Cure Sickle Cell Initiative

Natural History Data Resource

Operating Procedures

January 2022

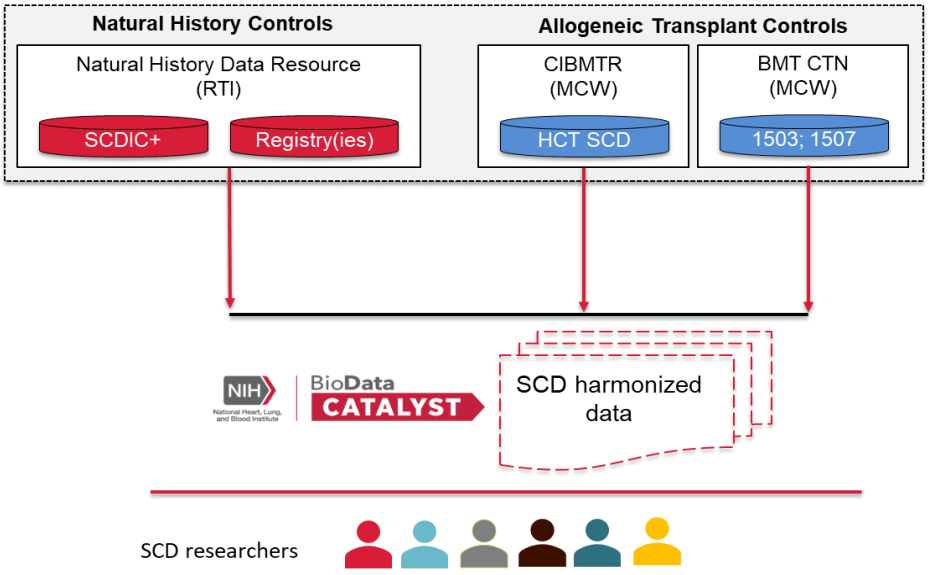
V2.0

1. Background and Rationale

More than 100,000 individuals in the United States are currently living with sickle cell disease (SCD). Approximately 1,600 infants are born each year with SCD, and nearly half will not survive past their fifth decade given the availability of current treatments. Gene therapies offer promising possibilities for a cure, which has led to the initiation of the Cure Sickle Cell Initiative (CureSCi) by the National Heart, Lung, and Blood Institute (NHLBI). Currently, gene therapy trials in SCD patients are in Phases 1 and 2, with small numbers evaluated for safety and efficacy. As trials move to Phase 3, a double-blind study is impossible and a study where the treatment is randomized may be unethical and unlikely to lead to an appropriate (unbiased) untreated comparison group that completes long term follow-up under controlled study conditions. Nonetheless, controls that do not undergo gene therapy are needed in order to compare the outcomes of the trial.

To that end, the NHLBI is leveraging existing SCD programs to build a contemporaneous control data resource that will be available to qualified researchers through the NHLBI data repositories: Biologic Specimen and Data Repository Information Coordinating Center (BioLINCC) and BioData Catalyst. As illustrated in Figure 1, the resource will include natural history data collected from SCD registries, clinical trial long term follow-up data from the Center for International Blood and Marrow Transplant Research (CIBMTR) and clinical trial data from the Blood and Marrow Transplant Clinical Trials Network (BMT CTN) SCD clinical trials in hematopoietic stem cell therapy. The data will be harmonized to enable unbiased control cohorts of transplanted and non-transplanted individuals to be formed using tools available within the BioData Catalyst cloud-based ecosystem.

Figure 1. Building a Contemporaneous Control Data Resource of Harmonized SCD Data



Natural history data collected by SCD registries will be combined to form a Natural History Data Resource (NHDR) under the operating procedures described in this document. RTI International (RTI) will coordinate and manage all NHDR activities. The Sickle Cell Disease Implementation Consortium (SCDIC) Registry will provide the framework for the NHDR; other registries that meet the NHDR requirements will be added to enrich the SCDIC cohort. Each registry will contribute Core Data Elements that the CureSCi Data Standard Working Groups and the ASH/FDA Working Groups identified as essential/required information based on the current clinical research best practices.

