering to hear them talk that way because they deserve a high quality of life, just like everyone else."*
>
> ***– Key Informant***

Improving the representation of key PROs in SCD data sources. The environmental scan and key informants highlighted that existing data sources often do not capture PROs or other patient-contributed data of interest to SCD researchers and patients alike. While clinical data are important, PROs support a more comprehensive understanding of people's experience and ability to function with SCD,[46] providing key information on QoL, pain, fatigue, or other symptoms that interfere with activities of daily living. One key informant also emphasized the need for more validated measures of intermediate outcomes, such as missed work or school days and time away from caregiving responsibilities. Some key informants noted the challenges involved in collecting PRO or other patient-contributed data, which can be labor-intensive for patients. Further, some PRO measurement instruments are difficult for patients to use, since they are not written in accessible language.[47] A key informant also noted that there is some uncertainty in the field as to whether PRO measures are sensitive to treatment, meaning that they can effectively detect changes in a patient's health status in response to an intervention. Despite these challenges, key informants emphasized the importance of PROs to the broader SCD community. One key informant highlighted the Adult Sickle Cell Quality of Life Measurement Information System (ASCQ-Me)[48] as a valuable yet underused tool that measures multiple dimensions of disease impact, including pain, emotional well-being, and social functioning, and advocated for its broader adoption in research to generate more meaningful, patient-centered data.

To address this challenge, researchers can leverage remote data collection to increase the representation of PROs and other patient-contributed data in SCD data sources. One key informant described how sleep is impacted for many SCD patients, which can have additional effects on pain levels and mental health. Remote collection was highlighted as a key opportunity for collecting patient-contributed data on the number of hours slept or data from a patient's pulse oximeter, which can help improve the understanding of nocturnal hypoxemia. Environmental scan findings also highlighted the use of electronic pain diaries as a potential method for collecting and validating PRO data.[49] Further, a key informant discussed how data collected remotely from wearable devices could be useful in establishing a baseline functional assessment for patients with SCD, which would be beneficial for assessing the impact of treatment or other interventions on physical function, pain, and QoL. As a foundational consideration, researchers should consider efforts to strengthen interoperability between health systems to ensure this cutting-edge data is effectively accessed and used by varied members of a clinical care team. Effective data exchange also lays the foundation for the future linkage of PRO data to other data sources, providing opportunities for more robust analyses.

Promoting consistent use of standardized data definitions to make existing datasets easier to use in SCD-focused research. Key informants shared that the usefulness of existing data sources can be limited by inconsistencies, errors, or a lack of standardization. For instance, inaccuracies in coding patient data lead to challenges in assessing the SCD population, such as whether patients truly have SCD, which SCD genotype a patient has, or whether patients are receiving SCD treatment, all critical factors that can impact health outcomes and treatment efficacy. Informants also highlighted inconsistencies arising from organizations using genotype-based definitions to determine whether a patient has SCD. Further, key informants described how the absence of a standardized set of data elements complicates collection of key SCD-relevant outcomes across data sources, limiting researchers' ability to fully leverage the data to generate actionable information.

To address this challenge, stakeholders can leverage existing data standardization efforts to advance the SCD data infrastructure. Key informants said it is critical for the field to leverage existing data standardization efforts relevant to SCD, especially at the point of care.[43] To more accurately identify patients with SCD in datasets, key informants encouraged the use of the administrative claims case definition used by the CDC's SCDC program. According to this definition, classifying an individual as an SCD patient should be based on at least three diagnoses of SCD over a five-year period within administrative claims data.[50,51] Given that this definition was developed specifically for administrative data, future research could

Data Standardization Efforts Relevant to SCD-Focused Research

Common Data Models

- PCORnet Common Data Model
- Observational Medical Outcomes Partnership (OMOP) Common Data Model
- Sickle Cell Data Collection Program Common Data Model
- Algorithms to Support Identification of Patients with SCD in Electronic Health Record Data

Terminology Standard

- Logical Observation Identifiers, Names, and Codes (LOINC)

Measure Development Efforts

- i2b2 Common Data Model
- PhenX Toolkit
- Adult Sickle Cell Quality of Life Measurement Information System (ASCQ-ME)
- U.S Core Data for Interoperability
- Sickle Cell Data Collection Program Administrative Claims Case Definition
- FDA-ASH Consensus Recommendations for SCD End Points
- National Institute of Standards and Technology Genome Editing Consortium

explore its application to other data types, such as EHRs. The environmental scan and key informants also identified several standardization efforts, listed in the column at right, that can be leveraged to consistently represent outcomes in data sources. These include common data models, such as the PCORnet Common Data Model and algorithms developed to identify patients with SCD in EHR data;[52,53] terminology standards, such as Logical Observation Identifiers, Names, and Codes (LOINC); and measure development efforts, such as the U.S. Core Data for Interoperability and the PhenX toolkit, which all support the standardized collection and sharing of key data points. All publicly available, these efforts include multiple variables relevant to SCD-focused research, including health outcomes, risk factors and biomarkers, and treatments and therapies. These efforts are described in additional detail in **Appendix F**. Though only five were developed specifically for SCD, all could serve as useful tools for conducting efficient, detailed research analyses.

5.2 Limitations of this Study

This report is subject to some limitations. First, while a significant amount of literature was reviewed to identify SCD data sources, some sources may be missing or represented more than once if the meta data source is also listed (e.g., NIH-funded study data). As a result, this report does not reflect a systematic review or compilation of all data sources relevant to SCD-focused health outcomes research. Further, just over half of all SCD clinical trials are conducted in the U.S.[54] This report does not comprehensively reflect potential data sources leveraged internationally. Second, in some cases, the report's findings reflect the limited publicly available information for certain data sources. Not all sources were accompanied by accessible data dictionaries describing their contents, format, and periodicity. Due to this barrier, the level of detail provided for each data source is inconsistent and may be incomplete. Finally, the subject matter experts who informed this report by participating in key informant interviews represented two major stakeholder groups: federal stakeholders and SCD researchers. As such, these experts may not represent the views of other key SCD stakeholder groups, such as clinicians.

6. Conclusion

The environmental scan of the current data landscape for SCD-focused health outcomes research found both encouraging progress and pressing challenges. While many robust data sources already exist, persistent gaps remain, including the absence of a nationally representative data source that enables longitudinal analyses; a lack of standardization in variables that limits linkages; and the underrepresentation of PROs and other relevant variables, such as psychological risk factors, medication adherence, reproductive health, and emerging therapies, which inhibit SCD-related research. These issues may be further compounded by barriers to data access and use, including burdensome approval processes and a lack of standardization across datasets.

