Statistical Programming for Dummies

PhUSE University Day, 15th October 2013
Sven Greiner, Nicola Tambascia
Statistical Programming, Accovion GmbH

Contents

Statistical Programming in Pharma

- Statistical Programmers are involved in the evaluation of clinical trials
- Main task: Technical implementation of statistical analyses
- Where can I work as a Statistical Programmer?
  - Pharmaceutical company: Bayer, Lilly, Roche, …
  - CRO: Quintiles, Parexel, Accovion, …

The drug development process

Phase I - III trials

Molecule discovery → **Pre-clinic** →

Clinical Trial Protocol

Clinical trial conduct

Data Management

Statistical Programming & Biostatistics

Medical Writing

Phase IV trials

← **Drug approval** ← **Marketing trial**

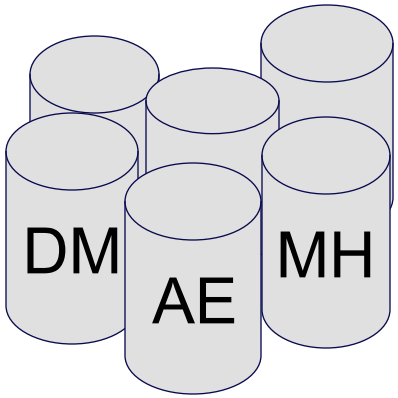
WE 1: Trial evaluation – Getting started

- Example of a trial evaluation for an ongoing trial
- Status:
  - Clinical Trial Protocol is available
  - Original datasets (ODS) provided by Data Management
  - (Draft) Statistical Analysis Plan (SAP) is available
- Goal: Create tables, figures and listings to analyse the trial endpoints

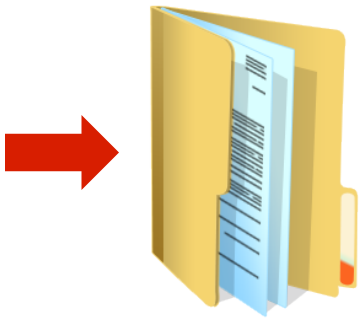
WE 1: Trial evaluation – ADS

- ✓ Create analysis datasets (ADS)
- ✓ About 10-30 ADS per trial evaluation

„data as collected“



ODS (SAS datasets)



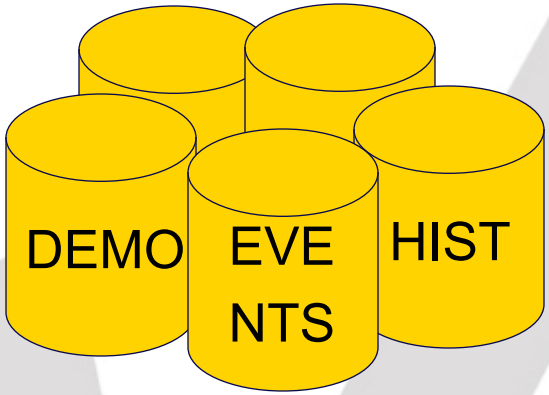
Specifications

```
DATA dm;
  SET ods.dm;
RUN;

* Create ADS;
DATA ads.demo;
  SET dm;
RUN;
```

SAS program

„data for analysis“



ADS (SAS datasets)

WE 1: Trial evaluation – ADS

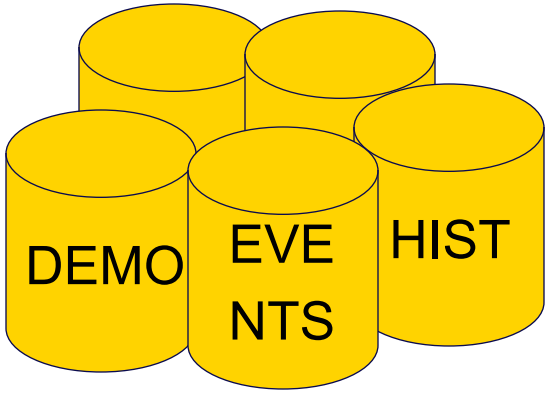
- Example of an ADS (ADSL)

