Data Lifecycle Plans

The Structure

Kit Howard, Kestrel

DLP Contents

- Purpose
- Regulatory Considerations
- Protocol Definition
- Statistical Analysis
- Displays
  - Patient Listings
  - Summaries
  - Graphics
- Data Capture
- Data Quality
  - Monitoring
  - Computer Checks
  - Manual Checks
- Database Structure
  - Input
  - Analysis
- References

DLP Purpose

- Defines the reason for collecting the data
- Should be succinct, a **brief** idea of what is being collected and why the data would be useful
- Also lists any costs or restrictions on the data, e.g., copyrighted instrument, or not validated

DLP Regulatory Considerations

- Defines if the data required by regulation
- What regulations or guidelines specifically apply to these data
- What role do the data typically play in a submission

DLP Protocol Definition

- Lays out the information that the protocol will contain pertaining to this DLP
- Can include specific wording for a protocol or a description of the points that should be in the protocol
- If aspects of the procedure can vary from study to study (e.g., timing of the procedure), it can be defined here

DLP Statistical Analysis

- Describes generally how the data are to be used in statistical analyses
  - what types of analyses are appropriate for these data
  - whether there are specific assumptions that must be met (e.g., normality, no missing data)
  - the role the data play in analysis (e.g., are they themselves analysis variables, or are they used as covariates in statistical models, etc.).
- Literature references for analytical methods can be here, or in the References section

DLP Displays – Patient Listings

- Describes how the data will be presented in patient listings
- Provides layout examples
- These designs will work best with DLPs that define a coherent set of data (such as a psychiatric test)

DLP Displays – Summary Tables

- Describes how the data will be summarized
- Provides layout examples
- Assumptions and calculations generally required by these summaries should be defined, e.g.,
  - how will missing data will be handled
  - how each N is defined

DLP Data Capture

- Contains a representation of the recommended layout for the CRF module(s) collecting the data contained in this DLP
- It may contain multiple layouts for paper or electronic capture

DLP Data Quality - Monitoring

- Defines the monitoring guidelines
- Defines source documentation
- Lists data consistency checks within or between data points for which the site monitor will be responsible
- States the rationale for why the check should be performed at the site as opposed to having it done elsewhere in the process

DLP Data Quality – Computer Checks

- Defines the computer logic checks used to verify the data
- Contains a tabular presentation of the standard computerized checks
- Definitions are programming language- and structure-independent, allowing for adaptation to any system
- Study-specific checks must be documented in the Project Manual that is created for each study

DLP Data Quality – Manual Checks

- Defines any additional data quality checks that are performed
- May include instructions for how to look for outliers, or comparisons usually made to other data

DLP Database Structure - Input

- Includes an annotated version of the data collection form
- Identifies the fields to be databased
- Showing the field names
- Contains pointers to any standard field naming conventions for the organization, as well as any other structural information of which the database programmers should be aware

DLP Database Structure - Analysis

- Presents the technical details for all analysis datasets
- Includes the source of the variables and what transformations occurred to result in this structure
- May be actual datasets, or they may be views
- Dataset/view names are defined, along with variable names and characteristics
- Defines any calculations needed to complement the fields being collected

DLP References

- Lists all relevant references
  - Regulatory
  - Statistical
  - Medical
  - Technical
- This section effectively defines what knowledge was gathered from other sources vs. generated internally

Odd Order of Sections?

- Purpose
- Regulatory Considerations
- Protocol Definition
- Statistical Analysis
- Displays
  - Patient Listings
  - Summaries
  - Graphics
- Data Capture
- Data Quality
  - Monitoring
  - Computer Checks
  - Manual Checks
- Database Structure
  - Input
  - Analysis
- References

Begin with the end in mind!