Clinical Data Standards
&
Automation

Pharmaceutical Users Software Exchange

PhUSE 2007

Author: **Giri Balasubramanian**

Company Name**: Kinship Technologies Pvt. Ltd.,**

Table of Contents

Executive Summary

Clinical Data Standards in the recent past has been given a wide importance and necessary publicity has been created by the CDISC - Clinical Data Standards Consortium. The Clinical Data Standards are not new to the world of Clinical Research & Development. It has been adopted by various pharmaceuticals and CROs in their native form. The generalization of standards has taken shape with the understanding of time saving at the reviewers end, which leads to huge saving of investment for the Pharmaceuticals and CROs which naturally leads to low cost of the drug upon approval. These standards has been widely spoken at various stages namely, Data Capture (ODM), Laboratory Results capture (LAB), Analysis Data Creation (ADaM) and final submission set (SDTM) with the regulatory agencies.

The above standards such as ODM, LAB if implemented at the Data Capture will lead to automation of Submission Dataset Creation and generation of Listings, Tables and Figures. Most of the Biostatistical needs across SAPs under different therapeutic areas are similar in terms of presenting the Listings, Tables and Figures. The CDISC Team has generalized this input based on thorough analysis of many SAPs across therapeutic areas. This has resulted in Automation of Biostatistical Analysis to a good extent.

1.0 Introduction

The Evolution of Clinical Data Standards is a blessing in disguise for the small and medium sized pharmaceutical organizations. The large sized pharmaceuticals already have implemented their own standards and procedures, which are followed for their submission. The standards proposed which is in the process of acceptance by FDA have gained momentum in usage among all the pharmaceuticals since it benefits both the parties namely, the Pharmaceutical and Regulatory Departments.

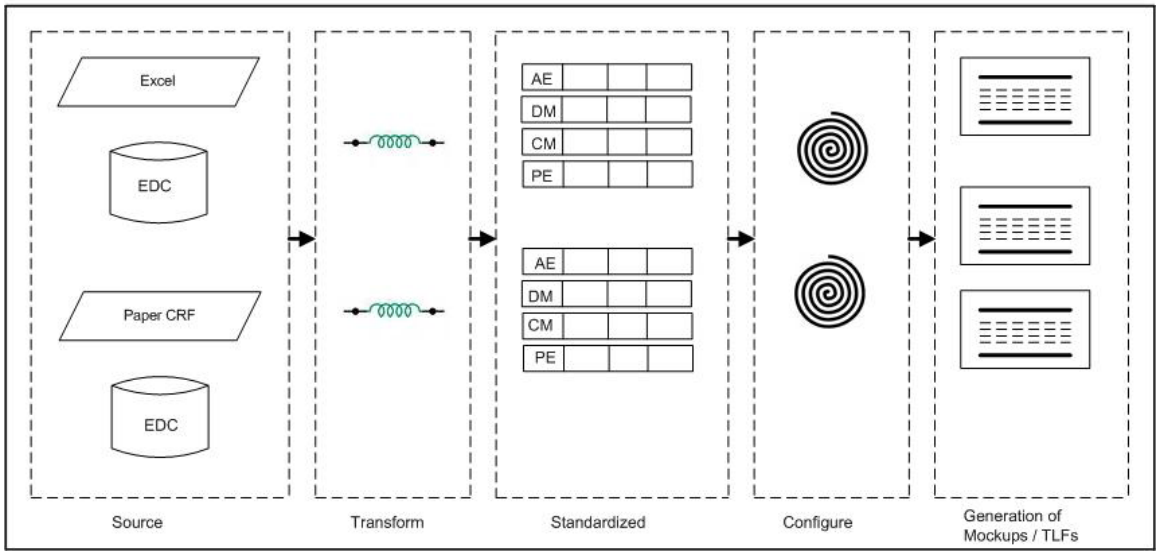
The proposed standards are for capture, analyses and submission related. These standards have initiated the process of Automation in post Data Capture scenario for the pharmaceutical organization. In the last couple of decades, the Automation in some form has been adopted by the large sized pharmaceutical organizations based on their own defined standards. In the recent past, the evolution of CDISC Standards has resulted in Automated solutions that are universal to all the Pharmaceutical companies who are in the process of Submission and seeking approval with the regulatory agencies.