The NHDR Operating Procedures described in this document provide the requirements for study participation, data submission, data governance and data use. In addition, it describes the data management, harmonization, and data submission to the NHLBI repository that will be conducted by the RTI data coordinating center (DCC). *This document does not address patient recruitment, data collection, or other activities at the research center level that are conducted as part of the participating registry, research questions that could be addressed with NHDR data, or specific statistical plans for how data may be analyzed.* These topics are addressed in the individual study/registry protocols that have undergone review and approval by local IRBs. In Section 2, we provide a justification for sample size and statistical power that may be achieved based on prevalence outcomes ranging from 2-20%.

2. Objectives, Benefits, and Intended Use of the NHDR

The NHLBI initiated the NHDR with the following objectives:

- Leverage existing SCD studies to build a data resource to provide contemporaneous control cohorts for gene therapy trials
- Make patient-level de-identified data available to all qualified investigators in a secure environment and in a timely manner at no cost
- Build on the Sickle Cell Disease Implementation Consortium (SCDIC) Registry platform and infrastructure (>2,400 patients)
- Automate EHR data extraction, to the extent possible
- Identify core data that can be obtained from all participating studies
- Share de-identified data through NHLBI Data Repositories, BioLINCC and BioData Catalyst

2.1 Resource for Contemporaneous Controls

The SCDIC provides a source of non-transplanted contemporaneous controls for ongoing gene therapy trials and other studies of people with SCD. For example, a 2019 analysis of the SCDIC subject population identified eligible people that met study inclusion and exclusion criteria for 7 gene therapy trials listed in *clinicaltrials.gov*. Out of the 2,423 patients in the SCDIC database, and depending on the gene therapy trial criteria, 88 to 987 patients (4 – 41%) would be eligible given the SCDIC data available. Eligible numbers across the trials mostly varied based on the differences in inclusion criteria for age, genotype and whether failure of hydroxyurea was required. Still, these numbers are likely sufficient for a matched case-control study of a small trial as the recommended number of matched controls per case does not exceed five.

2.2 Power and Sample Size Considerations

The potential uses of the NHDR are broad, including providing matched controls for clinical trials, observing a cohort of untreated patients as part of a natural history study, or conducting comparative effectiveness research on a subpopulation with selected co-morbidities or drug treatment profiles. This section focuses on the precision of prevalence estimates and the range of effect sizes that can be detected with respectable statistical power, given the numbers that are expected to be included in the NHDR.

Let's say for example, the expected outcomes of interest have prevalence rates of 2%, 10% and 20%. Given these rates and a projected enrollment into the NHDR of 1500 to 2000 patients with SCD, we set out to determine the precision (defined as width of the 99% confidence interval) available for estimating outcomes ranging from a prevalence of 2% to 20%. We used 99% confidence limits in place of the typical 95% confidence limits in recognition of the fact that we are looking to estimate multiple co-primary outcomes, which typically inflates Type I error and may provide overly optimistic precision using 95% confidence limits.

The results for analysis are presented in **Figure 2** and **Table 1** below. These calculations show that the precision, or 99% confidence interval width for an assumed outcome rate of 2% and 20% varies from 0.019 and 0.054, respectively, for a sample size of 1500 to 0.017 and 0.047, respectively, for a sample size of 2000. Thus, precision for a rare outcome is relatively low, with the 99% confidence interval for an assumed event rate of 2% going from 1.15% to 2.85% for an enrollment of 2000 SCD patients. On the other hand, precision is much higher for a more common outcome, with the 99% confidence interval for an assumed event rate of 20% going from 17.65% to 22.35% for an enrollment of 2000 SCD patients.