At the same time, promising opportunities are emerging. Key informants identified actionable strategies to strengthen the SCD data infrastructure, such as prioritizing and linking existing datasets, leveraging ongoing standardization efforts, and potentially using remote technologies to collect more meaningful patient-generated data. Importantly, key informants emphasized the need for sustainable and coordinated efforts by various partners, including federal and state agencies, researchers, and patient advocates anchored in collaboration with the SCD community, to identify core data elements, link registries to reduce burden on patients and providers, and promote data collection across the lifespan. Addressing these challenges is essential not only for improving the quality and relevance of health outcomes research but also for ensuring that individuals with SCD are represented in research in ways that reflect their experiences and needs.

References

[1] Centers for Disease Control and Prevention. Data and statistics on sickle cell disease. May 15, 2024. Accessed May 2, 2025. Retrieved from: https://www.cdc.gov/sickle-cell/data/index.html

[2] Brandow AM, Liem RI. Advances in the diagnosis and treatment of sickle cell disease. *J Hematol Oncol*. 2022;15(1):20. doi:10.1186/s13045-022-01237-z

[3] Jiao B, Johnson KM, Ramsey SD, Bender MA, Devine B, Basu A. Long-term survival with sickle cell disease: a nationwide cohort study of Medicare and Medicaid beneficiaries. *Blood Adv*. 2023;7(13):3276-3283. doi:10.1182/bloodadvances.2022009202

[4] American Society of Hematology. Quantifying the life expectancy gap for people living with sickle cell disease. 2023. Accessed May 2, 2025. Retrieved from: https://www.hematology.org/newsroom/press-releases/2023/quantifying-the-life-expectancy-gap-for-people-living-with-sickle-cell-disease

[5] National Heart, Lung, and Blood Institute. Sickle cell disease treatment. September 30, 2024. Accessed May 2, 2025. Retrieved from: https://www.nhlbi.nih.gov/health/sickle-cell-disease/treatment

[6] Williams H, Tanabe P. Sickle cell disease: a review of nonpharmacological approaches for pain. *J Pain Symptom Manage*. 2016;51(2):163-177. doi:10.1016/j.jpainsymman.2015.10.017

[7] Centers for Medicare and Medicaid Services. Cell and gene therapy access model. Accessed May 2, 2025. Retrieved from: https://www.cms.gov/priorities/innovation/innovation-models/cgt

[8] Desai P, George A, Ahmed F, et al. Access to gene therapy in sickle cell disease. *Blood*. 2024;144(1):7674. doi:10.1182/blood-2024-209630

[9] American Society of Hematology. ASH sickle cell disease research priorities, 2024 update. 2024. Accessed May 2, 2025. Retrieved from: https://www.hematology.org/research/sickle-cell-disease-and-sickle-cell-trait - duplicate reference with below

[10] Afolabi T. Leveraging data to address sickle cell disease. FedScoop. June 19, 2020. Accessed May 2, 2025. Retrieved from: https://fedscoop.com/leveraging-data-address-sickle-cell-disease/

[11] Kanter J, Meier ER, Hankins JS, Paulukonis ST, Snyder AB. Improving outcomes for patients with sickle cell disease in the United States: making the case for more resources, surveillance, and longitudinal data. *JAMA Health Forum.* 2021;2(10):e213467. doi:10.1001/jamahealthforum.2021.3467

[12] Centers for Disease Control and Prevention. Sickle Cell Data Collection Program report: data to action – gaps in knowledge. June 29, 2023. Accessed May 2, 2025. Retrieved from: https://archive.cdc.gov/www_cdc_gov/ncbddd/hemoglobinopathies/data-reports/2018-summer/knowledge-gaps.html

[13] Smeltzer MP, Reeves SL, Cooper WO, et al. Common data model for sickle cell disease surveillance: considerations and implications. *JAMIA Open*. 2023;6(2):ooad036. doi:10.1093/jamiaopen/ooad036

[14] National Heart, Lung, and Blood Institute. About the NHLBI. Accessed May 5, 2025. Retrieved from: https://www.nhlbi.nih.gov/

[15] Lanzkron S, Manwani D, Desai P, Kanter J, Little J. GRNDaD: big data and sickle cell disease. *Blood Adv*. 2022;6(3):1088. doi:10.1182/bloodadvances.2021005282

[16] Globin Research Network for Data and Discovery. GRNDaD: about us. Accessed May 2, 2025. Retrieved from: https://grndad.org/

[17] American Society of Hematology Research Collaborative. The Data Hub: a source for real-world hematology data. Accessed May 2, 2025. Retrieved from: https://www.ashresearchcollaborative.org/data-hub/

[18] Wood WA, Marks P, Plovnick RM, et al. ASH Research Collaborative: a real-world data infrastructure to support real-world evidence development and learning healthcare systems in hematology. *Blood Adv*. 2021;5(23):5429-5438. doi:10.1182/bloodadvances.2021005902

[19] Glassberg JA, Linton EA, Burson K, et al. Publication of data collection forms from NHLBI-funded sickle cell disease implementation consortium (SCDIC) registry. *Orphanet J Rare Dis*. 2020;15(1):178. doi:10.1186/s13023-020-01457-x

[20] Agency for Healthcare Research and Quality. Consumer Assessment of Healthcare Providers and Systems (CAHPS): surveys and tools to advance patient-centered care. Accessed May 2, 2025. Retrieved from: https://www.ahrq.gov/cahps/index.html

[21] Cronin RM, Yang M, Hankins JS, et al. Association between hospital admissions and healthcare provider communication for individuals with sickle cell disease. *Hematology*. 2020;25(1):229-240. doi:10.1080/16078454.2020.1780737

[22] Centers for Disease Control and Prevention. National Health Interview Survey: about NHIS. November 20, 2024. Accessed May 2, 2025. Retrieved from: https://www.cdc.gov/nchs/nhis/about/index.html

[23] Gyamfi J, Tampubolon S, Lee JT, et al. Characterisation of medical conditions of children with sickle cell disease in the USA: findings from the 2007-2018 National Health Interview Survey (NHIS). *BMJ Open*. 2023;13(2):e069075. doi:10.1136/bmjopen-2022-069075

