PROTOCOL

PROTOCOL TITLE: Comprehensive Sickle Cell Centers Collaborative Data Project

CSCC PROTOCOL NUMBER: Version 5.3

DATE: July 3, 2008

CSCC PROTOCOL CHAIR:
Susan Lieff, PhD
(919) 408-8000 ext. 467
slieff@rhoworld.com

CSCC PROTOCOL TEAM:

Ronald W. Helms, PhD
(919) 408-8000 ext. 222
helms@rhoworld.com

Kwaku Ohene-Frempong, MD
(215) 590-3423
ohene-frempong@email.chop.edu

Karen Kesler, PhD
(919) 408-8000 ext. 244
kkesler@rhoworld.com

Carlton Dampier, MD
(215) 427-5261
carlton.dampier@drexel.edu

Samir Ballas, MD
(215) 955-8485
samir.ballas@mail.tju.edu

Win Wang, MD
(901) 495-3497
winfred.wang@stjude.org

Cage Johnson, MD
(232) 442-1259
cagejohn@usc.edu

Zora Rogers, MD
(214) 648-3896
zora.rogers@utsouthwestern.edu

George Buchanan, MD
(214) 648-8594
george.buchanan@utsouthwestern.edu

Matthew Heeney, MD
(617) 355-7700
matthew.heeney@tch.harvard.edu

Marsha McMurray
(919) 408-8000 ext. 279
mmcmurray@rhoworld.com

Petra LeBeau, ScD
(919) 408-8000 ext. 27
plebeau@rhoworld.com

STUDY SITES: Southwestern Comprehensive Sickle Cell Center (CSCC); Bronx CSCC; Duke-UNC CSCC; CSCC Statistics & Data Management Center (SDMC); Children's Hospital of Philadelphia CSCC; Northern California CSCC; Boston Medical Center CSCC; St. Jude Children's Research Hospital CSCC; Cincinnati CSCC; Marian Anderson Sickle Cell Anemia Care and Research Center

Note: This protocol originated under the Comprehensive Sickle Cell Centers (CSCC) program. The CSCC program has been replaced by the Basic Science and Translational Research Program (BTRP). The NHLBI mandated that new enrollment into the C-Data protocol cease as of March 31, 2008, the end of the funding cycle. Therefore, all references to the CSCC Centers and sites refers to the program under which subjects were enrolled.

This protocol is sponsored by the National Heart, Lung, and Blood Institute.

1 SYNOPSIS

Title:	Comprehensive Sickle Cell Centers Collaborative Data Project (C-Data)
Overview:	The Comprehensive Sickle Cell Centers (CSCC) C-Data project will establish a comprehensive database of individuals from participating Centers who are potentially eligible for inclusion in any sickle cell research study. Such studies include observational (incidence/prevalence, cohort, case-control, cross-sectional) as well as interventional (randomized clinical trials) investigations. Possible study endpoints include traditional clinical and therapeutic measurements, health resource utilization, and patient-reported outcomes. The C-Data project encompasses these clinical and outcomes databases within the Patient Database component of the project, and includes additional components of specifically developed tools and methodologies required for practical, quality data capture, management, and utilization. A secure web-based Electronic Data Capture (EDC) system is provided to the Centers to allow data submission to the Statistics and Data Management Center (SDMC). The National Heart, Lung, and Blood Institute (NHLBI) and the CSCC Steering Committee fully support public dissemination of results generated by the C-Data project following completion of this project. Future, collaborative efforts between CSCC and non-CSCC investigators will be encouraged, especially for large proposed projects, but will not be required for all proposed projects. Adherence to all Federal guidelines regarding privacy and confidentiality is and will be rigorously maintained.
Objectives:	The primary goals of the CSCC C-Data project are to: 1) Provide a complete list and general overview of the CSCC patient population. 2) Provide an information base for designing collaborative CSCC studies, especially sizes of subgroups of potential patients with specific characteristics, such as headache, acute chest syndrome, or priapism, or who meet specific inclusion/exclusion criteria. 3) Support identification (by a local site) of specific individuals eligible for observational as well as clinical studies, including treatment specific Health-Related Quality of Life (HRQoL) and outcome studies. 4) Provide basic HRQoL data on a large sample of adults and children with sickle cell disease (SCD) for comparison with population norms. The secondary goals are to: 1) Provide the opportunity for exploration of relationships between patient characteristics and various clinical events, and allow measurement and reporting of frequencies and patterns of specific sequelae associated with SCD. 2) Provide a SAS™ database with a single structure for all sites, designed specifically for collection of clinical and other data for individuals with SCD.
Hypotheses / Estimates:	Initially, C-Data project data will not be used to test specific research hypotheses. Future use of these data for hypothesis testing or other purposes will require submission of proposals to the CSCC Steering Committee and will be subject to the standard review process.
Study Population:	All adult and pediatric patients who have been seen within the last 24 months in the hospital or clinical setting, and are expected to return episodically or regularly for care at one of the CSCCs, are eligible for inclusion in the C-Data project. Deceased patients, and patients considered inactive, will be excluded.
Sample Size:	Enrollment of a minimum of 4,000 participants is expected by the end of the funding cycle, March 31, 2008.
Data Management:	Statistical and data management support, as well as validation, will be provided by the SDMC.
Human Subjects:	There are no known risks to the subjects participating in this study.

