Rho, inc.

6330 Quadrangle Dr., Ste. 500 | Chapel Hill, NC | 27517

(919) 408-8000 FAX (919) 408-0999

Clinical Data Management Plan

Treatment of Pulmonary Hypertension and Sickle Cell Disease with Sildenafil Therapy (walk-PHaSST)

Version Number and Date:	Version 0.1 28FEB2008
Protocol Number:	V 5.0
Sponsor:	National Heart Lung and Blood Institute
Rho, Inc. Data Management Contact(s):	Vickie Coble

CLINICAL DATA MANAGEMENT PLAN

Treatment of Pulmonary Hypertension and Sickle Cell Disease with Sildenafil Therapy (walk-PHaSST)

Protocol Number V. 5.0

BB-IND 77461

INITIAL APPROVAL SIGNATURES

Nancy Yovetich, Principal Investigator — Date: 22 Oct 2008

Vickie Coble, Project Data Manager, Rho, INC. — Date: 21 Oct. 2008

CLINICAL DATA MANAGEMENT PLAN

Treatment of Pulmonary Hypertension and Sickle Cell Disease with Sildenafil Therapy (walk-PHaSST)

Protocol Number V. 5.0

BB-IND 77461

FINAL APPROVAL SIGNATURES

Nancy Yovetich, Principal Investigator	Date
Vickie Coble, Project Data Manager, Rho, INC.	Date

REVISION HISTORY

Version 0.1 (Draft) (28Feb2008)

Section	Change	Requestor	Reason for Change
All	N/A	N/A	Initial Draft

Version X.X (DDMMMYYYY)

Section	Change	Requestor	Reason for Change

TABLE OF CONTENTS

LIST OF APPENDICES

1 Introduction

This document describes the policies and procedures Rho®, Inc. (Rho) will employ to manage clinical trial data collected during the National Heart Lung and Blood Institute (NHLBI) sponsored walk-PHaSST trial. This study is a Randomized, Double-Blind, Placebo-Controlled, Phase II/III, Multi-Center Study for Treatment of Pulmonary Hypertension and Sickle Cell Disease with Sildenafil Therapy, and will measure exercise capacity by the distance traveled by patients during a six minute walk.

This is a dynamic document that may change over the course of the study to incorporate revisions to policies and procedures that reflect emerging study needs. Users of this document should note the version number and date printed on the title page and at the bottom of each subsequent page. Initial approval signatures (form included) should be obtained early in the data collection phase. Final approval signatures (form included) should be obtained prior to database lock.

The data management process used in this study will be governed by the standard operating procedures (SOPs) of Rho or will be appropriately documented as deviations to SOP as outlined in Rho's SOPs. Rho's SOPs will be available for inspection by the sponsor throughout the data management period of this study and while data from this clinical study are under review by regulatory authorities.

2 The Clinical Database

The Rho Electronic Data Capture System (RhoEDC) is validated software developed at Rho for data collection and query management for EDC trials. The current RhoEDC version will be used for electronic data collection and query management for this study. RhoEDC has been validated to 21 CFR 11 standards. The validation documentation is available for review at Rho's headquarters in Chapel Hill, NC. A study-specific database will be developed within RhoEDC. All electronic data entry will take place on an internet browser connected to Rho via a 128-bit encrypted internet connection. Rho's secure-access web and database servers run from a restricted access co-location facility with local mirroring/backup services. The entered and submitted data will be stored in an Oracle® version 9ir2 relational database. Any data management needs that cannot be satisfied within the automated RhoEDC will use SAS® version 9.1. Rho will produce SAS® datasets for reporting and analysis of the data for this trial.

2.1 Database Structure

Rho will create the following project components:

- A computer directory structure: The computer directory structure adheres to the standard naming conventions at Rho and will maintain each study's computer files separately from those of other studies. Study files will be logically organized, with separate and well-defined locations for programs, data, output, log files, queries, documentation, and results of analyses. User network privileges will be used to control access to project directories. Access to project directories will be granted as appropriate, only to Rho staff members who are actively engaged in project work and to Rho staff members who have managerial or support responsibilities for this study.

