



Introduction to study documents

Release Date: DD-Mmm-YYYY

Study Planning

- Potential issues with clinical trials
 - Bias
 - Confounding
- Control bias and confounding
 - Concurrent controls
 - Randomization
 - Blinding
- Study Designs
- Type I and Type II errors
- Sample Size and Power

Study Planning: Bias

- Bias: Any procedure which tends to produce results or conclusions that differ systematically from truth
- Caused by prejudice in the planning, conduct, analysis, or interpretation of clinical trials
(e.g., subject selection
allocation of subjects to treatments
measurement/evaluation of response
knowledge of measurements)

Study Planning: Bias

Example: Consider a single blinded trial of a cholesterol reducing drug

- The endpoint is mean percent reduction in LDL cholesterol from baseline
- Since the investigator knows the treatments of patients, he/she may judge those patients on active drug to have more drug-related side effects.
- Intentionally or unintentionally, the investigator has introduced bias into the study.

Pharma Academy of Learning and Sharing

Study Planning: Confounding

- Definition: The effect of the treatment cannot be definitively distinguished from an effect that may have been caused by other factors.
- Confounding is possible, if:
 - Factor has an effect on the outcome
 - AND
 - Distribution of patients over the treatment groups is not the same for each level of the factor ("stratum")

Study Planning: Confounding

- **Example:** NIH funded stroke study comparing treatment A to treatment B in preventing stroke.

Outcome	Trt A	Trt B	Total
Favorable	40 (40%)	40 (40%)	80
Not Favorable	60 (60%)	60 (60%)	120
Total	100	100	200

Study Planning: Confounding

Example (cont.)

- NIH score ≤ 10

Outcome	Trt A	Trt B	Total
Favorable	25 (62.5%)	15 (75%)	40
Not Favorable	15 (37.5%)	5 (25%)	20
Total	40	20	60

- NIH score >10

Outcome	Trt A	Trt B	Total
Favorable	15 (25%)	25 (31.25%)	40
Not Favorable	45 (75%)	55 (68.75%)	100
Total	60	80	140

Study Planning: Confounding

Example (cont.):

- What appears to be no treatment effect is really an effect, but is lost by failing to stratify on NIH score.
- Those receiving Trt B do better than those on Trt A within each NIH stroke group, but unequal A:B sample size ratio creates impression of no difference.

Study Planning: Controlling Bias and Confounding

Ways to control Bias and Confounding:

- Concurrent controls
- Blinding
- Randomization



Study Planning: Concurrent Controls

- Control group receives control treatment at about the same time as experimental treatment group
 - Prevents certain types of confounding (e.g., confounding due to time)
- Sets a benchmark response or provides a comparator for the response in the experimental treatment group

Study Planning: Concurrent Controls

- Placebo control
 - 'Gold standard' for efficacy for a new indication
 - Easy interpretation, basis for drug registration
- Active control (established treatment)
 - Placebo not ethically acceptable
 - Efficacy well established in placebo-controlled trials

Study Planning: Blinding

- Blinding or masking is intended to limit the occurrence of conscious and unconscious bias in the conduct and interpretation of a clinical trial arising from the influence which the knowledge of treatment may have on:
 - Recruitment and allocation of subjects and their subsequent care
 - Attitudes of subjects to the treatments
 - Assessment of endpoints,
 - Handling of withdrawals, the exclusion of data from analysis, etc.
- The essential aim is to prevent identification of the treatments until all such opportunities for bias have passed.

Study Planning: Blinding

- Single Blinding

- The subject does not know which treatment he/she was given. The investigator knows.
- Possible bias in data collection and assessment.

- Double Blinding

- The investigator and subject do not know which treatment the subject was given
- This level of blinding is maintained throughout the conduct of the trial
- When the data are cleaned to an acceptable level of quality, will appropriate personnel be unblinded
- Sometimes bioanalytical scientist, auditors or personnel handling SAE reporting may need to unblind subject
- The sponsor should have SOP to guard against inappropriate dissemination of treatment codes.

Study Planning: Blinding

- Difficulties in achieving the double-blind ideal can arise:
 - Treatment may be of a completely different nature, e.g., surgery and drug therapy
 - Two drugs may have different formulations – partial solution to this “Double Dummy” technique
- Double-blind nature may be partially compromised by apparent treatment induced effect. For example, immunosuppressive drug may reduce the lymphocyte counts dramatically
- If a double-blind trial is not feasible, then the single-blind option should be considered
- In some cases only an open-label trial is practically or ethically possible.

Study Planning: Randomization

- Randomization introduces a deliberate element of chance into treatment allocation
- **Purpose**: to balance groups on both observed and unobserved factors
- Measured and unknown prognostic factors will balance out among the treatment groups
- Any observed effect in the response can be attributed to the treatment because the only systematic difference in the groups is the treatment
- In multicenter trials, a separate random scheme for each center should be available, i.e. to stratify by center or to allocate several whole blocks to each center.