2 TABLE OF CONTENTS

3 ACRONYMS

CSCC	Comprehensive Sickle Cell Centers
C-Data	Collaborative Data Project
CRF	Case Report Form
DSMB	Data Safety Monitoring Board
EDC	Electronic Data Capture
GCP	Good Clinical Practice
HIPAA	Health Insurance Portability and Accountability Act
HRQoL	Health-Related Quality of Life
ICH	International Conference on Harmonisation
IRB	Institutional Review Board
MOS	Medical Outcome Study
NCHSR	National Center for Health Services Research
NHLBI	National Heart, Lung, and Blood Institute
NIH	National Institutes of Health
PD	Patient Database
PedsQL 4.0	Pediatric Quality of Life Inventory, Version 4.0
PRC	Protocol Review Committee
PSQ	Patient Satisfaction Questionnaires
RhoFED	Rho Federal Systems Division, Chapel Hill, NC
SDMC	Statistics and Data Management Center, located at Rho Federal Systems Division, Chapel Hill, NC
SCD	Sickle Cell Disease
SOP	Standard Operating Procedure

4 STUDY OBJECTIVES AND PURPOSE

The Comprehensive Sickle Cell Centers Collaborative Data Project (CSCC C-Data) will establish a comprehensive database of individuals from participating Centers who are potentially eligible for inclusion in any sickle cell research study. Such studies include observational (incidence/prevalence, cohort, case-control, cross-sectional) as well as interventional (randomized clinical trials) investigations. Possible study endpoints include traditional clinical and therapeutic measurements, health resource utilization, and patient-reported outcomes.

The C-Data Project encompasses the individual components of this comprehensive Patient Database, as well as the specifically developed tools and methodologies required for practical, quality data capture, management, and utilization.

The primary elements of the CSCC C-Data Project include the following:

1. A database of basic clinical and diagnostic information on individuals with SCD.

2. A database of Health-Related Quality of Life (HRQoL) and health resource utilization information on individuals with SCD.

3. Electronic Data Capture (EDC) System, a data entry system available online to serve as the sole information-capturing entity.

4. A SAS™ analysis database of the clinical patient data with a single structure for all sites, for individual Centers' use beyond the current CSCC funding period.

Goals of the C-Data Project are as follows:

<u>Primary</u>

1. Provide a complete list and general overview of the Centers' patient population.

2. Provide an information base for designing collaborative CSCC studies, especially sizes of subgroups of potential patients with specific characteristics, such as headache, Acute Chest Syndrome, or priapism, who meet specific inclusion/exclusion criteria.

3. Support identification of specific individuals eligible for observational as well as clinical studies, including treatment-specific HRQoL and outcome studies.

4. Provide basic HRQoL data on a large population of adults and children with SCD for comparison with population norms.

Secondary

1. Provide the opportunity for exploration of relationships between patient characteristics and various events, and allow measurement and reporting of frequencies and patterns of specific sequelae, healthcare utilization, and economic burden associated with SCD and its complications.

2. Provide a SAS™ database with a single structure for all sites, designed specifically for collection of clinical and other data for individuals with SCD.

5 BACKGROUND AND RATIONALE

5.1 Introduction

Broadly defined, a patient database or medical registry is a systematically gathered set of clearly defined health and demographic data for individuals with specific health characteristics, located in a central database for a predefined purpose (Solomon 1991). The dramatic increase in the number of such patient registries in recent years is likely the result of many factors, including progress in information technology, increased demand for accountability, increased pressure from health-related interest groups focused on a particular disease or condition, and public health agency emphasis on surveillance and research (Arts 2002, Sorenson 2001, Solomon 1991). These databases may serve as tools for monitoring and improving quality of care, to facilitate policy and administrative decisions or a resource for clinical and epidemiological research (Mitri 2002, Wilson 1993, Schulman 1993). A good patient database or disease registry can be a vital resource for researchers in providing an efficient way to identify specific subsets of patients for both observational studies and clinical trials.

A patient database may be constructed to include a limited or more extensive amount of patient data. A limited database would contain only basic demographic and diagnostic data and serve solely as a source population for sampling subjects. A more elaborate database would include current and previous treatment and medical conditions and would support very specific identification of potential study subjects, as well as provide the opportunity for preliminary exploration of relationships between patient characteristics and various events.

With a limited database, data can be stratified by age and diagnosis, although these data may not adequately identify the subgroup truly eligible for study participation. While this approach will present a critical first step in providing a count of adults or children, or men or women of a specific age or diagnosis, use of these data exclusively will necessitate additional screening to assess study inclusion and exclusion criteria and suitability. This Collaborative Data Project protocol outlines an intermediate approach to developing a patient database between the two scenarios described above. As outlined in this proposal, Patient Database information can be used to assess feasibility of potential future studies, as well as to identify Centers and individuals within Centers who can be seen realistically, as potential participants.

Several steps were taken to obtain information important for planning the content of the database and the process of data collection.

5.2 Survey of CSCC Centers

During the spring of 2002, all Center Directors or designees were surveyed to assess each Center's current electronic patient database utilization, the content of the database and extent of completeness, and the type of software used. Directors were also asked to provide an estimate of the number of individuals included in the database considered active patients, and to send the SDMC an unused copy of their data collection forms or screens, or list of data

fields. Specific concerns about or perceived obstacles to participating in a collaborative multi-center patient database were solicited as well. An overview of this assessment is presented in Table 5.2.1 below.