Figure 2. Width of Confidence Interval vs Proportion of Outcome by Sample Size

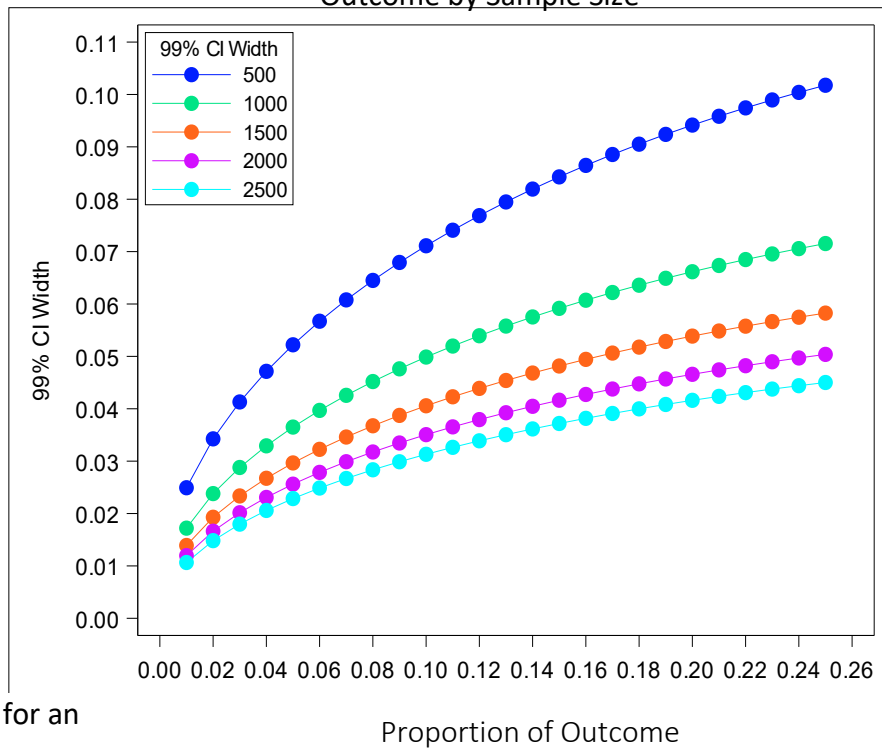


Table 1. Width of Confidence Interval vs Proportion of Outcome by Sample Size

	99% CI Width by Sample Size				
Outcome Risk/ Event Rate	*500*	*1000*	*1500*	*2000*	*2500*
1.0%	0.025	0.017	0.014	0.012	0.011
2.0%	0.034	0.024	0.019	0.017	0.015
10.0%	0.071	0.050	0.041	0.035	0.031
15.0%	0.084	0.059	0.048	0.042	0.037
20.0%	0.094	0.066	0.054	0.047	0.042

3. Organization of the NHDR

The organization of the NHDR, illustrated in Figure 1, currently includes one multi-center registry with its own DCC that will also operate as the NHDR DCC, and the funding institute, NHLBI. Additional SCD registries will be added as they meet inclusion requirements.

3.1 SCDIC+ Registry

The SCDIC Registry was initiated in 2016 and has enrolled over 2,400 patients with SCD, age 15-45 years at the time of enrollment, from eight treatment centers across the US (Augusta University, University of Illinois at Chicago, St. Jude Children's Research Hospital, Washington University School of Medicine, Icahn School of Medicine at Mount Sinai, Duke University, UCSF Benioff Children's Hospital Oakland, and Medical University of South Carolina). RTI serves as the DCC. Extensive baseline data were collected on each study participant which included medical record abstraction on the diagnosis of SCD, history of transfusions and other treatment, healthcare utilization, vital signs, complications, co-morbidities, and recent laboratory results. In addition, baseline surveys were completed by the patients which captured their patient-reported outcomes related to SCD and quality of life (pain, sleep, fatigue, anxiety, depression, cognition, social and emotional support, physical function), side effects of SCD medications, and their pregnancy history. Similar surveys were completed annually, and a follow-up medical record abstraction was obtained 2-3 years after enrollment. **Table 2** below presents a summary of the available data from the SCDIC as of June 2021.