2.0 Why Automation?

Automating a high majority of the Mockups & TLFs generation process will expedite the analysis of clinical data to a higher efficiency level. It frees programmers and statisticians from SAS limitations, and helps them stay focused on their business objectives. Resources being a constraint, and efficiency a challenge, the tool AutoGen TsLs will satisfy more than these needs.

The practices followed earlier lead to inefficiencies from constant re-programming and re-validation, as well as duplication of effort. The goal of this project is to make the process of tables and listings generation more efficiently by standardizing and automating a high majority of the process. This document takes us through the automation and benefits derived out of it and screen snapshots for better understanding.

FLOW OF TRIAL DATA



This productivity framework allows clinical programmers to easily, securely and effectively produces TLFs, leading to improved productivity and increased quality of submission to regulatory department. It allows a dramatic decrease in the amount of code needed to produce listings and tables, and programs are both written and validated much more quickly. It allows producing any kind of data listing and summary table, regardless of input data structure, output layout, and customer reporting process.

Some of the Key Benefits are:

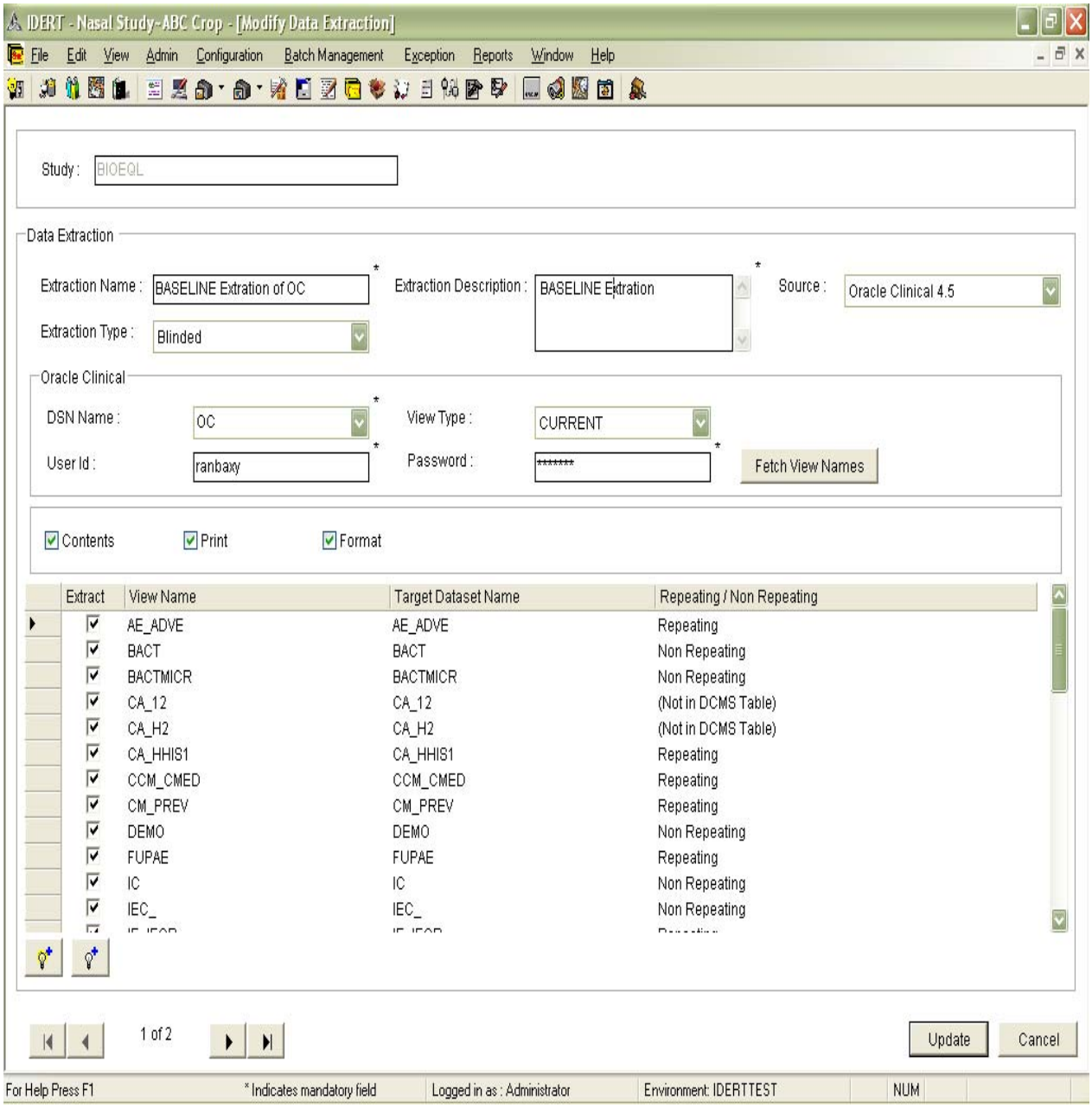
- Standard Analysis Set
- Reduction in validation time for Standard Analysis Set
- Key resource to focus on Exploratory Analysis
- Adhere or move towards FDA future submission needs
- Reduction in Time to Submission