Table 5.2.1 Status of Sickle Cell Patient Data in Electronic Format by CSCC Centers/Sites

CSCC Center	Geographic Location(s)	Computerized Database	Age Group	# of Patients
Boston Medical Center	Boston Medical Center	MSAccess 97	Adult	125
			Peds	200
	Brigham and Women's	Handibase	Adult	100
	Univ MS Medical Center Jackson, MS	MSExcel	Adult	200
	Children's Hospital Jackson, MS	No database	Peds	400
Bronx		Unknown		
Children's Hospital of Philadelphia	Philadelphia One project in Ghana	FileMaker Pro	Peds	500
Cincinnati	Children's Hospital University Hospital	MSAccess 97	Peds	270
Duke-UNC	Duke University – Duke Sickle Cell Clinic (adults)	MSExcel	Adult	450
	Duke University – Duke Pediatric Sickle Cell Clinic; Satellite clinics – Fayetteville, Greensboro	MSAccess 97	Peds	450
	University of North Carolina at Chapel Hill Sickle Cell clinic	MSAccess 97	Adult	264
			Peds	190
Marian Anderson	Thomas Jefferson Hospital	None	Adult	
	St. Christopher's Hospital	MSExcel	Peds	440
	Louisville, KY			
Northern California	Children's Hospital and Research Center at Oakland	Palm Pilot Data	Peds	
		None	Adult	
St. Jude Children's Research Hospital	Pediatric Hematology Center of Memphis (2 sites) St. Jude Children's Research Hospital	MSAccess 97	Peds	750
University of Southern California	Los Angeles, CA	MSExcel SAS		
University of Texas Southwestern	Children's Medical Center of Dallas Outreach Clinics, Tyler, Paris	MSAccess 97	Peds	600+

Collectively, approximately 3400 pediatric and 1140 adult patients were included in participating Center databases at the time of the survey. These figures were used to estimate projected enrollment. Since that time, the delayed date of study initiation, need to develop

standard development for prospective data collection and limited site resources led to the establishment of more realistic target enrollment goals of a minimum of 4,000 patients.

At the time of the original survey, one Center expressed concern that the task of enrolling adults would be very time consuming, and 3 centers mentioned HIPAA compliance as a potential problem we would face in implementing the Patient Database.

Also, Data fields were compared across Centers that provided this information, and the number containing each field tallied. The separate data fields were then grouped into general categories to describe the type of information included. Table 5.2.2 below describes the number of CSCC Centers with all, some, or no electronic data in each category by pediatric or adult site.

Table 5.2.2 Number of Centers Collecting Specific Types of Data in Electronic Format

Type of Information	PEDS				ADULT			
	Number of Centers With This Digital Information				Number of Centers With This Digital Information			
	All	Some	None	DK	All	Some	None	DK
Basic Demographic	7			1	2		1	1
Demographic II	3	2	2	1	2		1	1
Demographic III	1	2	4	1	2		1	1
Family History	4	2	1			2	1	1
Current Medication	7			1	2		1	1
Medical History	6	1		1	2		1	1
Recent Medical History	7			1	2		1	1
Prior Medical History	6	1		1	2		1	1
Study Participation	3		4	1	1	1	1	1
Transfusion	7			1	2		1	1
History of Painful Episodes	5	2		1	2		1	1
Lab	4	2	1	1	2		1	1
Surgical History	5	2		1	2		1	1
Reproductive Health	1	1	5	1	2		1	1
Psychosocial							3	1
Health Care Utilization	5	2	0	1	2		1	1
Secondary Health Care Utilization	0	0	7	1			3	1

5.3 Review of Eligibility Criteria in the Initial Multi-Center Proposals

In order to begin the process of developing the C-Data Patient Database content, 17 clinical proposals initially submitted for consideration as CSCC multi-center studies were reviewed, and a comprehensive list of inclusion/exclusion criteria were compiled. As presented in table 5.3 below, a total of 62 items were listed as eligibility criteria. As written, these criteria are not necessarily mutually exclusive and include some degree of overlap, e.g., “lung impairment” and “asthma/wheezing.” Twenty-seven criteria were listed for one study only; the remaining 35 were listed for 2 or more studies. It is also important to note that, while inclusion of various lab parameters in a patient database may assist in identification of

potential subjects, the need to re-evaluate these parameters at the time of study enrollment may be necessary to assess current eligibility.

Table 5.3 Inclusion/Exclusion Criteria for the Initial 17 CSCC Proposed Multi-Center Studies

Variable	# studies using criteria
Age	15
SC diagnosis	14
Ability to give Informed Consent	8
Acute chest syndrome, date	8
Pain episodes, dates in last year	6
Acute events, dates in last year	5
Previous clinical stroke, date	5
Vaso-occlusive episode, date	5
Splenic sequestration, date	5
Pain classification, date in last year	4
HU therapy status current, dose	4
Chronic transfusion, current status	4
Previous sub clinical stroke, date	4
Pregnancy status	4
Renal insufficiency	4
Priapism episodes, in last year	3
Current contraceptive practice	3
Chronic transfusion, prior	3
Hospitalization reasons, last 2 years	3
Total hemoglobin concentration	3

Variable	# studies using criteria
Malignancy	3
Retinopathy	3
Dactylitis, date	3
Hepatic impairment	3
Lung impairment	3
Diabetes mellitus	3
Hypertension	3
GI ulceration or bleeding	3
Avascular necrosis	3
Pain treatment per episode	2
Serum creatinine	2
HU therapy status, prior, dose	2
Hospitalizations, last 2 years	2
Alcohol/Drug use	2
Compliance with care	2
Sex	1
Serum transaminase	1
Serum albumin	1
Platelet count	1
ALT	1
Baseline coagulation profile	1
Alloimnunization to erythrocyte antigens	1
Pitted red cell count	1
WBC count	1

