**Data Quality in Clinical Trials**

**DIA Working Group**

Richard Chamberlain, PhD, Chair, President, ECS LLC

Yvonne McCracken, MPH, Clinical Operations & Compliance, RxTrials

Teri Stokes, Ph.D., Founder/Director, GXP International

Contents

Page

[1. Introduction 1](#_Toc443251663)

[2. Study Start-up 3](#_Toc443251664)

[2.1 Risk Analysis 4](#_Toc443251665)

[2.2 Data Management Plan Example 5](#_Toc443251666)

[2.3 Actual Study Example 6](#_Toc443251667)

[2.4 Equipment 9](#_Toc443251668)

[2.5 Audit or Pre-Study Site Visits (PSSV) 10](#_Toc443251669)

[3. Study Execution 11](#_Toc443251670)

[4. Study Close-out 15](#_Toc443251671)

[5. In Summary 19](#_Toc443251672)

[6. References 22](#_Toc443251673)

# Introduction

These chapters describe the procedures that are used during the conduct of a clinical trial to manage the Quality and Integrity of the data that is collected during the trial. This emphasizes the documentation that is prepared ahead of the start of the study as well as documentation that is prepared during the conduct of the trial to identify problems that might occur during the study and to perhaps identify them before they are too late to correct, and then when the study is complete, reports that can be developed to document the study Quality.

This documentation also applies to the instruments used to make the necessary observations. There needs to be evidence that those instruments are being used properly and are working correctly.

For our purposes a clinical trial will be divided into three parts;

Part 1 – Study Start-up. This starts when the study is being organized and ends when the first patient is enrolled in the study

Part 2 – Study Execution – This starts when the first patient is enrolled in the study and ends when the last patient has their last visit.

Part 3 – Study Close-out. This starts when the last patient has their last visit and ends when the study Report is complete and signed off.

During each of these three parts it will be necessary to identify procedures and documentation to show the quality of the clinical data.

These procedures and documentation will help to

Document what needs to be done

Document what was done

Document that these two are the same, if there is a difference, fix it.

Another characteristic of Quality is that it is *measurable*. The procedures here will show how the quality of data in a clinical trial can be measured.

# Study Start-up

There are some documents that are vital to determining the quality of data in a clinical trial. These are

Study Protocol

The study protocol will contain a detailed list of all of the information that will be collected during the trial. See ICH E6(1)

Data Management Plan

The DMP should list all the fields that will be used in the study. They should be listed by Case Report Form (CRF) page or Remote Data Entry (RDE) system screen.

It should also include patent consent procedures and drug randomization procedures if appropriate.

A working group within the DIA has prepared a template for a Data Management Plan. See DIA document for DMP(2)

Statistical Analysis Plan

There should also be a plan to show what Statistical Analyses will be performed on the data. The Data Management group should review this document and check to see if there are any problems based on what they know about the data and how it will be managed.

Supporting Documents

There will be a variety of other documents that are necessary but are typically not specific to a study. For example, virtually all of the data observations are made using an instrument of some kind. Those instruments will all have instruction documents, some might have records of calibrations and other maintenance records. These documents can all have an impact on data quality.

The biggest task for Data Management during this step is getting ready for the data entry. This means

1. Defining the Data Entry details for entering the study data in the database. This will include things like;

Data fields

* Identifying all of the fields to be collected
* The characteristics of each field – size, format, etc.
* Edit checks for each field
* The location and organization of all fields on the CRF page or entry screen.
* Any other directions associated with the entry of the field.

Recording directions

* Documenting practices for recording the observations, e.g. sitting blood Pressure, Fasting Urine Specimen
* Who should be recording the observation

1. Training and Procedure for those who are doing the entry or those who will need to train the sites.

Training for those making the observation

* The patient if they are using an APP to record the observation
* The nurse during a visit
* A diagnostic center such as for an x-ray.
* Training the study monitors to train each site.