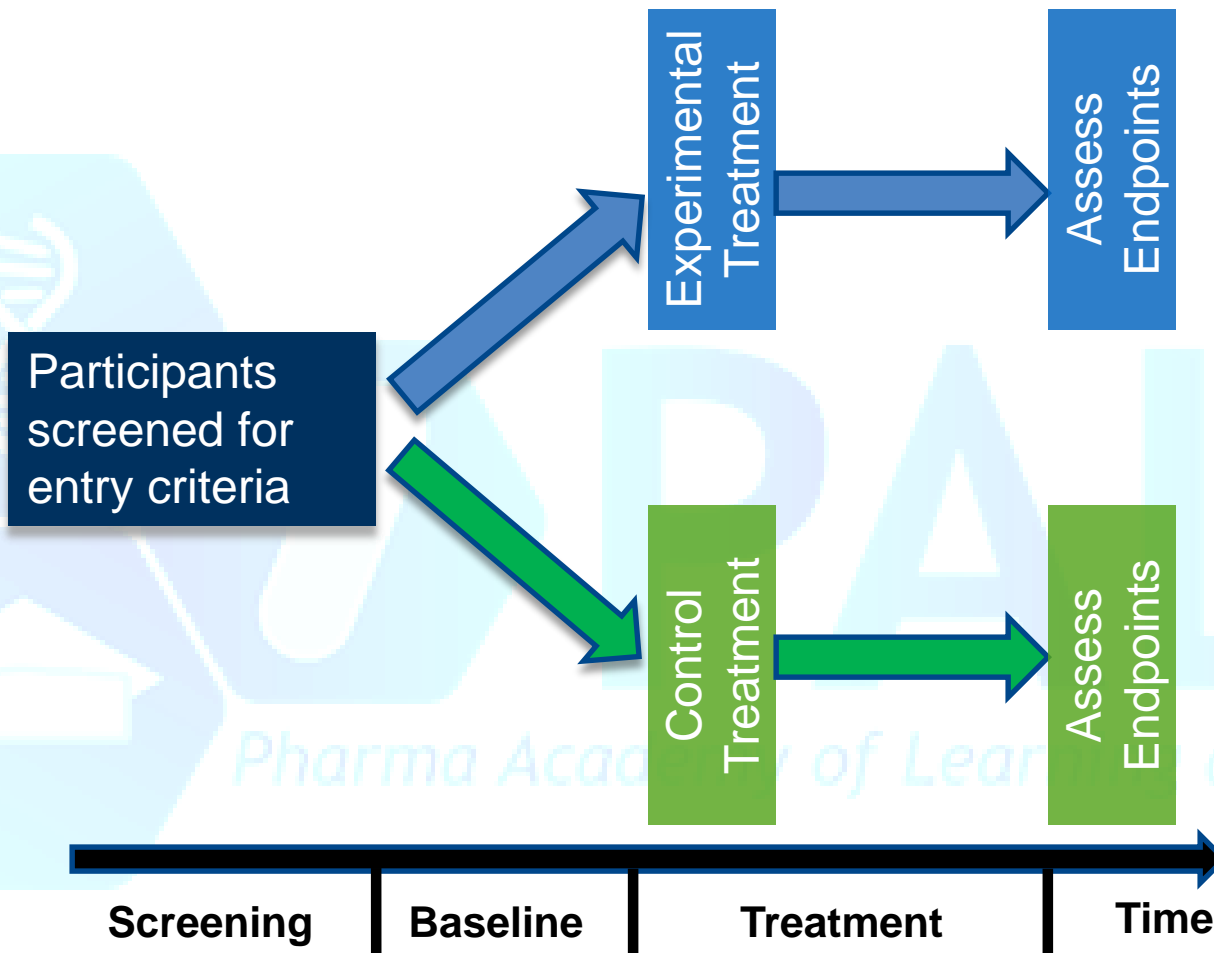
Study Planning: Randomization

- Stratification by important prognostic factors measured at baseline (e.g. severity of disease, age, sex, etc.) may sometimes be valuable to promote balanced allocation within strata
- The use of more than two or three stratification factors is rarely necessary, is less successful at achieving balance and is logistically troublesome
- A dynamic allocation procedure may help to achieve balance across a number of stratification factors
- Factors on which randomization has been stratified should be accounted for later in the analysis.