[24] Agency for Healthcare Research and Quality. Medical Expenditure Panel Survey (MEPS): survey background. April 22, 2019. Accessed May 2, 2025. Retrieved from: https://meps.ahrq.gov/mepsweb/about_meps/survey_back.jsp

[25] Johnson KM, Jiao B, Ramsey SD, Bender MA, Devine B, Basu A. Lifetime medical costs attributable to sickle cell disease among nonelderly individuals with commercial insurance. *Blood Adv*. 2023;7(3):365-374. doi:10.1182/bloodadvances.2021006281

[26] Krishnamurti L, Arnold SD, Haight A, et al. Sickle Cell Transplantation Evaluation of Long-Term and Late Effects Registry (STELLAR) to compare long-term outcomes after hematopoietic cell transplantation to those in siblings without sickle cell disease and in nontransplanted individuals with sickle cell disease: design and feasibility study. *JMIR Res Protoc*. 2022 Jul 6;11(7):e36780. doi: 10.2196/36780

[27] RAND Corporation. Pediatric Quality of Life Inventory (PedsQL Measurement Model). Accessed May 19, 2025. https://www.rand.org/health-care/surveys_tools/pedsql.html

[28] Hankins JS, Estepp JH, Hodges JR, Villavicencio MA, Robison LL, Weiss MJ, Kang G, Schreiber JE, Porter JS, Kaste SC, Saving KL, Bryant PC, Deyo JE, Nottage KA, King AA, Brandow AM, Lebensburger JD, Adesina O, Chou ST, Zemel BS, Smeltzer MP, Wang WC, Gurney JG. Sickle Cell Clinical Research and Intervention Program (SCCRIP): A lifespan cohort study for sickle cell disease progression from the pediatric stage into adulthood. Pediatr Blood Cancer. 2018 Sep;65(9):e27228. doi: 10.1002/pbc.27228. Epub 2018 May 24. PMID: 29797644.

[29] National Institutes of Health. PROMIS® Pediatric Instrument Banks. Accessed May 19, 2025. https://www.niams.nih.gov/grants-funding/niams-supported-research-programs/pediatric-patient-reported-outcomes-chronic

[30] National Institutes of Health. Child Health Questionnaire. Accessed May 19, 2025. https://datashare.nida.nih.gov/instrument/child-health-questionnaire

[31] Beverung LM, Strouse JJ, Hulbert ML, et al. Health-related quality of life in children with sickle cell anemia: impact of blood transfusion therapy. *Am J Hematol*. 2015;90(2):139-143. doi:10.1002/ajh.23877

[32] Eckman JR, Hassell KL, Huggins W, Werner EM, Klings ES, Adams RJ, et al. Standard measures for sickle cell disease research: the PhenX toolkit sickle cell disease collections. *Blood Adv*. 2017;1(27):2703–2711. doi: 10.1182/bloodadvances.2017010702

[33] Treadwell MJ, Hassell K, Levine R, Keller S. Adult sickle cell quality-of-life measurement information system (ASCQ-me): conceptual model based on review of the literature and formative research. C*lin J Pain*. 2014;30(10):902–914. doi: 10.1097/AJP.0000000000000054

[34] HealthMeasures. Patient-Reported Outcomes Measurement Information System (PROMIS): why use PROMIS? March 27, 2023. Accessed May 2, 2025. Retrieved from: https://www.healthmeasures.net/explore-measurement-systems/promis

[35] Njoku F, Pugh N, Brambilla D, et al. Mortality in adults with sickle cell disease: Results from the Sickle Cell Disease Implementation Consortium (SCDIC) registry. *Am J Hematol*. 2024;99(5):900-909. doi:10.1002/ajh.27279

[36] National Institutes of Health. All of Us Research Hub: about the Research Hub. Accessed May 2, 2025. Retrieved from: https://www.researchallofus.org/about-the-research-hub/

[37] Jubilee IT, Evans LK, Lemieux KP. Evaluation of the association between sickle cell disease and breast cancer using the All of Us Research Program dataset. *Cancer Research*. 2024 Mar 22;84(6_Supplement):3470-

[38] Dain AS, Diorio C, Fisher BT, Hankins JS, Witmer CM, Boustany M, Burton M, Ferrolino J, Sadaf S, Ross HS, Maron G; Pediatric COVID-19 US Registry. Description of a national, multi-center registry of patients with sickle cell disease and SARS-CoV-2 infection: Data from the Pediatric COVID-19 United States Registry. Pediatr Blood Cancer. 2024 Jun;71(6):e30909. doi: 10.1002/pbc.30909. Epub 2024 Mar 12. PMID: 38469996; PMCID: PMC11039375.

[39] National Academies of Sciences, Engineering, and Medicine; Health and Medicine Division; Board on Population Health and Public Health Practice; Committee on Addressing Sickle Cell Disease: A Strategic Plan and Blueprint for Action, Martinez RM, Osei-Anto HA, McCormick M, eds. *Addressing Sickle Cell Disease: A Strategic Plan and Blueprint for Action*. Washington (DC): National Academies Press (US); September 10, 2020.

[40] Snyder AB, Lakshmanan S, Hulihan MM, et al. Surveillance for sickle cell disease — Sickle Cell Data Collection Program, two states, 2004–2018. *MMWR Surveill*. 2022;71(No. SS-9):1–18. doi: http://dx.doi.org/10.15585/mmwr.ss7109a1

[41] National Institutes of Health. Cooperative Study of Sickle Cell Disease. Accessed July 15, 2025. Retrieved from: https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs002362.v1.p1

[42] Sick Cells. Identifying warrior values in SCD research. Accessed April 23, 2025. Retrieved from: https://sickcells.org/blog/identifying-warrior-values-in-scd-research/

[43] ASH Research Collaborative. Accelerating innovations for sickle cell disease with real-world evidence: workgroup report – coordinated registry network. January 2023. Accessed May 12, 2025. Retrieved from: https://hematology.my.salesforce.com/sfc/p/#6A000000uIj9/a/3u000000MgQI/69J.PGurkrmqVUS1SS1r46mK_5qz7g.bCCIxl1qGHOM

[44] Lee L, Smith-Whitley K, Banks S, Puckrein G. Reducing health care disparities in sickle cell disease: a review. *Public Health Rep*. 2019;134(6):599-607. doi:10.1177/0033354919881438

[45] Centers for Disease Control and Prevention. Sickle Cell Data Collection program. August 28, 2024. Accessed July 15, 2025. Retrieved from: https://www.cdc.gov/sickle-cell/scdc/index.html

[46] Dierick K, Rodriguez-Grande B, Navarro-Aragall AG, Beraud M. Quality of life in people with sickle cell disease treated with automated red blood cell exchange. *Vox Sang*. 2025;120(1):4-12. doi:10.1111/vox.13757

[47] Rao N, Perdomo S, Jonassaint C. A novel method for digital pain assessment using abstract animations: human-centered design approach. *JMIR Hum Factors*. 2022;9(1):e27689. doi:10.2196/27689

[48] Treadwell MJ, Hassell K, Levine R, Keller S. Adult sickle cell quality-of-life measurement information system (ASCQ-Me): conceptual model based on review of the literature and formative research. The Clinical journal of pain. 2014 Oct 1;30(10):902-14.

