Persons using assistive technology may not be able to fully access information in this file. For assistance, e-mail niddk-cr@imsweb.com. Include the Web site and filename in your message.

HALT-C Data Management Overview

I. Introduction

The purpose of the HALT-C Trial Data Management System (DMS) is to assist clinical sites in entering trial data, organizing trial visits, monitoring trial protocols, and tracking shipped specimens. Personnel at the clinical sites complete paper data forms. Data from the forms are entered into the HALT-C Trial DMS at each clinical site via the World Wide Web. Each response is validated as it is entered into the DMS, and edit reports provide immediate feedback.

The DMS is menu driven and easy to use. There are a large number of features in this system, however, and site Data Management (DM) personnel will need to devote some time to becoming familiar with them.

A. Data Management System

The HALT-C Trial uses a Web-based data management system, known as ADEPT (**A**dvanced **D**ata **E**ntry and **P**rotocol **T**racking). The data that are entered at each clinical site are stored in the trial database, which is located at the Data Coordinating Center (DCC) at New England Research Institutes (NERI).

One of the key design features of the HALT-C Trial DMS is its focus on protocol management at the clinical sites. It is expressly designed to facilitate clinical site task completion. The system provides support for study visit scheduling, data entry, edit and form tracking, specimen shipment preparation, and study status and protocol monitoring. The Web-based system allows for immediate feedback to the clinical sites, which greatly improves the quality of multi-center data.

The structure and content of the DMS are stored in an Oracle database. The DMS is menu driven and provides the following features:

- Complete overview of each patient's HALT-C Trial status, including participation in any Ancillary Studies;
- Protocol tracking for each patient;
- Automated patient randomization into treatment and control groups, as well as tracking week 20 responders and patients who are unable or unwilling to be randomized;
- Data entry with on-line editing and context-based help;
- Form status information in the trial database permitting real time tracking;
- Edit, form, patient, clinical, and site status reports designed to facilitate patient and clinical site management and monitoring;
- Visit Control Sheets that provide an individualized blueprint for patient study visits;
- Shipping database that tracks collection, processing, shipping, and status of all specimens shipped to the central Repository and to central labs for ancillary studies;
- A patient appointment scheduling and reporting system;
- Internet connection between each site and the DCC, allowing for remote data entry, and an email connection for reporting problems;
- Data entry certification and double data entry of specified forms for quality control;
- Automatic audit trail capable of tracking changes to data, monitoring system usage and user logons;
- Password protection of the DMS as a whole and individual security levels within the DMS; and
- Real-time accrual reports.

B. Clinical Center Data Management Responsibilities

Clinical center data management personnel are critical to ensuring a smooth operation and high quality of data in the HALT-C Trial. At some sites there will be a full time data manager; other sites will have a number of people sharing data management tasks. Clinical center data management personnel, together with site Study Coordinators and Principal Investigator, develop a list of responsibilities that cover the needs of each individual site, as well as the more general needs of the HALT-C Trial. The expectations of data management staff are explained below. This list applies whether there is a single Data Manager or several people sharing data management responsibilities at a site.

Site data management responsibilities include:

- Preparing visit packets for patient study visits, including Visit Control Sheets and data forms labeled with HALT-C patient ID labels;
- Tracking the status of all forms related to the trial;
- Reviewing forms prior to data entry for adherence to study protocol;
- Data entering forms (this task may be shared with a data entry clerk);
- Running edit reports, following-up with the data collector, and entering corrected data;
- Maintaining patient records to ensure that:
  - Paper copies of data forms are properly filed after entry,
  - All documents are accounted for,
  - Data forms are accessible for timely quality control (QC) activities, and
  - Data forms are stored and handled in a manner that maintains patient confidentiality.
- Preparing and distributing DMS reports as needed;
- Following up with DCC on any issues or problems;
- Maintenance of all correspondence with the DCC Data Manager via email, problem reports, memos, or phone calls;
- Participation in data management and training conference calls;
- Training data entry personnel as needed;
- Sending copies of forms for QC double data entry of forms as required;
- Resolution of discrepancies identified during QC process;
- Participation in on-site QC measures during DCC site monitoring visits; and
- Demonstration of QC methods during the DCC site monitoring visits.

Depending on the clinical site, data management responsibilities may also include:

- Preparing and finalizing shipments to the Repository;
- Maintaining the supply of patient ID and aliquot labels; and
- Labeling aliquot tubes as part of preparation for study visits.

These data management tasks are explained in more detail in Section VII below, “Routine Data Management Tasks”.

C. Hardware and Software

i. System Requirements for accessing ADEPT

For adequate performance when accessing ADEPT data systems the local workstation should have the following:

- PC Running Windows OS or Mac running OS 9 or OS X. ADEPT can be run on Macs, but with some limitations. For questions contact your project data manager. Currently ADEPT does not support access from UNIX or Linux machines.
- 500 MHz or faster microprocessor
- 64 megabytes of RAM
- Monitor capable of 800x600 pixel resolution
- High speed internet connection

- Appropriate Browser (see below)

Systems with slower microprocessors, less RAM will still be able to access ADEPT systems but performance will not be ideal. Systems supporting lower-resolution displays will also be able to access the system - but the visual interface is designed for 800 x 600 resolution. A high speed Internet connection (cable modem, DSL, or TI) is highly recommended for high volume of data entry. Slower dial-up access may be adequate for occasional access or reviewing reports.

ii. Browser Requirements

PC System running any Windows Operating System:

- Netscape browsers 4.5x-4.8x
- IE 5.0, 5.5, or 6.0

Mac

- Netscape browsers 4.7x (OS 9 only)
- IE 5.2 (OS X only)

D. Types of HALT-C DMS Accounts

There are two types of HALT-C Data Management System user accounts:

- **Full User:** Users who complete data entry certification will be Full Users. Full Users have access to all facilities of the data management system appropriate for their site. These users will be able to enter patient data forms, view data and resolve edits for their site, and run reports. In addition, users at clinical sites where patients are seen can schedule and modify appointments, print visit control sheets, and use the specimen shipping and tracking system.

- **Limited User:** Users who have not completed data entry certification will be Limited Users. Limited accounts are available to users who only need to run reports. Limited users are not able to directly view or modify patient data, schedule appointments, print visit control sheets, or use the shipping and tracking system. No data entry certification is required to become a limited user.

E. Data Entry Certification Goals

The goals of the HALT-C data entry certification program are to ensure:

- Data are entered accurately into the HALT-C Data Management System (DMS).

- Users are familiar with interpreting and resolving edit reports, know how to correct data entry errors, and how to deal with missing or out of range information on data forms.

- Users are familiar with the features of the HALT-C DMS, including printing Visit Control Sheets, setting appointments, running reports, and monitoring visits and the status of data entry of forms.