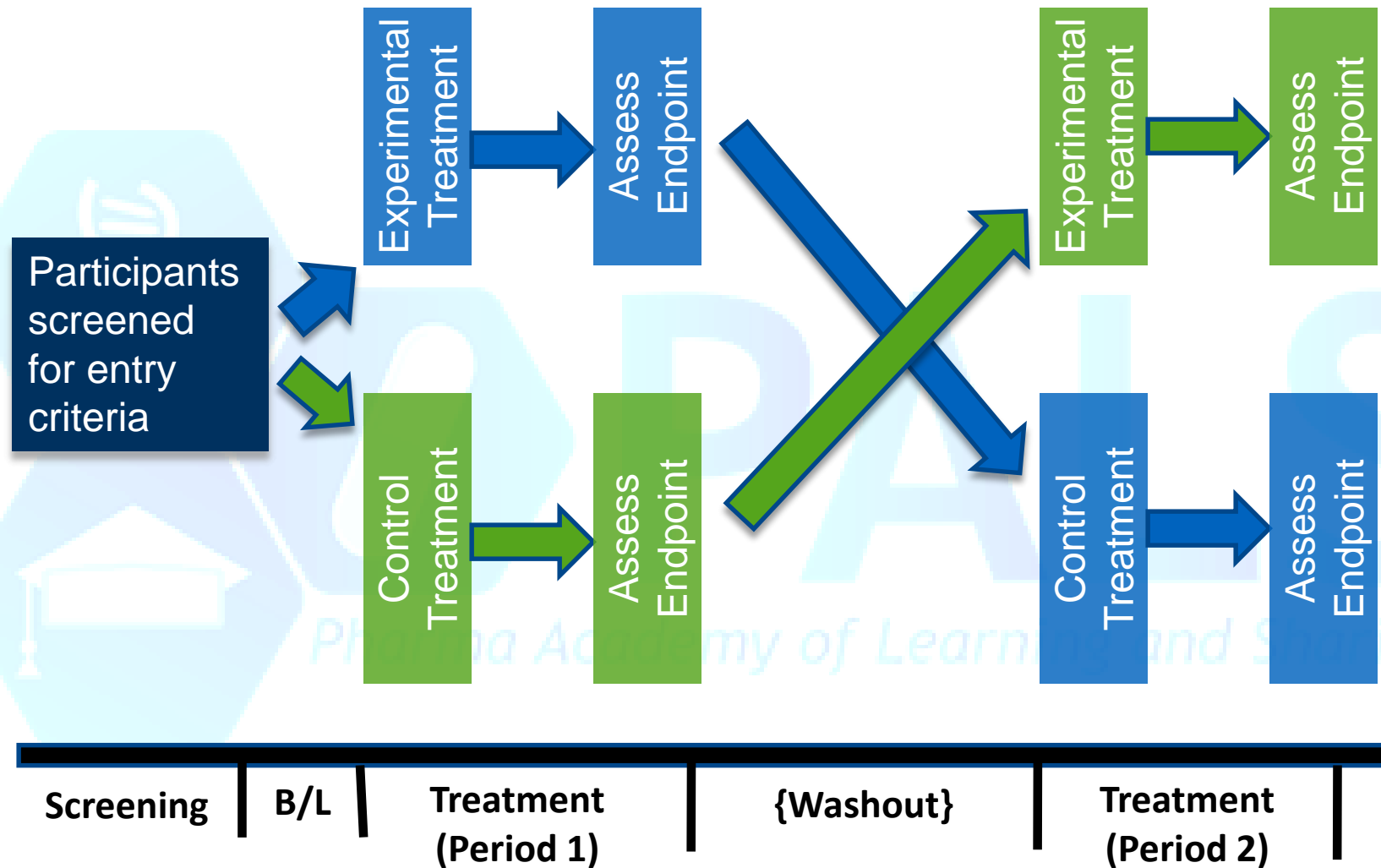
Study Planning: Study Designs

- Select a study design which will allow us to draw valid statistical conclusions from the data (i.e., avoids bias and confounding)
- Two common study designs are
 - Parallel
 - Crossover