Table 2. Data in the SCDIC Registry by Type and Number of Patients

Completed Data Collection Form	Number of Patients
Consent and Registration Form	2,440
Baseline medical record abstraction with retrospective history	2,282 (94%)
Baseline laboratory data	2,195 (90%)
Baseline pregnancy history	2,395 (98%)
Baseline Patient Survey –patient-reported outcomes (PROs)	2,420 (99%)
Patient Survey – 1st follow-up	1,867 (77%)
Patient Survey – 2nd follow-up	1,275 (52%)
Patient Survey – 3rd follow-up (ongoing)	253
Patient Survey – baseline plus at least one follow-up	1,963 (80%)
Follow-up medical record abstraction (ongoing)	832
Follow-up laboratory (ongoing)	1,163
Off Study	174

Table 3 presents a summary of the demographics of the 2,440 patients enrolled in the SCDIC Registry that were age 15-45 years of age at the time of enrollment. The people enrolled in the SCDIC Registry have a close relationship with staff at the enrolling centers. This allows for continuous longitudinal data collection and implementation of new surveys since they are in regular contact with the center at which they receive care.

When the original nine NHLBI-funded grants that support the SCDIC Registry end in June 2022, RTI will fund the participating SCDIC centers through their CureSCi OTA agreement with NHLBI. At this time, the name of the Registry will be changed to the **SCDIC+ Registry** because additional core data elements will be added. In addition, patients will be reconsented and the data collection for the SCDIC+ Registry will include annual medical record abstraction and extraction, when feasible, of clinical, laboratory and imaging data, and completion of an annual Patient Survey. A Central IRB will review and approve the human subjects activities being conducted by the individual sites and the DCC for the SCDIC+.

Table 3. SCDIC Selected Demographics

Variable	Frequency (%)
Age at Enrollment	
15-25	1018 (41.7)
26-35	916 (37.5)
36-45	506 (20.7)
Gender	
Male	1051 (43.1)
Female	1389 (56.9)
Race	
Black	2327 (95.4)
Other race	113 (4.6)

External Monitoring. An NHLBI-appointed Observational Study Monitoring Board (OSMB), comprised of a multi-disciplinary group of experts, will perform independent monitoring of overall progress of data collection and patient safety for the SCDIC+ Registry. The OSMB makes recommendations to NHLBI regarding appropriate protocol and operational changes. Any decision to modify the protocol or significantly change study operations may have a substantial effect upon the study. Thus, the OSMB plays an essential role in assuring quality research.

3.2 Other Potential Registries

The NHDR will consider inclusion of other registries that meet the criteria described in Section 4. The GRNDaD Registry is a candidate that can enrich the NHDR with subjects who were enrolled in 2020 or earlier that were 8-15 years old at the time of enrollment. This is a population that is not covered by the SCDIC+ Registry. GRNDaD was initiated in 2017 and has enrolled over 1100 patients of all ages using a web-based REDCap data collection system. Patients were recruited and consented from 11 SCD treatment centers (Medical College of Wisconsin, Case Western Medical Center, Ohio State, Washington University, University of Birmingham, Johns Hopkins All Childrens Hospital, University of North Carolina at Chapel Hill, Beth Israel, Albert Einstein College of Medicine, Benioff Children's Hospital of Oakland, Johns Hopkins Hospital). Johns Hopkins also serves as the GRNDaD DCC.