	Subject Identifier for the Study	Age	Age Units	Age Group 1	Age Group 1 (N)	Sex	Sex (N)	Race	Race (N)	Ethnicity	Ethnicity (N)	Smoking History	Smoking History (N)
1	11001	51	YEARS	>40 - 55	2	M	1	WHITE	5	NOT HISPANIC/LATINO	0		.
2	11002	47	YEARS	>40 - 55	2	M	1	WHITE	5	NOT HISPANIC/LATINO	0	CURRENTLY SMOKES	2
3	11003	49	YEARS	>40 - 55	2	M	1	WHITE	5	NOT HISPANIC/LATINO	0	CURRENTLY SMOKES	2
4	11004	52	YEARS	>40 - 55	2	M	1	WHITE	5	NOT HISPANIC/LATINO	0	CURRENTLY SMOKES	2
5	11101	56	YEARS	>55	3	M	1	WHITE	5	NOT HISPANIC/LATINO	0		.
6	11102	44	YEARS	>40 - 55	2	M	1	WHITE	5	NOT HISPANIC/LATINO	0	EX-SMOKER	1
7	11103	44	YEARS	>40 - 55	2	M	1	WHITE	5	NOT HISPANIC/LATINO	0	EX-SMOKER	1
8	11104	43	YEARS	>40 - 55	2	M	1	WHITE	5	NOT HISPANIC/LATINO	0	CURRENTLY SMOKES	2
9	11105	58	YEARS	>55	3	M	1	WHITE	5	NOT HISPANIC/LATINO	0	EX-SMOKER	1
10	11201	45	YEARS	>40 - 55	2	F	2	WHITE	5	NOT HISPANIC/LATINO	0	CURRENTLY SMOKES	2
11	11202	43	YEARS	>40 - 55	2	M	1	WHITE	5	NOT HISPANIC/LATINO	0	CURRENTLY SMOKES	2
12	11203	34	YEARS	18 - 40	1	F	2	WHITE	5	NOT HISPANIC/LATINO	0	CURRENTLY SMOKES	2
13	11204	44	YEARS	>40 - 55	2	M	1	WHITE	5	NOT HISPANIC/LATINO	0	EX-SMOKER	1
14	11301	53	YEARS	>40 - 55	2	M	1	WHITE	5	NOT HISPANIC/LATINO	0	EX-SMOKER	1
15	11302	47	YEARS	>40 - 55	2	M	1	ASIAN	2	NOT HISPANIC/LATINO	0	CURRENTLY SMOKES	2
16	11303	56	YEARS	>55	3	M	1	WHITE	5	NOT HISPANIC/LATINO	0	EX-SMOKER	1
17	11401	40	YEARS	18 - 40	1	F	2	WHITE	5	NOT HISPANIC/LATINO	0	EX-SMOKER	1
18	11402	62	YEARS	>55	3	M	1	WHITE	5	NOT HISPANIC/LATINO	0	EX-SMOKER	1
19	11403	54	YEARS	>40 - 55	2	M	1	WHITE	5	NOT HISPANIC/LATINO	0	CURRENTLY SMOKES	2
20	11404	54	YEARS	>40 - 55	2	M	1	WHITE	5	NOT HISPANIC/LATINO	0	EX-SMOKER	1
21	11405	56	YEARS	>55	3	F	2	WHITE	5	NOT HISPANIC/LATINO	0	EX-SMOKER	1
22	11406	26	YEARS	18 - 40	1	M	1	ASIAN	2	NOT HISPANIC/LATINO	0		.
23	11407	54	YEARS	>40 - 55	2	M	1	WHITE	5	NOT HISPANIC/LATINO	0	EX-SMOKER	1
24	12101	54	YEARS	>40 - 55	2	M	1	WHITE	5	NOT HISPANIC/LATINO	0	CURRENTLY SMOKES	2
25	12102	50	YEARS	>40 - 55	2	M	1	WHITE	5	NOT HISPANIC/LATINO	0	CURRENTLY SMOKES	2
26	12103	34	YEARS	18 - 40	1	F	2	WHITE	5	NOT HISPANIC/LATINO	0	CURRENTLY SMOKES	2

WE 1: Trial evaluation – Reporting

- Reporting starts after the first ADS are programmed
- Up to several thousand outputs per trial

SAP &
Table Shells



ADS

```
DATA all;
   SET ads.demo;
RUN;

* Print output;
PROC REPORT;
RUN;
```

Reporting program (SAS)

Statistical Output

WE 1: Trial evaluation – Reporting

- Example of an efficacy table

Table 2 Progression-free survival (primary analysis) - investigator assessment / randomised set

	Drug A N (%)	Drug B N (%)
Number of patients [N (%)]	612 (100.0)	615 (100.0)
Number of patients with PFS° event [N (%)]	465 (76.0)	345 (56.1)
Progression-free survival time [months]		
P25	32.4	35.9
Median	38.7	40.8
P75	42.1	52.0
Drug A vs. Drug B		
Hazard ratio#		0.64
(95% CI)		(0.55 , 0.73)
p-value		0.0175