F. The Full User Data Entry Certification Process

All training and certification exercises are done in the Development & Training System only. **Never** use the Production System for training. See MOO Section J-III, QC/QA HALT-C Data Entry Certification Program for detailed information on becoming certified to use the DMS.

G. Development and Production

ADEPT has two versions – Development and Production. All development, testing and training is done in Development. Actual patient data should not be entered in Development. Production is used to enter all study data. It is very important that only study data be entered in Production. Utmost care must be taken not to perform any testing or data entry practice in Production. To minimize risk of an error, each site has a different URL (web address). The training login page also has a red border around the data entry boxes. This is to help you remember that you are in Development, not Production. Also, the menu section of Development is green, while it is blue in Production. This is another way to remind you which version of ADEPT you are using.

To access ADEPT **Development** type in the following address on your browser:

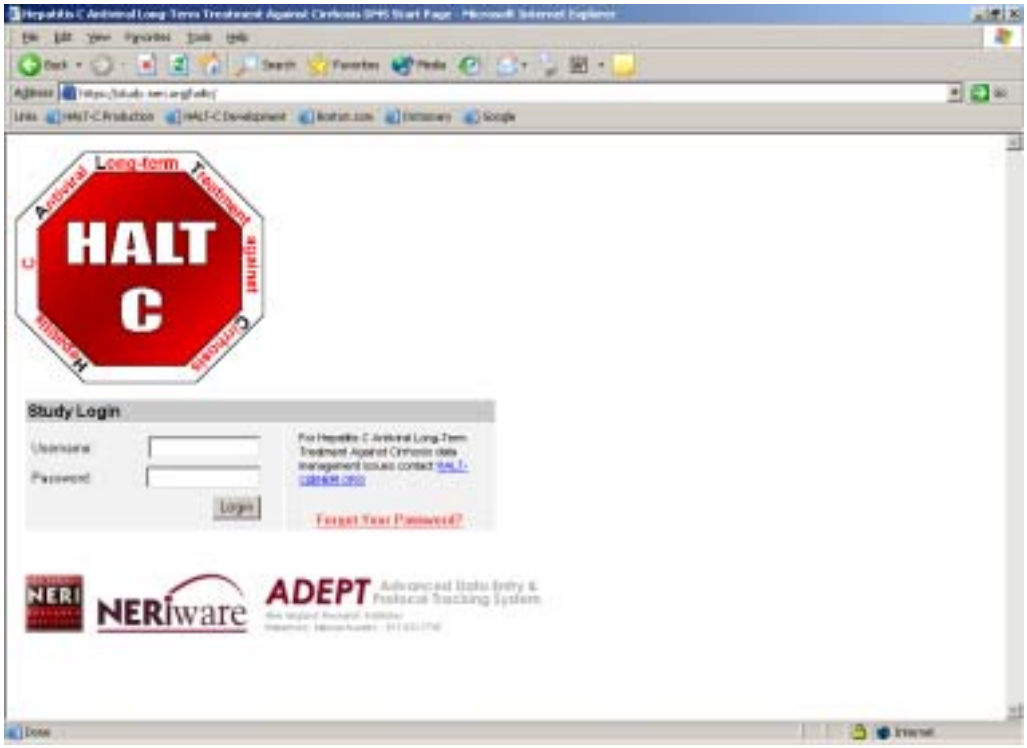
https://studydev.neri.org/haltcdev/

You will then see the following login screen:

To access ADEPT Production type in the following address on your browser:

https://study.neri.org/haltc/

You will then see the following login screen:



II. HALT-C Site Numbers and Trial IDs

A. Site numbers

For the HALT-C Trial, each clinical site has a 2-digit number, between 11 and 20. The site numbers are:

Site #	Name
11	University of Massachusetts & University of Connecticut
12	Saint Louis University
13	Massachusetts General Hospital
14	University of Colorado Health Sciences Center
15	University of California, Irvine & VA Medical Center, Long Beach
16	University of Texas Southwestern
17	University of Southern California
18	University of Michigan
19	Virginia Commonwealth University Health System
20	NIDDK Liver Disease Section

B. HALT-C Trial IDs (Patient IDs)

All patients who enter pre-screening for the HALT-C Trial are assigned a trial ID, also referred to as the patient ID. The Screening Log sheet should be used to assign a trial ID to each patient. This ID will stay with the patient for the duration of the trial. For more information on assigning IDs using the Screening Log, see MOO Section D-1, Screening.

Each trial ID is six digits long. The first two digits refer to the site number. The next three digits are a unique number, from 001 to 999, assigned to each patient at a given site. The final digit is a computer assigned check digit designed to catch data entry errors. For example, for patient ID 11-032-4: 11 is the site number (University of Massachusetts); 032 is the patient number; 4 is the check digit.

III. Protocol Tracking and the DMS

- The HALT-C DMS, implemented in ADEPT, is designed to capture all the data required by the HALT-C protocol. There are a number of ways that the HALT-C DMS tracks the protocol:
- Screening and eligibility criteria are programmed into the DMS;
- Expected visits and visit windows are programmed into the DMS;
- The DMS tracks data forms expected at particular study visits;
- Completed data forms and protocol-related events, such as randomization, are triggers for other expectations in the DMS;
- Lab tests and procedures (e.g. ultrasound, biopsy, endoscopy) expected at each study visit are programmed into the DMS; and
- The DMS tracks specimens to be shipped to the central Repository – from collection to receipt at the final destination.

Individualized Visit Control Sheets listing expected data forms, lab tests, and Repository specimens can (and should) be printed for each study visit for each patient. (See Visit Control Sheets, in Section E, below.)

A. Expected Study Visits

All study visits expected by the protocol have been programmed into the HALT-C DMS. These are the visits listed in the HALT-C Visit Schedule (MOO, Section B-5). The visits programmed are the Screening visit (S00), Lead In Phase visits (W00-W24), Randomized Phase visits (M09-M54), visits for Week 20 Responders (W30-W72), and the Randomization visit (R00).

B. Expected Data Forms and the Visit Schedule

Per protocol, a set series of data forms is expected for each study visit for the main HALT-C Trial. The Visit Schedule, in Section B-5 of this manual, lists all expected main trial forms, and the visits at which they are expected. The Visit Schedule in Section B-5 also lists the ancillary study forms that are to be completed for all patients in the HALT-C Trial.