3.0 Approach

The proposed approach will be interesting to Biostatisticians and the Programming Team who are supporting the Data Extraction, Biostatistical Department and Data Standards & Regulatory Affairs of any pharmaceutical companies.

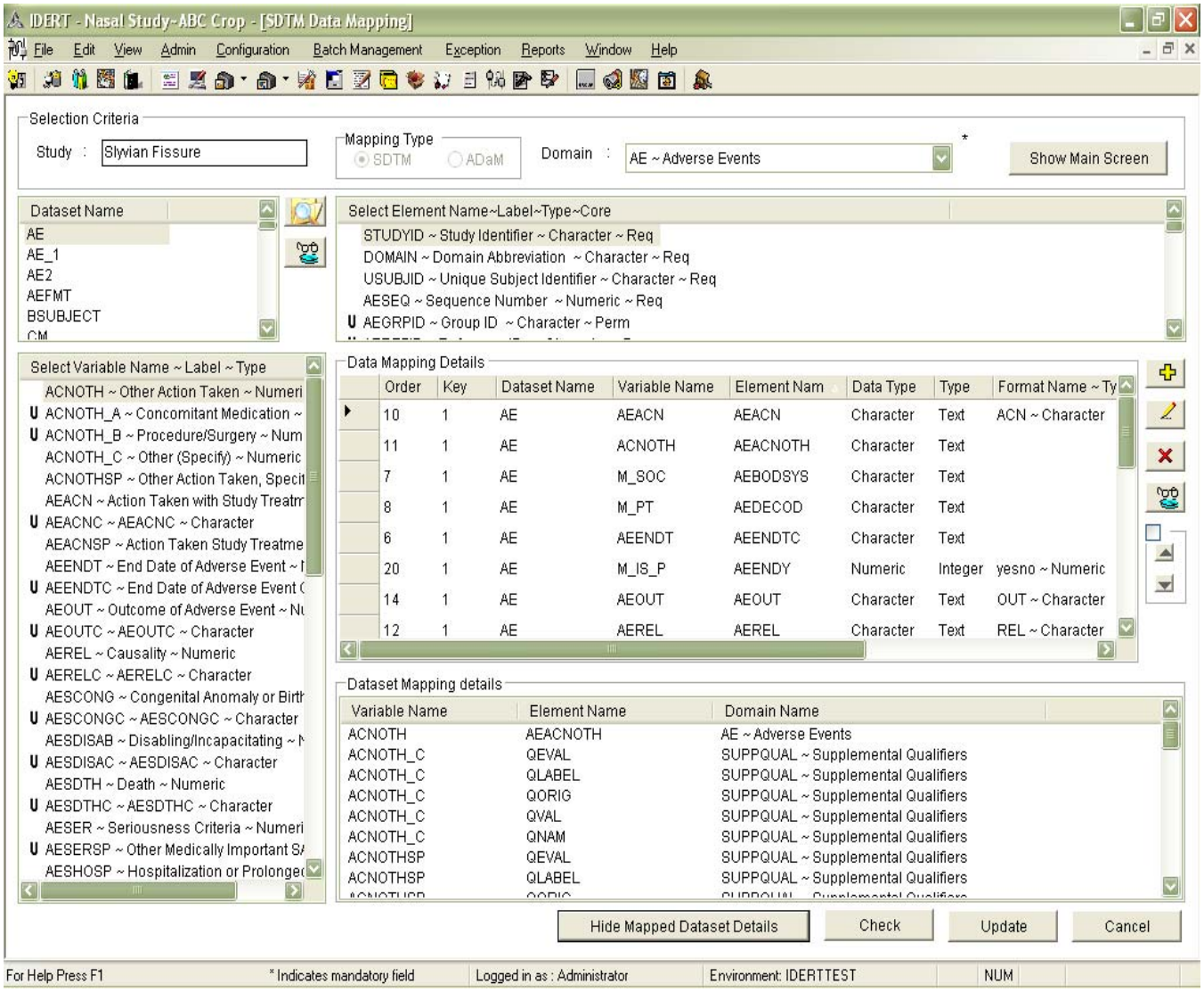
Phase – I

This phase is focused to bring the required clinical trial data into the Automation Solution. The data is being extracted from various sources such as EDCs, Excel, XPTs and so on and converted to SAS Datasets.


IDERT - Nasal Study-ABC Crop - [Modify Data Extraction]
File Edit View Admin Configuration Batch Management Exception Reports Window Help
Study : BIOEQL
Data Extraction
Extraction Name : BASELINE Extration of OC
Extraction Description : BASELINE Extration
Source : Oracle Clinical 4.5
Extraction Type : Blinded
Oracle Clinical
DSN Name : OC
View Type : CURRENT
User Id : ranbaxy
Password : *******
Fetch View Names
Contents
Print
Format
Extract
View Name
Target Dataset Name
Repeating / Non Repeating
AE_ADVE AE_ADVE Repeating
BACT BACT Non Repeating
BACTMICR BACTMICR Non Repeating
CA_12 CA_12 (Not in DCMS Table)
CA_H2 CA_H2 (Not in DCMS Table)
CA_HHIS1 CA_HHIS1 Repeating
CCM_CMED CCM_CMED Repeating
CM_PREV CM_PREV Repeating
DEMO DEMO Non Repeating
FUPAE FUPAE Repeating
IC IC Non Repeating
IEC_ IEC_ Non Repeating
1 of 2
Update
Cancel
For Help Press F1
* Indicates mandatory field
Logged in as : Administrator
Environment: IDERTTEST
NUM