[49] Bakshi N, Gillespie S, McClish D, McCracken C, Smith WR, Krishnamurti L. Intraindividual pain variability and phenotypes of pain in sickle cell disease: a secondary analysis from the Pain in Sickle Cell Epidemiology Study. *Pain*. 2022;163(6):1102-1113. doi:10.1097/j.pain.0000000000002479

[50] Snyder AB, Zhou M, Theodore R, Quarmyne MO, Eckman J, Lane PA. Improving an administrative case definition for longitudinal surveillance of sickle cell disease. *Public Health Rep*. 2019;134(3):274-281. doi:10.1177/0033354919839072

[51] Singh A, Sontag MK, Zhou M, et al. Evaluating the discriminatory ability of the Sickle Cell Data Collection Program's administrative claims case definition in identifying adults with sickle cell disease: validation study. *JMIR Public Health Surveill*. 2023;9:e42816. doi:10.2196/42816

[52] Michalik DE, Taylor BW, Panepinto JA. Identification and Validation of a Sickle Cell Disease Cohort Within Electronic Health Records. Acad Pediatr. 2017 Apr;17(3):283-287. doi: 10.1016/j.acap.2016.12.005. Epub 2016 Dec 13. PMID: 27979750.

[53] Singh A, Mora J, Panepinto JA. Identification of patients with hemoglobin SS/Sβ^0 thalassemia disease and pain crises within electronic health records. *Blood Adv*. 2018;2(11):1172-1179. doi:10.1182/bloodadvances.2018017541

[54] Costa E, Ware RE, Tshilolo L, Makani J, Leufkens HGM, Luzzatto L. Globalization in clinical drug development for sickle cell disease. *Am J Hematol*. 2025;100(1):4-9. doi:10.1002/ajh.27525

Appendix A. Environmental Scan Methods

Exhibit A1. Websites Prioritized During Web Scan

HHS Agency Websites
Center for Medicare & Medicaid Services' (CMS's) data.cms.gov
Assistant Secretary for Technology Policy/Office of the National Coordinator for Health Information Technology (ASTP/ONC)
DATA.GOV
Health Resources and Services Administration's (HRSA's) Maternal and Child Health Bureau (MCHB)
Agency for Healthcare Research and Quality (AHRQ)
The National Heart, Lung, and Blood Institute (NHLBI) – Sickle Cell Branch
The National Human Genome Research Institute (NHGRI)
SCD-Relevant Websites
American Society of Hematology (ASH)
Sickle Cell Disease Association of America (SCDAA)
American Academy of Pediatrics (AAP)
American Public Health Association (APHA)
American Society of Pediatric Hematology/Oncology (ASPHO)
Sickle Cell Disease Foundation
Sick Cells
National Clinical Research Networks
PCORnet
NIH RePORTER

Exhibit A2. Literature Search Terms

Search Term Category	Example Terms
SCD	Sickle Cell Disease, Sickle Cell Anemia*, Sickle Cell Trait, SCD management, SCD treatment, SCD outcomes
Health Outcomes Research	Patient-Centered Outcomes Research, Patient-Reported Outcome*, patient engagement, shared decision-making, health-related quality of life, patient satisfaction, patient experience, patient-reported outcomes, patient-centered, healthcare research, health outcomes research, clinical effectiveness research
Data and Data Infrastructure	Data [title/abstract], Database; dataset; data source; clinical registry; data tools; electronic health records [MeSH], Data Collection/methods [MeSH]; Database Management Systems [MeSH]; data infrastructure, data linkage
Data Gaps and Opportunities	Data limitations; data challenges

Exhibit A3. Literature Search Inclusion and Exclusion Criteria

Category	Inclusion Criteria	Exclusion Criteria
Publication Year	• Peer-reviewed journal articles: 2019-present • Grey literature: 2019-present	• Peer-reviewed journal articles: Before 2019 • Grey literature: Before 2019
Document Type	• Peer-reviewed journal articles: Theoretical articles, primary and secondary data analyses, scoping review, meta-analyses/systematic reviews • Grey literature: Reports, working papers, evaluation studies, white papers, conference proceedings, presentations, case studies, fact sheets, issue briefs, and government documents	• Grey literature: Opinion pieces
Language	• English	• Non-English
Source	• Academic, expert, evaluator	• News outlet
Population	• SCD patients	• Non-SCD patients
Focus	• Research with or concerning a study for the specified population that uses quantitative data and defined process and/or outcome measures	• Non-US-based studies • Studies or databases that only focus on genetic or biomedical research and do not include a discussion of patient-centered outcome measures, or studies that only involve qualitative data • Focus primarily on non-SCD chronic conditions without direct relevance to SCD comparisons

Exhibit A4. Fields Included in Full Article Abstraction Form

Field Type/Category	Specific Fields
Resource Metadata	• Article name • Article type (e.g., prospective study, grey literature, etc.) • Literature source • Full citation with hyperlink to DOI, if available
Organizations Involved	• Stakeholders referenced • Collaborations/research initiatives referenced
Study Details	• Research question/study aim • Main findings • SCD population characteristics (e.g., adults of a certain age) • Intervention(s) assessed or compared, if any • Health outcomes research measures included in study

Field Type/Category	Specific Fields
Data-Related Information	• Data types utilized (e.g., prospective data from patients, claims, electronic health records, registry, survey) • Names of unique databases/data sources identified
Data Gaps and Opportunities for Health Outcomes Research	• Identified gaps in the data infrastructure • Areas for improvement • Research questions proposed as future directions within the study/article