- Project data entry screens: Rho will create an electronic Case Report Form (e-CRF) to collect clinical data required for the protocol. The screens will include data entry input fields, descriptive field labels and access to online completion assistance.

- Annotated e-CRF: A copy of the e-CRF will be annotated with dataset names, field names, and values to be entered for multiple-choice and check box fields based on the specifications of RhoEDC.

- Project data dictionary: A project data dictionary that describes the data to be processed will be created. The data dictionary will include field names, descriptive field labels, and field attributes (such as length, type, position, ranges, and limits) for univariate error detection.

3 Validation Plan

This section describes the computerized portion of the edit check validation plan employed by RhoEDC.

3.1 Univariate Alerts

RhoEDC allows valid-value, valid-range, and missing-value alerts to be provided and specified independently for each electronically captured field. These three kinds of alerts are referred to collectively as univariate alerts.

During electronic data collection, all values will be entered per the source document. During entry, RhoEDC checks each value in each field to confirm that it passes the univariate check(s) for that field. The program alerts the personnel entering data of any potential discrepancies. (See attachment C for a complete list of univariates checks)

3.1.1 Valid Value Alerts for Multiple-choice Fields

Data entry screens generally include a variety of fields for which a finite number of responses are valid and all other values are disallowed. Standard practice is to perform a valid-value check for every multiple-choice field included in the e-CRF. Examples include the following:

- Yes/No fields
- Normal/Abnormal fields
- Check box fields, in which the user is presented with a set of boxes with instructions to check one of them or all that apply
- Coded fields, in which the user records a coded value taken from a list that is usually presented on the page or in an instruction guide.

3.1.2 Valid Range Alerts for Date/Time Fields

Standard practice is to perform a valid-range check for every element of every date and time field included in the e-CRF. Valid ranges for year, month, day, hour and minute are standard logic (e.g., month must be between 1 and 12, minute must be between 0 and 59) but, where appropriate, are further specified in accordance with the study protocol and timetable.

3.1.3 Valid Range Alerts for Numeric Variables

Standard practice is to perform a valid range check for every numeric field included in the e-CRF for which such a range can be defined. Valid ranges for numeric fields will be chosen with the following two goals in mind: 1) to generate queries for values that are impossible or unlikely and 2) to avoid generating queries for normal values. Except in cases where the Principal Investigator or designee provides valid ranges, Rho will use past experience or published standard references to determine valid ranges for numeric fields.

3.1.4 Missing Value Alerts

Standard practice is to perform missing value checks on all fields except those that are pre-planned to be missing on some occasions.

Univariate alerts will be custom-designed to meet the needs of this trial. The data dictionary document defining the valid value, valid range and missing value alerts will be forwarded to the Principal Investigator or designee for final approval. After this final approval is obtained, Rho will create univariate checks that will generate alerts for all instances in which the identified conditions are violated.

3.2 Multivariate Alerts

In addition to the univariate alerts described in Section 3.1, RhoEDC allows for the specification of relational alerts among groups of variables. Such alerts are referred to as multivariate alerts.

Common types of multivariate alerts include the following:

- Confirming that diastolic blood pressure reading is less than the associated systolic blood pressure reading.
- Confirming that "Other, specify" is completed when "Other" is marked.

Multivariate alerts will be custom-designed to meet the needs of this trial. A document listing the multivariate alerts and cross-module alerts will be forwarded to the Principal Investigator or designee for final approval. After this final approval is obtained, Rho will create multivariate checks that will generate alerts for all instances in which the identified contingencies are violated. The project statistician and/or the Principal Investigator will review the multivariate checks to confirm that the important endpoints are being subjected to sufficient validation. (See attachment C for a complete list of multivariate checks).

4 Data Entry and Data Cleaning

4.1 Data Entry

Trained site personnel will be responsible for entering most clinical data in the clinical database via RhoEDC, as outlined in the study Manual of Operations.