Written Procedures

* How to make the necessary observations
* How to enter and manage the data
* How to Monitor the study

These activities must all be completed before the study Start-up step can be complete and in most cases these activities will need to be documented. In most cases this can be done with a simple log that is signed by the proper authority.

## Risk Analysis

It is also likely that it would be good to do a risk analysis of some of the information above. In other words,

What could go wrong?

How severe is it if it does go wrong?

How likely is it to go wrong?

What can be done to fix it if it does go wrong?

One key area where this would typically be very useful is with the edit checks being done on the entry. Ask the four questions above when an edit check is tripped.

There has recently been an effort to do Risk Based Monitoring RBM (3). Most larger sponsors/CROs are now using RBM so it has become a daily reality for sites.

Looking at the data and the procedures mentioned here can play a major role in that effort.

There will need to be documentation of each of the entries above. Below are some simple examples of such tables from the DIA template for a DMP, and others are tables from an actual recent study.

## Data Management Plan Example

The DIA has a template for a Data Management Plan (DMP) that is very useful.

Template 2. Real Time (Front-end) Edits in EDC

The examples used in this template matrix are not meant to represent real data or situations. This matrix can be created in Excel, MS Word table or any other equivalent system.

| Edit Number | CRF | Visit/Event | Fields | Specification | Message\* | Action |
| --- | --- | --- | --- | --- | --- | --- |
| 1 | MedHist | Visit 1 | SEX  PREG | If SEX is male than PREG should be blank | If SEX is male PREG should be blank |  |
| 2 | AE,  Study Completion | End of Study | SAE  REASON | If REASON for dropping out of study is *serious adverse event* then subject should have an AE record with an adverse event designated as an *SAE* | Reason for dropping out is serious adverse event but no AE record designated as an SAE |  |
| 3. | Physical | Visit 1 | SYSTOLIC | Systolic cannot be over 120 | Blood pressure is over 120 protocol violation | Automatic email sent out to PI and Project manager |

* If an error message identifies a protocol deviation, this should be stated in the message.

## 2.3 Actual Study Example

Following are some examples of similar documentation actually generated during a study. The study was a large international end-point study that lasted for several years.

The study had 20,000 patients and lasted almost 9 years. The web site for more details is

[ascotstudy.org](http://ascotstudy.org)

There was a lot of data collected over the life of the study. There were some complexities that made it necessary to actively monitor the data management.

The data entry system had nearly 500 edit checks programmed in. The edit checks were executed as the data was entered.

By entry screen the number of edit checks was:

|  |  |
| --- | --- |
| Screening | 201 |
| Randomization | 94 |
| Follow-up | 139 |
| Adverse Event | 43 |
| Medication Forms | \*19 |
|  |  |
| Total | 496 |

These edit checks were processed as the data was being entered. When an edit check was tripped, the person doing the entry was faced with the following;

The individual could then respond by either of the following:

a. The value entered is correct regardless of the edit check

b. I will check the value and enter it later

If there is no value, the individual can respond similarly:

a. There will never be a value – e.g. the test was not performed

b. I will enter the value later

For study closeout and database locking, it is important to understand that the record in the database is signed *after* the edit checks have been processed and the person doing the entry has been notified and been given a chance to respond.

The edit checks programmed into the Data Entry System will be rerun on all patient data. Reports will be generated by the monitors that display the results of the edit checks by site and by patient. The monitors will review the listings and address any incomplete records.

Because of the size of the study, the amount of data, and the number of edit checks (500), the edit checks were prioritized in to 4 different “types”. There was a great deal of data collected on each patient during the study and some data was more important than others. For example, there was a series of “Quality of Life” questions that were very general and difficult to really edit. One question was “How many cigarettes do you smoke each day?” Most of these were considered type 4. Other questions such as the patient’s age, name or the visit name, were considered type 1, where type 1 was rated as most important.

Then to start the monitoring of the data, reports were generated that showed where there were potential problems by error type. In other words, were there regions where there were a large number of type 1 edit checks tripped? Then more detailed reports were generated to follow up on where the problem might be.