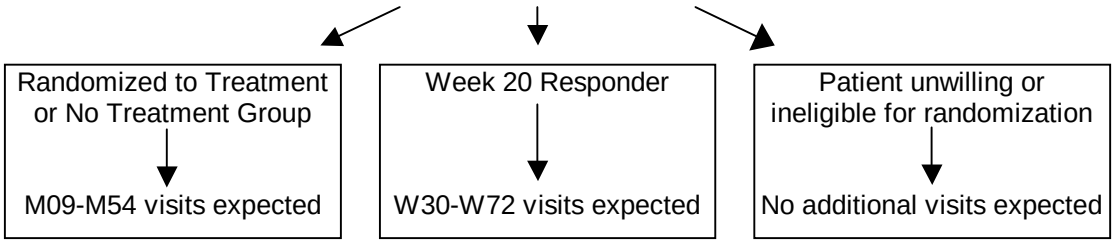
In addition, there are several ancillary studies that some patients are participating in, depending on study site and other ancillary study inclusion criteria. Some of these ancillary studies require additional patient consent. (For more information, see the ancillary study protocols in Section K of this manual.) If a patient is participating in one or more ancillary studies, all the data forms for those ancillary studies will be expected by the DMS for that patient. All expected ancillary study forms, as well as main trial forms, will appear on the patient's Visit Control Sheet for each study visit. (See Visit Control Sheets, in Section E, below.)

C. Major Protocol Triggers

Protocol triggers are occurrences (e.g. study visits completed, forms data entered) that cause the DMS to change its state (e.g. expect additional study visits or data forms). The required sequence of events for the trial is programmed as a set of DMS triggers. The major protocol triggers for the HALT-C Trial are shown in the figure, below.

Major HALT-C Protocol Triggers

Trigger	DMS result
1. Data entry of Form #1 (Trial ID Assignment) →	• Adds a new Patient ID to the DMS • Screening visit (S00) is now expected • Form #3 (Screening Medical History Interview) now expected
2. Data entry of Form #3 →	• Sets screening to visit window (up to 14 weeks) • All other screening forms are now expected. As a result, **no screening forms can be entered before Form #3.**
(Complete screening forms, with Form #4, Screening Checklist, last)	
3. Data entry of Form #4 →	• Triggers expectation of the baseline visit (W00) **Note: No baseline forms can be entered until Form #4 is data entered.**
(Form #8, Baseline Visit Date, should be the first form entered.)	
4. Data entry of Form #8 →	Triggers visit windows and the expectation of visits from week 2 – week 24 (W02-W24)
(Enter expected forms for visits W02-W20)	
(Between the W20 & W24 visits, collect week 20 HCV RNA status, Ishak Fibrosis Score from central review of the screening biopsy, week 20 ultrasound results, week 20 AFP result, week 12 and week 20 CTP score - so that Form #21, Randomization Checklist, can be entered.)	
5. Data entry of Form #21 →	Triggers the randomization program for all patients.



D. Visit Windows

Visit windows are the time frame (determined by protocol) in which study visits should be completed. The length of each visit window (number of days or weeks) is based on the number of weeks between study visits. Thus, visit windows early in the Lead-in Phase are shorter than visit windows during the Randomized Phase or Week 20 Responder Phase. HALT-C trial visit windows are listed below:

- Screening: Up to 14 weeks between the first screening visit and the baseline (W00) visit.
- Lead-in Phase: W02, W04 visits: The visit window is +/- 3 days.
- Lead-in Phase: W08 – W24 visits: The visit window is +/- 1 week.
- Randomized Phase: M09 – M54 visits: The visit window is +/- 3 weeks.
- Week 20 Responder Phase: W30 – W48 visits: The visit window is +/- 1 week.
- Week 20 Responder Phase: W60, W72 visits: The visit window is +/- 3 weeks.

Visit windows are programmed into the HALT-C DMS. They are displayed on Visit Control Sheets and the Summary of Study Visit screen (see below for more information on these DMS screens). Visit windows are also programmed into appropriate date validations for form data entry (see below for information on form validations).

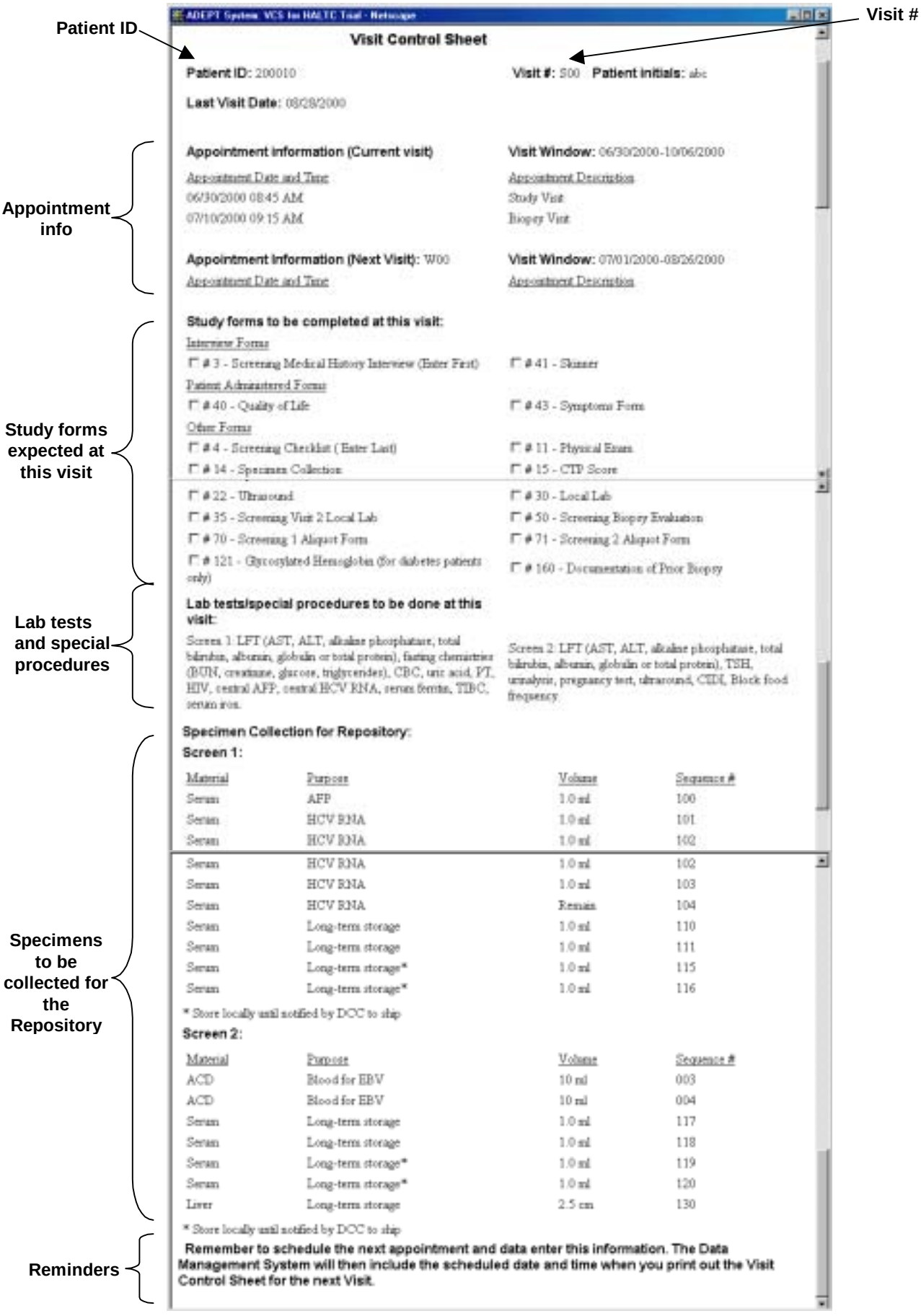
E. Visit Control Sheets

A Visit Control Sheet (VCS) is a blueprint for a patient study visit. Visit Control Sheets are available from within the HALT-C DMS, and contain individualized patient information pertaining to a particular study visit.