4.2 Completion Guidelines

- Study-specific, form-specific completion guidelines (instructions) will be prepared by the lead data manager or designee. These instructions will be provided electronically as an online document, accessible during the EDC process. Instructions will be evaluated and updated on an as needed basis.

- General "help" documentation will be provided to assist the user to navigate RhoEDC. This document will be provided electronically as an online document, accessible during the EDC process.

- Special data fields will be set up in RhoEDC as needed to identify reasons for missing values, such as "unknown", "not done", or "not applicable".

4.3 Data Security

- Each RhoEDC user will be assigned access to the study website by the Rho website administrator. Each user must successfully complete a training module administered by Rho, prior to access.

- If a user does not require RhoEDC data entry access, the data entry link will not appear on the user's web page.

- Users will be able to view/access data only from sites that they are authorized to access.

- Data are submitted to Rho's secure web server using Secure Sockets Layer (128 byte public key encryption methodology) and are stored in the study's "operational database" each time a form is created or updated.

- No data are ever stored on the user's personal computer or local network.

- The database is backed up nightly, and backup tapes are saved in a secure, off-site location.

4.4 Quality Control Procedures

- Subject ID numbers are assigned centrally through the ID registration component of RhoRAND. Each ID number is created with a "check digit," used to detect transcription errors. Each ID entered in RhoEDC is checked to ensure it is a valid ID number per RhoRAND.

- The Subject ID number is displayed on each page of the e-CRF, as an added assurance that the user always knows which subject's data are displayed.

- Status symbols appear next to each form name on an individual subject's menu, to indicate the validation status of each form.

  ! (Red) Indicates the form was submitted with one or more fields that failed a validation test and the validation failure remains unresolved.

  ✓ (Green) Indicates the form was submitted, and there are no fields that have either failed validation tests or have unresolved validation failures.

- RhoEDC includes a complete audit trail of all transactions, with an electronic time/date stamp, and capture of user login of the person creating or updating each form.

4.5 Status Flags

RhoEDC will create an associated status flag for every variable defined in the project data dictionary. The status flags will allow separate status information to be maintained for every data value entered into the project database. Status flags may be used in three ways:

- As a data management tool, to assist in tracking the progress of electronic data entry and query resolution.

- As an analysis tool, to assist in creating appropriate analysis files from incompletely-processed datasets (e.g., for interim analyses). These status flags will allow statisticians to obtain up-to-the-minute information about which data values are "clean" and which should be considered "questionable."

- As permanent project documentation, to record information about how each data value was validated.

Status Flag values are detailed below:

Status Flag Values

Status Flag	Interpretation of Status Flag
3	The data value has failed a univariate (valid value or valid range) test. This status is also used for data values that are missing (or blank) but are not permitted to be missing (or blank).
5	The data value has passed a univariate test(s) but has failed a multivariate test.
6	The data value has passed a univariate test(s) and has not failed a multivariate test.
8	The site or sponsor has documented this data value as being correct, and the value is equal to the value on the source document or file.

4.6 Query Generation

4.6.1 EDC-generated Queries

Univariate validity checks are defined in the data dictionaries for the study. Multivariate edit checks are programmed as part of the set-up of the project in RhoEDC. Both types of checks generate automatic queries during electronic data entry.

When RhoEDC detects a possible error during electronic data entry, an alert will appear as the user exits the involved field and/or when the form is submitted. The field(s) where the error(s) occur is highlighted in pink as part of this alert. The alert gives the user three options: 1) to correct the entry so it no longer fails the validity check; 2) to continue without correcting the entry so the failed validity check remains unresolved; or 3) to "override" the validity check that triggered the alert.

If the user overrides the validity check, he/she will be required to supply a reason for the override. For example, a univariate check for weight may have been set up to expect a range of 22 to 120 kilograms. An entry of 130 kilograms would trigger an alert. If, upon examination of the source documentation, the user determines 130 is a correct entry, the user would need to override the validity check and provide a reason for the override. The reason for the override may be "Subject was weighed on several occasions and weight of 130 is correct." If an automatic query is overridden, the reason is stored in the audit trail, along with the user ID of the person supplying the reason.

