

Automatic Version Control and Track Changes of CDISC ADaM Specifications for FDA Submission

Min Chen, Xiangchen (Bob) Cui
Vertex Pharmaceuticals, Cambridge, MA



ABSTRACT

CDISC ADaM specifications documentation is the core of the programming activities in FDA submission. It serves as the primary source for ADaM programming and validation, define.xml and reviewer guide, and also helps in establishing metadata traceability. It is preferable to create ADaM specifications in Word® Format to facilitate both reviewing and approving of specifications documentation by statisticians and/or validators. Since derivation rules may be complex and subject to constant change during the whole ADaM programming activities, it is desirable to automatically keep track of different versions of ADaM programming specifications. It is more beneficiary for sponsors to keep track of different versions when ADaM programming is outsourced to external vendors. This paper introduces a SAS macro to automatically detect and report any revisions between the new version and the old versions of programming specifications. It helps to generate necessary documents for version control, establishes traceability and achieves high efficiency in FDA submission.

Goal:

- Track any changes from any old versions of specifications
- Help to finalize the programming specification at the very early stage in an efficient way
- Ensure the traceability and clear documentation for FDA submission
- Achieve Cost-Effectiveness and Efficiency

Introduction: Programming Specification for ADaM Dataset(s)

1.1.1 ADSL: Subject Level Analysis Dataset Domain Information Table

Dataset	ADSL
Program Name	Adsl.sas
Description	Subject-Level Analysis Data
Unique identifier Variables	usubjid
General Class	Special Purpose
Structure	One record per subject
Input Datasets	DM, DS, EX, VS, DC, HC, AE
Notes	Includes all subjects enrolled.

Variable Name	Variable Label	Type	Length	Controlled Terms or Formats	Origin	Role	Comments	Core
STUDYID	Study Identifier	Char	20		DM.studyid	Identifier	Constant Value: "ABC-ZZZ-XXX"	Req
USUBJID	Unique Subject Identifier	Char	40		DM.usubjid	Identifier	Equivalent to studyid "-" strip(siteid) "-" strip(subid)	Req
SUBJID	Subject Identifier for the Study	Char	20		DM.subjid	Identifier	(e.g. 102130)	Req
SITEID	Study Site Identifier	Char	8		DM.siteid	Record Qualifier		Req
AGE	Age	Num	8		DM.age	Record Qualifier		Req
AGEGR1	Pooled Age Group 1	Char	20		Derived	Record Qualifier	Equals to DM.age	Perm
AGEGR1N	Pooled Age Group 1 (N)	Num	8		Derived	Synonym Qualifier	Category derived if age non-missing. Equals 1, if age <= 45 2, if 45 < age <= 65 3, if age > 65	Perm

Transfer from Word to SAS Format Variable Information Table

	DOMAIN	VARNUM	VARIABLE	LABEL	TYPE	LENGTH	ORIGIN	TERM	CODELIST	ROLE	COMMENT	CORE
1	ADSL	99	AEANYPFL	Any AEs during Overall Phase	Char	2	Derived	YESNOF	(1)Y (2)N	Analysis	If one subject has at least one...	Perm
2	ADSL	98	AEANYTFL	Any AEs during TVR Phase	Char	2	Derived	YESNOF	(1)Y (2)N	Analysis	If one subject has at least one...	Perm
3	ADSL	111	AEDSCFPL	Any AEs Leading to Perm TRT DSC Overall	Char	2	Derived	YESNOF	(1)Y (2)N	Analysis	If one subject has at least one...	Perm
4	ADSL	110	AEDSCFPL	Any AEs Leading to Perm TRT DSC TVR	Char	2	Derived	YESNOF	(1)Y (2)N	Analysis	If one subject has at least one...	Perm
5	ADSL	115	AEDTHPFL	Any AEs Leading to Death during Overall	Char	2	Derived	YESNOF	(1)Y (2)N	Analysis	If one subject has at least one...	Perm
6	ADSL	114	AEDTHPFL	Any AEs Leading to Death during TVR	Char	2	Derived	YESNOF	(1)Y (2)N	Analysis	If one subject has at least one...	Perm
7	ADSL	107	AEINTOFL	Any AEs Leading to Dose INT Overall	Char	2	Derived	YESNOF	(1)Y (2)N	Analysis	If one subject has at least one...	Perm
8	ADSL	106	AEINTVFL	Any AEs Leading to Dose INT TVR	Char	2	Derived	YESNOF	(1)Y (2)N	Analysis	If one subject has at least one...	Perm
9	ADSL	25	AGE	Age at Enrollment	Num	8	DM.AGE			Record Qualifier	age = DM.AGE	Req
10	ADSL	26	AGEGPFL	Age Category	Char	20	derived			Record Qualifier	<=45, if age <= 45;>45 and <...	Perm

Version Control - Macro %read_spec

%macro read_specs(indir=,specsnm=,outdir=,newdtm=,runorder=);

Where

INDIR: Full Path for ADaM programming specification.

SPECSNM: Name of ADaM programming specification.

OUTDIR: Full Path for output SAS dataset which contains the attributes of ADaM variables.

NEWDTNM: A valid SAS dataset name for SAS dataset containing current specifications information.

RUNORDER: A valid numeral, defining the order for a specific domain to run (in the final run).

Function:

- Read information from individual ADaM specifications in CSV format
- Store both word version and SAS dataset of the programming specifications with time stamp in a study subfolder, named as ‘history’.