The following table represents some of the data cleaning and coding that was to be performed on the data in the data entry database.

| **No.** | **Edit Check** | **Response Description** |
| --- | --- | --- |
| 1 | Unsigned Records | Only records that have been signed by the investigator in the Country 1 and the Monitor in Country 2 will be used in the final database. |
| 2 | Incomplete Fields  (Signed Records) | The data is to be reviewed to see if records that have “Incomplete” values can be completed. |
| 3 | Questions of Visit Names/Dates | Errors in Visits names and dates are to be queried and corrected. There should be no erroneous Visit Names or Dates.  Patient status should be known at each planned visit. When the status of a patient is not known, extra steps must be taken to determine the status. |
| 4 | Range edit checks | In addition, certain extreme values for some of the data are to be queried against the values in the table below.  Where appropriate, the limits used for the range checks should be the same as those used by the RDE system. *Additional* “impossible” values may be used if necessary, such as 0 for a blood pressure. |
| 5 | AE is Causally Related | AE was caused by study drug and was serious enough to lead to discontinuation of study medication. |
| 6 | AE Terms (Causally related). | All “Causally-related” AE terms are to be coded and queried.  Queries will come from the coding staff at XXX.  (Procedure to be determined this fall) |
| 7 | AE Terms (Non-Causally Related) | All Non-serious AE terms are coded automatically; those that do not auto-encode will be coded manually. Those that cannot be coded manually, will be reviewed at XXX. Where they still could not be coded, the code will be set to missing. That is, no querying of Non-serious AE terms will be done. |
| 8 | Concomitant Medications | All Concomitant Medications are coded automatically; those that do not auto-encode will be coded manually. Those that cannot be coded manually, will be reviewed at XXX. Where they still could not be coded, the code will be set to missing. That is, no querying of Concomitant Medications will be done.  Medication Dates and Doses? |

**Edit checks:**

| **Check** | **Limits** |
| --- | --- |
| 1. Blood Pressure | As in RDE System |
| 2. SBP | Value of 0 |
| 3. DBP | Value of 0 |
| 4. HR | <30 or >160 |
| 5. Weight | <30 or >200 |
| 6. Height | <100 or >230 (Same as RDE System) |
| 7. Number of cigarettes per week | >500 (Same as RDE System) |
| 8. Total Cholesterol | <2 (Same as RDE System) |
| 9. HDL Cholesterol | <0.2 or >3 (Same as RDE System) |
| 10. Ratio: Total Cholesterol / HDL Cholesterol | <0.5 or >20 (Same as RDE System) |
| 11.The maximum difference between the 3 SBP measurements at the same visit | Δ SBP > 70 |
| 12.The maximum difference between the 3 DBP measurements at the same visit | Δ DBP> 60 |
| 13.The maximum difference between the 3 HR measurements at the same visit | Δ HR> 50 |
| 14.The maximum difference between SBP measurements for each patient over all visits | Δ SBP > 140 |
| 15.The maximum difference between DBP measurements for each patient over all visits | Δ DBP> 100 |
| 16.The maximum difference between HR measurements for each patient over all visits | Δ HR > 80 |
| 17. Weight change max-min  over all visits | Δ Weight > 70 |
| 18.The difference between a specific SBP & DBP measurement at the same visit | < 10 or > 160 |
| 19A. Number of days between predicted date of visit and date of visit | 6 Weeks: ± 6 weeks  3 Month: -6 weeks or + 3 months  6 Months: - 3 months or + 6 months  Other visits: ± 6 months |

## Equipment

Since virtually all of the data above requires some kind of instrument to make the observation, there should be procedures for the following;

1. Training for the operator.
2. Power replacement for batteries
3. Calibration procedures if necessary
4. Supplies such as ink or paper if that is necessary

These things can all be documented with a simple log near the instrument.

As the study progresses reports by site will be produced for the monitors to follow up on.

## Audit or Pre-Study Site Visits (PSSV)

It will likely be necessary to do a Pre-Study Site Visit or conduct an audit before the study actually starts. According to the ICH guidelines on investigator selection (5.6) the sponsor is responsible for selecting qualified investigators. This should include a documented review of qualifications.