Visit Control Sheets contain the following information:

- HALT-C patient ID
- Visit number for this VCS
- Patient initials, as recorded on Form #1
- Date of the last study visit
- Appointments (date, time and type of visit) for the current visit that have been scheduled using the appointment setting feature in the DMS.
- The visit window for the current study visit and the next study visit.
- A checklist of all study forms that are expected for the current study visit, grouped as: interview forms, patient administered forms and other forms. Forms required by the main HALT-C protocol are listed. In addition, if the patient is participating in any ancillary studies that require completing data forms at this visit, those ancillary study forms will be listed.
- A list of lab tests and special procedures for this study required for the main HALT-C protocol and any ancillary studies in which the patient is participating.
- Non-trial medications that the patient was taking at the last study visit. This list will be used in completing the Medications Interview form (#12). Note: This list is available on the VCS beginning with W04.
- Adverse events that were entered as ongoing at the last study visit. All ongoing adverse events are listed on the VCS, so they can be tracked and updated at the next study visit.
- A list of all the specimens (specimen type, purpose, volume, and sequence #) that need to be collected and shipped to the Central Repository for this study visit. Specimens to be shipped to the Repository for the main trial and ancillary studies are listed. See Specimen Shipping and Tracking in Section C-5 for more information.
- Reminders of special procedures, such as ultrasound, endoscopy, or biopsy, that may need to be scheduled for the next study visit.

A sample Visit Control Sheet, for screening, is included on the next page.


Patient ID
Visit #
Appointment info
Study forms expected at this visit
Lab tests and special procedures
Specimens to be collected for the Repository
Reminders
ADEPT System: VCS for HALTC Trial - Netscape
Visit Control Sheet
Patient ID: 200010
Visit #: S00
Patient initials: abc
Last Visit Date: 08/28/2000
Appointment information (Current visit)
Visit Window: 06/30/2000-10/06/2000
Appointment Date and Time
Appointment Description
06/30/2000 08:45 AM
Study Visit
07/10/2000 09:15 AM
Biopsy Visit
Appointment Information (Next Visit): W00
Visit Window: 07/01/2000-08/26/2000
Appointment Date and Time
Appointment Description
Study forms to be completed at this visit:
Interview Forms
3 - Screening Medical History Interview (Enter First)
41 - Skinner
Patient Administered Forms
40 - Quality of Life
43 - Symptoms Form
Other Forms
4 - Screening Checklist (Enter Last)
11 - Physical Exam
14 - Specimen Collection
15 - CTP Score
22 - Ultrasound
30 - Local Lab
35 - Screening Visit 2 Local Lab
50 - Screening Biopsy Evaluation
70 - Screening 1 Aliquot Form
71 - Screening 2 Aliquot Form
121 - Glycosylated Hemoglobin (for diabetes patients only)
160 - Documentation of Prior Biopsy
Lab tests/special procedures to be done at this visit:
Screen 1: LFT (AST, ALT, alkaline phosphatase, total bilirubin, albumin, globulin or total protein), fasting chemistries (BUN, creatinine, glucose, triglycerides), CBC, uric acid, PT, HIV, central AFP, central HCV RNA, serum ferritin, TIBC, serum iron.
Screen 2: LFT (AST, ALT, alkaline phosphatase, total bilirubin, albumin, globulin or total protein), TSH, urinalysis, pregnancy test, ultrasound, CIDI, Block food frequency
Specimen Collection for Repository:
Screen 1:
Material | Purpose | Volume | Sequence #
Serum | AFP | 1.0 ml | 100
Serum | HCV RNA | 1.0 ml | 101
Serum | HCV RNA | 1.0 ml | 102
Serum | HCV RNA | 1.0 ml | 102
Serum | HCV RNA | 1.0 ml | 103
Serum | HCV RNA | Remain | 104
Serum | Long-term storage | 1.0 ml | 110
Serum | Long-term storage | 1.0 ml | 111
Serum | Long-term storage* | 1.0 ml | 115
Serum | Long-term storage* | 1.0 ml | 116
* Store locally until notified by DCC to ship
Screen 2:
Material | Purpose | Volume | Sequence #
ACD | Blood for EBV | 10 ml | 003
ACD | Blood for EBV | 10 ml | 004
Serum | Long-term storage | 1.0 ml | 117
Serum | Long-term storage | 1.0 ml | 118
Serum | Long-term storage* | 1.0 ml | 119
Serum | Long-term storage* | 1.0 ml | 120
Liver | Long-term storage | 2.5 cm | 130
* Store locally until notified by DCC to ship
Remember to schedule the next appointment and data enter this information. The Data Management System will then include the scheduled date and time when you print out the Visit Control Sheet for the next Visit.

F. Forms Completed with Central Lab Data – Not Entered at Clinical Sites

Several data forms for the main trial and ancillary studies are completed at central labs, or at the DCC. Note: Even though these forms are entered centrally they are expected at study visits, and become part of the patient data record in the same manner that site-entered data forms do. These centrally entered forms can be viewed at the clinical center once they have been entered.

Examples of forms that are entered centrally, but become part of the patient's DMS record are:

- Form #31, Central Lab - HCV RNA: This form is expected at several study visits. HCV RNA is tested and data entered centrally, at the Virology Lab at the University of Washington.
- Form #32, Week 20 Quantitative HCV RNA and Form #33, Central Lab - HCV Genotype: These forms are data entered at the central Virology lab.
- Form #51, Central Review of Pathology: This form is completed centrally, following central biopsy reviews by Pathology Committee Pathologists. The form is data entered at the DCC.

G. HCV RNA Email for Screening and Week 20 Visits

At several HALT-C study visits, HCV RNA testing is done by the Central Virology Laboratory, at the University of Washington (See the Visit Schedule, MOO Section B-5). The result of HCV RNA testing is needed quickly at screening (to determine eligibility for the trial) and at week 20 (to determine eligibility for randomization). To ensure that Study Coordinators get these results as quickly as possible, the sites receive email notifying them of the results. This email is sent automatically by the ADEPT system when the Virology Lab data enters Form 31 (Central Lab HCV RNA). An additional report can be printed out for the patient's medical record.