Sample Code:

```
*** Get date of programming running as time stamp ***;
data _null_;
    day = compress(put(today(),yymmdd10.),'-');
    call symput('day',strip(day));
run;
*** Output a SAS data of specification with time stamp to a subfolder ‘history’ ***;
data history.&newdtm._&day.;
    set __vars;
run;
*** Copy a Word Version with time stamp to a subfolder ‘history’ ***;
options noxwait NOXSYNC;
%let specdoc = %scan(&specsnm.,1);
X "%str(copy %"&indir.&specdoc..doc%" %"&outdir.history%\&specdoc._&day..doc%")";
```

An Example of Version Control for ADSL Specification Document

Name	Date modified	Type	Size
ADSL_20110913	9/12/2011 9:08 AM	Microsoft Word Document	341 KB
ADSL_20110915	9/15/2011 9:51 AM	Microsoft Word Document	340 KB
ADSL_20110919	9/19/2011 4:11 PM	Microsoft Word Document	346 KB
ADSL_20111007	9/22/2011 4:35 PM	Microsoft Word Document	342 KB
ADSL_20111018	10/14/2011 9:41 AM	Microsoft Word Document	337 KB
adsl_vars_20110913	9/13/2011 11:15 AM	SAS Data Set	1,921 KB
adsl_vars_20110915	9/15/2011 10:18 AM	SAS Data Set	1,921 KB
adsl_vars_20110919	9/19/2011 4:12 PM	SAS Data Set	1,921 KB
adsl_vars_20111007	10/7/2011 10:31 AM	SAS Data Set	1,889 KB
adsl_vars_20111018	10/18/2011 12:50 PM	SAS Data Set	1,873 KB

Tracking Changes - Macro %track_change

%macro track_specs(newdir=,olddir=,newdtm=,predtnm=,outdir=);

Where

NEWDIR: Full Path for new ADaM programming specification.

OLDDIR: Full Path for old ADaM programming specification.

NEWDTNM: A valid SAS dataset name for SAS dataset containing current specifications information.

PREDTNM: A valid SAS dataset name for SAS dataset containing old specifications information.

OUTDIR: Full Path for reports of specifications changes.

Function:

Report changes from any previous version of specifications in ‘history’ folder

Facilitates reviewing the new ADaM specification.

Output – Typical Report of ADaM Specification Changes

Typical Report of Adding Variables

The following Variables Were Added in the New Version of Specs.!

Variable	Variable Attribute	Comment
ARMCD	ARMCD label='Arm' length=\$8	DM.ARMCD

Typical Report of Attributes Changing

The following Variables with Attributes Changed!

Variable	Variable Number	Source	Variable Attribute
COHORT	7	New Specs	COHORT label='Cohort' length=\$40
		Old Specs	COHORT label='Cohort' length=\$8

Typical Report of Deleting Variables

The following Variables Were Deleted in the New Version of Specs.!

Variable	Variable Attribute	Comment
AGEC	AGEC label='Enrollment Age Category' length=\$40	if . < age <= 45 then, agec = "AGE <= 45 years"; else if 45 < age <= 65 then, agec = "45 < AGE <= 65 years"; else if 65 < age then, agec = "AGE > 65 years";

Typical Report of Comments Changing

The following Variables with Comment Changed!

Variable	Variable Number	Source	Comment
COHORT	10	New Specs	For XX1. if ADHC.svrfl="Y" then cohort="COHORT A"; else cohort="COHORT B"; For XX2. if ADSL.svz4def="Y" then cohort="COHORT A"; else cohort="COHORT B".
		Old Specs	If OUTCOME.outcome="SVR" then cohort="COHORT A"; else cohort="COHORT B".

Typical Report of Origin and/or Controlled Term or Format Changing

The following Variables with Origin and/or Controlled Terminology/Format Changed!

Variable	Variable Number	Source	Origin	Controlled Term or Format	Codelist
AGEU	22	New Specs	DM.AGEU	AGEU	(1)YEARS
		Old Specs	DM.AGEU		

Post-Delivery Changes and Tracking Changes

- Tracking change function will be triggered for any changes in specifications after a Clinical Study Report (CSR) delivery
- Reports with changes from the last version of specifications are generated as documentation for audit.
- Example: comments column is updated after the delivery due to editorial change, ONLY one report is generated: the report of change of comments.

The following Variables with Comment Changed!

Variable	Variable Number	Source	Comment
COHORT	10	New Specs	For XX1. if ADHC.svrfl="Y" then cohort="COHORT A"; else cohort="COHORT B"; For XX2. if ADSL.svz4def="Y" then cohort="COHORT A"; else cohort="COHORT B".
		Old Specs	If OUTCOME.outcome="SVR" then cohort="COHORT A"; else cohort="COHORT B".

Conclusion

- The macro makes it possible to track any changes of the specifications in an efficient way.
- The traceability can be achieved with the storage of the previous version of specifications.
- The documentation for audit can be automatically generated for post delivery changes.
- The tool achieves the high quality, cost-effectiveness, and the efficiency.

Acknowledgement

Appreciation goes to Tathabbai for his review and comments. Special thanks to Stacy for her supporting.

Disclosure

Authors of this presentation have the following to disclose concerning possible financial or personal relationships with commercial entities that may have a direct or indirect interest in the subject matter of this presentation:
Min Chen, Xiangchen (Bob) Cui: Employees – Vertex Pharmaceuticals