Variable	# studies using criteria
WBC, previous	1
Pitted red cell, total Hgb, WBC at 18. 24 months	1
Abnormal blood flow velocity, MCA, ICA	1
Pulse oximetry, current	1
HLA identical sibling donor	1
ABO incompatibility	1
Cognitive impairment	1
Corticosteroid tx, prev 2 weeks	1
Systemic bacterial infection	1
Acute asthma, wheezing	1
Infarct, hemorrhage on MRI	1
Hx head injury resulting in neurological symptoms	1
Abnormal neurological exam, focal findings	1
Chronic illness not SCD related	1
Mini Mental Status Exam	1
POMS score	1
History of claustrophobia	1
History metallic implants	1

5.4 Support from the Broader Scientific Research Community

In November 2003, the conference "New Directions for Sickle Cell Therapy in the Genome Era" was held at the NIH in Bethesda, MD. The conference was supported by the National Human Genome Research Institute, the National Heart, Lung, and Blood Institute, the National Institute of Diabetes and Digestive and Kidney Diseases, the Office of Rare Diseases, the Fogarty International Center, and the Foundation for the National Institutes of Health. The goal of this meeting was to consider how new genomic methods might be used to better understand the biology of Sickle Cell Disease as well as develop more effective therapeutic and preventive strategies. Among the more than 30 priorities that emerged from this meeting, the establishment of ***"an innovative multidisciplinary Sickle Cell Disease Research Network with a central prospective registry of several thousand well phenotyped patients"*** was given the highest priority. The CSCC Center Directors endorsed formation of a committee of investigators from within and outside of the CSCC system, to begin examination and definition of sickle cell subphenotypes based on current pathophysiological knowledge and available technologies, which would form the basis of the database for this network. The second step in establishing such a Research Network will be planning for development of a repository of biological specimens including plasma, mRNA, transformed cell lines and DNA samples linked to that database. While these goals are clearly beyond the scope of the Collaborative Data Project as it is currently defined, and their achievement will require additional resources, establishment of the CSCC Patient Database will initiate the process by developing the necessary infrastructure within the CSCC network, at the SDMC and Clinical Centers and set the stage for this more comprehensive future effort.

5.5 Health-Related Quality of Life

The CSCC RFA charges the SDMC with establishing a sickle cell Health-Related Quality of Life and Health Care Utilization/Burden of Care database. The SF-36®, the Pediatric Quality of Life Inventory, and the Patient Satisfaction Questionnaire will be used to achieve this goal. The SF-36® is the most frequently administered health status measure currently in use. RAND Health® developed the 36-Item Short Form Health Survey 1.0 as a part of the Medical Outcome Study (MOS) to explain patient outcome variations. The SF-36® is derived from 8 health concepts: physical functioning, bodily pain, physical role limitations, emotional role limitations, emotional well-being, social functioning, vitality, and general health perceptions. The recently revised version (SF-36v2®), developed by QualityMetric Inc., offers a number of improvements including updated population norms, norm-based scoring, increased range and precision, and major enhancements in item wording and formatting.

The Pediatric Quality of Life Inventory (PedsQL 4.0) is a multi-dimensional child self-report and parent proxy report developed as core measures to be integrated with disease specific modules. The PedsQL 4.0 has four domains: physical functioning, emotional functioning, social functioning, and school functioning. The PedsQL 4.0 has child self-report formats for age groups 5-7, 8-12, and 13-18, and parallel parent proxy reports for

the same age groups. It also has a parent proxy report for toddler range (2-4, no parallel child report) with 3-point Likert scale and a happy-to-sad faces scale. The PedsQL can be administered in five minutes, and can be self-administered by children as young as five years old.

The National Center for Health Services Research (NCHSR) created the Patient Satisfaction Questionnaires (PSQs) to provide a patient satisfaction survey appropriate for general population studies that would produce reliable measures of concepts important to the preparation, management, and evaluation of health services delivery programs and address doctor/patient communication, affected resource availability, and specific hospital experiences. The PSQ-18 takes approximately 3-4 minutes to complete.

5.6 CSCC Collaborative Data Project Subcommittee

A CSCC subcommittee of Center investigators assisted in data report form development and revision to ensure inclusion of a sufficiently broad scope of readily available diagnostic, lab, treatment, intervention, and utilization data while simultaneously limiting data collection and its effect on clinical practice. This committee has accomplished its work via conference calls, brief 2-3 hour in-person meetings held prior to quarterly Steering Committee meetings, and daylong work sessions. This work will continue over the course of the project.

5.7 Electronic Data Capture and Electronic Data Transfer

During the survey process, many Center investigators and data management staff members expressed concerns related to the collection, entry, and submission of data for the C-Data Project. Therefore, a web-based Electronic Data Capture (EDC) System will be used for this project. Section 6.2.2 further describes the EDC System.

5.8 The SAS™ Patient Database with a Single Structure

This component was included as part of the Collaborative Data Project in response to requests from Center Directors for access to and use of their patient clinical data.

6 STUDY METHODS

6.1 Patient Database Structure and Content

As stated above, the C-Data Project encompasses individual components including the clinical and Quality of Life databases, as well as the specifically developed tools and methodologies required for practical, quality data capture, management, and utilization. This methods section will focus on the Patient Database component. The descriptions of the data collection instruments have been grouped to parallel the timing of use of these forms.