Study Planning: Parallel Group Designs

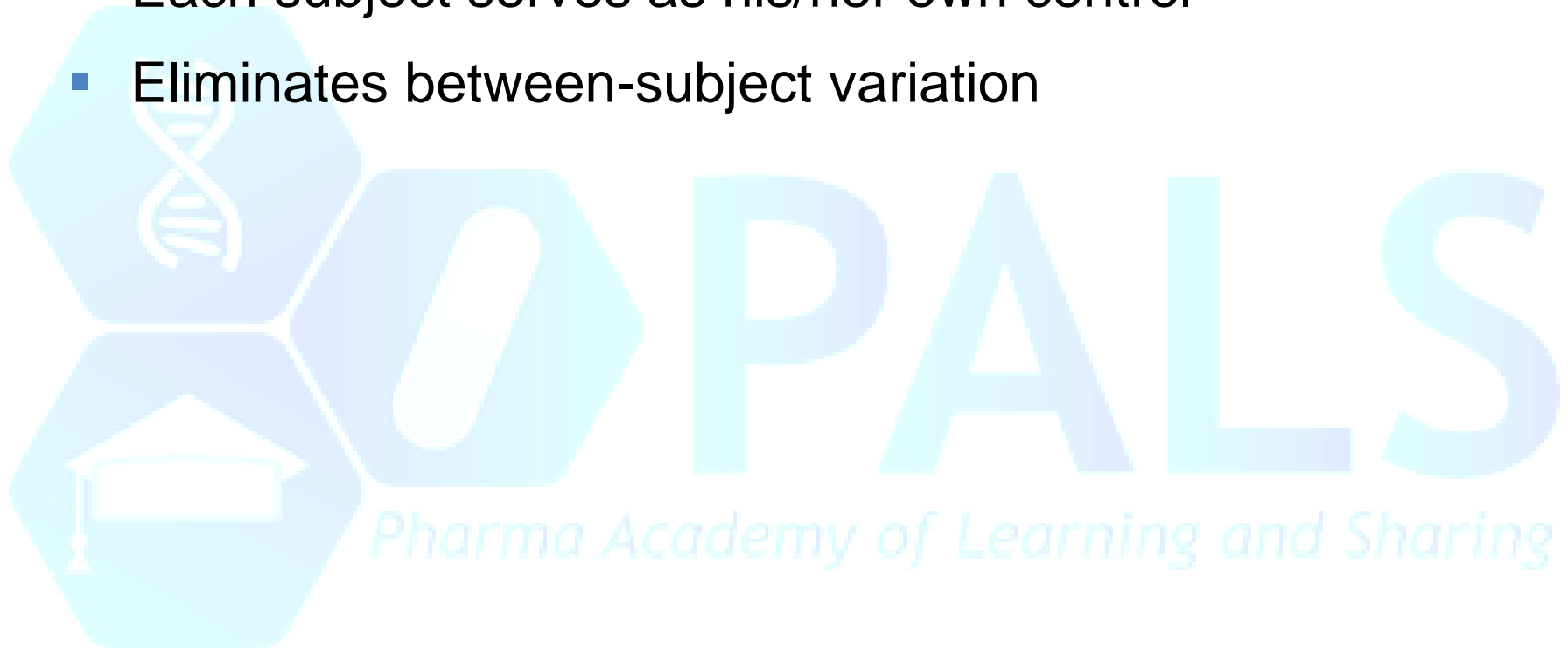


Study Planning: Crossover Designs



Study Planning: Crossover Designs

- Each subject is randomized to a sequence of treatments with a wash-out period between treatments
- Each subject serves as his/her own control
- Eliminates between-subject variation



Study Planning: Crossover Designs

- Needs smaller sample size than parallel designs
- Wash-out period needs to be long enough so that there is no carryover effect
- Not appropriate when
 - Treatment period is long
 - Treatment intended to be curative
 - Washout period is long (drugs with long half-life)
 - Carryover exists

Study documents

- Protocol
- Case Report Forms (CRFs)
- CRF Completion Guideline
- Annotated CRFs
- Statistical Analysis Plan
- Database Management
- Specifications
- Mock Shells
- ADS
- TLFs
- Report/Manuscript

Protocol

- The study protocol is the blueprint that all researchers follow.
- A study protocol is a document that describes, in detail, the plan for conducting the clinical study. The study protocol explains the purpose and function of the study as well as how to carry it out.
- Typically Protocol includes following:
 - ✓ The reason for the study
 - ✓ The number of participants
 - ✓ Eligibility and exclusion criteria
 - ✓ Details of the intervention or therapy the participants will receive (such as frequency and dosages),
 - ✓ What data will be gathered
 - ✓ What demographic information about the participants will be gathered
 - ✓ Steps for clinical caregivers to carry out
 - ✓ Study endpoints.

Contents of Protocol (as per ICH guidelines)

General Information	
Background Information	
Trial Objectives and Purpose	
Trial Design	
Selection and Withdrawal of Subjects	
Treatment of Subjects	
Assessment of Efficacy	
Assessment of Safety	
Statistics	
Direct Access to Source Data/Documents	
Quality Control and Quality Assurance	
Ethics	
Data Handling and Recordkeeping	
Financing and Insurance	
Publication Policy	
Supplements	

Protocol Development

- Objectives
- Definition of Endpoints
- Study Design
- Sample size
- Analysis Plan

PI



BS

PI: Principal/Clinical Investigator; BS: Study Biostatistician

Biostatistician – Protocol Writing

- Objectives (input)
- Endpoints (input)
- Study Design (input)
- Randomization procedures (writes)
- Allocation concealment (input)
- Sample size (writes)
- Analysis plan (writes)

Biostatistician – Protocol Review

- It is highly recommended that the Biostatistician reviews the full protocol prior to final review/sign-off:
 - Clarity
 - Completeness
 - Consistency
 - Data quality issues
 - Feasibility

Biostatistician- Objectives and endpoints

Based on the objective/clinical research question

- Valuable information on key design inputs
- Clear specification of hypotheses to be tested (or parameters to be estimated)
- Selection and definition of endpoints (Primary outcome variable)

Pharma Academy of Learning and Sharing

Biostatistician- Study Design

- Is the study design appropriate to provide the data needed to answer the objectives?
- Eligibility criteria
- Selection/recruitment of Participants:
 - Define procedures for minimizing selection bias
 - If an RCT, define randomization procedures (sequence generation, and allocation concealment), blinding
- Length of follow up and frequency of contacts