3.3 RTI International

RTI is an independent non-profit research organization that serves as the NHDR Data Coordinating Center (DCC). In this role, RTI will obtain and de-identify core data from the SCDIC+ registry sites before depositing the data into the NHLBI data repository for public access. If other registries join, RTI will assure de-identification, and will also harmonize the registry data with the SCDIC+ data before submission to the NHLBI data repository. RTI has served as the DCC for the SCDIC Registry since 2016. In this role, RTI has managed the collection of clinical and survey data from over 2,400 patients with SCD from eight treatment centers through a centralized REDCap system hosted by RTI. In addition, RTI has been responsible for the SCDIC data analysis in support of presentations and manuscripts for publication. Under continuation funding from NHLBI to support the NHDR, RTI will continue to serve as the DCC for the SCDIC+ Registry as well as for the NHDR.

3.4 National Heart, Lung, and Blood Institute (NHLBI)

The NHDR was established in 2021 and funded by the NHLBI. The NHLBI is responsible for organizing the overall program, providing direct oversight of the NHDR, overall monitoring of completion of milestones and the timeline. They will collaborate with the DCC and the NHDR investigators in the development and implementation of the NHDR. They will also oversee the review and approval of requests for access to the NHDR public use data.

4. Requirements for Inclusion in the NHDR

4.1 Subject Inclusion Criteria

For the NHDR, the NHLBI is interested in longitudinal core data on subjects that are 8 years of age and older to correspond to gene therapy trial inclusion criteria. Section 5 describes the core set of data that will be collected. Eligible subjects must be in an established IRB approved Registry and meet the following criteria for inclusion in the NHDR:

- Enrolled in the registry as of May 2020
- Signed Informed consent, which includes permission to share data through the NHLBI Data Repository
- Age 8 – 50 years at time of enrollment
- Confirmed SCD diagnosis (any genotype)
- Complete core baseline data from medical records
- Anticipated completion of annual surveys of patient reported outcomes (PROs), unless the subject is off-study or has rescinded their data sharing permission
- Availability and access to the subject's medical records for collection of follow-up core clinical data

Identification of subjects eligible for the NHDR will occur at the individual registry level and will be verified by the NHDR DCC prior to depositing the data into the NHLBI Data Repository.

4.2 Registry Inclusion Criteria

Registries participating in the NHDR must meet the following criteria:

- Utilize existing infrastructure and platforms for collecting and managing data that are already in use at the respective registries
- Auto-extract select clinical core data from the EHR, to the extent possible
- Obtain annual follow-up of clinical data and patient-reported outcomes on all patients
- Submit a copy of the Registry Protocol and Informed Consent documents to the NHDR DCC
- Provide documentation to the NHDR DCC that the registry is following all federal and local regulations regarding the collection and sharing of data, including documentation of annual IRB approval.
- Submit a codebook and data frequencies of Core baseline data to the NHDR DCC for review prior to being accepted as a participating registry in the NHDR
- Submit complete set of clean Core baseline data to the NHDR DCC within 6 months of acceptance in to the NHDR
- Submit follow-up Core data to the NHDR DCC on a quarterly schedule

4.3 Human Subjects and Informed Consent

Transfer of data for the NHDR will not begin until all involved IRBs have approved the protocol for the respective registry and subjects have provided informed consent for the study in which they are enrolled. This may include a single IRB, or multiple IRBs as determined by NHLBI and the participating institutions. Approval of the NHDR Protocol by an IRB is at the discretion of each institution as the data sharing component should already have been approved by the IRB. Documentation of annual IRB review and approval by participating registries must be provided to the NHDR DCC, as well as any significant changes to the study protocol that were approved and that affect the NHDR.

All data, including personal health information (PHI), will be de-identified before submission to the NHDR DCC. Full names and other identifying information will be retained by the enrolling Center. Participants' data will be labeled and stored at the NHDR DCC with coded identification numbers that can be linked to names only by the enrolling Center. The NHDR DCC will not be able to link an individual to their identifying information. More about the transfer of data to the NHDR DCC and the data security is included under Section 6.6.