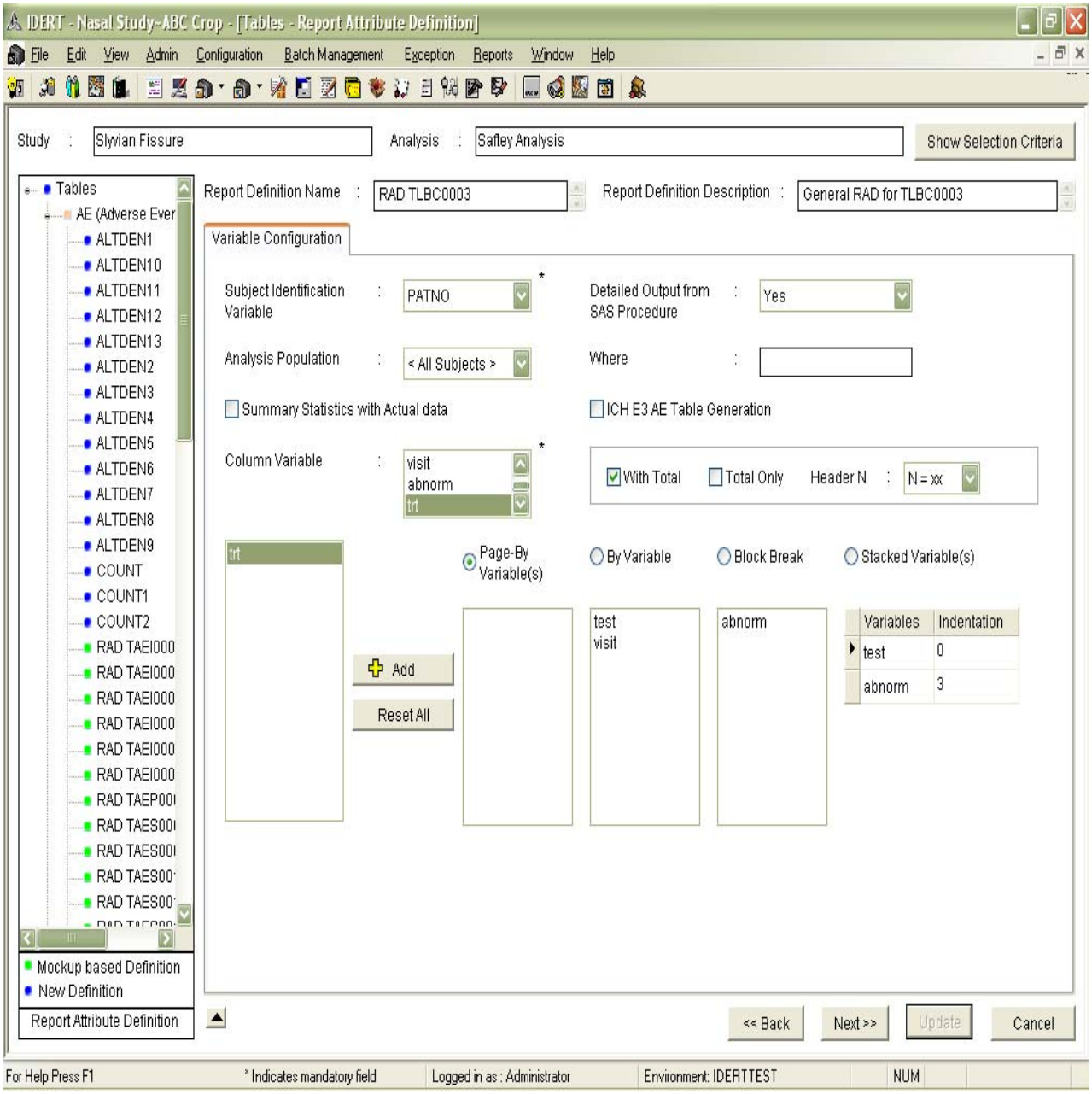
Phase - II

This phase covers the Scope of Clinical Data Standards Implementation. The Implementation can be done through CDISC conversion or if the Clinical Data Standards has been adopted at EDC, then the conversion may not be required at this stage. The CDISC Conversion is the process in which the Source or Trial Data is converted to CDISC Standards using a set of configurable features provided through the GUI. These conversions are vital for automation of work, since the automation is driven by rule based approach. The conversion is facilitated by a set of screens that capture data for Value List & Imputation, Mapping and Conversion execution interfaces. This complete process gets automated with less human intervention especially for RELREC, SUPPQUAL creations. This again is validated using the WebSDM Rules that are available.



