

5 Data processing tables, worksheets, and checklists

WS 5.1 Data system worksheet (DataSys.WS)

When: Prior to start of data collection

Who: Personnel in the coordinating center

Purpose: To establish the data system for the trial

Definitions

centralized data system - A data system established and maintained at a central site (e.g., the data coordinating center) in a multicenter study used for data entry and data capture.

distributed data system - A data system consisting of component parts that are established and maintained at individual data collection or generation sites for data capture.

A. Identifying information

1. Study name: _____
2. Form completed by: _____
3. Date completed (day-month-year) _____

B. Data system

4. Type of data system
 - Distributed
 - PC-based
 - Web-based
 - Centralized
 - Other (specify) _____
5. Systems custodian
 - Coordinating center
 - Contract research organization
 - Other (specify) _____
6. System server vendor and location

Vendor: _____

Location: _____
7. Backup frequency and location of backup files

Frequency: _____

WS 5.1 Data system worksheet

Location of backup files: _____

8. Data security protections (check all that apply)

- Central authority (typically the coordinating center) for authorizing access to the data system via issue of access password
 - User password for access to data system
 - Transmission encryption
 - Training and certification of persons authorized to access the data system
 - Other (specify)
-

9. System applications (check all that apply)

- Treatment assignment
 - Treatment unmasking
 - Appointment schedules
 - Data collection reports
 - Other (specify)
-

WS 5.2 Data access worksheet (DataGet.WS)

When: Early in the course of data collection

Who: Coordinating center personnel in conjunction with other study leaders

Purpose: To specify who has access to study data during and after the trial

A. Identifying information

1. Study name: _____
2. Form completed by: _____
3. Date completed (day-month-year) _____

B. Data access policy during the trial

4. Access to interim treatment results
 - () Restricted to people in the data center and to members of the treatment effects monitoring committee
 - () Other (explain)

5. Access to interim baseline results by treatment group
 - () Restricted to study investigators
 - () Access allowed external to investigator group on request (explain)

6. Access to interim results for control-assigned group (check one)
 - () Limited to people in the data center and to members of the treatment effects monitoring committee
 - () Summary results periodically presented to study investigators
 - () Data available to study investigators to produce "natural history" papers on control-assigned group
 - () Access allowed external to investigator group on request (explain)

WS 5.2 Data access worksheet**C. Data access policy internal to the study investigatorship after completion of data collection**

7. Access to finished dataset?

 No (justify) Yes

If yes, indicate conditions for access (check all that apply)

 Signed statement accepting dataset Statement indicating assurance to respect patient privacy Statement indicating willingness to maintain chain of custody of dataset IRB approval Other (specify)

8. Access to finished dataset for paper writing (check one)

 Access proscribed Unrestricted access Access subject to approval of study leaders; use for paper writing subject to approval of study leaders and review prior to submission for publication Other (specify)**D. Access by parties external to the study investigatorship**

9. Access to treatment results during the trial (check one)

 Proscribed Other (specify)

10. Access to baseline results during the trial (check one)

 Proscribed Other (specify)

11. Access to results for control-assigned group during the trial (check one)

 Proscribed Other (specify)

12. Data access policy after completion of the trial (check any that apply)

 No access provided

WS 5.2 Data access worksheet

- () Access provided on a case-by-case basis after review and approval of request by study leaders
 - () Requests for special analyses considered; if approved by study leaders analyses done by study coordinating center and results provided to requesting party
 - () De-identified dataset on deposit for unrestricted use
 - () Other (specify)
-

E. Sign-off approval

13. Name of review and approving authority: _____

14. Date of sign-off (day-month-year) - - - - -

WS 5.3 Data editing and auditing worksheet (DataEdit.WS)

When: Early in the course of data collection

Who: Coordinating center personnel

Purpose: To outline plans for data editing and auditing

Definitions

data audit - 1. The comparison of data in a source document with those in a secondary document as a means of checking for discrepancies. 2. A comparison of data in a study form with those recorded in a medical chart for discrepancies. 3. A comparison of specified data elements on a study form with the corresponding elements in an electronic file produced from the study form for discrepancies. *Usage note:* Not to be confused with data edit. See usage note for data edit.

data edit - An instance of a change to data as a result of an edit check. *Usage note:* Not to be confused with data audit. Typically, data audits in studies involving keyed data forms consist of comparison of what is on the study form with what is keyed; done to find discrepancies between the two sources and to resolve the discrepancies. Data edits involve checks of the information recorded for inconsistencies and correction based on the checks performed.