# Study Execution

As the study is executed the results of the edit checks should be tracked and corrections made wherever possible, particularly early in the study.

This might take the form of an audit of the study progress. This is particularly true if the study is being conducted by an outside organization such as a CRO.

There will also need to be continued training as new people are involved in the study.

There will need to be some form of “Change Control” procedures as things change during the study. It is not unusual after the study starts to find things that are not working properly and they need to be changed. It might also be that a regulation would change and adjustments would need to be made. It is critical to make sure the changes cause minimal impact on the quality and integrity of the data. The data entry system should have written procedures for managing any changes to the system to ensure the accuracy and integrity of the data is maintained during these changes.

When the data is being entered into the computer during the visit instead of using paper and entering the data after the study is complete, it should be possible to produce reports that show the current status of the edit checks.

For example, in this example the entry screens can be grouped based on their use during the study. The following report shows the number of visits with Errors, the number of visits that are incomplete, and the number of visits that are missing for the Screening visits, the Randomization Visits, and the Adverse event Visits.

The following report is one of a series of reports designed to be used to indicate where there might be problems managing the data. This report shows by country and edit check type, how many edit checks were tripped. Obviously, more detailed reports were necessary to identify any specific problems.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Errors** | | |  | | | | | | |  | | |  | |  | | |  |  | |  |  |  |  | |  | | | |  | | |  |  | | |  | | | |  | |
|  | **Screening** | | | | | **Randomization** | | | | | | | | **Adverse Event** | | | | | | | |
|  | **Country 1** | | | **Country 2** | | **Country 1** | | | **Country 2** | | | | | **Country 1** | | | | **Country 2** | | | |
| **No Visits** | **9096** | | | **10146** | | **9096** | | | **10146** | | | | | **87368** | | | | **48840** | | | |
| Type 1 | 23 | 0.25% | | 99 | 0.98% | 17 | 0.19% | | 178 | | 1.75% | | | 21 | 0.02% | | 19 | | | 0.04% | |
| Type 2 | 6 | 0.07% | | 83 | 0.82% | 0 | 0.00% | | 0 | | 0.00% | | | 0 | 0.00% | | | 0 | | 0.00% | |
| Total 1+2 | 29 | 0.32% | | 182 | 1.79% | 17 | 0.19% | | 178 | | 1.75% | | | 21 | 0.02% | | | 19 | | 0.04% | |
| Type 3 | 17 | 0.19% | | 74 | 0.73% | 19 | 0.21% | | 138 | | 1.36% | | | 58 | 0.07% | | | 257 | | 0.53% | |
| Type 4 | 275 | 3.02% | | 588 | 5.80% | 12 | 0.13% | | 106 | | 1.04% | | | 0 | 0.00% | | | 0 | | 0.00% | |
| Total | 321 | 3.53% | | 844 | 8.32% | 48 | 0.53% | | 422 | | 4.16% | | | 79 | 0.09% | | | 276 | | 0.57% | |
|  | | |  | | | | |  | |  | |  | | |  | | |  | |  | |  | | | | |  | |  | |  |  | | | |  | | |  | | | | |  | |  | |
| **Incomplete** | | | | | | | |  | |  | |  | | |  | | |  | |  | |  | | |  | | |  | | | | |  | |  | | |  | |  | |  | | | | |  | |
|  | **Screening** | | | | | **Randomization** | | | | | | | | **Adverse Event** | | | | | | | |
|  | **Country 1** | | | **Country 2** | | **Country 1** | | | **Country 2** | | | | | **Country 1** | | | | **Country 2** | | | |