5. Data Collection

5.1 Core Data

The NHDR Working Group (WG) identified a set of core common data elements to be included in the NHDR. The WG included representatives from the SCDIC and GRNDaD Registries, RTI and the NHLBI. The NHDR WG considered the core data elements that were recommended for SCD genetic studies by the CureSCi, ASH and FDA working groups. In addition, the NHDR Working Group included data elements considered clinically important and that could reasonably be collected as part of an observational natural history study. The CureSCi WGs categorized the common data elements (CDEs) as Core, Highly Recommended, Supplemental, and Exploratory. Core CDEs included demographics, physical exam, medical history, medications, transfusion history and other treatments, routine laboratory results, cardiopulmonary, renal and cerebrovascular assessments, genetic assays, side effects of gene therapy, and patient reported outcomes. The NHDR WG agreed that participating registries would provide most of the clinical Core CDEs that were routinely collected as part of clinical care and all of the Core PROs, plus additional PROs the group thought were important for a SCD natural history study. The SCDIC Registry was already collecting the majority of the data to be included in the NHDR. Participating registries may need to enhance their data collection forms to include additional core CDEs, as needed, to create the set of core CDEs for the NHDR. The full list of the Core CDEs to be included in the NHDR is provided in **Appendix A**. To the extent possible, Core CDEs will be collected using standard definitions as described by the CureSCi Data Standards.

5.2 Data Sources and Tools

The data that are being included in the NHDR are from the patient's medical record and from patient-completed surveys. Data from the medical record will be manually abstracted or obtained through automated data extraction from the electronic health record (EHR), to the extent possible. There is a separate CureSCi effort supporting the identification and piloting of algorithms that can be applied to EHR data to identify select patient characteristics, such as diagnoses, laboratory results, procedures and treatments. This effort is focused on reducing the cost of data collection while maintaining the same standards of data accuracy that occur with manual abstraction and review of

EHR data prior to recording on a case report form. Once these algorithms have been validated, they will be available to the NHDR registries to support more efficient data collection. The DCC will support the extraction of data through open-source tools and applications used in other NIH-funded efforts.

Clinical and Laboratory Data from Medical Record. The medical record will be the source for the following Core CDEs collected at baseline and annual follow-up. Data not obtained/recorded as part of routine care will be missing:

- Diagnoses – co-morbidities and complications related to SCD
- Treatments – current and past including transfusions and medications
- Vital signs
- Healthcare utilization related to pain and other SCD issues
- Results from procedures – MRI, Echo, TCD, biopsies, PFT, transplants, etc.
- Laboratory results

Patient Completed Surveys. Surveys will be completed by the patient on an annual basis and will include information about medications and side effects, tobacco and alcohol use, and PROs. The PROs are key data elements that can only be provided by the patient and are crucial for assessing quality of life. Patients have the right to refuse to answer survey questions, but every effort should be made to collect as many of the PROs as possible. These include:

- Pain frequency and severity (ASCQ-Me Pain Episode)
- Pain intensity (ASCQ-Me Pain Episode NRS)
- Pain interference (PROMIS)
- Anxiety (PROMIS)
- Depression/Emotional impact (PROMIS)
- Fatigue (PROMIS)
- Sleep (ASCQ-Me)
- Physical Function (ASCQ-Me Stiffness)
- Social Function (ASCQ-Me)
- Global Health-10 (PROMIS), first four general questions

5.3 Baseline Data Collection

Baseline Core Data for each subject will largely already have been collected as part of each registry. Additional baseline clinical or laboratory data for the NHDR may need to be obtained from the medical record if it was not part of the original registry data collection plan and it is on the Core Data Element list and available in the record. These additional baseline data will be obtained from the medical record by the end of 2022.

5.4 Frequency of Follow-up Data Collection

At least annually, and as part of participation in the NHDR, follow-up Core Clinical and Laboratory Data are expected to be obtained from the medical record on each patient and a patient survey of PROs is expected to be completed. Annual Core Clinical Data collection will cover the time between the most recent encounter in the medical record and the last encounter through which the previous baseline or follow-up data were collected. Follow-up Patient Surveys will represent the time frames referenced in the survey questions (e.g., last 12 months, last 7 days etc.). Before joining the

NHDR, it is expected that participating registries will have collected at least 2-3 years of follow-up clinical, laboratory and PROs. These data will be available for inclusion in the NHDR in 2023.