° PFS = progression-free survival.
\# If hazard ratio is below 1 then favours drug B

table2.sas 01JAN2013

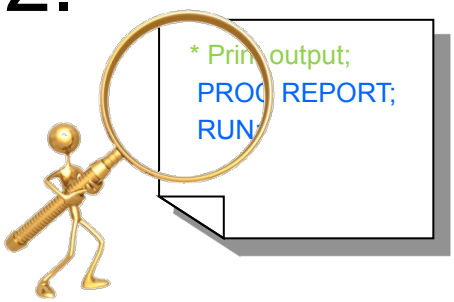
WE 1: Trial evaluation – Validation

- Validation is part of quality control
- "The process of evaluating software during or at the end of the development process to determine whether it satisfies specified requirements."

1.

Output review

2.



Code review

3.

Double programming

WE 1: Trial evaluation – Delivery

- ✓ Final statistical outputs in different formats
- ✓ Final ADS and specifications

Recipient: Medical Writing, Regulatory, Medical, Management

WE2: CDISC data conversion

- Current trend to have clinical data in standardized format
  - Driven by authorities that ask for standardized data
- Original data needs to be converted into standard format
- Standard currently recommended by authorities is provided by the Clinical Data Interchange Standards Consortium (CDISC)
  - SDTM for raw data
  - ADaM for derived data used for Analysis

WE2: Why CDISC SDTM/ADaM?

- General Benefits of Standards
  - Improve communication
  - Time/resource savings
  - Comparison across submissions

- FDA ‘s eStudy Data Draft Guidance (Feb 2012)
  - Written official recommendation that sponsors submit study data in a **standardized electronic format**

WE2: What is SDTM?

→**S**tudy

→**D**ata

→**T**abulation

→**M**odel

SDTM defines a standard structure for study data tabulations that are to be submitted as part of a product application to regular authority such as FDA. It based on material prepared by the Submissions Data Standard (SDS) Team of the CDISC.

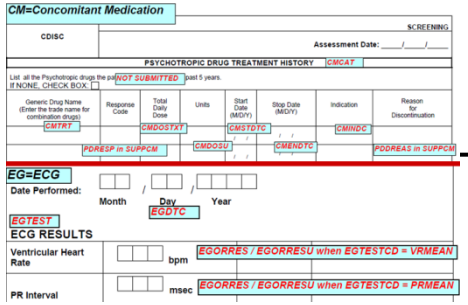
SDTM V1.2
Study_Data_Tabulation_Model_v1.2.pdf

SDTM IG V3.1.2
SDTM_Implementation_Guide_V3_1_2.pdf

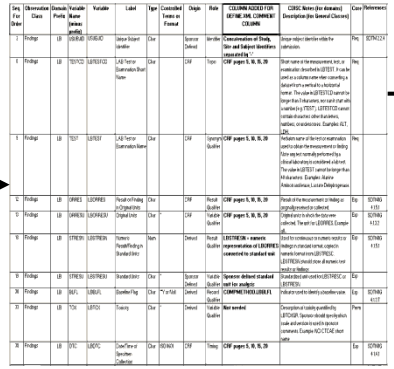
CDISC Strength Through Collaboration

WE2: Process Flow – (Legacy) Data Conversion

Annotated CRF



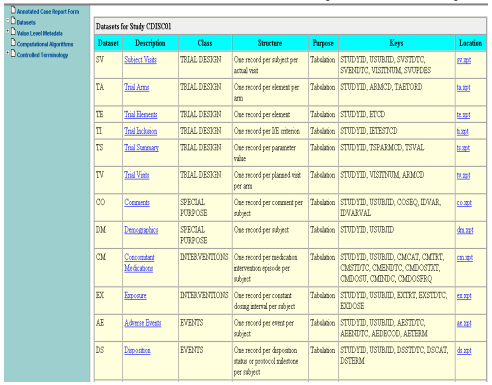
Specifications



Mapping programs

SDTM

define.xml (SDTM)