Phase – III

This phase is critical to the Biostatistical Team, who are involved in the Analyses of Trial Data using various statistical techniques. The statistical techniques are grouped under parametric and non-parametric test, which covers various PROCs supplied by SAS. This analyses module is executed in SAS Platform. A complete automation can be achieved, if the source data are converted to CDISC standards and the mockups pre-configured through GUI. The source or trial data are mapped to the selected mockups for the specific study. The user configures rest of the items such as reporting and layout at the initial stages of configuration, which is applicable for all TLFs for a specific study.


IDERT - Nasal Study-ABC Crop - [Tables - Report Attribute Definition]
File Edit View Admin Configuration Batch Management Exception Reports Window Help
Study : Slywian Fissure
Analysis : Saftey Analysis
Show Selection Criteria
Tables
AE (Adverse Ever
ALTDEN1
ALTDEN10
ALTDEN11
ALTDEN12
ALTDEN13
ALTDEN2
ALTDEN3
ALTDEN4
ALTDEN5
ALTDEN6
ALTDEN7
ALTDEN8
ALTDEN9
COUNT
COUNT1
COUNT2
RAD TAEI000
RAD TAEI000
RAD TAEI000
RAD TAEI000
RAD TAEI000
RAD TAEI000
RAD TAEP00I
RAD TAES00I
RAD TAES00I
RAD TAES00
RAD TAES00
Mockup based Definition
New Definition
Report Attribute Definition
Report Definition Name : RAD TLBC0003
Report Definition Description : General RAD for TLBC0003
Variable Configuration
Subject Identification Variable : PATNO
Detailed Output from SAS Procedure : Yes
Analysis Population : < All Subjects >
Where :
Summary Statistics with Actual data
ICH E3 AE Table Generation
Column Variable : visit abnorm trt
With Total
Total Only
Header N : N = xx
trt
Page-By Variable(s)
By Variable
Block Break
Stacked Variable(s)
Add
Reset All
test visit
abnorm
Variables Indentation
test 0
abnorm 3
<< Back
Next >>
Update
Cancel
For Help Press F1
* Indicates mandatory field
Logged in as : Administrator
Environment: IDERTTEST
NUM

4.0 Validation

Validation of automated solution is the key to its success. Validation needs to be taken care with atmost importance. FDA considers software validation to be "*confirmation by examination and provision of objective evidence that software specifications conform to user needs and intended uses, and that the particular requirements implemented through software can be consistently fulfilled*." In practice, software validation activities may occur both during, as well as at the end of the software development life cycle to ensure that all requirements have been fulfilled. Since software is usually part of a larger hardware system, the validation of software typically includes evidence that all software requirements have been implemented correctly and completely and are traceable to system requirements. A conclusion that software is validated is highly dependent upon comprehensive software validation, inspections, analyses, and other verification tasks performed at each stage of the software development life cycle.

Validation of such automated solutions should follow GAMP CAT 4 Model of Development and Validation.

5.0 Regulatory Compliance

Any solution requiring the use of -computerized systems to create, analyze and maintain source data of patients participating in clinical trials must satisfy the FDA requirements to assure the data quality in the same manner as paper records. Because these data are integral to any FDA assessment of the safety and effectiveness of the new human drugs, FDA has published a final guidance document, entitled *Guidance for Industry on Computerized Systems Used in Clinical Investigations.*

Automated Solution using the proposed phases should meet this guidance and also adhere to the Part 11, Electronic Records.

6.0 Conclusion

The approach presented in this paper will lead to much more automation in the immediate future and will change the way the Biostatistical analysis are being done with the Clinical Data Standards in place across all the Pharmaceutical organization. This will lead to many solutions from the lead vendors, which will all carry an immense value to the Biostatistical Team in speeding up the process of analyses and focus more on the exploratory analyses.