Table 6.1 Study Flow Sheet

Form	Study Visit/ Enrollment[1]	…	…	Annual History	Annual Interview	…
Consent Authorization[2]	✓					
Enrollment[2]	✓					
History[2]	With Enrollment					
HRQoL/PSQ	Once during the first year, at a well visit prior to completion of the 1st Annual Interview, and possibly a second and third time during subsequent years.					
Annual History[3]		Annually beginning one year after Enrollment		✓		
Annual Interview[3]		Annually beginning one year after Enrollment			✓	
Termination or Transfer		At any time in the event of a patient's death, termination, or transfer				

1. Enrollment can occur at a patient visit or at a special study visit.
2. Complete one time only.
3. All Annual forms and Annual Interviews are to be completed within the context of routine patient visits: any sickle cell-related, in-person interaction between a patient and Sickle Cell Center clinical staff.

6.1.1 Patient Screening/Enrollment Log

All patients with sickle cell disease who are regularly seen at participating sites should be considered for enrollment into the Collaborative Data Project. After a site has been enrolling subjects for at least three months, the study coordinator will capture brief, non-identifying data on each patient seen during two non-consecutive weeks each year. These data will be captured (to the extent allowed by local IRBs) and provided to the SDMC via the Patient Screening/Enrollment Log. Sites not permitted to submit the form per their IRB will be asked to provide aggregate data. The DSMB will review this information to ensure there is no systematic exclusion of specific patient subgroups that would create bias in the database.

6.1.2 Enrollment and Initial Data Collection (Informed Consent/Authorization, Enrollment, Medical History, Health-Related Quality of Life, Encounter Forms)

Subject enrollment can occur either at the time of a special recruitment event (see Section 6.3) or during a routine clinic visit, following appropriate completion of Informed Consent and Authorization materials. Participating centers will utilize their own respective recruitment method for adult and pediatric patient participation. Completion and data entry of an Enrollment Form will indicate an individual has been formally included in the CSCC Patient Database. At this time s/he is assigned a CSCC Patient ID that will "follow" him/her during the entire CSCC 5-year period. This Patient ID will be used as a designation for this individual for the Patient Database, as well as for any other CSCC multi-center or single-center studies of which s/he is a part. The Enrollment Form is constructed very simply and includes 7 data fields pertaining to the individual and 2 additional fields for the date and initials of the person providing the data. Both the Center and Hospital codes will appear on the screen automatically, or will be preprinted on the data forms. All information on the Enrollment Form can be obtained from the medical record.

The Medical History Form is completed one time only, at the time of Enrollment. This form has been designed to provide an overview of the individual's medical history relevant to Sickle Cell Disease. The Medical History Form has been subdivided into two parts. Part I contains data obtained from medical records, Part II from patient interview. For ease of use, Part II comes in two formats - IIA for interviews directly with the patient and IIB for interviews with a parent/accompanying adult and patient. All interviews can be conducted in approximately five minutes. Completion of Part I may require review of medical records from other medical facilities. Therefore, all participants should have current, signed medical release forms on file.

The SF-36® and PSQ-18, or the PedQL will be administered at the enrollment visit if the patient is well, or at the first available well visit thereafter. These are the only self-administered data collection tools used as part of the Collaborative Data Project and if necessary may be administered at any well visit during the first year following enrollment prior to completion of the Annual Interview. Completion of these forms will occur in the context of a clinic or hospital visit.

If enrollment occurs at a clinic or hospital visit that results in an admission, the dates of admission and discharge and primary discharge diagnoses are recorded on the Semi-Annual Form. The type of visit and facility, and up to 3 discharge diagnoses per admission, can be selected from "drop-down" or check boxes in the EDC system. These data will allow the SDMC to generate reports regarding the utilization of health care services and the economic burden associated with sickle cell disease and its sequelae. Although some underreporting may occur, these data will indicate the frequency and type of services utilized, and number and length of hospital admissions for a large portion of sickle cell patients from these Centers. Small ancillary validation studies may be used to assess the extent of underreporting.

6.1.3 Post-Enrollment Data Collection (Annual Form, Annual Interview)

The Annual Form will be completed annually beginning twelve months following enrollment. This form contains data obtained from medical records and includes brief transfusion, antibody, medication, lab, and diagnostic test data, as well as whether the patient is enrolled in other research studies. The Annual Interview will be completed annually beginning 1 year following enrollment. The Interview comes in two formats - IIA for interviews directly with the patient and IIB for interviews with a parent/accompanying adult and patient. The Part II interview includes brief demographic, recent transfusion and (if adult) tobacco and alcohol use. Both interviews can be conducted in approximately 5-10 minutes. These annual data are to be collected in the context of a patient encounter and therefore may occur before or after the actual annual date. The Part II interview may also be completed via telephone if the patient does not return to the clinic within fourteen months of enrollment or previous annual interviews.

In addition, following standard approval of separate supplementary protocol, C-Data participants may be asked for consent (separate consent) to provide a blood sample for inclusion in a sample repository for future research.

6.1.4 Termination/Transfer Form

In the event that a patient dies, transfers to another medical facility, or terminates with a Center site, a Termination/Transfer Form will be completed. An additional Termination/Transfer Form may be completed if the subject was not deceased when first form was completed.