Appendix B. Key Informant Discussion Protocol

1. Please introduce yourself and tell us about your role within [*organization*].
2. Can you briefly describe your involvement in data collection and/or research related to sickle cell disease?
3. We initially inventoried 24 sources with individual-level data on the population with sickle cell disease, spanning registries, surveys, administrative data (e.g., within claims) and other categories (e.g., those developed using a mix of various data types). These are included in the tab "Individual Data" on the spreadsheet. We also identified a few sources of aggregate data, for example, via tools and dashboards. These are included in the tab "Aggregate Data."
    a) Based on your general knowledge about sickle cell disease data sources, does our inventory look comprehensive to you, or are there any important data sources, data tools, or dashboards that we are missing?
4. Are there any data sources from this list, or data types more generally, that you use most often in your research or that you know are used often by researchers at large to assess outcomes for populations with sickle cell disease?
    a) Conversely, are there emerging sources of federal or other data, or data types that hold potential but are not yet widely used for sickle cell disease-focused PCOR?
5. Which items on the list of data infrastructure gaps stood out to you?
    a) Would you add anything to this list?
6. Do you have any concerns related to the timeliness (i.e., recency), standardization, and/or other aspects of the quality of data on sickle cell disease?
7. Moving on to the list of data infrastructure opportunities, which items resonated with you?
    a) Would you add anything to this list?
    b) [*To federal informants*] What data-oriented collaborations across HHS do you think would enhance the ongoing work in sickle cell disease conducted by your agency?
8. What were your first impressions of the visual?
9. When looking at this visual, do you come away with a big-picture understanding of the data infrastructure for sickle cell disease?
    a) Are there aspects of the visual that can be improved for clarity?
10. Focusing now on the left panel in yellow, are there other entities not depicted in the diagram that have funded sickle cell disease data collection efforts? If so, who are they and what did they fund?
11. We already talked about data sources in depth earlier in this discussion, so we can skip that portion of the visual. Moving now to the bottom of the middle panel, we've listed a few data standardization efforts relevant to sickle cell disease data. Are there other efforts that are missing from this list?
12. Considering the green section of the visual on the right – SCD research areas – are there any priority gaps or opportunities that come to mind for improving data infrastructure for these specific research areas that have not already been mentioned?
13. More broadly, what aspects of sickle cell disease would you like to see PCOR researchers focus on in the next five to ten years?
    a) What types of data sources or linkages are needed to explore this type of research?
14. Before we wrap up, we wanted to ask if there are there any specific reports or other resources you think would be helpful for us to review, or any additional considerations to understand the current PCOR data infrastructure for sickle cell disease and opportunities for enhancement.

Appendix C. American Society of Hematology (ASH) SCD Health Outcomes Related Priorities – 2024 Update Mapped to Research Topics

ASH SCD Research Priority – 2024 Update[1]	Identification and Assessment of Disease-Related Risk Factors & Biomarkers	Identification and Assessment of Health Outcomes	Assessment of SCD Treatments/ Therapies
In-depth studies of adolescents and young adults including clinical phenotyping, biomarkers, and patient-reported outcomes; prospective collection of outcomes to address the contribution of social determinants of health to the pathophysiology of SCD and organ dysfunction.	Sociodemographic characteristics Biomarkers	Organ function Patient-reported outcomes	
Inclusion of long-term follow-up of organ function in clinical trials of new therapies for sickle cell disease.		Organ function	Treatments/therapies
Determine the natural history and modifiable risk factors (e.g., social, environment, psychological, biological) for the development of varied SCD pain phenotypes that include acute and chronic pain. Examine the effect of social determinants of health, gender, and organ damage on the transition from acute to chronic pain.	Sociodemographic characteristics Environmental risk factors Psychological risk factors	Pain Organ function	

[1] American Society of Hematology. ASH Sickle Cell Disease Research Priorities – 2024 Update. 2024. Retrieved from: https://acrobat.adobe.com/id/urn:aaid:sc:VA6C2:62fb4a73-ee25-4fc2-85e8-576a19c9581a

ASH SCD Research Priority – 2024 Update[1]	Identification and Assessment of Disease-Related Risk Factors & Biomarkers	Identification and Assessment of Health Outcomes	Assessment of SCD Treatments/ Therapies
Investigate the use of existing pharmacological (e.g., opioids, nonsteroidal anti-inflammatory drugs, serotonin and norepinephrine reuptake inhibitors, gabapentinoids, ketamine, cannabinoids) and non-pharmacologic therapies (e.g., behavioral health interventions, integrative approaches, physical therapy) for the treatment of acute and chronic pain and understand how to optimize the use of these for individualized pain treatment. Determine which treatment has the best evidence for efficacy and effectiveness, whether there are genetic factors that influence response to analgesics (e.g., opioids), and ensure risks, benefits, and long-term effects are clearly understood.	Risk factors Genotype	Pain	Non-pharmacological therapies Pharmacological therapies
Assess the impact of pharmacologic (e.g., hydroxyurea, glutamine, voxelotor, crizanlizumab, etc.), curative and potentially curative therapies (i.e., bone marrow transplantation, gene therapy), and integrative approaches on the development of chronic pain.		Pain	Treatments/therapies Non-pharmacological therapies Pharmacological therapies Gene therapy Transplant regimens
Longitudinal studies, including the use of real-world data, to determine the long-term effects of transfusions, hydroxyurea, and the newer disease modifying therapies, including gene therapy or hematopoietic stem cell transplant on preservation or restoration of organ function. Consider the establishment of long-term registries, especially for the newer medications, across lifespan and global settings.		Organ function	Treatments/therapies Pharmacological therapies* (hydroxyurea) Gene therapy Transplant regimens Blood transfusions
Clinical trials to modify disease altering co-morbidities, such as kidney disease, obstructive lung disease, pulmonary hypertension, and cardiovascular injuries (e.g., cardiac and brain).		Comorbidity progression	