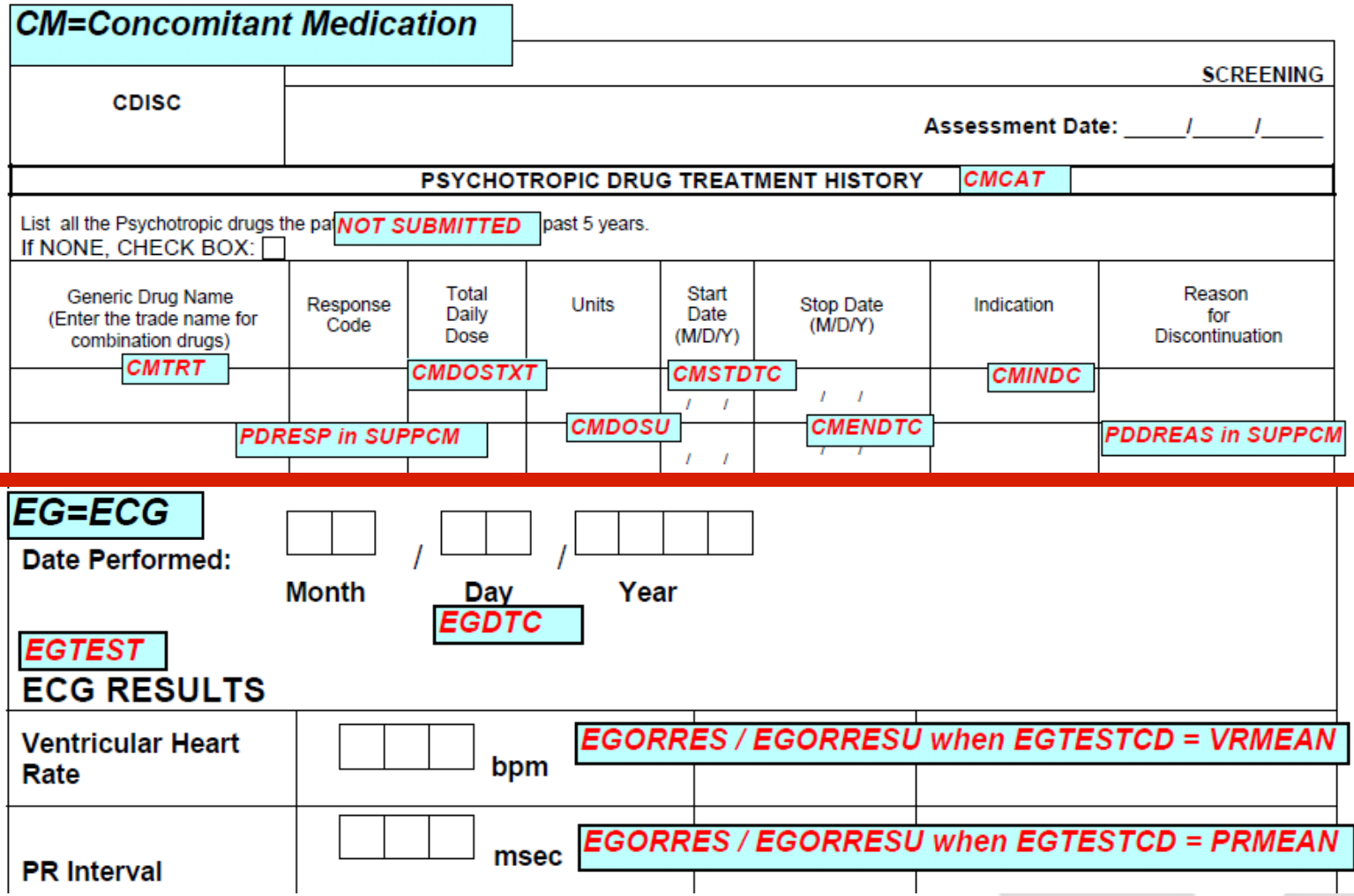


Datasets for Study CDISC01

Dataset	Description	Class	Structure	Purpose	Keys	Location
SV	Subject Visits	TRIAL DESIGN	One record per subject per actual visit	Tabulation	STUDYID, USUBJID, SVSTDTC, SVENDTC, VISITNUM, SVUPDES	sv.xpt
TA	Trial Arms	TRIAL DESIGN	One record per element per arm	Tabulation	STUDYID, ARMCD, TASTORD	ta.xpt
TE	Trial Elements	TRIAL DESIGN	One record per element	Tabulation	STUDYID, ETCD	te.xpt
TI	Trial Inclusion	TRIAL DESIGN	One record per IE criterion	Tabulation	STUDYID, IETESTCD	ti.xpt
TS	Trial Summary	TRIAL DESIGN	One record per parameter value	Tabulation	STUDYID, TSPARMCD, TSVAL	ts.xpt
TV	Trial Visits	TRIAL DESIGN	One record per planned visit per arm	Tabulation	STUDYID, VISITNUM, ARMCD	tv.xpt
CO	Comments	SPECIAL PURPOSE	One record per comment per subject	Tabulation	STUDYID, USUBJID, COSEQ, IDVAR, IDVARVAL	co.xpt
DM	Demographics	SPECIAL PURPOSE	One record per subject	Tabulation	STUDYID, USUBJID	dm.xpt
CM	Concomitant Medications	INTERVENTIONS	One record per medication intervention episode per subject	Tabulation	STUDYID, USUBJID, CMCAT, CMTRT, CMSTDTC, CMENDTC, CMDOSTXT, CMDOSU, CMDOSFRQ, CMDOSFRQ	cm.xpt
EX	Exposure	INTERVENTIONS	One record per constant dosing interval per subject	Tabulation	STUDYID, USUBJID, EXTRT, EXSTDTC, EXDOSE	ex.xpt
AE	Adverse Events	EVENTS	One record per event per subject	Tabulation	STUDYID, USUBJID, AESTDTC, AEENDTC, AEDECOD, AETERM	ae.xpt
DS	Disposition	EVENTS	One record per disposition status or protocol milestone per subject	Tabulation	STUDYID, USUBJID, DSSTDTC, DSCAT, DSTERM	ds.xpt

WE2: CRF Annotation Guidelines

- See CDISC Metadata Submission Guidelines (SDTM)