IV. Main Trial Randomization and the DMS

- Form #21, Randomization Checklist, is used to trigger the DMS randomization system, which will set a patient's randomization status. Form #21 must be completed between W20 and W24 visits for all HALT-C patients, whether or not the patient will enter the Randomized Phase of the trial.

- In order for a patient to be eligible for randomization, the following criteria must be true. The patient must:
  - be HCV positive;
  - have at least one CTP score < 7;
  - have AFP level $\leq$ 1000;
  - have no evidence of HCC on the week 20 ultrasound; and
  - have completed all other week 20 testing.

  (See the HALT-C Protocol for additional details on eligibility requirements for randomization.)

- Randomization for the HALT-C Trial is done between weeks 20 and 24, after the results of week 20 tests have been obtained. Randomization needs to be done early enough, so that there is time to prepare for the week 24 visit. Since time is very limited between week 20 and week 24, Form #21, Randomization Checklist, was developed to streamline the randomization process. Although week 20 results must be reviewed before randomization, it is not necessary that all week 20 forms be entered before randomization.

- Once all week 20 test results are available, Form #21 should be entered. The patient automatically receives a randomization status when the form is entered. The patient's randomization status is displayed at the top of the patient data entry screen and an automatic email containing the randomization status is sent to specified study personnel at the sites.

- Randomization status: After Form #21 is data entered, a patient will be assigned to one of four possible randomization categories:
  - Randomized to Treatment - will enter the Randomized Phase (M09 - M54 visits)
  - Randomized to No Treatment - will enter the Randomized Phase (M09 - M54 visits)
  - Week 20 Responder – will enter the Week 20 Responder Phase (W30 – W72 visits)
  - Not Eligible or Not Willing to be Randomized – No further study visits

Randomization problems: If a site's Internet connection is not available, or there are other reasons that the patient cannot be randomized by entering Form #21 at the clinical site, contact DCC Data Management personnel for further assistance.

V. Tracking Forms by Data Entry Status

The DMS tracks the status of every form. Each data form in the HALT-C DMS has a single status at any one point in time. This status changes as a form goes through the entry and editing processes. Possible form statuses are:

- Expected
- Complete
- Pending edits
- Missing

Forms that are required per the protocol, but not yet data entered, are listed in the DMS as *Expected.* As the forms are data entered into the HALT-C DMS, on-line checks for data entry errors, missing data, and invalid or out-of-range values occur. When a form is entered without any errors, it will become *Complete*. If there are validation errors or missing fields upon completion of data entry, the DMS will generate an edit report and the form will have a *Pending Edits* status. Resolution of the entire edit report will convert a form with pending edits into a *Complete* form. If data for an expected form are not collected, and never will be, the form can be set to *Missing***.**

Data management reports (Detail Form Status Report, Detail Form Status by ID) should be run periodically to track form status.

VI. Recording Clean Data: Validations, Edits and Edit Reports

A. Validations

All data that are entered into the HALT-C DMS are checked by on-line validations. Validations are question-by-question checks that give immediate feedback (during form entry) to help catch entry errors or form completion errors. If data entered for a question fall outside the validation range, a validation error box is displayed. The validation error box provides information on the reason that the validation failed and, depending on the question, explains options for continuing with data entry. Validation information is also available on the help screens associated with each question during data entry.

Each question on every data form has been programmed with a validation that is tailored to catch problems with that particular question. For example, questions regarding lab results would have validations based on the expected range of values for those tests. A question of

patient's date of birth would check that a reasonable date is entered; for example, dates after the current date would cause a validation error. Subsequent questions pertaining to dates, such as the date the patient began prior treatment with Interferon, would be programmed so that dates prior to the patient's date of birth cause validation errors.

Validation Rigidity

Validations have different levels of rigidity, depending on the type of question and the particular validation.

The levels of validation rigidity are:

- Cannot proceed – It is not possible to continue entering data until the validation error is corrected.
- Cannot be complete – Data entry can continue, but the form cannot have Complete status unless the validation error is corrected.
- Override acceptable with reason – Values outside the validation range may be acceptable, but an explanation and initials of the data collector are required.
- Override acceptable as is – Values outside the validation range are acceptable. No explanation is required.

The process for correcting validation errors is explained below, in the section, "The Data Entry / Edit / Data Cleaning Process". Section C-4, Data Entering Study Visit Forms, describes data entry and help screens in detail.

B. Missing Information on Data Forms

There can be several reasons why a question on a data form is missing information at the time of entry. The data collector may overlook the question on the form or the data may be truly missing (e.g. lab test not done). For interview forms or patient administered forms, the patient may refuse to answer a question or not know the answer to a question. A hand written answer to a question may be illegible, so that it is not possible to enter the data. Or, in rare cases, the question may not apply to a specific situation.