6. Data Management

Currently the NHDR consists of only the SCDIC Registry, for which the data already reside at the DCC, RTI. When additional registries join the NHDR, the following sections describe the general procedures for submission of their data to RTI, the de-identification of SCDIC data already hosted by RTI, the harmonization of the data from multiple registries, edit checks to be performed, and the transfer of the combined datasets to the NHLBI Data Repository. Section 6.6 also describes the data security measures RTI will employ to protect the confidentiality of the data while at RTI.

6.1 Overlapping Patients Between SCDIC and Other Registries

RTI will work with other registries to identify the likelihood of duplicate enrollments between registries (e.g., the same center is participating in more than one registry). If duplicates can be identified ahead of time, data for those patients will be linked and updated for the NHDR only under the SCDIC+ Registry. Identification of other duplicates will be the responsibility of the other registry and it will omit data for these subjects from the NHDR submission to RTI. A variable in the NHDR will identify the study each subject is participating in.

6.2 Submission of Other Registry Data to RTI

If other registries join the NHDR, they will be responsible for submitting, at least quarterly, one or more files of cleaned, de-identified Core Data (See Appendix A) in Excel or .CSV format to RTI through a secure portal. The most recent codebook/data dictionary which clearly documents what is in each datafile will also be included in the transfer. Variable names should follow the naming scheme and format laid out in the data dictionary. Each subject's record must be labeled with a unique identification number that is maintained across the datafiles. The same identification number must be used for each data submission so that longitudinal data can be linked to the appropriate individual.

De-identification of the data will follow the NHLBI guidelines (https://www.nhlbi.nih.gov/grants-and-training/policies-and-guidelines/guidelines-for-preparing-clinical-study-data-sets-for-submission-to-the-nhlbi-data-repository) except that the original registry's subject ID will be maintained for ease in linking follow-up data by the NHDR DCC. RTI will replace the registry ID number with an NHDR ID prior to submission of the data to the NHLBI data repository.

The first submission of data will include the baseline core data on the previously enrolled patients with complete medical record data. These data will include extant core data that was collected at the time of enrollment into the other registry, as well as any supplemental core data from the medical record that needed to be subsequently collected as part of the NHDR requirements. It is understood that core patient-reported outcomes (PROs) that were not previously part of another registry's data collection package will not be included in the first submissions of data but will be included in future submissions of longitudinal follow-up data.

6.3 SCDIC+ Data

All SCDIC+ data, whether they are part of the NHDR Core CDEs to be placed into public access or not, are maintained at RTI. These include data elements that are considered personal health information (PHI) and possibly other data that is of interest to SCDIC researchers only (e.g., pregnancy

data). In preparation of data submission to the NHLBI Data Repository for the NHDR, RTI staff will de-identify the data according to NHLBI guidelines and rename and reformat, as needed, variables according to the NHDR Core Data Dictionary.

6.4 Quality Control of Data

Checks for internal consistency with respect to dates, acceptable ranges, required items, and skip patterns will be performed by all participating registries under the data management of their respective protocols.

For SCDIC, these checks are set up as validation at the time of data entry into REDCap. In batch, the data undergo additional automated, electronic edits that could not efficiently be included in the data entry screens, such as cross-form editing. Edit checks that are performed in real time are replicated by the SCDIC DCC in batch. In addition, complex within-form and across-form consistency and logic checks are applied. These checks are based upon the specifications appropriate to each of the data collection forms. Any failures are reported to the Center as error resolution reports. Center personnel enter the corrections for the keyed data into the data management system, and an audit trail of corrections are maintained.

As the DCC for the NHDR, RTI will apply reasonable range and logic checks to other registry data prior to submission to the NHLBI repository.