ASH SCD Research Priority – 2024 Update[1]	Identification and Assessment of Disease-Related Risk Factors & Biomarkers	Identification and Assessment of Health Outcomes	Assessment of SCD Treatments/ Therapies
Consideration and investigation of the effect of existing co-morbidities on the potential use, efficacy, or safety of existing therapies to clearly understand the indications and approaches to differences in the appropriate dosing, safety, and efficacy across the lifespan.	Comorbidities		Pharmacological therapies
Research on the effect and role of disease modifying therapies on reproductive health concerns and outcomes, and how to minimize the risks, for individuals with SCD including pregnancy, lactation, menstruation, and fertility for men and women.		Reproductive health and fertility outcomes	Treatments/therapies
Include assessment of patient-reported outcomes as a clinical endpoint for all studies investigating the effects of disease modifying therapies.		Patient-reported outcomes	Treatments/therapies
Pharmacogenomic studies investigating inter-patient variability in response or toxicity for existing disease modifying therapies to identify which individuals are more or less likely to benefit from specific medications and experience potential toxicity.	Risk factors		Treatments/therapies
Define additional biomarkers (e.g., blood, metabolite, imaging) for detecting, longitudinal tracking and predicting utility (i.e., responsiveness), and impact of these disease modifying agents on long-term health and quality of life.	Biomarkers Laboratory test results Imaging results	Patient-reported outcomes	Treatments/therapies
Design and execute prospective clinical trials to determine the efficacy of hydroxyurea in individuals with HbSC and other less common SCD genotypes. Considerations should include investigation of predictors of response, preliminary research utilizing existing real-world or registry data, research focused on the suboptimal medication adherence to hydroxyurea and other disease modifying pharmacologic therapies, and exploration of specific markers or definitions of treatment futility.	Genotype	Medication adherence	Pharmacological therapies

ASH SCD Research Priority – 2024 Update[1]	Identification and Assessment of Disease-Related Risk Factors & Biomarkers	Identification and Assessment of Health Outcomes	Assessment of SCD Treatments/ Therapies
Research focused on short- and long-term effects of blood transfusions on organs, with a focus on challenges that limit the safety or efficacy of transfusions, including mechanisms behind hyperhemolytic transfusion reactions and focused research to reduce the risk of alloimmunization and iron overload.		Organ function	Blood transfusions
Research focused on novel approaches to identifying and addressing barriers to medication adherence as well as monitoring medication adherence, including development and evaluation of digital health tools and community-based interventions.		Medication adherence	
Consider repurposing FDA-approved drugs and new therapeutics for SCD complications such as stroke, renal, and cardiovascular disease.		Comorbidity progression	Treatments/therapies
Establish long-term follow up of all individuals with SCD, including those treated with novel drugs, cellular, and gene therapies and track the correlation between the percentages of corrected cells (chimerism), durable efficacy, side effects, toxicity, and safety of transformative therapies. Using comprehensive robust prospective longitudinal registries for SCD clinical trials addressing current gaps will align with recognized practices for other inherited diseases.		Health outcomes	Treatments/therapies Pharmacological therapies Gene therapy
Identify biomarkers to guide providers and assist individuals with SCD in making informed decisions about the best curative therapy option (personalized medicine), and support studies comparing curative therapies with disease modifying therapies including evaluation of quality of life and impacts on daily life.	Biomarkers	Patient-reported outcomes	Treatments/therapies
Study pregnancy-related complications for individuals with SCD.		Pregnancy outcomes	

Appendix D. Acronyms Used in Visual Framework of SCD Data Infrastructure

Agency for Healthcare Research and Quality (AHRQ)

American Society of Hematology (ASH)

American Thrombosis and Hemostasis Network (ATHN)

Biologic and Data Repositories Information Coordinating Center (BioLINCC)

Consumer Assessment of Healthcare Providers and Systems (CAHPS)

Centers for Disease Control and Prevention (CDC)

Center for International Blood and Marrow Transplant Research (CIBMTR)

Centers for Medicare & Medicaid Services (CMS)

Chesapeake Regional Information System for our Patients (CRISP)

Cure Sickle Cell Initiative (CureSCi) Metadata Catalog

The Database of Genotypes and Phenotypes (dbGaP)

Dissemination and Implementation of Stroke Prevention Looking at the Care Environment (DISPLACE) Study

Global Action Network for Sickle Cell & Other Inherited Blood Disorders (GANSID)

Global Blood Therapeutics (GBT)

Globin Research Network for Data and Discovery (GRNDaD) Registry

Health Resources and Services Commission (HRSA)

Hydroxyurea to Prevent Organ Damage in Children with Sickle Cell Anemia (BABY HUG) Clinical Trial

Indiana Network for Patient Care Research Database (INPCR)

Indiana Sickle Cell Data Collection (IN-SCDC) Dashboard

Medical Expenditure Panel Survey (MEPS)

National Health Interview Survey (NHIS)

National Heart, Lung, and Blood Institute (NHLBI)

National Inpatient Sample (NIS)

National Institutes of Health (NIH)

PA Patient & Provider Network (P3N)

Patient-Centered Outcomes Research Institute (PCORI)

Recipient Epidemiology and Donor Evaluation Study (REDS) Program

Silent Cerebral Infarct Transfusion (SIT) Clinical Trial

Surveillance Epidemiology of Coronavirus Under Research Exclusion (SECURE)-SCD Registry

Science, Technology and Research Partnership (STAR) Clinical Research Network (CRN) Data

Sickle Cell Disease (SCD)

Sickle Cell Disease Association of America (SCDAA)

The Sickle Cell Transplantation Evaluation of Long-Term and Late Effects Registry (STELLAR)

Transformed Medicaid Statistical Information System (T-MSIS)

Trans-Omics for Precision Medicine (TOPMed)