| **No Visits** | **9096** | | | **10146** | | **9096** | | | **10146** | | | | | **87368** | | | | **48840** | | | |
| Type 1 | 0 | 0.00% | | 0 | 0.00% | 0 | 0.00% | | 23 | | 0.23% | | | 671 | 0.77% | | | 149 | | 0.31% | |
| Type 2 | 1 | 0.01% | | 0 | 0.00% | 0 | 0.00% | | 17 | | 0.17% | | | 0 | 0.00% | | | 0 | | 0.00% | |
| Total 1+2 | 1 | 0.01% | | 0 | 0.00% | 0 | 0.00% | | 40 | | 0.39% | | | 671 | 0.77% | | | 149 | | 0.31% | |
| Type 3 | 1 | 0.01% | | 15 | 0.15% | 0 | 0.00% | | 2 | | 0.02% | | | 0 | 0.00% | | | 0 | | 0.00% | |
| Type 4 | 4 | 0.04% | | 14 | 0.14% | 0 | 0.00% | | 264 | | 2.60% | | | 72 | 0.08% | | | 11 | | 0.02% | |
| Total | 6 | 0.07% | | 29 | 0.29% | 0 | 0.00% | | 306 | | 3.02% | | | 743 | 0.85% | | | 160 | | 0.33% | |
|  | | |  | | | | |  | |  | |  | | |  | | |  | |  | |  | | |  | | |  | | | | |  | |  | | |  | |  | | | | |  | |  | |
| **Missing** | | |  | | | | |  | |  | |  | | |  | | |  | |  | |  | | |  | | |  | | | | |  | |  | | |  | |  | | | | |  | |  | |
|  | **Screening** | | | | | **Randomization** | | | | | | | | **Adverse Event** | | | | | | | |
|  | **Country 1** | | | **Country 2** | | **Country 1** | | | **Country 2** | | | | | **Country 1** | | | | **Country 2** | | | |
| **No Visits** | **9096** | | | **10146** | | **9096** | | | **10146** | | | | | **87368** | | | | **48840** | | | |
| Type 1 | 0 | 0.00% | | 3 | 0.03% | 0 | 0.00% | | 25 | | 0.25% | | | 2194 | | 2.51% | | 2956 | | 6.05% | |
| Type 2 | 4 | 0.04% | | 53 | 0.52% | 0 | 0.00% | | 7 | | 0.07% | | | 0 | | 0.00% | | 0 | | 0.00% | |
| Total 1+2 | 4 | 0.04% | | 56 | 0.55% | 0 | 0.00% | | 32 | | 0.32% | | | 2194 | | 2.51% | | 2956 | | 6.05% | |
| Type 3 | 1 | 0.01% | | 16 | 0.16% | 0 | 0.00% | | 5 | | 0.05% | | | 0 | | 0.00% | | 0 | | 0.00% | |
| Type 4 | 1 | 0.01% | | 67 | 0.66% | 3 | 0.03% | | 784 | | 7.73% | | | 150 | | 0.17% | | 760 | | 1.56% | |
| Total | 6 | 0.07% | | 139 | 1.37% | 3 | 0.03% | | 821 | | 8.09% | | | 2344 | | 2.68% | | 3716 | | 7.61% | |

These reports and the detail that generated them was given to the monitors (and sites) to be followed up on.

As a step in RBM, reports such as these are prepared and sent to the sites to follow up on before the monitor arrives for a visit.

Below is another example of two reports that shows status of data at a much lower level. It shows the results by edit check.

Edit checks for adverse event screen in Country 1 (87,368 Visits)

|  |  |  |
| --- | --- | --- |
| Error No. | No of Pat. | Error Description |
| 2 | 2 | Adverse Event is given and "No Adverse Event" box is ticked. |
| 6 | 26 | Date of onset is later than date of visit (1:st column) |
| 7 | 4 | Stop date should not be earlier than the date of onset (1:st column) |
| 10 | 14 | "No action" box is ticked and some of the other actions is also ticked (1:st column) |
| 12 | 21 | Duplicate entry of same Adverse Events with same onset date between visits (1:st column) |
| 18 | 8 | Duplicate entry of same Adverse Events with same onset date within this visit (1:st column) |
| 19 | 143 | Averse Event resolved at previous visit but Adverse Event Present at previous Evaluation ticked (1:st column) |
| 22 | 8 | The same Adverse Event given with resolved date on the first same as date of onset for the second but the severity has not changed |
| 48 | 4 | Other action (column 1) is not ticked but is specified |