6.2 Data Management

6.2.1 Submission of Data

Centers will be expected to utilize electronic data submission to decrease lag time between patient identification and enrollment and to facilitate overall usefulness of the Patient Database.

6.2.2 Web-Based Data Capture

EDC, Rho's internet-based remote data entry system, will be used to capture the data for the C-Data Project. Using this system, the clinic's study coordinator or data coordinator uses an internet browser (Internet Explorer or similar) to key data into electronic case report forms. Univariate data validation tests are performed as the data are keyed and most implausible data values are resolved immediately. Data are not stored on the site's computer. At the end of each "page," data are submitted to Rho's secure web server using SSL (128 byte public key encryption methodology) and stored in the study's

"operational database." (The database used for capturing, validating, updating, and storing the data is called an "operational database.") The database is backed up nightly; backup tapes are saved in a secure, off-site location. At any time site personnel may log in to the system, review and correct previously entered data, or key additional data.

The pages will be accessible via the CSCC website and require Center-specific user ID/password privileges. The data will be converted to intermediate datasets prior to incorporation into the C-Data Project format (SAS datasets).

6.2.3 Data Quality Assurance and Quality Control

CSCC Collaborative Data Project Procedures for Quality Assurance and Quality Control:

1. Train site personnel to use CSCC EDC system.

2. Train site personnel about C-Data protocol, data standards, and procedures.

3. Set up C-Data EDC system to perform real-time error detection and correction procedures.

4. Centrally monitor site data entry and correction processing. Intervene when site performance is inadequate.

5. Prepare reports of site-specific performance. Communicate with the PI as needed and with the CSCC Steering Committee semi-annually.

6.3 Patient Recruitment and Enrollment

Given the sporadic, irregular, and often unplanned patterns of utilization of health care and treatment services by many sickle cell patients, a recruitment and enrollment approach based on Center visits may result in extremely slow patient accrual and a database that is too small to be of use during this funding period. Therefore, if necessary, the SDMC will assist each Center in planning an individually tailored Patient Enrollment Initiative designed to contact, recruit and enroll a large proportion of Center patients within a specific timeframe. The specific goals of the initiative are to inform as many patients as possible about the Collaborative Data Project, request informed consent and authorization, have the HRQoL and Patient Satisfaction questionnaires completed, and submit the data for these individuals. The Initiative might be developed as a mass mailing followed up by phone calls to schedule patients to come in to discuss the C-Data Project and possibly agree to participate. Depending upon the Center and its patient population, the Initiative might be planned as part of an educational or social event. Specific funds have been included in the proposed C-Data Project budget to support such a Patient Enrollment Initiative at each Center. Patients will be compensated for their time after completing the Informed Consent, Authorization, Medical History Part II, HRQoL, and Patient Satisfaction questionnaires.

Clearly, all Center patients will not be reached through an enrollment effort like this, and recruitment will continue for an extended period of time. Once the Collaborative Data Project has begun, Centers will begin enrolling patients as they return for care. The SDMC will create and print a series of standard brochures and flyers to assist the Centers with patient recruitment and enrollment.

Participating centers will utilize their own respective recruitment method for adult and pediatric patient participation. Enrollment of a minimum of 4,000 participants is expected by the end of the funding cycle, March 31, 2008.

6.4 Collaborative Data Project Timeline

The suggested Collaborative Data Project Timeline is provided below.

Steering committee meeting approval to proceed	September 2003
Subcommittee develop data items, data definitions	October 2003 – December 2003
Steering committee approves final protocol	January 2004
SDMC begin dialogue/planning w/centers for data transfer	February 2004 – August 2004
PRC approval	April 2004
Complete training materials	March 2004 – April 2004
Distribute IRB packets to sites	March 2004 – April 2004
Development and planning of training	April 2004
Central training for site data coordinators	April 2004
DSMB approval	May 2004
Begin site submission of IRB materials locally	December 2004
Initiate pilot test of EDC	December 2004 – February 2005
Refine EDC system	February 2005
Begin patient registration, Patient Database	March 2005
First Subject Enrolled	March 2005
Conduct training update	April 2005
Begin data transfer process	June 2005
First data transfer	November 2005
1000th Subject Enrolled	February 2006
Revision of annual forms	March 2006 – August 2006
Begin monitoring visits	May 2006
DSMB approval of amendment (Protocol v5.2)	September 2006
Initiation of new sites (UAB, UOC)	October 2006
First semi-annual data collected	October 2006
Conduct training update	October 2006
Conduct training update	October 2007
Revision of semi-annual forms	May 2008
Approval of amendment (Protocol v5.3)	July 2008

7 THE STATISTICS AND DATA MANAGEMENT CENTER

The SDMC will provide operational support for all C-Data Project related activities, including development of the subcommittee work group meetings and conference call arrangements, agendas, minutes and the project Manual of Procedures. Rho personnel will provide technical support needed for developing communication, data management, and software development and support for individual Centers.

8 COMMUNICATION

A well-developed and easily used communication system will facilitate successful achievement of the Collaborative Data Project objectives. A web-based Network Communication System (RhoNET™) developed by Rho is currently being used to support CSCC work. The SDMC will establish and maintain a C-Data area of this CSCC website to facilitate ongoing day-to-day communication between the SDMC, Center data coordinators and data entry staff, Steering Committee and Patient Database subcommittee, and the National Heart, Lung, and Blood Institute (NHLBI) Project Office regarding all aspects of the C-Data Project. The C-Data area of the website will provide features currently available on the CSCC site, with special emphasis on the following:

- A general information area where C-Data Project goals and objectives, SOPs, the Manual of Procedures, etc., are posted. Directories of contact information for all Center Data Coordinators, other Center personnel, and SDMC Collaborative Data Project staff.