Biostatistician- Sample Size

- Justification in terms of power or precision for the primary endpoint
- Method used to calculate the sample size
 - Should be consistent with the primary method for data analysis, and appropriate for the study design
- Historical data to support the assumptions
- Justification in terms of feasibility

Statistical Programmers - Read and Understand study Protocol

- ✓ Protocol Synopsis
- ✓ Study Design
- ✓ Inclusion and Exclusion criteria
- ✓ Schedule of assessment
- ✓ Study Flow Chart
- ✓ Endpoints
- ✓ Statistical Methods

Game

- Divide the team into groups and ask them to read the study protocol and make presentation/questions on
 - (1) Study Design
 - (2) Inclusion Exclusion Criteria
 - (3) Primary/ secondary endpoints
 - (4) Visit schedule
 - (5) Statistical methods

Study documents

- Protocol
- Case Report Forms (CRFs)
- CRF Completion Guideline
- Annotated CRFs
- Statistical Analysis Plan
- Database Management
- Specifications
- Mock Shells
- ADS
- TLFs
- Report/Manuscript

Case Report Forms

- A case report form (CRF) is a data collection tool used in clinical trials to support investigators and coordinators in capturing all protocol-required information.
- The International Conference on Harmonization Guidelines for Good Clinical Practice define the CRF as: A printed, optical or electronic document designed to record all of the protocol – required information to be reported to the sponsor on each trial subject.
- A well-designed CRF should represent the essential contents of the study protocol and in an ideal situation, CRF is designed once the study protocol is finalized.
- The data from the CRFs is compiled as a dataset into a database and validated prior to statistical analysis.
- Paper CRF is the traditional way of data capture and a better option if studies are small or vary in design, whereas eCRFs are considered if studies are large with similar designs.
- eCRFs are preferred over paper CRFs as they are less time-consuming, and also encourage the sponsor/pharmaceutical company to carry out large multicentric studies at the same time due to the ease of administration.

Pharma Academy of Learning and Sharing

Examples of CRF

DEMOGRAPHY	
Date of birth (DD/MM/YYYY)	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>
Gender	Male <input type="checkbox"/> 1 Female <input type="checkbox"/> 2
Height (cm)	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>
Weight (kg)	<input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>
Smoker	Yes <input type="checkbox"/> 1 No <input type="checkbox"/> 2
Family history	Yes <input type="checkbox"/> 1 No <input type="checkbox"/> 2

Coding on the case report form module

ADVERSE EVENT		
Did any adverse event occur after the last visit?—Indicator Question		Yes <input type="checkbox"/> 1 No <input type="checkbox"/> 2
If "Yes", please provide the information in the adverse events (AE) page (page no. XX) and give details below:—Skip		
AE page number	AE serial number	Did any unscheduled visit happen after the last visit? (Please provide details on page no. YY)—Skip
<input type="text"/> <input type="text"/>	<input type="text"/> <input type="text"/>	Yes <input type="checkbox"/> 1 No <input type="checkbox"/> 2
<input type="text"/> <input type="text"/>	<input type="text"/> <input type="text"/>	Yes <input type="checkbox"/> 1 No <input type="checkbox"/> 2

Poorly designed v/s Well-designed CRFs

E.g.	Poorly Designed CRF Data Fields	Well-Designed CRF Data Fields												
1.	<div>Result</div> <div>Neutrophils <input type="text"/></div> <div>Lymphocytes <input type="text"/></div>	<table border="1"> <thead> <tr> <th></th> <th>Result</th> <th>Unit if different</th> </tr> </thead> <tbody> <tr> <td>Neutrophils (%)</td> <td></td> <td></td> </tr> <tr> <td>Lymphocytes (%)</td> <td></td> <td></td> </tr> </tbody> </table>		Result	Unit if different	Neutrophils (%)			Lymphocytes (%)					
	Result	Unit if different												
Neutrophils (%)														
Lymphocytes (%)														
2.	<div># Result</div> <div>1 <input type="text"/> Comment: _____</div> <div>2 <input type="text"/> _____</div> <div>3 <input type="text"/> _____</div>	<div># Result</div> <div>1 <input type="text"/> Comment: <div></div></div> <div>2 <input type="text"/></div> <div>3 <input type="text"/></div> <table border="1"> <thead> <tr> <th>#</th> <th>Result</th> <th>Comment</th> </tr> </thead> <tbody> <tr> <td>1</td> <td></td> <td></td> </tr> <tr> <td>2</td> <td></td> <td></td> </tr> <tr> <td>3</td> <td></td> <td></td> </tr> </tbody> </table>	#	Result	Comment	1			2			3		
#	Result	Comment												
1														
2														
3														

CRF completion Guideline

- The guidelines help to bridge the gap between the study protocol and the users in regards to CRF completion, correction, signing and handling procedures.
- Data formats for appropriate response fields, a data correction guide, how to handle unknown or unavailable data, and a retrieval schedule for completed CRFs can be outlined.
- instruct how to handle missing or unavailable data. If a required piece of information or entire section cannot be retrieved, the use of “NA”, “ND” or “UNK” can be defined to avoid ambiguous responses.
- The guidelines can also demonstrate how to complete CRF pages for unscheduled assessments or in the cumulative log.
- Furthermore, CRF retrieval procedures including how to handle CRFs for subjects who have discontinued during the study can be listed.

