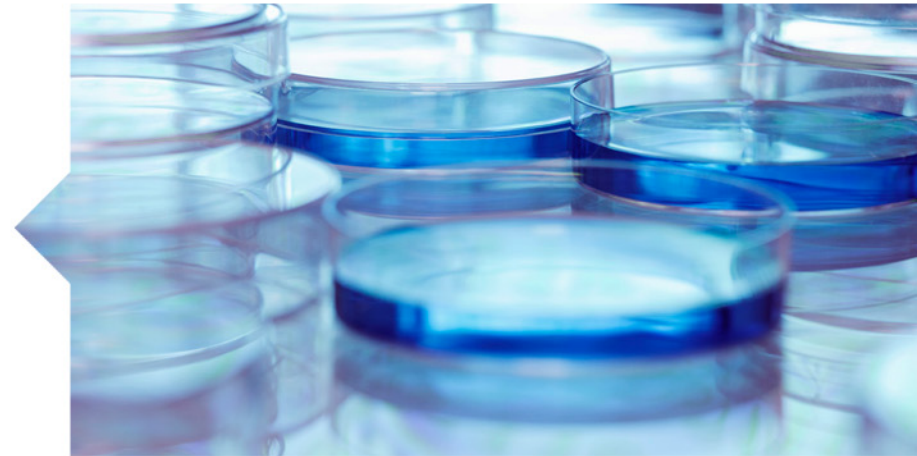




Late Stage Programming and Validation

23rd February 2013



Ramkumar M.Rajendiran



All materials in this slides are the opinion of the speaker, and do not reflect the views of Quintiles.

In a fast-paced production environment of a clinical trial (especially in Phase I and II)

- Timelines are short and timely delivery is very important
- Not getting the datasets early from the DM team
- Last minute changes and frequent transfers
- Working more than 2 studies simultaneously
- Actual programming started only at the time of the soft lock or dry run of datasets.
- Complete information about the subjects will be available only at the end of the study hard lock.
- Reusing similar study programs due to short time lines.

Hence there is lack of knowledge about the source SDTM datasets before ADaM datasets or TLF generation.

- Handling the value level changes over data transfers like typo error and deviating from the standard terminology etc
- Handling of dropped/added datasets or variables
- More number of transfers and hence rechecking the outputs after every rerun
- Not checking the derivations properly when the programs are copied from some other study.

Hence it will be easy, If we get

- Clear picture of the changes between the data transfers from SDTM Team
- Detailed information about the data transfer in dataset level, variable level, value level
- The unique values within each variable and its frequency.

QProbe used simple methods to get the results

- ODS tagsets.excelxp
- Sashelp.vcolumn, sashelp.vtable
- Proc compare
- Proc freq
- Proc report

Keep Ready:

SDTM or ADaM or sponsor defined standards in imported dataset

Important features of QProbe



Generates two excel files

Findings:

- Dataset Level
- Variable Level
- Value Level

Summary :

- Unique variable values and its counts for each domain

Macro Calling :

```
%QProbe(nrwlib=xxx/xxxx/proj00x/study00x/data/transfer/15jan2013,  
    orwlib=xxx/xxxx/proj00x/study00x/data/transfer/01jan2013,  
    outpath=xxx/xxxx/proj00x/study00x/data/transfer/compres,  
    stdtype=SDTM);
```

Findings - Dataset Level



Domain Name	Data Set Label	Date Created	Date Modified	Number of Variables	Number of Logical Observations	Longest variable name	Longest label	Comparison Result
AE	Adverse Events	17DEC12:00:53:34	17DEC12:00:53:34	20	64	8	38	FAIL
CM	Concomitant Medications	18DEC12:09:53:36	18DEC12:09:53:36	18	7	8	38	FAIL
DM	Demographics	18DEC12:06:40:09	18DEC12:06:40:09	17	22	7	33	FAIL
DS	Disposition	18DEC12:09:35:34	18DEC12:09:35:34	12	131	8	39	PASS
EX	Exposure	18DEC12:10:45:56	18DEC12:10:45:56	13	261	8	31	PASS
IE	Inclusion Exclusion Exceptions	02NOV12:06:02:53	02NOV12:06:02:53	13	0	8	40	Empty dataset
LB	Laboratory Test Results	18DEC12:14:09:38	18DEC12:14:09:38	33	7750	8	40	FAIL
MH	Medical History	18DEC12:09:58:06	18DEC12:09:58:06	10	37	8	40	FAIL
PE	Physical Examinations	17DEC12:00:54:02	17DEC12:00:54:02	16	610	8	38	FAIL
QS	Questionnaires	02NOV12:07:26:22	02NOV12:07:26:22	19	990	8	38	Dropped in current transfer
RELREC	Related Records	18DEC12:10:39:17	18DEC12:10:39:17	7	15	8	27	PASS
SE	Subject Elements	18DEC12:13:21:32	18DEC12:13:21:32	8	109	7	26	Newly added in current transfer
SUPPAE	Supplemental Qualifiers	18DEC12:09:48:46	18DEC12:09:48:46	10	316	8	27	PASS
SUPPCM	Supplemental Qualifiers	18DEC12:09:53:42	18DEC12:09:53:42	10	27	8	27	PASS
SV	Subject Visits	18DEC12:14:42:01	18DEC12:14:42:01	7	0	8	25	Newly added in current transfer
TA	Trial Arms	18DEC12:10:02:53	18DEC12:10:02:53	10	5	8	27	PASS
TE	Trial Elements	18DEC12:10:53:41	18DEC12:10:53:41	7	5	7	27	PASS
TI	Trial Inclusion/Exclusion	18DEC12:10:41:48	18DEC12:10:41:48	5	28	8	30	PASS
TS	Trial Summary	18DEC12:10:55:27	18DEC12:10:55:27	6	23	8	34	PASS
TV		18DEC12:10:54:10	18DEC12:10:54:10	7	5	8	19	PASS
VS		31JAN13:04:57:51	31JAN13:04:57:51	21	4001	8	40	FAIL