- Patient Database subcommittee materials including in-person and conference call meeting schedules, agendas, and minutes.

- C-Data Project materials including training materials, protocols, forms, calendars of important dates, and consent and authorization materials.

The C-Data website area may be structured in many different ways. One possible format is as follows: After successful login and selection of the C-Data protocol, the main C-Data screen will appear and users will be presented with a list of options. The main screen options include sections for EDC Links (to enter the EDC data entry system), General Protocol Documents, Case Report Forms, Teleconference Minutes, and Training and Help Documents. Choosing to enter the EDC system will in turn offer options to enter any type of patient information (Medical History, Semi-Annual, Annual, etc.), or submit additions or corrections of previously entered data.

Conference calls for Center Data Coordinators and/or protocol team will be scheduled bi-monthly, to discuss C-Data Project progress and issues.

9 COLLABORATIVE DATA PROJECT SUBCOMMITTEE

The CSCC Patient Database Subcommittee will function in the same manner as a protocol study committee, working closely with the SDMC during the 5-year funding cycle. The subcommittee will develop procedures to guide investigator use of C-Data Project data. When indicated, subcommittee recommendations will be referred to the CSCC Steering Committee for approval.

10 PARTICIPATING CENTERS AND POPULATION FOR INCLUSION

All Centers and all clinical sites within each Center are expected to participate in the Collaborative Data Project. All adult and pediatric patients seen within the last 24 months in the hospital or clinical setting at a CSCC site, and expected to return episodically or regularly for care, are eligible for inclusion. Enrollment will require compliance with appropriate IRB and HIPAA requirements. Deceased patients and patients considered inactive will be excluded.

11 HUMAN SUBJECTS PROTECTION

Inclusion in the CSCC Collaborative Data Project or Patient Database in no way implies that a patient will be included in any other CSCC studies. Once a patient has been identified as part of a potential pool of study subjects for a CSCC study, he/she will be presented with the opportunity to participate in a manner consistent with compliance with Good Clinical Practice (GCP) guidelines for the conduct and monitoring of research, and the ethical and regulatory requirements presented in ICH E6, Good Clinical Practice: Consolidated Guideline.

11.1 HIPAA

Sites will be responsible for maintaining adherence to HIPAA regulations.

11.2 Informed Consent

Each patient must grant Informed Consent for his/her data to be used in the Collaborative Data Project. Each Center must determine whether (a) previously granted informed consent is sufficient or (b) a new Informed Consent Form must be signed.

11.3 IRB Review

IRB approval must be obtained individually by each Center (and each site within multi-site Centers). Since inclusion in the C-Data Project itself will not involve participation in any medical activities, it is possible that expedited IRB review and approval may be obtained.

The investigators and institutions affiliated with this study will permit monitoring, audits, IRB/IEC review, and regulatory inspection(s) by providing direct access to source documents.

Site investigators are responsible for annual renewal of IRB approval and all required communications with their institution's IRB.

11.4 Confidentiality

Subject confidentiality will be maintained by the investigator, the investigator's associates and co-workers, and by all administrators who are part of the CSCC project. Confidentiality will be maintained according to ICH E6; 4.8.10, part O: "Records identifying the subject will be kept confidential and, to the extent permitted by the applicable laws and/or regulations, will not be made publicly available. If the results of the trial are published, the subject's identity will remain confidential."

Additionally, the SDMC has obtained a Certificate of Confidentiality. These are issued by the NIH to protect the privacy of research subjects by protecting investigators and

institutions from being compelled to release information that could be used to identify subjects with a research project. Certificates of Confidentiality are issued to institutions or universities where the research is conducted. They allow the investigator and others who have access to research records to refuse to disclose identifying information in any civil, criminal, administrative, legislative, or other proceeding, whether at the federal, state, or local level.

Providers must feel comfortable sharing data and the clinical practices of their site or center. The Certificate of Confidentiality will be used to make sure there are no biases or unfavorable comparisons between centers. This will also serve as another layer of protection for subjects, especially with regard to potentially sensitive information such as alcohol, drug history, etc.

11.5 Strategies to Protect Privacy and Confidentiality

Since participation in the C-Data Project involves inclusion in a database only, there are no potential risks to subjects that may result from clinical treatment or intervention. The risks to an individual's privacy or confidentiality due to unauthorized or inappropriate sharing or release of medical data are no greater than that encountered when participating in any other CSCC-supported investigation. However, since participant data will be continually incorporated into the database over an extended period of time, it is especially important that the SDMC utilize several strategies to minimize this risk. As indicated previously, these strategies include:

- SDMC provision of centralized training to all CSCC Center staff prior to patient recruitment and enrollment, and at regular intervals during the remaining project period.
- Providing technical assistance related to all aspects of assuring patient privacy, confidentiality and protection.
- Providing a data management system fully validated to 21CFR, Part 11 requirements.

Additional steps include implementation of appropriate management tools such as the standard operating procedures (SOPs) already in use by the SDMC at Rho, Inc., and assuring all electronic resources utilized for the C-Data Project are protected with physical security measures and user access controls.