**Edit checks for followup CRF in Country 2 (32605 visits)**

|  |  |  |
| --- | --- | --- |
| Error No. | No of Patients | Error Description |
| 3 | 5 | Date of this Follow-Up Visit is before Previous Visit |
| 11 | 4 | '1:st Systolic blood pressure' is outside limits 80-280 |
| 12 | 9 | '1:st Diastolic blood pressure' is outside limits 40-135 |
| 13 | 114 | '1:st Heart rate' is outside limits 40-150 |
| 14 | 4 | '2:st Systolic blood pressure' is outside limits 80-280 |
| 15 | 8 | '2:st Diastolic blood pressure' is outside limits 40-135 |
| 16 | 117 | '2:st Heart rate' is outside limits 40-150 |
| 18 | 7 | '3:st Diastolic blood pressure' is outside limits 40-135 |
| 19 | 130 | '3:st Heart rate' is outside limits 40-150 |
| 20 | 2 | 'Mean Systolic blood pressure' is outside limits 80-280 |
| 21 | 3 | 'Mean Diastolic blood pressure' is outside limits 40-135 |
| 22 | 87 | 'Mean Heart rate' is outside limits 40-150 |
| 23 | 3 | Difference between '1:st Systolic blood pressure' and '1:st Diastolic blood pressure' can not be less than 10 |
| 24 | 3 | Difference between '2:nd Systolic blood pressure' and '2:nd Diastolic blood pressure' can not be less than 10 |
| 25 | 3 | Difference between '3:rd Systolic blood pressure' and '3:rd Diastolic blood pressure' can not be less than 10 |
| 26 | 25 | Mean value of 'Systolic blood pressure' is not correct |
| 27 | 21 | Mean value of 'Diastolic blood pressure' is not correct |
| 28 | 28 | Mean value of 'Heart rate' is not correct |
| 37 | 1 | Has the patient been hospitalized since previous visit' is not Yes, but information about hospitalization is given |
| 40 | 4 | Stop date of hospitalization is after Date of visit.( 1:st line) |
| 66 | 2 | Weight outside limits 30-538 |
| 67 | 1 | Alcohol units outside limits |
| 71 | 1 | 'Smoking habits' is not previous smoker but stopped smoking since last smoke question is Yes |

It was also possible to prepare reports that show all “missing” data by subject for an individual site. Such a report was prepared for the study monitors for each of their sites so that when they went for a monitoring visit they would have a report to go over with the site staff.

How can the errors reported be corrected?

Do the errors indicate that the site is not performing the study properly?

The point being that during the study it is important to verify that the study is being conducted according to the specifications mentioned during the study start-up. It is not good to finish a study and then find out that something was done wrong. Experience would indicate that in clinical trials, where there is smoke, there is fire!!

# Study Close-out

During study close-out presumably all of the data has been entered. When paper CRFs were in use the CRFs were usually collected and sent to the study sponsor who entered all of the data and generated Data Clarification Forms (DCFs) that contained lists of any errors found during preparation for entry or during entry.

These forms were often returned by the monitor to the site for resolution which could take months.

When using a data entry computer system typically there is an attempt to get all edit checks performed during the study so after the last visit there are no more edits to perform. This can save months on a large study.

When all of the edits are resolved that can be resolved then it is possible to generate reports that list the errors in the data. These reports show the “Quality of the Data”. One golden rule of Quality Assurance is that the *Quality is measurable*.

The reports above could potentially be used although it is likely that ones with more detail will be used, based on the type of study.

The focus of the activities here has been

* + To give the highest quality treatment to the subjects enrolled in the study.
  + The information stored and managed in the computer during the trial will be used by the Statisticians to produce the required analyses.
  + The Clinical portion of the final report will also be based on much of the information stored in the database.