Annotated CRFs

- What This is a blank CRF annotations that document the location of the data with the corresponding names of the datasets and the names of those variables included in the submitted datasets
- Annotations are meant to help the FDA reviewer find the origin of data variables included in the submitted datasets
- Annotations should reflect the data submitted within the SDTM

Examples of Annotated CRF

Subject identification DOMAIN=[DM] [SC]		Ethnicity: ETHNIC <input type="checkbox"/> Hispanic or Latino <input type="checkbox"/> Not Hispanic or Latino	
Subject number: SUBJID	BIRTHDT	Race: <input type="checkbox"/> Asian <input type="checkbox"/> Black or African American <input type="checkbox"/> White <input type="checkbox"/> American Indian or Alaska Native <input type="checkbox"/> Native Hawaiian or Other Pacific Islander Other, specify: _____	RACE
Date of birth:	d d m m y y y y	SCORRES where SCTESTCD="RACEOTH"	
Gender: <input type="checkbox"/> male SEX <input checked="" type="checkbox"/> female			

Group: Physical examination: Base	
OID=IG_PE_BASE, Repeating=No	
Height VSORRES Alias: SDTM: VSORRES where VSTESTCD=HEIGHT	<input type="text"/> (in)
Weight VSORRES Alias: SDTM: VSORRES where VSTESTCD=WEIGHT	<input type="text"/> (lbs)
Systolic blood pressure VSORRES Alias: SDTM: VSORRES where VSTESTCD=SYSBP	<input type="text"/> (mm Hg)
Diastolic blood pressure VSORRES Alias: SDTM: VSORRES where VSTESTCD=DIABP	<input type="text"/> (mm Hg)
Blood pressure ratio @SDSVarName Not Set	<input type="text"/> . <input type="text"/>
Does the subject feel dizzy when standing up from a sitting position @SDSVarName Not Set	<input type="radio"/> No <input type="radio"/> Yes

Method: Systolic blood pressure divided by diastolic blood pressure
Skip condition: Needs not be collected when the diastolic blood pressure is 90 or higher

Statistical Programmers: CRF Design/CRF review/ Read and Understand CRF

- CRF/eCRF design
 - ✓ Since statistical SAS programmers have an in-depth knowledge of each data point utilized in the programming, their involvement can help to ensure that critical data points are collected efficiently via CRF/eCRF.
 - ✓ Example, concomitant therapy reporting, in order to differentiate a prior therapy from a concomitant therapy, either a date or a flag needs to be collected via CRF/eCRF. Without that, it will be extremely difficult to differentiate prior therapies from concomitant therapies in the programming.
- Read and understand CRF, CRF completion guidelines
 - ✓ Check if all required information is captured on CRF
 - ✓ Read CRF completion guideline to understand how data is captured on CRF
 - ✓ Try to visualize mapping of variables in different SDTMs

Statistical Programmers: Annotation of CRF/Review of CRF

- In addition to CRF/eCRF design, statistical SAS programmers are also often expected to play a significant role in mapping each and every CRF/eCRF data point in to the database.
- To perform this activity, they should be familiar with industry quality standards, guidelines and procedures. For example, a good knowledge of the Study Data Tabulation Model (SDTM), which defines a standard structure for study data tabulations that are to be submitted as part of a product application to a regulatory authority such as the United States Food and Drug Administration (FDA), is definitely helpful.
- How to Annotate?
 - ✓ Make a list of SDTMs/ Source datasets
 - ✓ For each variable captured on CRF page, try to map to a SDTM/Source data set
 - ✓ Create a excel sheet which will have a variable name from CRF and a destination dataset name

Exercise

- Annotation of CRF page or provide annotated CRF pages to understand the annotations



Study documents

- Protocol
- Case Report Forms (CRFs)
- CRF Completion Guideline
- Annotated CRFs
- **Statistical Analysis Plan**
- Database Management
- Specifications
- Mock Shells
- ADS
- TLFs
- Report/Manuscript

Statistical Analysis Plan

- A statistical analysis plan (SAP) describes the planned analysis for a clinical trial.
- Protocol outlines the analysis, SAP describes the statistical techniques for study analysis in detail.
- The SAP defines all the statistical output which will be included in the clinical study report.
- Shell tables, figures and sometimes listings are usually attached to the SAP although they should not be formally part of the SAP.
- Usually the SAP is written by the trial or project statistician by using a template.
- According to ICH E9 the statistical analysis is planned a priori.

Statistical Analysis Plan (Contd...)