Appendix E. Dashboards Identified in the Environmental Scan

Exhibit E1. ASH Research Collaborative Data Hub Site Portal

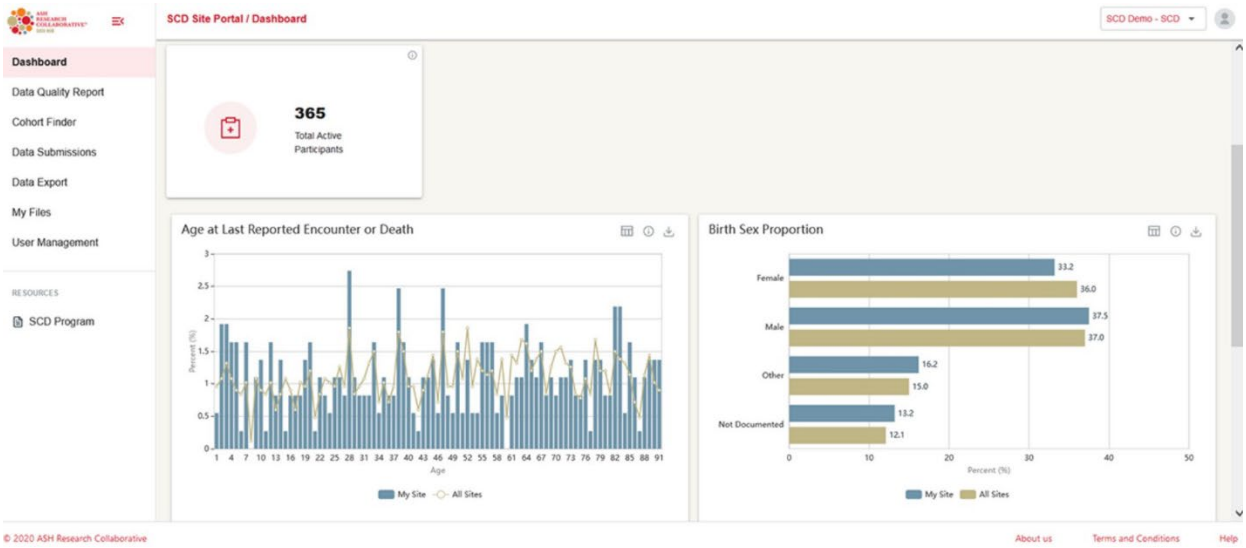


Exhibit E2. Pediatric COVID-19 US Registry Aggregated Data Summary

Back to Main | Demographics | Underlying Conditions | Symptoms | Imaging | Treatment | Outcomes

Last updated: 10 Aug 2024

Underlying/Pre-existing Conditions for ALL Cases

12925

Number of Completed Cases

System	General Pediatric	HCT or CT Recipient	Immunocompromised (Non-Transplant)	Solid Organ Transplant	Total
⊞ None specified	7454	5	7	26	7492
⊞ Pulmonary	1831	9	57	30	1927
⊞ Neurologic	870	8	44	14	936
⊞ GI/Hepatology	731	12	49	38	830
⊞ Cardiac	490	9	42	79	620
⊟ Obesity	481	3	13	18	515
Obesity	481	3	13	18	515
⊞ Hematologic	417	11	50	3	481
⊞ Endocrine	401	14	36	25	476
⊞ Renal	191	4	26	69	290
⊟ Cancer - Leukemia or Lymphoma	10	38	158	9	215
Acute lymphoid leukemia	4	18	103	1	126
Other	2	7	23	6	38
Acute myeloid leukemia		12	10	1	23
Hodgkin's lymphoma	2		13		15
Chronic myeloid leukemia	1		4		5
Non-hodgkin's lymphoma			4	1	5
Unspecified	1		1		2
Myelodysplastic syndromes		1			1
⊟ Cancer - Solid Tumor	38	27	137	3	205
Other	14	13	52	1	80
Brain tumor	7	6	31		44
Bone/soft tissue sarcoma	9	1	31		41
Neuroblastoma	5	7	10		22
Wilm's tumor			7		7
Retinoblastoma	2		4		6
Hepatoblastoma			1	2	3
Unspecified	1		1		2
⊟ Rheumatological/inflammatory conditions	28		40	2	70
Inflammatory bowel disease	9		13	1	23
Unspecified	7		9	1	17
JIA	10		6		16
Lupus	2		12		14