12 ACCESS TO THE PATIENT DATABASE

Access to the PD will be limited to participating CSCCs and designers of multi-center clinical trials. During the funding cycle, investigators will be limited to use of their own site data. Aggregate (no patient or site ID) data are available by request to the Protocol Development Committee, for CSCC investigators. Once this funding cycle is complete, requests for de-identified, anonymous data will be made to the NHLBI and a review committee.

12.1 Disclosure of Data

Data from the C-Data Project will be used and/or disclosed only as approved by NHLBI Project Office.

12.2 Publication of Research Findings

Proposed manuscripts or abstracts developed as a result of the C-Data Project efforts must be approved in accordance with the Guidelines and Standard Operating Procedures outlined by the CSCC Publications and Presentations Committee and included in the CSCC Manual of Procedures.

13 SUBJECT COMPENSATION

Individuals who provide Informed Consent and Authorization for the C-Data Project, and complete the Medical History Interview and the Health-Related Quality of Life and Patient Satisfaction questionnaires will receive a one-time $50.00 reimbursement to compensate for transportation and parking costs.

Individuals who are unable to complete the Health-Related Quality of Life and Patient Satisfaction questionnaires due to a language barrier or literacy issues will receive the $50.00 reimbursement.

14 PROTOCOL SIGNATURE PAGE

I, _____, MD agree to conduct:

The Comprehensive Sickle Cell Centers Collaborative Data Project.

I understand that no deviations from this protocol, Version 5.3, July 3, 2008 may be made without the written permission of the NHLBI CSCC Protocol Chair, except where necessary to eliminate immediate hazard(s) to study subjects, or when the change(s) involve only logistical or administrative aspects of the study.

Name: _____

Signature: _____

Site: _____

Date: _____

15 LIST OF INVESTIGATORS

CENTER NAME	INVESTIGATOR	ADDRESS AND TELEPHONE OF STUDY CENTER
Southwestern Comprehensive Sickle Cell Center	George Buchanan, MD	The University of Texas Southwestern Medical Center at Dallas Pediatrics Department 5323 Harry Hines Boulevard Dallas, TX 75390-9063 (214) 648-8594
Bronx Comprehensive Sickle Cell Center	Mary Fabry, PhD	Albert Einstein College of Medicine Division of Hematology Ullman Building, Room 915 1300 Morris Park Avenue Bronx, NY 10461 (718) 430-3753
Duke-UNC Comprehensive Sickle Cell Center	Marilyn Telen, MD	Division of Hematology Box 2615 Duke University Medical Center Durham, NC 27710 (919) 684-5378
Children's Hospital of Philadelphia Comprehensive Sickle Cell Center	Kwaku Ohene-Frempong, MD	34th Street & Civic Center Blvd. Philadelphia, PA 19104 (215) 590-3423
Northern California Comprehensive Sickle Cell Center	Elliott Vichinsky, MD	Children's Hospital & Research at Oakland Dept. of Hematology/Oncology 747-52nd Street Oakland, CA 94609 (510) 428-3651
University of Southern California Comprehensive Sickle Cell Center	Cage S. Johnson, MD	Department of Medicine RMR 304 2025 Zonal Avenue Los Angeles, CA 90033 (323) 442-1259
Boston Medical Center Comprehensive Sickle Cell Center	Martin Steinberg, MD	One Boston Medical Center Place, FGH-2 Boston, MA 02118 (617) 414-1020
St. Jude Children's Research Hospital Comprehensive Sickle Cell Center	Winfred Wang, MD	St. Jude Children's Research Hospital Department of Hematology/Oncology 332 North Lauderdale Bldg R-6010 Mail Stop Code 763 Memphis, TN 38105 (901) 495-3497
Cincinnati Comprehensive Sickle Cell Center	Clinton H. Joiner, MD, PhD	Children's Hospital Medical Center Division of Hematology/Oncology 3333 Burnet Avenue Cincinnati, OH 45229-3039 (513) 636-4541
Marian Anderson Sickle Cell Anemia Care and Research Center	Marie J. Stuart, MD	Thomas Jefferson University Hematology Division, Pediatrics Dept. College Bldg., Suite 727 1025 Walnut Street Philadelphia, PA 19107 (215) 955-9820

16 REFERENCES

Arts DG, De Keizer N, Scheffer GJ. Defining and Improving Data Quality of Medical Registries: A Literature Review, Case Study, and Generic Framework. *J Am Med Inform Assoc* 2002; 9:600-11.

Mitri W, Sandridge AL, Subhani S, Greer W. Design and development of an Internet registry for congenital heart defects. *Teratology* 2002; 65:78-87.

Schulman J, Edmonds LD, McClearn AB, Jensvold N, Shaw GM. Surveillance for and comparison of birth defect prevalences in two geographic areas--United States, 1983-1988. *MMWR Morbid Mortal Wkly Rep*, Centers for Disease Control 1993; 42:1-7.

Solomon DJ, Henry RC, Hogan JG, Van Amburg GH, Taylor J. Evaluation and implementation of public health registries. *Public Health Rep* 1991; 106:142-50.

Sørensen HT Johnsen SP, Nørgard B. Methodological issues in using prescription and other databases in pharmacoepidemiology. *Norwegian Journal of Epidemiology* 2001; 11:13-8.

Wilson PD, Correa Villasenor A, Loffredo CA, Ferencz C. Temporal trends in prevalence of cardiovascular malformations in Maryland and the District of Columbia, 1981-1988. The Baltimore-Washington Infant Study Group *Epidemiol* 1993; 4:259-65.