- Clinical trial statistician needs to ensure that the SAP is carefully reviewed and approved prior to unblinding of the study.
- For open label studies the SAP should be reviewed and approved prior to database lock approval.
- The SAP is a document which is submitted to regulatory authorities as part of a submissions package.
- The SAP is part of the appendix of a clinical study report.
- The SAP is stored in the trial master file and it is used during audits to check if statistical programming followed exactly the descriptions in the SAP.
- Besides the technical statistical details it should contain brief descriptions and summaries of the protocol, rather than just referring to the protocol.

Contents of SAP

- Brief description of the study and purpose – e.g. description of conduct details which are important for analysis, study objectives and variables
- Statistical methods to be used – e.g. summary statistics of subject data (means, standard deviations, extreme values, counts with corresponding percents), statistical tests (analysis of variance, t-tests,)
- Description of analysis populations – e.g. safety set, per protocol set, full analysis set, Data handling rules – e.g. imputation rules, algorithms for derived variables
- Complete table of contents with all TFLs to be produced by the statistical programmer as attachment to the SAP text (i.e. not as an appendix or any other formal and official part of the SAP)
- Shell TFLs to be produced by the statistical programmer to define the layout of the TFLs as attachment to the SAP text (i.e. not as an appendix or any other formal and official part of the SAP)

Biostatistician - Statistical Analysis Plan

Summary

- Provides the statistical methodology for the assessment of the primary objective(s):
 - Statistical Hypotheses and testing procedures
 - Primary Analysis Population
 - Secondary analysis
 - Sensitivity analysis
- Missing data
- Multiplicity of testing
- Subgroup analyses, covariates
- Statistical method(s) for Secondary or Exploratory objectives(s)

Biostatistician- Statistical Analysis Plan Summary

- Discusses statistical methods to be used in planned interim analyses, if any (Role of a DSMB).
- Purposes:
 - Assuring objectives can be achieved
 - Justifying design and data collection
 - Justifying the statistical test and procedure

Biostatistician- Data Analysis - 1

- Write a detailed analysis plan:
 - Define Primary and Secondary endpoints
 - All hypothesis to be tested and testing procedures (or parameters to estimate)
 - Analysis populations
 - Interim analysis and adjustments to type I error
 - Hierarchy of analysis
 - Handling of missing data
 - Subgroup analysis
 - Exploratory analysis, if any

Biostatistician- Data Analysis - 2

Variable Types and Derivations to be described in the Analysis Plan

- Categorical (Binary or dichotomous)
 - Sex (Male, female), diabetes (type 1, type 2)
- Ordinal
 - Attitudes (strongly disagree, disagree, neutral, agree, strongly agree)
 - Cancer stage (I, II, III, IV)
- Numerical
 - Discrete (Number of abnormal cells)
 - Continuous (Age; Serum Creatinine level)
- Derived
 - Absolute % change, Time to response, AUC

Biostatistician- Data Analysis - 3

- Should be written prior to un-blinded review of the data or even prior to data collection (First Subject in):
 - Helps to prepare for report and manuscript writing
 - Helps in the validity and credibility of the results

Statistical Programmers- Read and Understand SAP

- ✓ Study Plan
- ✓ Study design and objectives
- ✓ Analysis Population
- ✓ Analysis of Endpoints
- ✓ Baseline Definitions
- ✓ Imputation rules , if any
- ✓ Data Handling rules , if any
- ✓ Statistical methods
- ✓ Interim Analysis, if applicable
- ✓ Database Lock

Game : Reading SAP

- Read study SAP and Discuss Key points



Study documents

- Protocol
- Case Report Forms (CRFs)
- CRF Completion Guideline
- Annotated CRFs
- Statistical Analysis Plan
- Database Management
- Specifications
- Mock Shells
- ADS
- TLFs
- Report/Manuscript

Database Management

- Biostatisticians and Statistical programmers are involved in providing inputs to Database creation
- In next few slides we will learn about different activities which they support



Biostatistician: Input to Data Management

- CRF Development:
 - Content
 - Design
- Dataset specification:
 - Annotation of CRFs
 - Record Layout
- Validation:
 - Error checking specification
 - Test data

Statistical Programmers: Database set up

SAS Programmers may be involved in Database set up activities to support Clinical Data Management team. Some of the activities can be :

- ✓ Edit checks programming
- ✓ Metadata Setup- Phase, treatment groups, Baseline flags etc
- ✓ Raising Data issues identified during programming

Statistical Programmers: Database UAT

- A good database is critical for final reporting.
- Statistical SAS programmers are often expected to be involved in the database UAT because of the close relationship between database quality and accurate reporting.
- Database UAT is done after the database design using mock data to check each and every data point is collected properly.
- A key data point missing or incorrect collection can have a negative effect on lots of other seemingly-unrelated data points and also negative effect the programming.
- For example, a missing study medication date can effect phasing. A good statistical SAS programmer can definitely contribute a lot in the establishment of a database in high quality.