6.5 RTI Data Security

All subject direct identifiers will be kept solely at the clinical sites. The files maintained at RTI contain limited identifying information and protect subject confidentiality to the extent possible. Safeguards are in place to greatly decrease the chances that characteristics of a case can be linked to the individual participating in the study. These include de-identification of data, transfer of datasets through secure portals, and storage of datasets on HIPAA-compliant servers at RTI.

Access to datasets at RTI will be authenticated for users based on the definition of study roles. These roles allow access only to those systems needed by a user to carry out the responsibilities of his/her role. Role-based access protects data confidentiality and security and increases data integrity.

The secure RTI portal and servers can only be accessed by secure web connection. To log-in, users require a username and password that is unique to each individual. Authorized requests for access to the RTI portal will be filled by designated RTI staff. Any direct, administrative access to the system and database by RTI staff is both limited by location (IP address) and to specific computers using cryptographic key pairs.

6.6 Data Transfer to NHLBI Data Repository

Preparing the Datafiles. At least annually, the NHDR DCC will create de-identified datasets of the clinical and PRO data following the NHLBI guidelines referenced in Section 6.2. Specific file formats for transfer will be determined, but it is expected that one datafile of each data type will be prepared for each update, and these files will include the combined harmonized data from both registries. Each record will contain a new unique random ID, which replaces the original subject ID, and a flag for which study the subject belongs. The random ID will be assigned once to each subject and then used for the lifetime of the NHDR and each update. The link between the random ID and the study's subject ID will

be maintained by the DCC as a protected file. The DCC will also update documentation of datasets according to NHLBI/BioLINCC instructions.

Data Transfer for Public Access. Following NHLBI Data Repository guidelines for data transfer, the DCC will deliver the datasets, data forms and documentation to the Data Repository (i.e., BioData Catalyst) as well as the codebooks and annotated forms, which will be updated as needed. Metadata will also be made available through the CureSC MetaData Catalog at https://curesicklecell.rti.org/.

All data transfer will be conducted through a secure password-protected portal with access only to authorized users. Once transferred, the responsibility for data security of the files belongs to NHLBI.

7. Data Sharing and Governance

The primary goal of the NHDR is to establish a data inventory that will be shared with qualified researchers interested in identifying contemporaneous controls for their studies or in studying various aspects of the natural history of SCD. De-identified patient-level data will be made available to researchers through an application and approval process as part of the NHLBI Data Repository, BioData Catalyst (https://biodatacatalyst.nhlbi.nih.gov/) and BioLINCC (https://biolincc.nhlbi.nih.gov/home/). To protect the confidentiality and privacy of the subjects, investigators granted access to the data must adhere to strict requirements incorporated into a standard Data Use Agreement. In accordance with NHLBI policy, researchers will also be required to submit an approval from their Institutional Review Board (IRB).

8. Schedule and Timeline

Task	Start or Due date	Responsible party
Pilot NHDR Core Data collection at 3 SCDIC Registry sites under the original protocol	October 2021 – April 2022	RTI
Submit NHDR Procedures to CureSCi leadership	January 2022	NHLBI
Revise SCDIC protocol and consent to include NHDR Core Data and data sharing procedures (and rename as SCDIC+ Registry)	February 2022	RTI, SCDIC+ centers
Evaluate results from pilot data collection	April 2022	RTI
Submit SCDIC+ to Central IRB	May 2022	RTI
Approval from Central IRB for SCDIC+	June 2022	IRB
Reconsent SCDIC subjects for SCDIC+	June – October 2022	SCDIC+ centers
Collect additional baseline Core Data added specifically for the NHDR	June – December 2022	SCDIC+ centers
Collect prospective data for NHDR	Starting July 2022	SCDIC+ centers
Submit NHDR data to Contemporaneous Control Data Resource (BioLINCC and BDC) semi-annually	Starting January 2023	RTI