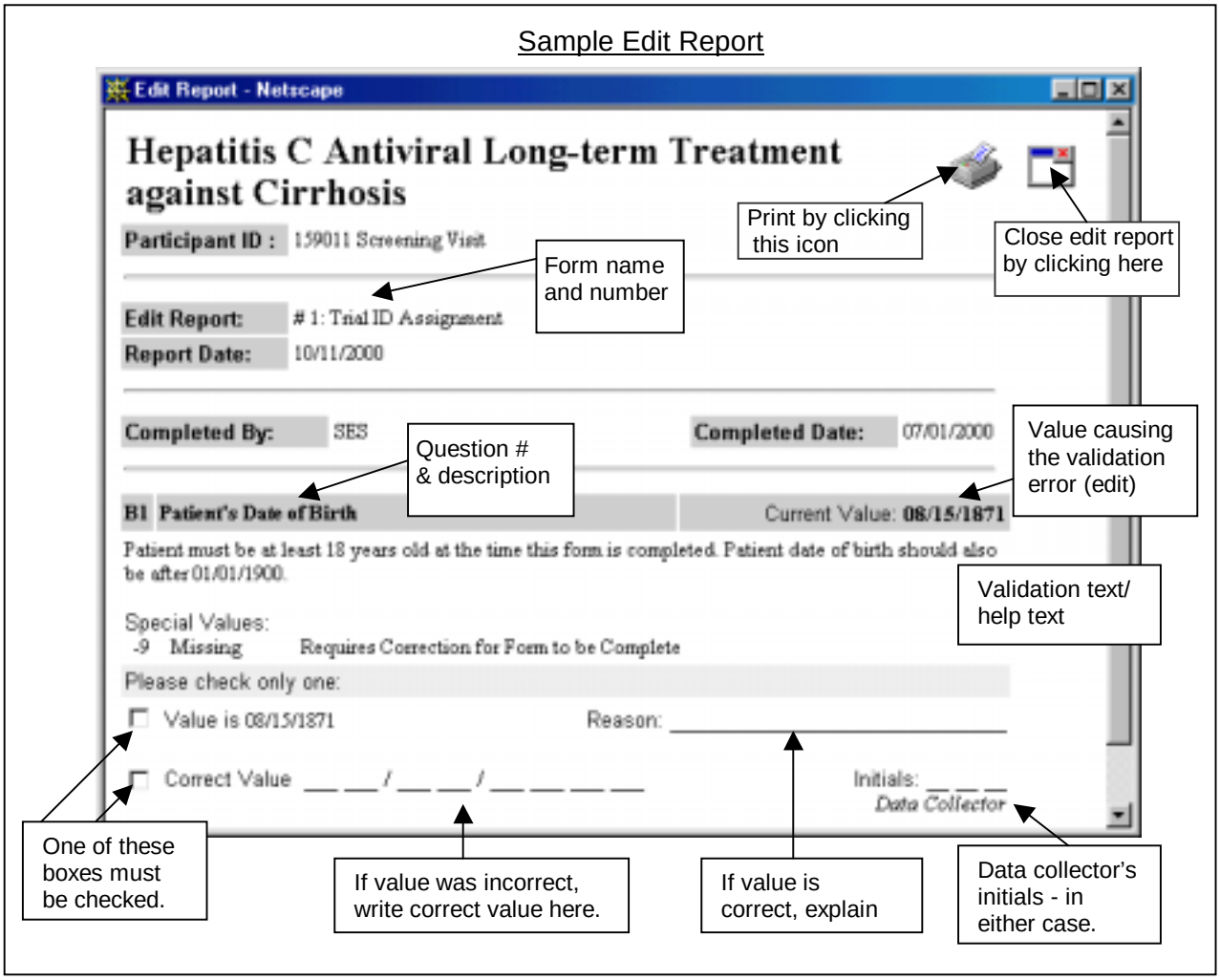
Depending on the particular form and question, missing data will require different mechanisms for data entry. As with validation errors, there is a "rigidity" associated with missing information. It ranges all the way from situations in which form entry cannot proceed if information is missing, to cases where missing information is acceptable without further explanation. During data entry, the help screen that is associated with each question provides information on types of missing data and missing data rigidity for that question.

The process for dealing with missing information is explained below, in the section, "The Data Entry / Edit / Data Cleaning Process". Section C-4, Data Entering Study Visit Forms, describes data entry and help screens in detail.

C. The Data Entry / Edit / Data Cleaning Process

An edit occurs when data fall outside of acceptable ranges. Edits will also occur when fields are entered as missing. All edits require clarification. An edit report (a paper trail of all edits) will be displayed, automatically, if there are any outstanding edits when data entry reaches the end of a form. If this occurs, the edit report should be printed.

Sample Edit Report

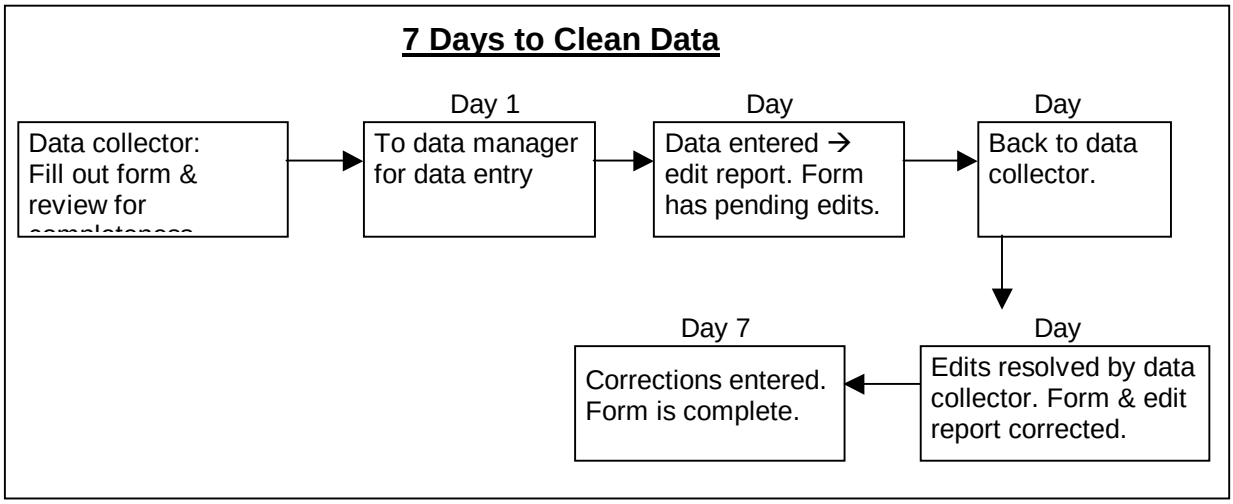


After data entry, any edit reports that are produced should be returned to the data collector for resolution. An edit report outlines all the validation errors that occur in a form and indicates which type of validation each listed value failed. Edits require an explanation or verification of the answer provided. If a value missing at the time of original entry can be determined during the editing process, then the correct information should be written in the space next to “Correct Value” on the edit report. If a value fails a validation and cannot be reconciled, the blank line that appears on the edit report labeled “Reason” should be used to justify or explain the inconsistency. If required data is unavailable, this should be explained on the line provided on the edit report. Each explanation or validation on the edit report should be dated and initialed by the person resolving the edit. It is important to remember that these reasons should be as informative as possible (within the space restrictions) since they will be used to explain missing or out of range data when the data is analyzed and incorporated into trial publications.

Once the edit is clarified on the edit report, the resolution and associated information must be written onto the original paper form and corrected with the initials of the person resolving the

edit and date of the correction. **The edit report should be stapled to the back of the original form and saved as a paper record of changes made.** The corrected values should then be entered into the DMS. If, however, exceptional or missing values have been verified and explained, an **override** can be entered to resolve the edit. Once all edits are cleared, the DMS changes the form's status from *Pending Edits* to *Complete*.

Below is a timeline for the data entry/ edit/ data cleaning process. The goal is to resolve edits as quickly as possible, so that edits do not pile up, and the information is still fresh (i.e. records still available). Quicker edit turn-around results in cleaner data.



D. Avoiding and Resolving Edit Reports

Several things can be done to streamline the process of collecting clean data.

Data collectors can (should):

- Explain exceptional (or missing) values in the margins of data forms, with date and initials. This will allow data managers to enter these values, with override reasons, date and initials, during initial data entry.
- Write clearly when filling out forms.
- Use standard procedures when making corrections on forms. Use a single line to cross out incorrect values. Write the correct values next to the question, with date and initials.
- Review forms prior to data entry.
- Do not leave blank, non-skipped questions on forms.
  - → If the patient refuses to answer a question on an interview or patient administered form, write "RF", the date and your initials in the margin of the form.
  - → If the patient does not know the answer to a question on an interview or patient administered form, write "DK", the date and your initials in the margin of the form.
  - → If a value is missing, and will not be available, write "Missing", the reason, the date and your initials in the margin of the form.