ACCOVION

WE2: Sample SDTM Domain Specifications (EG)

Seq. For Order	Domain Prefix	Variable Name	Variable Label	Type	Length	Controlled Terms or Format	Origin	Role	Comments (define.xml)	Mapping Information (Legacy data => SDTM)	Core
1	EG	STUDYID	Study Identifier	Char	12		CRF Page Cover	Identifier		STUDYID = "SampleXY"	Req
2	EG	DOMAIN	Domain Abbreviation	Char	2		Derived	Identifier	DOMAIN = "EG"	DOMAIN = "EG"	Req
3	EG	USUBJID	Unique Subject Identifier	Char	14		Derived	Identifier	USUBJID = Concatenation of STUDYID, DM.SUBJID separated by "_"	USUBJID = STUDYID || "_" || <name of subject variable in legacy format>	Req
4	EG	EGSEQ	Sequence Number	Num	8		Derived	Identifier	Unique sequence number within the dataset per USUBJID	1. Sort data set by key variables USUBJID, EGTESTCD, VISITNUM, EGTPTNUM and EGDTC 2. Assign a unique number within the subject	Req
5	EG	EGGRPID	Group ID	Char				Identifier		Not needed	Perm
6	EG	EGREFID	ECG Reference ID	Char				Identifier		Not needed	Perm
7	EG	EGSPID	Sponsor-Defined Identifier	Char				Identifier		Not needed	Perm
8	EG	EGTESTCD	ECG Test or Examination Short Name	Char	7		CRF Pages 5, 45	Topic		Transpose ECG parameter variables in <source dataset> to records and apply CDISCc Terminology for test coeds and names as follows: <source variable>: EGTESTCD, EGTEST): HR: "HR", "Heart Rate" PR: "PR", "PR Duration"	Req
9	EG	EGTEST	ECG Test or Examination Name	Char	25		Derived	Synonym Qualifier	Long name or label associated with EGTESTCD (see value level metadata for details)	See derivation of variable EGTESTCD	Req
10	EG	EGCAT	Category for ECG	Char	11	EGCAT	CRF Pages 5, 45	Grouping Qualifier		EGCAT = "12-LEAD ECG"	Perm