If a flagged entry is not either corrected or overridden, that field is marked within the system as having failed a validity test and will be tracked as an outstanding query.

4.6.2 Manual Queries

Manual queries (also known as Data Clarification Memos, DCMs) will be used to request clarification about data entered on an e-CRF when RhoEDC is not suited to automatically generate the necessary query. These situations might include pages that have not been completed or completed incorrectly and data inappropriately recorded in text fields. Some manual queries may be issued to resolve data inconsistencies between forms or for multivariate checks as described in section 3.2.

DCMs may be written at any time during the data collection process. All DCMs will be followed and documented. The DCM will request that the site make necessary changes in RhoEDC to resolve the data issue or to explain why the data issue does not require a data change. DCMs will be returned to Rho along with an explanation of how the data issue was resolved. Rho will confirm that the data issue is resolved.

5 Loading Electronic Files

BNP 5.1 data transfer will be received from the NHLBI central lab every 6 months beginning 6 months after first patient enrollment. Laboratory results will be transferred from the site or central laboratory to Rho via electronic data transfer. The data forwarded from the site or central laboratories will be cumulative.

When Rho receives a database electronically from the central laboratory via email from a secure server, the Rho Study Coordinator or designee scans the database for known viruses. If any viruses are found, the Rho Data Manager or designee will not load the database into the network and will contact the sender so that the sender may resend the database without the virus.

The Rho Data Manager or designee uploads the database into the electronic project files, sets the database to read only, and notifies the Study Statistician that the files have been uploaded. The Rho Data Manager logs in the receipt of the database in the study files.

The Statistical Programming Team Lead or designee is responsible for verifying that the database conforms to the expected structure. If laboratory data are not delivered as SAS datasets, statistical programming personnel convert the data to SAS datasets.

The content of the laboratory database is compared to specifications provided by the laboratory. Reports on any discrepancies are forwarded to the Lead Data Manager or designee. These discrepancies are also brought to the attention of the laboratory for possible resolution.

Central lab ECHO 5.2 data transfer will be received at Rho the first Monday of every month. The ECG Process is outlined in Appendix D.

6 Laboratory Normals Data

Not applicable for this study.

7 Medical Coding

7.1 The Rho® Coder Suite

E-CRFs capture data in a structured but flexible format in an attempt to collect a wide range of observations and data. While asking specific questions and allowing minimally structured responses ensures the widest possible collection of data, the method creates special challenges for data analysis, because Investigators may use a wide variety of descriptive terms to refer to the same kind of event. For example, on the Adverse Events form, Investigators may report headaches by using terms like "head hurts," "migraine," "forehead pain," "headache," etc. In order to determine whether the study drug causes headaches, it is necessary for statisticians to be able to tell that all of these disparate descriptions refer to the same thing, i.e., "headache."

In order to facilitate comprehensive data analysis of these disparate terms, the values in fields such as AEs and concomitant medications are often assigned standard codes. These codes make it possible to perform the necessary statistical analyses. The Rho® Coder Suite provides a tool to assign standard codes to the verbatim terms entered in RhoEDC based on a set of pre-determined, study-specific rules. The Rho® Coder Suite accomplishes this in two different ways.

An Auto-coding program attempts to map each verbatim term in a specified field to a standard code. The Auto-coder codes terms based on existing dictionary definitions. The Auto-coder performs the initial step in the coding process, coding any terms that exactly match entries in the dictionary specified for the field being coded. When a term cannot be Auto-coded, the Auto-coder indicates that the record was "not coded" and requires a "coding rule." A coding rule is a specification that maps a verbatim term to a standard dictionary term.

After the Auto-coder runs, a human coder reviews the results of the Auto-coder and makes any coding decisions that the Auto-coder could not. The Rho® Coder Suite provides the coder with access to the same dictionary entries available to the Auto-coder, as well as access to e-CRF data. During a coding session, a coder creates coding rules for terms that did not Auto-code. The coder can also view rules created during previous coding sessions. By creating and manipulating coding rules, the coder can code records, uncode previously coded records, and split and unsplit terms from the clinical data.