Note: There are special data entry values, listed in the help text for each question during data entry, that explain how to enter questions that have "RF," "DK," or "Missing" comments. The section of this manual on data entry procedures, C-4, explains the procedures to follow when questions that are missing answers are encountered during data entry.

Data management personnel can:

- Give feedback to data collectors to avoid future mistakes.
- Run the Detailed Form Status Report, weekly, to identify all outstanding edit reports.
- Work with the DCC to refine edits.

- Establish procedures to facilitate timely resolution of edits. If necessary, establish regular meetings to discuss open edits.

VII. Routine Data Management Tasks

A. Introduction

This section provides an overview of the major data management tasks, including:

- Printing Visit Control Sheets for study visits;
- Preparing packets for patient study visits, including data forms labeled with HALT-C patient ID labels;
- Printing Upcoming Data Cleaning by ID report
- Maintaining a supply of patient ID and aliquot labels;
- Reviewing forms prior to data entry;
- Data entering forms, including adding log entries to trial medication and adverse event forms;
- Printing edit reports, following-up with the data collector, and entering corrected data;
- Tracking the status of all forms related to the HALT-C Trial;
- Maintaining patient records, including filing data forms;
- Preparing and finalizing shipments of specimens to the Repository;
- QC double data entry; and
- Resolving problems identified during QC process.

The above tasks are likely to occur at least weekly, or even daily. In addition, there are tasks that should be performed periodically. These include:

- Running reports;
- Participating in data management and training conference calls; and
- Following up with DCC on any issues or problems.

B. Printing Visit Control Sheets

A Visit Control Sheet (VCS) should be printed for each patient study visit. As explained above in the section on Visit Control Sheets, the VCS contains important individualized information for conducting a study visit. The VCS should be printed with enough lead-time before a study visit to allow for adequate preparation. See Section C.3 of this manual, “Using the Data Management System”, for instructions on printing a VCS for a particular patient visit.

C. Preparing Packets for Study Visits

Prior to a patient visit, consult the VCS for a list of data forms to be completed at that study visit. Each page of each data form should be labeled with a paper patient ID label.

D. Printing Upcoming Data Cleaning by ID report:

Prior to a patient visit, print the Upcoming Data Cleaning by ID report for that patient. If any items need to be resolved, include these in the packet. See Section C-6 of this manual, DMS Reports, for instructions on how to use this report.

E. Maintaining a Supply of Patient ID and Aliquot Labels

Paper patient ID labels and patient ID/visit # aliquot labels: Paper patient ID labels are used to label data forms and biopsy slides. Patient ID/visit # aliquot labels are used to label aliquot tubes and Vacutainer tubes that will be shipped to the Repository. Both of these types of labels are supplied by the DCC. Notify the DCC of the patient ID numbers of patients who are going to be screened for the HALT-C Trial. The DCC will send enough paper patient ID labels for all forms for screening and baseline visits. The DCC will also send enough patient ID/visit

aliquot labels for screening and baseline. You can request these labels in a batch by sending the DCC a list of patient ID numbers via email.

You will need additional labels for patients who complete screening and enter the Lead-in Phase. Labels for W02 – W24 visits will be sent automatically by the DCC, after data entry of Form #8, Baseline Visit Date. Labels for the later trial phases (Randomized Phase, Week 20 Responder Phase) will be sent automatically after data entry of Form 21 (Randomization Checklist), between W20 and W24.

Sample ID/sequence # aliquot labels: These labels are supplied in packets by the Repository (BBI). They are used to label aliquot tubes and Vacutainer tubes that will be shipped to the Repository. Sample ID/sequence # aliquot labels are shipped in bulk. Notify the DCC if your supply of these labels is running low.

There are separate packets of sample ID/sequence # aliquot labels that correspond with each of the aliquot forms. Each packet of these aliquot labels may only be used for a single patient visit; any labels not used should be discarded. See the section on Specimen Shipping and Tracking, below, for a full discussion of using aliquot labels.

F. Reviewing Forms Prior to Data Entry

Both the data collector and the person doing data entry should review forms. This will improve data quality and reduce the number of problems detected during data entry. The following should be reviewed:

- Completeness;
- Legibility; and that
- Corrections have been made properly (single black line to cross out, dated and initialed). See section C-2 of this manual, Data Collection, for a full explanation.

G. Entering Forms into the DMS

Forms should be entered as soon as possible after they are completed and reviewed. The bulk of data entry will be data from forms expected at particular study visits. In addition, log entries should be routinely entered for the log forms (trial medication tracking forms and Adverse Event Report), so that the DMS reflects the information collected at study visits and lab tests.

H. Printing Edit Reports, Follow-up With Data Collector, and Entry of Corrected Data

These tasks should be done on a regular basis. The goal is to resolve edits as quickly as possible, so that they do not accumulate and information on the data collected is still fresh (i.e. records are still available). See the section above on Recording Clean Data for a full explanation.

I. Tracking the Data Entry Status of Forms for the HALT-C Trial

Reports should be run regularly to monitor the status of data forms required for the trial. Forms that are expected at study visits (status ***Expected***) should be completed and data entered as soon as possible. Forms with pending edits due to missing data or values out of range (status ***Pending edits***) should be resolved as soon as possible. See the sections, Tracking Forms by Data Entry Status, above, and Section C-6, Running Reports, for more information.

J. Maintaining Patient Records with Paper Data Forms and Edit Reports

Paper copies of data forms should be properly filed after entry. Edit reports should be stapled to the forms that were corrected. Log forms should be kept up to date, and re-filed after data entry. Data forms should be accessible for timely quality control (QC) activities.

K. Preparing and Finalizing Shipments to the Repository

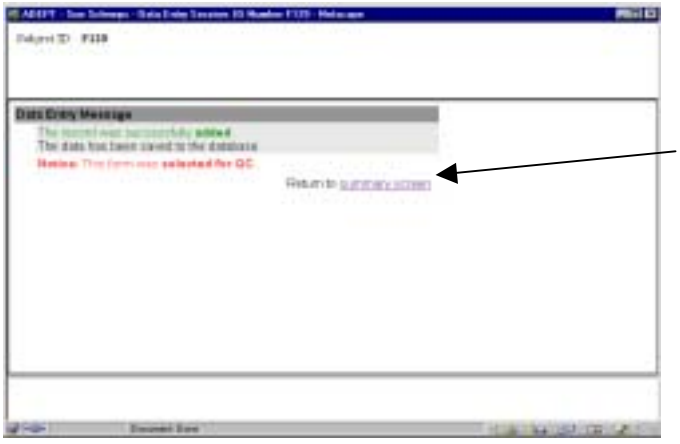
Manifests for all shipments to the Repository (BBI) must be prepared in the DMS prior to shipping. In addition, the information about the specimens being shipped must be verified and finalized in the DMS prior to shipping.