For Study Close-out the activities would be

| Activity | Description |
| --- | --- |
| Review Start-up Activities | Assure that the activities used to set up the study did not change or if they did change the changes were corrected properly.  It might be worthwhile to do a small internal audit of the procedures. |
| Prepare reports | Prepare reports that show the current status of edit checks to send to the sites for review and for the monitors to use to do final checks on some or all sites.  Consider using Risk Based Monitoring (RBM) to identify the particular edits checks that will be reported on. |
| Prepare reports | Prepare reports that show the current status of edit checks for the statisticians to use to prepare the Statistical Reports.  Missing data can weaken the strength of the statistical analyses in the study.  These same reports might be used for the Medical summary of the study. |
| Audit | Potentially conduct an audit of some of the sites or the site that is responsible for building the final database. |

The Study Quality

To measure the study quality more reports could be prepared that are similar to the ones above. For problems such as Missing Data or Out-of-Range data reports similar to the ones below could be prepared that would show where potential problems might occur.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Missing Data** | | |  | | | | | | |  | | |  | |  | |  |  | |  |  |  |  | |  | | | |  | | |  |  | | |  | | | |  | |
|  | **Screening** | | | | | **Randomization** | | | | | | | | **Adverse Event** | | | | | | |
|  | **Country 1** | | | **Country 2** | | **Country 1** | | | **Country 2** | | | | | **Country 1** | | | **Country 2** | | | |
| **No Visits** |  | | |  | |  | | |  | | | | |  | | |  | | | |
| Edit  Type 1 |  |  | |  |  |  |  | |  | |  | | |  |  |  | | |  | |
| Edit  Type 2 |  |  | |  |  |  |  | |  | |  | | |  |  | |  | |  | |
| Edit  Type 3 |  |  | |  |  |  |  | |  | |  | | |  |  | |  | |  | |
| Edit  Type 4 |  |  | |  |  |  |  | |  | |  | | |  |  | |  | |  | |
| Edit  Type 5 |  |  | |  |  |  |  | |  | |  | | |  |  | |  | |  | |
| Total |  |  | |  |  |  |  | |  | |  | | |  |  | |  | |  | |
|  | | |  | | | | |  | |  | |  | | |  | |  | |  | |  | | | | |  | |  | |  |  | | | |  | | |  | | | | |  |  | |
| **Out of Range** | | | | | | | |  | |  | |  | | |  | |  | |  | |  | | |  | | |  | | | | |  | |  | | |  | |  | |  | | | |  | |
|  | **Screening** | | | | | **Randomization** | | | | | | | | **Adverse Event** | | | | | | |
|  | **Country 1** | | | **Country 2** | | **Country 1** | | | **Country 2** | | | | | **Country 1** | | | **Country 2** | | | |
| **No Visits** |  | | |  | |  | | |  | | | | |  | | |  | | | |
| Edit  Type 1 |  |  | |  |  |  |  | |  | |  | | |  |  | |  | |  | |
| Edit  Type 2 |  |  | |  |  |  |  | |  | |  | | |  |  | |  | |  | |
| Edit  Type 3 |  |  | |  |  |  |  | |  | |  | | |  |  | |  | |  | |
| Edit  Type 4 |  |  | |  |  |  |  | |  | |  | | |  |  | |  | |  | |
| Ediit  Type 5 |  |  | |  |  |  |  | |  | |  | | |  |  | |  | |  | |
| Total |  |  | |  |  |  |  | |  | |  | | |  |  | |  | |  | |

Given that all edits are resolved then the next step is generally to create the analysis database. Usually a SAS Dataset.

The statisticians will want to see reports showing any errors in the data. They will have to deal with them in their analysis.

Obviously there are a lot of different reports that could be done to show the nature and magnitude of any errors or discrepancies in the final database.

If these reports were run at regular intervals during the study it would be possible to see if the number of errors increased or decreased during the study. Hopefully it could be shown that problems decreased as the study progressed.