Once any coding for a project has been completed, the Rho® Coder Suite creates a set of output datasets. These datasets combine all of the original data put into the Rho® Coder Suite with coding data. The output datasets can then be used for complex statistical analysis of the data that would have been impossible in their original state. Output datasets also include information about how each coding decision was reached. In addition to the output datasets, the Rho® Coder Suite generates a variety of reports throughout the coding process for internal and client review.

7.2 Adverse Events

Rho will code Adverse Events using the Medical Dictionary for Regulatory Activities (MedDRA™) Coding system, version v10.1. Coded transfers will occur at the end of each month, beginning of each month, and as requested for the DSMB.

Each adverse event will be assigned both a symptom coding symbol (term) and a MedDRA™ body system value. This MedDRA™ symbol assignment, also referred to as "mapping," will be performed by a medical coding specialist under the direction of the lead data manager, using computer software to assist in locating matches between e-CRF terms and coding symbols. Spelling errors may be corrected to facilitate mapping.

Each term listed on the e-CRF will be mapped separately. To facilitate appropriate mapping, a single e-CRF entry describing multiple adverse events will be duplicated, allowing each event to be coded separately.

Information from the comments and from other e-CRF pages for the same subject may be used to clarify the appropriate symptom code and body system. When this information is inadequate, sites may be asked to provide additional information to clarify the appropriate mapping.

Following mapping, a listing will be generated that includes the text from the e-CRF, any synonyms used to facilitate mapping, the preferred MedDRA™ term assigned, and the higher level terms. This listing will distinguish terms that match directly from those that match by using a synonym and those that did not find a match in the dictionary. A medical coding specialist will review this listing. The approved listing or data will be submitted for review by NHLBI medical personnel, blind to treatment assignment. When MedDRA™ coding is determined to be satisfactory, a Medical Coding Approval Form will be signed by the NHLBI personnel assigned to the study and stored as part of the official study documentation.

At times, multiple symptoms of a single adverse event are initially reported as separate Adverse Events, usually because the single diagnosis that covers the AE and its symptoms is not yet known. Over time, as a single diagnosis becomes clear, the AE record should be updated by the site to reflect it, removing the extra records.

7.3 Medications

Rho will code medications using WHODRUG (2007.03). Each medication will be assigned a drug classification value. This WHODRUG code assignment, also referred to as "mapping," will be performed by a medical coding specialist under the direction of the lead data manager, using computer software to assist in locating matches between e-CRF terms and coding symbols.

Each term listed on the e-CRF will be mapped separately. To facilitate appropriate mapping, a single e-CRF entry describing multiple medications will be duplicated, allowing each medication to be coded separately.

Information from the comments and from other e-CRF pages for the same subject may be used to clarify the appropriate drug classification. When this information is inadequate, a query will be issued to the site for additional information to clarify the appropriate mapping.

Following mapping, a listing will be generated that includes the text from the e-CRF, any synonyms used to facilitate mapping, and the preferred WHODRUG code assigned. This listing will distinguish between terms that match exactly from those terms that match by using a synonym and those that did not find a match in the dictionary. A medical coding specialist will review this listing. The approved listing or data will be submitted for review by NHLBI medical personnel, blind to treatment assignment. When WHODRUG coding is determined to be satisfactory, a Medical Coding Approval Form will be signed by the medical personnel assigned to the study and stored as part of the official study documentation.

8 Database Conversion and Transfer

Not applicable for this study.