A. Identifying information

1. Study name: _____
2. Form completed by: _____
3. Date completed (day-month-year) _____

B. Data editing

4. Edit checks for data to be admitted to the study database (check all that apply)
 - Check for permuted numbers in Id number via use of check digits
 - Check of one-to-one correspondence between Id number and name code
 - Check of visit number and time window for indicated visit
 - Other (specify)

5. Data entry and edit change rules (check all that apply)
 - Key what is recorded on study forms even if known to be wrong when keyed
 - Document changes to recorded data on study forms to maintain a one-to-one correspondence between what is recorded on study forms and what is keyed (e.g., by requiring study personnel to date and initial strike outs and to write in data replacing those struck)
 - No white out on study forms
 - No changes to data on study forms without bases for changes

WS 5.3 Data editing and auditing worksheet

- () Indelible audit trail of edit changes
 () Other (specify)
-

6. Types of edit checks performed (check all that apply)

- () Data entry with messages to indicate entry of inadmissible codes and to aid entry for skip patterns
 () Edits within forms
 () Edits across forms
 () Edits triggered by data analysis
 () Other (specify)
-

7. Edit change freezes (check all that apply)

- () Once a dataset is frozen for data analysis
 () Once a dataset has been distributed to study investigators
 () Once a dataset has been deposited in a public archive
 () Other (specify)
-

8. Describe process for dealing with data errors discovered in datasets supporting publications, including conditions under which editors are notified of errors

9. Describe process for dealing with data errors discovered in distributed datasets; include details as to how users are informed of errors and conditions under which revised datasets issued

C. Data auditing

10. Percentage of data forms routinely audited (if 0% explain) ____%

11. Frequency of routine data audits

- () Daily
 () Weekly
 () Monthly

WS 5.3 Data editing and auditing worksheet

- Once every 6 months
 - Yearly
 - Other (specify)
-

12. Method of selection of records for routine audits

- Random
 - Other (specify)
-

13. Person or group responsible for auditing

- Circuit rider
 - Data coordinating center personnel
 - Other (specify)
-

14. Site of routine data audits

- On-site
 - Off-site
 - Data coordinating center
 - Office of study chair
 - Office of the study sponsor
 - Contract research organization
 - Other (specify)
-

15. Trigger points for cause audits (check all that apply)

- High discrepancy rate in routine audits
 - Data discrepancies suggestive of data falsification
 - Large number of mismatches between what is keyed versus what is recorded on study forms
 - Other (specify)
-

16. Authorizing authority for cause audits (check any that apply)

- Coordinating center
- Study chair
- Study sponsor
- Study steering committee

WS 5.3 Data editing and auditing worksheet

() Other (specify)

17. Summary audit reports prepared and distributed to study investigator?

() Yes

() No (explain)

18. In multicenter trials, summary report distributed to all participating centers showing audit results by clinic?

() Yes

() No (explain)

WS 5.4 Data processing worksheet (DataKey.WS)

When: Prior to the start of data collection

Who: Data coordinating center personnel

Purpose: To outline data capture and processing procedures

Definitions

data harvest - The act of receiving and assembling data from data collection and generation sites, e.g., as accomplished by a data center in a multicenter trial by downloading data residing at study clinics or by receipt of electronic files from study clinics.

dependent double data entry - Double data entry with both entries performed by the same person.

independent double data entry - Data entry in which data are keyed by different people working independently of each other; as distinct from dependent double data entry.

A. Identifying information

1. Study name: _____
2. Form completed by: _____
3. Date completed (day-month-year) _____

B. Data capture and harvest

4. Principal mode of data capture (check one)
 - Keying directly from study forms at site of completion
 - Keying directly from study forms at central processing center
 - Keying from code sheets produced from study forms at site of data collection
 - Keying from code sheets produced from study forms at central processing site
 - Direct from screens presented on laptops or personal digital assistants during data collection
 - Electronic data capture (EDC)
 - Other (specify)

5. If the primary mode of data capture involves keying from completed data collection forms, indicate steps taken to minimize time lag from data collection to capture for harvest (check all that apply)
 - Dedicated data entry personnel
 - Monitoring performance to show differences in lag time from completion to data entry
 - Site visits to clinics with poor data entry performance

WS 5.4 Data processing worksheet

() Other (specify)

6. Frequency of data harvests

- () Daily
() Weekly
() Monthly
() Other (specify)

7. Data entry principles; check all that apply (skip if the principal mode of data capture is not from data collection forms)

- () Key directly from completed data forms with coding indicated on the data collection forms
() Instructions to data entry personnel to key what is recorded on forms even if known to be incorrect
() Maintenance of an electronic audit trail reflecting changes to data as results of data edits
() Double-dependent data entry
() Double-independent data entry
() Training and certification of data entry personnel
() Monitoring of data entry performance (by clinic in multicenter trials); including calculation of the median time to data entry and interquartile time range from completion of a data form to keying
() Other (specify)