L. QC Double Data Entry

The ADEPT Double Data Entry (DDE) system was developed to directly measure and monitor the accuracy of data entry. In this system, data forms are entered into the DMS and some forms are selected for QC.

Study forms are data entered by a DCC-certified data entry (DE) person under an assigned username. The paper copy of a form must match what is data entered. If an edit report for a form has been completed, the corrected information must appear both on the edit report and in the appropriate place on the original form. The QC system selects completed forms for DDE using a computerized algorithm, which is used to determine how often a particular form is selected for QC for each user. As a form is entered the second time, the DMS compares the value entered in each field during DDE against the value entered initially. If the information data entered in the DMS is not written on the paper copy, the DCC has no choice but to call it an error. (See Sections C-1 and C-4 of the Manual of Operations for information on data entry.) Errors are then used to calculate the user's error rate and to determine the percentage of forms selected for QC

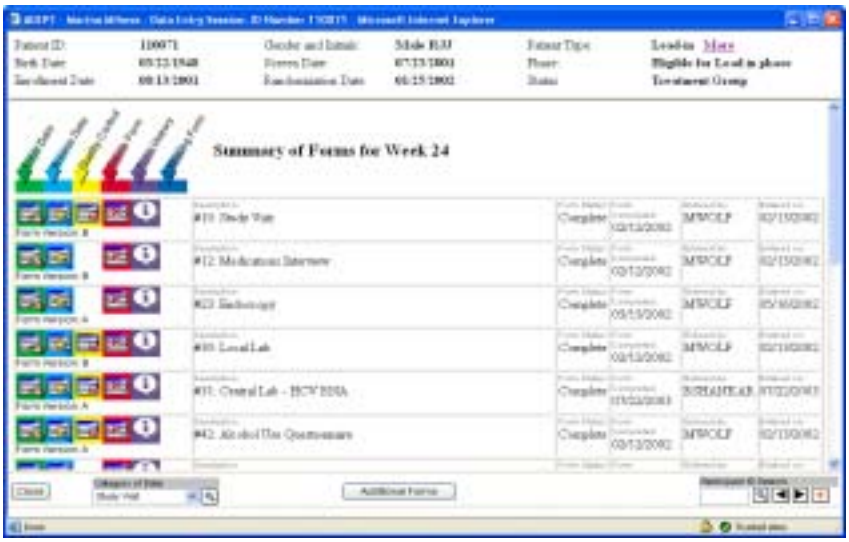
If a form is selected for QC, the DMS shows a message screen stating, "Notice: this form was selected for QC" after the form has been entered. Click the Summary Screen button to return to the **Summary of Forms** screen.



Upon returning to the summary screen, a yellow box appears to the left of the form.

At the **Summary of Forms** screen, note that there is a yellow icon under the "Quality Control" column. The green "Enter Data" icon no longer appears. No edits can be made to the form until the QC process has been

When DDE is completed by the DCC, the green "Enter Data" icon reappears on the Summary of Forms screen. Only then can your site or laboratory view and/or edit this form.



Since forms that are triggered for QC will not be available for edits until the DCC completes the DDE process, forms that require constant access are not selected for QC (e.g. adverse event, medication logs and shipping forms). Forms #31, #37, and #51 are selected 100% of the time due to their importance. All other forms are selected for QC based on past accuracy of data entry, as measured by forms data entered by a particular user. Since the percentage of the forms selected for QC is based on the most recently entered forms, fewer forms will be selected if a user's accuracy improves.

The DMS automatically re-calculates DE accuracy rates after completion of each DDE. Only the DCC can view these rates. The accuracy rate is defined as the total number of fields with initial DE errors divided by the total number of data entered fields. Overall site and form accuracy rates are calculated. An individual accuracy rate is also calculated on a form-by-

form basis for each username. If you give your username and password to another person, your accuracy rate will be affected!
Individual accuracy rates determine the percentage of forms selected for QC in the future. A new DE person experiences 100% QC selection until 10 of each form have undergone DDE. A DE person who has been consistently very accurate on a particular form will notice that few are selected for QC (the minimum is 7%). A DE person who consistently makes errors on a particular form will notice that many are selected for QC (the maximum is 75%). Fewer forms will be selected if accuracy rate improves for that particular form over time. More forms will be selected if accuracy rate decreases for that particular form over time. Forms with one year of excellent study-wide DE accuracy rates are removed from QC.

All forms selected for QC must be sent to the DCC. This should occur within 10 business days, and preferably occurs within a day or two. There are two possibilities for sending forms to the DCC: FedEx or fax. Copies of forms may be batched and sent via **Federal Express** to the DCC. If there is a completed Edit Report for a particular form, please send a copy with the form. The DCC Federal Express address is:

HALT-C Study
New England Research Institutes
9 Galen Street
Watertown, MA 02472

Forms may also be sent using NERI's HALT-C Federal Express account. Sites and labs should contact the DCC Project Director for permission and instructions. Alternatively, forms may be **faxed individually or in a batch** to the DCC. If there is a completed Edit Report for a particular form, please fax it with the form. The **QC fax number is: (617) 673-9526.** The QC fax cannot be used for any other purpose. A cover sheet is not required. The computerized QC fax accepts faxes 24 hours a day. The DCC prints QC faxes each morning. If any pages are missing or illegible, the DCC contacts the site to fax again.

QC forms are double data entered by DCC staff within 5 business days of receipt. The DMS automatically compares initial and second DE values and immediately notifies DCC staff of inconsistencies. DCC staff compare all inconsistencies to the copy of the form sent by the site. The form must match the data entry. DCC staff corrects all data entry errors.

The report **"Site Outstanding QC"** under Reports-Tracking lists all forms that have been selected for QC, but not yet double data entered at the DCC.

QC Steps

1. The forms are faxed to the DCC and are sorted through once they are printed out to ensure that all pages are present and legible. If there are no discrepancies, the form is QC'ed and the site can then access the form.
2. If there is a problem with a form (missing pages, illegible pages, missing information) then an email will be sent to the site indicating the form number, Pt ID, Pt initials, the visit and the problem, asking for the form to be resent in order for QC to take place.
3. The site is expected to respond within a reasonable amount of time. If nothing has been received within a week, the DCC will contact the site again.
4. The sites will notify the DCC when a form is being resent. The emails sent to the sites, from the DCC, should be used as the cover sheet of the fax to indicate that it is a re-fax.
5. If there are problems during DDE, such as discrepancies between the DMS and the paper form, illegibility or missing information, then DDE is stopped. The site will be emailed

indicating exactly what the discrepancy is and asking for the problem to be resolved within a reasonable amount of time. The process and time expectancy for receiving the refax is the same as if the problem was found prior to the attempt at double data entry.