# In Summary

The goal of having high quality data is to;

* Support Biostatistics and their analyses of the data
* Assure Data Management that all study information is accurate,
* Medical Writing when they prepare the study summary, and
* eClinical Technology to take advantage of the latest tools,

It is important to emphasize that the use of this data is to decide on patient safety, product efficacy, and the quality of the patient care experience. All three of these aspects will define the medical viability of a product in the market at large over time.

As was stated, this data is used in the statistics analyses that appear in the research summary. The fact is that the strength of the statistics will be a function of many of the numbers displayed above. The amount of missing data, data that is out of range, or whatever the problem might be can lessen what we know of the study compound and practice.

As the product goes from controlled trial conditions to general medical practice one key need is for accurate, trustworthy data as the foundation for successful product launch into much broader use.

The following tables summarize what goes on in each of the three steps in a clinical trial.

Study Start up

| Activity | Description |
| --- | --- |
| Study Preparation | Obtain IRB approval to conduct the study. |
| Study Protocol | Describes the data to be collected and its use. |
| Data Management Plan | Describes all data fields to be collected and how it will be managed, that is, entered, corrected, documented, and approved. This should also include a detailed list of all edit checks to be performed on the data.  This should also include drug randomization for a blinded study and patient consent. |
| Statistical Analysis Plan | Describes how all of the data will be analyzed. This includes how data will be handled that is missing or looks erroneous. |
| Instrumentation | Assure that all instrumentation and its use is properly managed  Assure that all equipment and instrumentation is adequately maintained. This include things like batteries and supplies. |
| Training | Train all users on the equipment and instrumentation they will use.  Train all users on how to enter the data and carry out any entry instructions included in the DMP.  Document the training. |
| Audit | Do a Pre-Study Site Visit or conduct an audit before the study actually starts. According to the ICH guidelines on investigator selection (5.6) the sponsor is responsible for selecting qualified investigators. This should include a documented review of qualifications. |

Study Execution

| Activity | Description |
| --- | --- |
| Monitor Start-up Activities | Assure that the activities used to set up the study do not change or if they do change the changes are correct and are done properly.  Verify   * Documentation * Instrumentation * Training * Randomization |
| Prepare reports | Prepare reports for the monitors to use on their visits to the sites that show the current status of edit checks.  Consider using Risk Based Monitoring (RBM) to identify the particular edits checks that will be reported on. |
| Review Edit Checks | Periodically review the data quality with others on the team based on the edit checks listed above. Ask if there are any changes that need to be made to improve the quality of the data. |
| Audit | Conduct whatever audits might be necessary to document the accuracy of the work being performed. |

Study Close-out

| Activity | Description |
| --- | --- |
| Review Start-up Activities | Assure that the activities used to set up the study did not change or if they did change the changes were corrected properly.  It might be worthwhile to do a small internal audit of the procedures. |
| Prepare reports | Prepare reports that show the current status of edit checks for the monitors to use to do final checks on some or all sites.  Consider using Risk Based Monitoring (RBM) to identify the particular edits checks that will be reported on. |
| Prepare reports | Prepare reports that show the current status of edit checks for the statisticians to use to prepare the Statistical Reports.  Missing data can weaken the strength of the statistical analyses in the study.  These same reports might be used for the Medical summary of the study. |
| Audit | Potentially conduct an audit of some of the sites or the site that is responsible for building the final database. |

**In General**

The process is the following;

1. Prepare by writing down procedures
   1. what the data to be collected is and its specifications.
   2. what the site is supposed to do to see the subject
   3. what the instrumentation to be used is and is it accurate
   4. Training everyone receives to be able to do their tasks
2. Perform the tasks described above, leaving a record
3. Periodically review what is being done and correct any deficiencies
4. Prepare final reports showing anything that might be missing or not consistent with the instructions above (Study Quality).
5. Prepare final study reports based on the final data

# References

1. ICH E6 – Guidance for Industry – E6 Good Clinical Practice: Consolidated Guidance
2. DMP Template
3. Risk Based Monitoring - Guidance for Industry - Oversight of Clinical Investigations — A Risk-Based Approach to Monitoring