WE2: CDISC SDTM Sample

- Selected variables from a SDTM questionnaire dataset

Unique Subject Identifier (USUBJID)	Domain Abbreviation (DOMAIN)	Sequence Number (QSSEQ)	Question Short Name (QSTESTCD)	Question Name (QSTEST)	Visit Name (VISIT)	Finding in Original Units (QSORRES)	Original Units (QSORRESU)	Numeric Finding in Standard Units (QSSTRESN)	Standard Units (QSSTRESU)	Derived Flag (QSDRVFL)	Date/Time of Finding (QSDTC)
01-701-1015	QS	5001	ACITM01	WORD RECALL TASK	BASELINE	3		3			2014-01-02
01-701-1015	QS	5002	ACITM02	NAMING OBJECTS AND FINGERS (REFER TO 5 C	BASELINE	1		1			2014-01-02
01-701-1015	QS	5003	ACITM03	DELAYED WORD RECALL	BASELINE	3		3			2014-01-02
01-701-1015	QS	5004	ACITM04	COMMANDS	BASELINE	0		0			2014-01-02
01-701-1015	QS	5005	ACITM05	CONSTRUCTIONAL PRAXIS	BASELINE	3		3			2014-01-02
01-701-1015	QS	5006	ACITM06	IDEATIONAL PRAXIS	BASELINE	0		0			2014-01-02
01-701-1015	QS	5007	ACITM07	ORIENTATION	BASELINE	1		1			2014-01-02
01-701-1015	QS	5008	ACITM08	WORD RECOGNITION	BASELINE	1		1			2014-01-02
01-701-1015	QS	5009	ACITM09	ATTENTION/VISUAL SEARCH TASK	BASELINE	23		23			2014-01-02
01-701-1015	QS	5010	ACITM10	MAZE SOLUTION	BASELINE	3	seconds	3	seconds		2014-01-02
01-701-1015	QS	5011	ACITM11	SPOKEN LANGUAGE ABILITY	BASELINE	1		1			2014-01-02
01-701-1015	QS	5012	ACITM12	COMPREHENSION OF SPOKEN LANGUAGE	BASELINE	1		1			2014-01-02
01-701-1015	QS	5013	ACITM13	WORD FINDING DIFFICULTY IN SPONTANEOUS S	BASELINE	1		1			2014-01-02
01-701-1015	QS	5014	ACITM14	RECALL OF TEST INSTRUCTIONS	BASELINE	1		1			2014-01-02
01-701-1015	QS	5015	ACTOT	ADAS-COG(11) Subscore	BASELINE			13		Y	2014-01-02

Source: CDISC SDTM/ADaM Pilot (qs.xpt)

WE2: Metadata – define.xml

- Annotated Case Report Form
- Datasets
- Value Level Metadata
- Computational Algorithms
- Controlled Terminology

Datasets for Study CDISC01

Dataset	Description	Class	Structure	Purpose	Keys	Location
SV	Subject Visits	TRIAL DESIGN	One record per subject per actual visit	Tabulation	STUDYID, USUBJID, SVSTDTC, SVENDTC, VISITNUM, SVUPDES	sv.xpt
TA	Trial Arms	TRIAL DESIGN	One record per element per arm	Tabulation	STUDYID, ARMCD, TAETORD	ta.xpt
TE	Trial Elements	TRIAL DESIGN	One record per element	Tabulation	STUDYID, ETCD	te.xpt
TI	Trial Inclusion	TRIAL DESIGN	One record per I/E criterion	Tabulation	STUDYID, IETESTCD	ti.xpt
TS	Trial Summary	TRIAL DESIGN	One record per parameter value	Tabulation	STUDYID, TSPARMCD, TSVAL	ts.xpt
TV	Trial Visits	TRIAL DESIGN	One record per planned visit per arm	Tabulation	STUDYID, VISITNUM, ARMCD	tv.xpt
CO	Comments	SPECIAL PURPOSE	One record per comment per subject	Tabulation	STUDYID, USUBJID, COSEQ, IDVAR, IDVARVAL	co.xpt
DM	Demographics	SPECIAL PURPOSE	One record per subject	Tabulation	STUDYID, USUBJID	dm.xpt
CM	Concomitant Medications	INTERVENTIONS	One record per medication intervention episode per subject	Tabulation	STUDYID, USUBJID, CMCAT, CMTRT, CMSTDTC, CMENDTC, CMDOSTXT, CMDOSU, CMINDC, CMDOSFRQ	cm.xpt
EX	Exposure	INTERVENTIONS	One record per constant dosing interval per subject	Tabulation	STUDYID, USUBJID, EXTRT, EXSTDTC, EXDOSE	ex.xpt
AE	Adverse Events	EVENTS	One record per event per subject	Tabulation	STUDYID, USUBJID, AESTDTC, AEENDTC, AEDECOD, AETERM	ae.xpt
DS	Disposition	EVENTS	One record per disposition status or protocol milestone per subject	Tabulation	STUDYID, USUBJID, DSSTDTC, DSCAT, DSTERM	ds.xpt

Excerpt from the CDISC Draft Metadata Submission Guidelines define.xml Sample

WE2: Review Tool: OpenCDISC (www.opencdisc.org)