Study documents

- Protocol
- Case Report Forms (CRFs)
- CRF Completion Guideline
- Annotated CRFs
- Statistical Analysis Plan
- Database Management
- Specifications
- Mock Shells
- ADS
- TLF
- Report/Manuscript

Statistical Programmers: Specifications

- Specifications are data derivation rules for creating variables required in Source datasets/ Analysis Datasets
- Specifications should provide details of attributes along with derivation rules for each variable in Source/Analysis dataset
- If you follow CDSIC data standards then you need to follow SDTM and ADaM Guideline while writing data derivation rules

Statistical Programmers: Table Shells

- Shells are layout of how you want to present data in a Table, Listing or Figure
- Shells should provide relevant Titles and Footnotes
- If Possible provide programming notes to clarify the information presented in Shell.

Statistical Programmers :Write Data Definition files

- Follow SDTM/ ADaM Guidelines/ Client specific Guidelines or work instructions as applicable
- Write data derivation rule in Simple English
- Do not provide programming statements as data derivation rules. E.g if age<50 then agegrp=1
- Do not give reference of Source dataset while writing data derivation rules of ADaMs ; instead give reference of CRF page
- Where ever required provide detailed Notes explaining derivation rule.

Pharma Academy of Learning and Sharing

Statistical Programmers: Create SDTM/ADaMs

- Follow Programming standards, as applicable
- Write Generic program, try to visualize data captured on CRF and write code which will take care of all data points
- Avoid Hard coding, in case you need to do hardcoding discuss with Client and get the reason for hardcoding documented
- Write programming comments , while you develop program
- Perform Self qc using qc checklist , if available. If no qc checklist is available then create a qc checklist
- Ensure program is creating required results, all variable attributes are as mentioned in define file
- If you are using standard macros for creating SDTM/ADaMs ensure you understand the macro documentation well before using them.

Statistical Programmers : Validate SDTM/ADaMs

- Follow Programming standards/SOP, as applicable
- If you are an independent programmer ,then you can start writing validation code once the data definition file is finalized.
- Objective of double programming is to create an independent code.
- Once the independent code is complete and the first line output is ready for validation, you can compare the datasets and try to identify reasons for differences, if any
- Ensure that both datasets (first line and validation) are run on same data and follow same version of data definition file

Pharma Academy of Learning and Sharing

Statistical Programmers : Create TLFs

- Follow Programming standards/SOP, as applicable
- Program should produce a report as per Table shell (including indentation, display of percentages, titles, footnotes etc)
- Read SAP/ programming standards to understand how percentages need to be calculated (based on N or based on non missing values of a category etc)
- While creating figures, make sure that size of the figure is not too small or too large.
- If you are using standard macros for creating TLFs ensure you understand the macro documentation well before using them.

Pharma Academy of Learning and Sharing

Statistical Programmers : Validate TLFs

- Follow Programming standards/SOP, as applicable
- Check for Report layout, title, footnotes, numbers presented in report, spellings etc
- For Listings which flow on to many pages, make sure that all pages are displaying required information.
- For Figures, check for correctness of numbers using the final dataset which is passed to a SAS procedure that creates figure.
- If possible , cross check Listings and summary tables of same domain. E.g AE listing and AE summary table

Pharma Academy of Learning and Sharing

Study documents

- Protocol
- Case Report Forms (CRFs)
- CRF Completion Guideline
- Annotated CRFs
- Statistical Analysis Plan
- Database
- Specifications
- Mock Shells
- ADS
- TLFs
- Report/Manuscript

Biostatistician: Report/Manuscript Writing

- Method Section:
 - Description of the data (design, endpoints)
 - Statistical Methodology
- Result Section:
 - Data présentation (tables, graphs, etc.)
- Discussion Section:
 - Appropriate interpretation of results.
- Usually the third author in papers

Summary

- Biostatistician Role
- Programmer Role



Biostatistician Role

- Protocol Development
- Data Management
- Study Implementation
- Study Monitoring
- Data Analysis
- Report/Manuscript Writing

Study Implementation

- Sampling Selection
- Implementation of Randomization procedures
 - Typically performed by an Independent Biostatistician with the instructions provided by the Study Biostatistician(e.g. Block size)

Study Monitoring

- Monitoring for Quality
- Monitoring for Safety/Efficacy:
 - Use of an Independent Biostatistician
 - Study Biostatistician still responsible for the content, analysis methodology, and data transfer coordination

Pharma Academy of Learning and Sharing

Biostatistician Role

- The Biostatistician should be involved in most aspects of a study
 - Plays a major role in protocol development,
 - Writes detailed statistical Analysis plan
 - DM review, interim reporting, and data analysis
- The Biostatistician ensures that ICH E9 guidelines on Statistical Principles are followed to:
 - Minimize Bias and Maximize Precision
- The Biostatistician cares about data because s/he does the analysis, so s/he should be involved in any/all aspect/decision related to the data

Programmers Role

- Read and Understand study Protocol
- Read and understand CRF, CRF completion guidelines
- Annotate CRFs(if required) or review Annotated CRFs
- Read SAP
- Database set up
- Database UAT
- Write Define files for SDTM/ADaMs
- Create SDTM/ADaMs
- Validate SDTM/ADaMs
- Create TLFs
- Validate TLFs

Thank You