Exhibit E3. SCD Implementation Consortium Dashboard

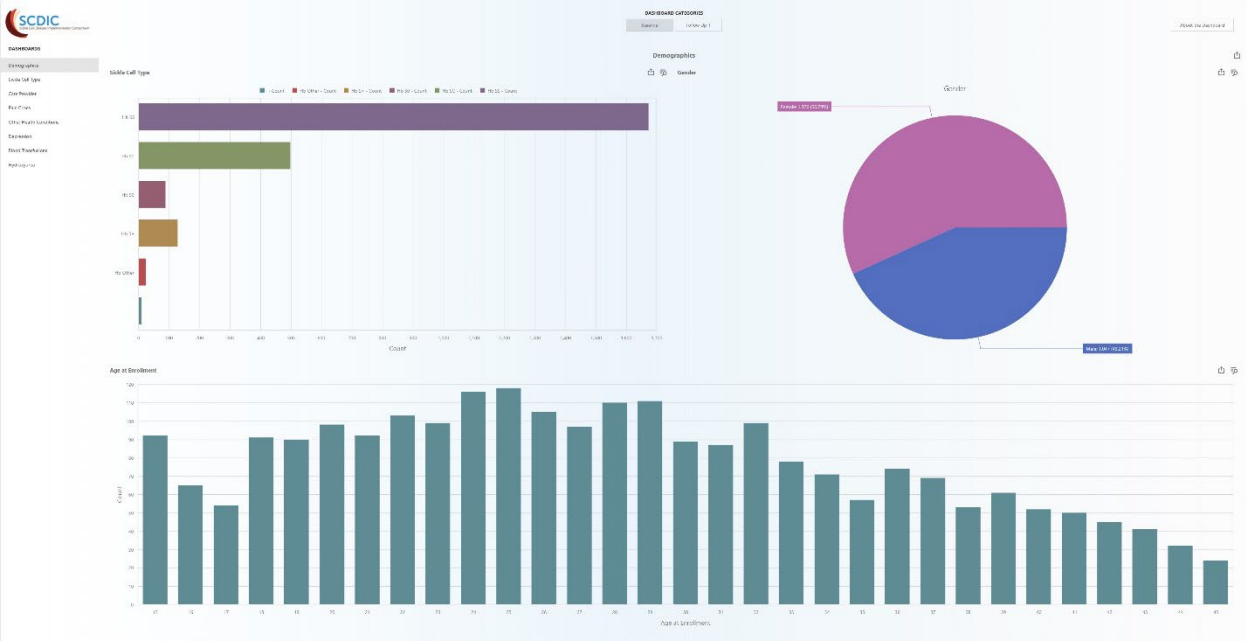
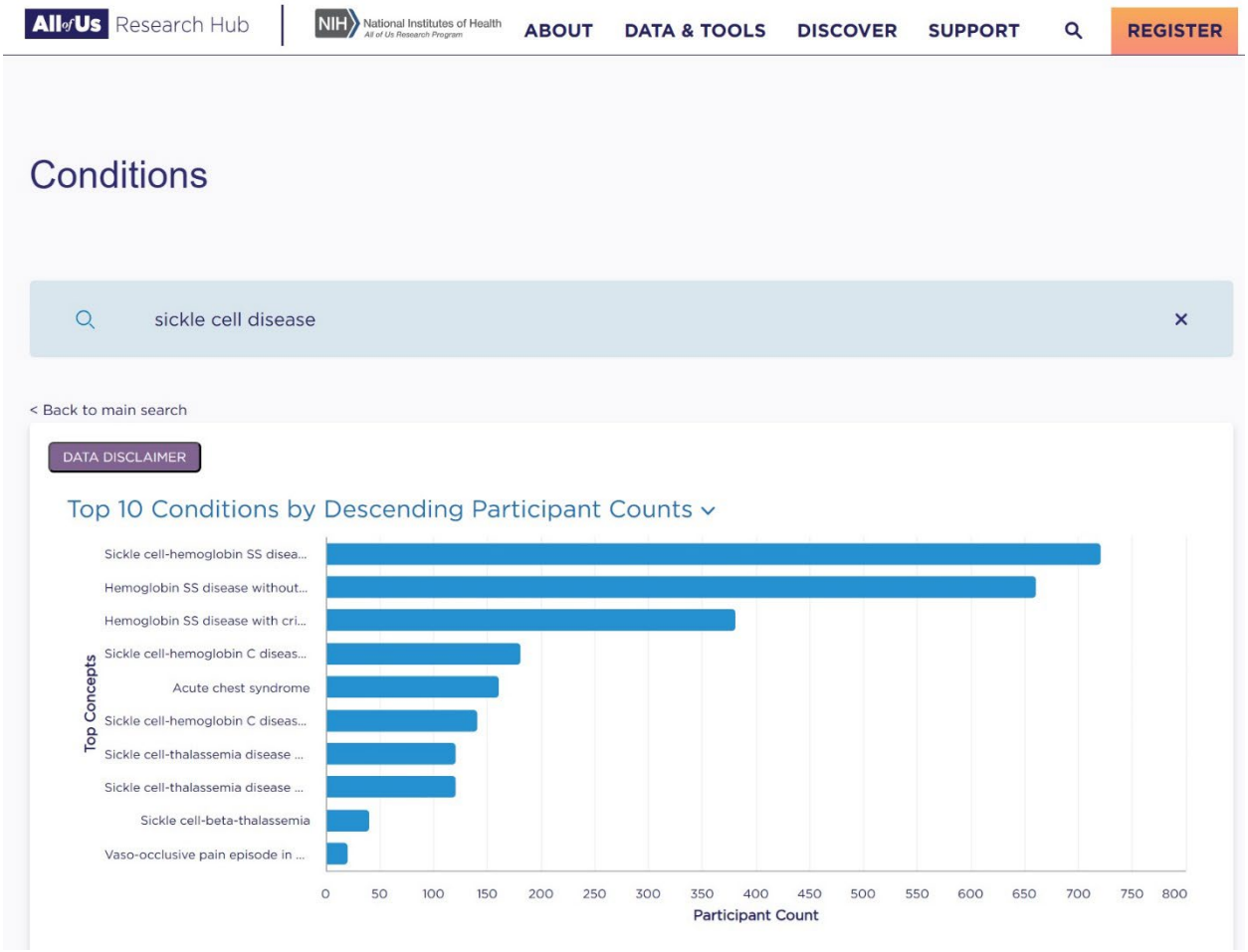


Exhibit E4. All of Us Research Hub Data Browser



Appendix F. Data Standardization Efforts Relevant to SCD-Focused Research Identified in the Environmental Scan

Type of Effort	Data Standardization or Measurement Effort Name	Brief Description	Included Variables Relevant to SCD-Focused Research
Common Data Model	PCORnet Common Data Model	Publicly available resource that standardizes data points from PCORnet participants' information systems which includes electronic health record data, claims data, and patient-reported data.	• Healthcare utilization • Medical complications • Medication adherence
	Observational Medical Outcomes Partnership (OMOP) Common Data Model	Open community data standard designed to standardize the structure and content of observational data.	• Healthcare utilization • Health insurance • Diagnosis
	i2b2 Common Data Model	Open-source data model used to store and integrate multiple types of data, which includes coded electronic health record and medical claims data, genomics, and clinical trial data.	• Diagnoses • Laboratory tests • Imaging results • Medications
	SCDC Common Data Model*	Data model used to standardize key data elements from states participating in the SCDC, simplifying reporting and analyses.	• Healthcare utilization • Health insurance • SCD genotype
Terminology Standard	Logical Observation Identifiers Names and Codes (LOINC)	A catalog of identifiers, names, and codes that supports the identification of health measurements and observations to promote interoperable data exchange.	• Laboratory tests • Vital signs

Type of Effort	Data Standardization or Measurement Effort Name	Brief Description	Included Variables Relevant to SCD-Focused Research
Measure Development	PhenX Toolkit	Public resource that includes select high-quality standards to establish a framework for data sharing across SCD research projects.	• Pregnancy outcomes • Pain type and intensity • Quality of life • Self-efficacy
	Adult Sickle Cell Quality of Life Measurement Information System (ASCQ-ME)*	A publicly available patient-reported outcome measurement system that evaluates and monitors the physical, mental, and social well-being of adults with SCD.	• Pain episodes • Emotional impact • Sleep impact • Social functioning
	U.S Core Data for Interoperability (USCDI)	A standardized set of health data classes and constituent data elements for nationwide, interoperable health information exchange.	• Healthcare encounter information • Health insurance • Health concerns • Functional status
	SCDC Administrative Claims Case Definition*	An approach to accurately identify adults with SCD using Medicaid claims data.	• Diagnosis of SCD
	FDA-ASH Consensus Recommendations for SCD End Points*	A joint effort led by FDA and ASH to develop consensus recommendations for the standardized measurement of clinical trial outcomes.	• PROs • Pain • Cure • Biomarkers
	National Institute of Standards and Technology Genome Editing Consortium	Cross-stakeholder consortium that works to develop measures and standards across the genome editing field.	• Genome editing tools and concepts
	Algorithms to Support Identification of Patients with SCD in EHR Data*	Algorithms developed and validated by the Medical College of Wisconsin to rapidly and accurately identify patients with SCD in available EHR data.	• Healthcare utilization • SCD genotype

** Standardization effort was developed specifically for SCD.*