CT0075	Value for IESTRESC not found in NY controlled terminology codelist	Variable values should be populated with terms found in 'No Yes Response' (C66742) CDISC controlled terminology codelist	IE	Terminology	Error	High	R5072
CT0076	Value for --STAT not found in ND controlled terminology codelist	Variable values should be populated with terms found in 'Not Done' (C66789) CDISC controlled terminology codelist	All	Terminology	Warning	Medium	IR5107
SD0001	No records in data source	Identifies domain table that has zero rows and therefore contains no data.	All	Presence	Warning	Medium	IR5000
SD0002	Null value in variable marked as Required	Required variables (where Core attribute is 'Req') cannot be null for any record.	All	Presence	Error	High	IR5001
SD0003	Invalid ISO 8601 value	Dates and times of day must conform to the ISO 8601 international standard.	All	Format	Error	High	IR5002
SD0004	Inconsistent value for DOMAIN	Domain Abbreviation (DOMAIN) variable should be consistent with the name of the dataset.	All	Consistency	Warning	Low	IR5003
SD0005	Non-unique value for SEQ	Identifies records where non-unique values for Sequence Number variable exist within a subject.	All	Consistency	Error	High	IR5004
SD0006	No baseline result in [Domain] for subject	All subjects should have at least one baseline observation (--BLFL = 'Y') in EG, LB, QS, and VS domains, except for subjects who failed screening (ARMCD = 'SCRNFAIL') or were not fully assigned to an Arm (ARMCD = 'NOTASSGN').	EG, LB, QS, VS	Presence	Warning	Low	IR5005
SD0007	Inconsistent value for Standard Unit	Identifies Short Name of Measurement, Test or Examination values where standard units value (Standard Units) is not consistent across all records	EG, LB, QS, VS	Consistency	Error	High	IR5006
SD0008	Invalid value for Preferred Term	Identifies records where the value for the Preferred Term could not be found in the MedDRA dictionary	AE	Terminology	Error	High	IR5007
SD0009	AE is Serious but no qualifiers set to 'Y'	Identifies records where Serious Event='Y' but none of Involves Cancer, Congenital Anomaly or Birth Defect, Persist or Signif Disability/Incapacity, Results in Death, Requires or Prolongs Hospitalization, Is Life Threatening, Other Medicaly Important Serious Event, or Occured with Overdose equals 'Y'	AE	Consistency	Warning	Medium	IR5008

ACCOVION

WE2: ADaM – General Principles

- Dataset design
  - Analysis ready
    → "One Statistical Procedure away“
- Main Input: SDTM
- Documentation via standard metadata concept
  - Similar but not equal to SDTM define.xml
  - Facilitate clear and unambiguous communication
    → Traceability (results => source data (SDTM) and vice versa)
    → Describe contents, source and quality
- Data & metadata
  - Usable by currently available software tools

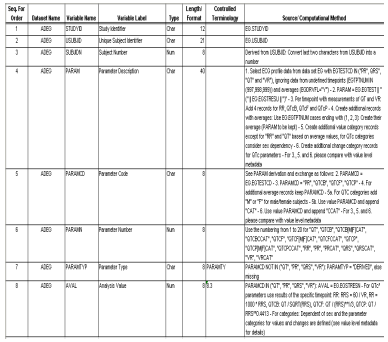
WE2: Process Flow: ADaM

SDTM

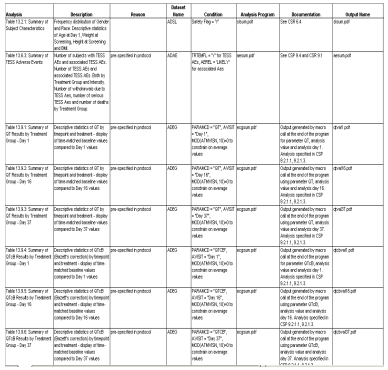
SAP

ADaM Specifications →

- Analysis datasets



- Analysis results



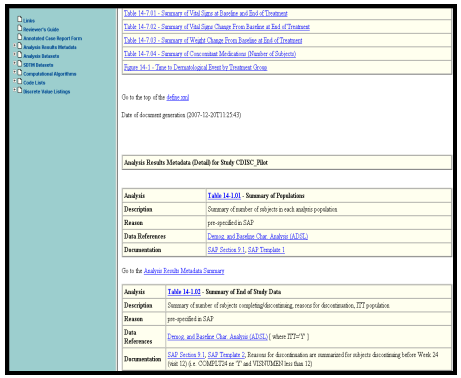
Datasets creation programs

ADaM

Results creation programs

Analysis results (TLGs)

