VIRAHEP-C Data Archive

The Virahep-C data archive contains the study protocol and related descriptive documentation; study forms; and data collected by the study.

The data files are organized into the following four directories:

1. **Documentation**
2. **Forms**
3. **Data**
4. **Dataset Integrity Check**

1. Documentation Directory

The Documentation directory contains documentation of the DPP protocol and descriptive documentation, including:

- **Virahep-C Summary**: a summary of the study
- **Manuals_of_Operations**: a subdirectory containing 29 files
  - 27 files contain the manual for data collection forms
  - two files contain the additional manual for MATI interview
- **Protocol**: a subdirectory containing the following three files:
  - **ViraheC_Protocol**: the study protocol
  - **Data collection Protocol**: an overview of data and sample collection for study forms by various study phases and time points
  - **Protocol_legend**: a legend for the "Data collection Protocol" document

2. Forms Directory

The Forms directory contains the 27 forms used for data collection and a history of updates to the forms (Form_updates_history.pdf). The forms are described in Tables 1-4.

Table 1: Screening Forms

Form		Keys	Description
SE	Screening Evaluation	ID, RSC	Initial screening form including demographics, baseline physical exam, baseline medications, symptoms, serologies, and labs
SC	Screening Criteria	ID, RSC	Checklist containing all inclusion/exclusion criteria to determine study eligibility.
PE	Pathology Evaluation	ID	Pathology data as completed by central pathologist. Section II is all initial scoring. Section III is additional scoring appended to the database in December 2004.
RA	Risk Assessment	ID	Obtains information on known risk factors associated with Hepatitis C, PI determination of infection source, and date of infection.

Table 2: Clinic Assessments Forms

Form		Keys	Description
TE	Treatment Evaluation	ID, TMPT	Patient evaluation during the treatment period including physical exam, patient reported symptoms, concomitant medications, depression, and most recent dose of study medications
FE	Follow-up Evaluation	ID, TMPT	Patient evaluation during the follow-up period including physical exam, patient reported symptoms, concomitant medications, depression, and most recent dose of study medication (only provided at follow-up week 4 visit).
UE	Unscheduled Visit	ID, DOEDATE	Patient evaluation if patient was seen at anytime during the study period outside of protocol timepoints including physical exam, patient reported symptoms, concomitant medications, and depression.
MA	MATI-HepC	ID	Checklist of specified educational topics to be covered by the coordinator with the patient at the baseline visit. Captures date and time of first doses of peg-interferon and ribavirin and prescribed dose of study medications.
MB	Brief MATI	ID, TMPT	Checklist of specified educational topics to be covered by the coordinator with the patient at specified timepoints during the treatment period.
LE	Laboratory Evaluation	ID, TMPT, DSMDATE	Captures all laboratory tests completed throughout the study.

Table 3: Patient Questionnaires

Form		Keys	Description
SQ	Screening Questionnaire	ID	Baseline patient information including demographics, education and work status, alcohol, and smoking history.
QL	Quality of Life	ID, TMPT	Questionnaire based on the SF-36 to measure the patient's quality of life through questions about the patient's physical activities and emotional health.
SS	Social Support and Self-efficacy	ID, TMPT	Designed to measure the amount of social support the patient receives and the patient's level of confidence in performing certain activities.
CD	CES-D	ID, TMPT	CES-D questionnaire which measures the level of depression at specified timepoints during the study.
SA	Symptom Assessment	ID, TMPT	Captures specific symptoms on a continuous VAS line (10 cm). Measurement is the distance, in centimeters, from the left axis to the point where the line was marked by the participant.
XM	Sexual Function: Male	ID, TMPT	Measures sexual functioning in males at specified timepoints throughout the study.
XF	Sexual Function: Female	ID, TMPT	Measures sexual functioning in females at specified timepoints throughout the study.

Form		Keys	Description
AD	Adherence Questionnaire	ID, TMPT	Questionnaire designed to monitor patient's adherence to study medications during the treatment period, e.g. study medication taking behavior, missed doses, and reasons for missed doses.
MM	MEMS Questionnaire	ID, TMPT	Questionnaire designed to monitor the patient's adherence to the study medication schedule through use of the MEMS cap.