9 SAE Reconciliation

A reconciliation of key data points for Serious Adverse Events (SAEs) will be conducted to ensure consistency in the number and medical concepts of SAEs contained in the safety database versus the clinical trial database. The following fields in the safety database and the clinical trial database will be compared:

- Subject ID
- Verbatim text
- Causality
- Onset date (within a 2-day window)
- Stop Date or its equivalent

An electronic comparison of the fields identified above will be conducted on a quarterly basis. The results of these comparisons will be reviewed, and any discrepancies will be resolved. Listings of the reconcilable fields in the safety database and all SAEs in the clinical trial database will be produced to assist in resolving discrepancies. Verbatim text must convey the same medical concept but are not required to electronically compare. All discrepancies will be brought to the attention of Rho safety personnel before any queries are issued. If efforts to explain discrepancies fail, queries will be issued to the site for the following:

- A medical concept exists in the clinical trial database but not in the safety database for an onset date within the acceptable time window
- Discrepant subject IDs
- Discrepant causality or relatedness
- Discrepant onset dates
- Discrepant outcome dates
- An outcome date present in the clinical trial database, but the SAE is considered to be ongoing in the safety database

10 Database Closure

10.1 Closure Checks

The data validation, data clarification and database modification activities described in the preceding sections will be repeated until the Rho Lead Data Manager and the sponsor representative consider the data clean. When clean data are achieved, the Rho Lead Data Manager will initiate steps to lock the clinical database. Prior to locking the database, Rho will perform a set of closure procedures to verify the integrity and completion of the database.

Computer programs will be used to do the following:

- Check that all queries have been resolved;
- Check for consistency of key variables;
- Check that the project database is consistent with the specifications in the project data dictionary;
- Check for duplicate observations;
- Ensure that all expected observations are present for each subject;
- Determine the status of each subject entered (i.e., excluded, ongoing, completed, withdrawn, lost to follow-up, and so on);
- Check status flags to ensure that all data points have been completely processed;
- Check for value formatting problems;
- Check for consistency between whole-date fields and associated part-date fields.

Manual procedures will be used to do the following:

- Check the output from all computerized closure checks to confirm that no problems remain.

10.2 Quality Assurance Audit and Database Lock

Data entered into RhoEDC will be verified against the source during the Site Monitoring Visit(s) as specified in the Clinical Monitoring Plan.

The Rho Lead Data Manager will declare the database locked and ready for analyses when monitoring is complete and closure procedures have been conducted.

11 Data Archiving

All study-specific paper documents deemed to be essential are maintained in archive status at Rho, according to current SOPs. Final versions of all electronic data are converted to a transportable format to accommodate future software changes. These include all management files used in set-up, as well as the data collected for the study. The transportable data files and electronic copies of the programs and specialized software that were used to create, edit and analyze the data are maintained in archive status at Rho, according to current SOPs.

Appendix D
Data Transfer Process
For Echocardiogram Central Review

Site to Vandana:

1) For the first 20 patients, sites will send the local echocardiogram immediately upon completion via RhoLAB (assigning a specimen number to the CD)
2) After the first 20 patients have been reviewed by Vandana, sites will send weekly batches to the central lab via RhoLAB
3) Rho will need to make sure that the local echocardiogram and the physical exam CRFs are completed in the EDC system (these forms will be needed to complete some of the calculations in Vandana's review)

Vandana to Rho:

1) Every first Monday of the month, Vandana will send an excel spreadsheet with the raw, central lab data
   - Email distribution should include: Rob Woolson, Janet Willis, Vickie Coble, Wendy McBane, Jamie Spencer, Christopher Woods
2) Once received at Rho, data management will convert the excel spreadsheet to SAS
3) Janet Willis will email Rob Woolson the path of the dataset
4) Rob Woolson will derive variables calculated in an analysis dataset
5) Specifications for the analysis dataset will be created by Biostatistics (based on algorithms sent by Vandana and in the MOO chapter for echocardiograms), programmed by Stats, and validated by Biostatistics
   - The data set will be put into a pdf that can be emailed or posted on the study website

Rho to Vandana:

1) Christopher Woods will post a data listing in pdf format of the echo lab data for her to review and compare to her central lab readings
2) The pdf'd data listing will provide a cumulative list from the local echocardiogram CRF so Vandana can review her central review with the local review
3) If required, Vandana will follow-up with the sites for additional feedback on the echocardiograms sent to her during the study