Findings - Variable Level

1	List of variables dropped in current transfer		
2			
3	Domain Name	Variable Name	Variable Label
4	PE	PESTAT	Completion Status
5			
6	List of variables newly added in current transfer		
7			
8	Domain Name	Variable Name	Variable Label
9	CM	CMENDTC	End Date/Time of Medication
10	DM	COUNTRY	Country
11	DM	SEX	Sex
12	MH	MHENRF	End Relative to Reference Period
13	PE	PECAT	Category for Examination
14	PE	VISITDY	Planned Study Day of Visit

Dataset Level

Variable Level

Value Level

Findings - Value Level

List of values in standard but not in dataset		
Dataset Name	Variable Name	Value of the Variable
AE	AESEV	SEVERE
LB	LBCAT	CHEMISTRY
LB	LBCAT	HEMATOLOGY
LB	LBCAT	URINALYSIS
VS	VSSTRESU	BEATS/MIN
VS	VSSTRESU	BREATHS/MIN
VS	VSSTRESU	kg/m*m
List of values in dataset but not in standard		
Dataset Name	Variable Name	Value of the Variable
DS	DSDECOD	ADVERSE EVENT
DS	DSDECOD	BIOMARKER INFORMED CONSENT OBTAINED
LB	LBCAT	BIOMARKER
LB	LBCAT	Chemistry
LB	LBCAT	Hematology
LB	LBCAT	Urinalysis
LB	LBNRIND	HIGH CRITICAL
LB	LBNRIND	LOW CRITICAL
SUPPAE	QNAM	AEACN2
SUPPAE	QNAM	AEREL2
VS	VSSTRESU	beats/min
VS	VSSTRESU	breath/min
VS	VSSTRESU	kg/(m*m)

28			
29	Information about No Data to Display		
30			
31	Dataset Name	Variable Name	Value of the Variable
32	AE	AESER	There are Serious Adverse Events Reported - Display the results
33	AE	AECAUDIS	No subject discontinued due to AE - No data to display
34	AE	AEOUT	No adverse event caused death - No data to display

Summary



1	VARNAME	VALUE	COUNT
2	AGE	19	2
3	AGE	20	2
4	AGE	22	1
5	AGE	23	2
6	AGE	30	1
7	AGE	36	2
8	AGE	43	2
9	AGE	44	2
10	AGE	47	2
11	AGE	49	2
12	AGE	52	1
13	AGE	54	2
14	AGE	55	1
15	AGEU	YEARS	22
16	ARM	XA/XB/XC	22
17	ARMCD	ABC	22
18	COUNTRY	USA	22
19	DOMAIN	DM	22
20	ETHNIC	HISPANIC OR LATINO	1
21	ETHNIC	NOT HISPANIC OR LATINO	21
22	INVNAM	DAVID R. MATHEWS	22
23	RACE	AMERICAN INDIAN OR ALASKA NATIVE	1
24	RACE	BLACK OR AFRICAN AMERICAN	10
25	RACE	WHITE	11
26	RFENDTC	2012-04-08T08:00	1
27	RFENDTC	2012-04-08T08:07	1

1	VARNAME	VALUE	COUNT
2224	VSTPT	1 HOUR 30 MINUTES POSTDOSE	1050
2225	VSTPT	1 HOUR POSTDOSE	168
2226	VSTPT	1 HOUR POST DOSE	228
2227	VSTPT	2 HOURS POSTDOSE	168
2228	VSTPT	24 HOURS POSTDOSE	216
2229	VSTPT	30 MINUTES POSTDOSE	168
2230	VSTPT	4 HOURS POSTDOSE	171
2231	VSTPT	8 HOURS POSTDOSE	171
2232	VSTPT	PRE-DOSE	90
2233	VSTPT	PREDOSE	1169

It helps to

- Get the overall information about the study datasets
- Status of the ADaM datasets validation without checking the tracking sheet.
- Estimate how many resources need and how much time will take to update the changes in the current transfer
- Summary excel file will be helpful during the senior review
- Senior review can be done only for the domains in which the changes happened in new transfer.

It helps to

- Get clear picture of the datasets in each level
- Understand the changes in datasets comparing to previous transfer
- Check and identify the dataset issues
- Get the unique values
- Get information about values deviating from standard like label length or variable length, and standard values of the variables.
- Get the list of information about “No data to display”
- Check the derivations with the unique values from summary excel file, if the derivations are copied from some other study. (ex. Study discontinued flag)

Future Implementations



- Complete automated report from Raw datasets to CDISC and ADaM submission
- ADaM - from the available comparison report(lst) or using the program batch file, the report can be generated.
- Check for the CDISC and ADaM rules like required and expected variables, Label level, length etc.
- Summary excel file: Checking data issues with filtered condition in SAS viewer or SAS while selecting the value in the excel sheet.



Questions?