Table 4: Other Study Forms

Form		Keys	Description
TW	Trackcap Worksheet	ID	Records all refill and non-event openings for distributed MEMS caps for each patient.
AE	Adverse Events	ID, PNUM, LNUM	Obtains information regarding adverse events that occurred during the study period.
DC	Dose Change	ID, PNUM, LNUM	Tracks physician prescribed changes to study medication dosages throughout the treatment period.
DF	Discontinuation Form	ID, DDCDATE	Documents patient discontinuation from treatment and/or study from the time of enrollment (start of treatment).
GF	Growth Factor	ID, PNUM, LNUM	Documents any use of growth factor use during the study period (off-protocol event).

3. Data Directory

The Data directory contains the first public release of the VIRAHEP-C dataset. It consists of the following files:

- **"data description":** subdirectory containing 4 data description files
  - **Dataset_description_JUNE2007.pdf** : details regarding each dataset
  - **Nam-Rowers Occupational Status Scores Full.pdf** : NAM powers scores used in Screening Questionnaire
  - **PROC CONTENTS JUNE2007.pdf**: proc contents for the datasets
  - **virahepc_variable_description.pdf**: extra information regarding variables
- **VHEPFMT.SAS**: a SAS format file for categorical variables
- **Form_data:** subfolder containing 24 SAS data files collected using the corresponding forms from the main study.
- **Analysis_data**: subfolder containing 10 SAS data files that are not associated with any form. The files are described in Table 5.

Table 5: Analysis Datasets

Dataset	Keys	Description
INELIG	ID	A record for each protocol exemption received.
SAE	ID, ONSDATE	Summary of Serious Adverse Events submitted to the Virahep-C Safety Officer. One record per event containing expedited review determination criteria as reported by the Safety Officer and the Clinical Center PI. Can be linked with the adverse event record (AE dataset) by onset date.
LABS	ID, TMPT, DSMDATE	One record per patient timepoint with lab results along with upper and lower normal ranges, specific to clinical center and date of test.
MEMSINT	ID	One record per PEG-interferon event (cap opening). Nonevents (cap openings for refills or accidental openings) recorded on the TW are removed. If a patient ever received a replacement cap, data was incorporated into the MEMSINT dataset under the original cap serial number.
MEMSRIB	ID	One record per ribavirin event (cap opening). Non-events (cap openings for refills or accidental openings) recorded on the TW are removed. If a patient ever received a replacement cap, data was incorporated into the MEMSRIB dataset under the original cap serial number.
GTYPE	ID, DSMDATE	Genotype results for screened patients. One record per patient. Genotype results are by VERSANT™ HCV Genotype Assay (Bayer, Tarrytown, NY) completed by BBI/Seracare. Dr. John Tavis (virology ancillary) provided results if the genotype required sequencing.
VLOAD	ID, DSMDATE	Quantitative and Qualitative test results.* One record per patient timepoint. Includes dilution results for quantitative testing and 'intent to treat' variable to accommodate one error with the originally reported Screen 1 viral level.
VQUAL	ID, DSMDATE	Qualitative test results by patient timepoint for every qualitative test performed by BBI/Seracare. Each record contains up to two qualitative results and the associated internal control results. Results with negative internal control results are included, thus explaining some of the retests. Qualitative tests are currently run in duplicate but testing procedures have changed throughout the study. There may be more than one record per patient timepoint.
IFN	ID, TMPT	Serum Interferon levels (pg/ml)** for each patient over the treatment period: day 0, 1, 2, 3, 7, 14, 28, and treatment week 8, 12, 24, and 48. Serum interferon levels were also measured at Follow-up week 12. Day 3 was not collected after 7/7/2003 due to protocol change.
OAS	ID, TMPT	The serum concentrations of 2'-5' OAS** are measured (pmol/dL) for each patient during the treatment period: day 0, 1, 2, 3, 7, 14, 28 and treatment week 8, 12, 24 and 48. Comment flags are BLQ = Below the Limit of Quantitation AAR = Above the accepted Range; S = Single value. Day 3 was not collected after 7/7/2003 due to protocol change.

Notes:
* Quantitative test: COBAS Amplicor Hepatitis C Virus Monitor Test, version 2.0 TM assay, (sensitivity 600 IU/ml: Roche Molecular Diagnostics, Alameda, CA)
Qualitative test: AmplicorTM assay (sensitivity 50 IU/ml. Roche)
** OAS and IFN assays were performed by MDS Pharma Services

4. Data Set Integrity Check (DSIC)

The Virahep-C Data Archive also contains a report of an examination of the repository for completeness by statisticians and quality control specialists at the Repository. The published data from the Virahep-C study were compared to values recalculated from the Virahep-C data in the NIDDK repository to check for dataset completeness.