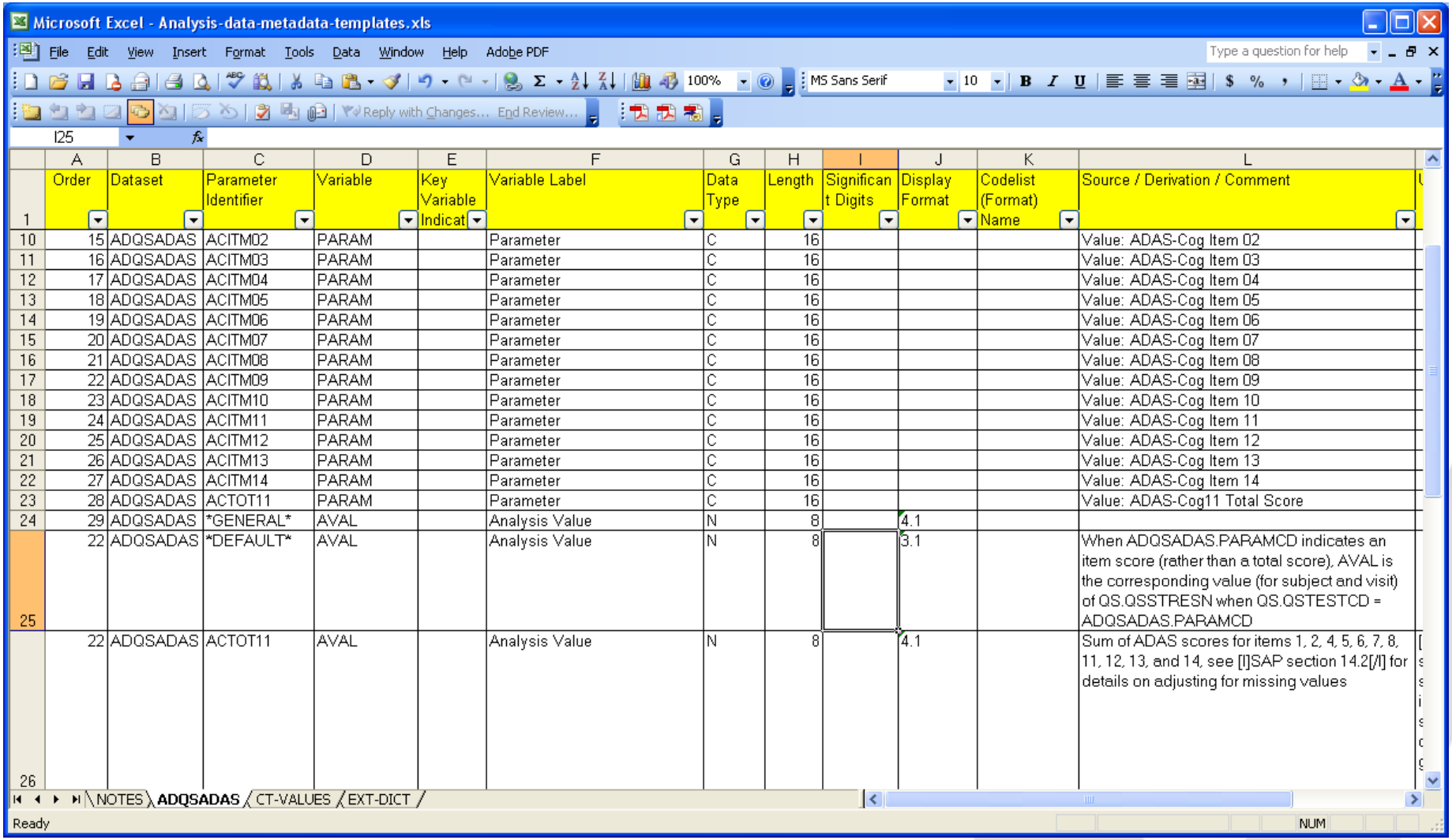
define.xml (ADaM)



WE2: Retrospective vs. Prospective

- Retrospective ADaM Generation
  - Reported results based on a different data format
  - Conversion to ADaM needed at all?
  - If yes,
    - → Focus on ADSL and key results + ensure reproducability
    - → Focus on data requirements for integrated analyses
    - → Good documentation/metadata to facilitate review
- Prospective ADaM Generation
  - Calculate study results based on ADaM datasets
  - 2 Major Requirements
    - → Efficient generation of Tables/Listings/Graphs (TLGs)
    - → Good documentation/metadata to facilitate review

WE2: Specifications/Metadata in EXCEL



Microsoft Excel - Analysis-data-metadata-templates.xls

	A	B	C	D	E	F	G	H	I	J	K	L
1	Order	Dataset	Parameter Identifier	Variable	Key Variable Indicat	Variable Label	Data Type	Length	Significant Digits	Display Format	Codelist (Format) Name	Source / Derivation / Comment
10	15	ADQSADAS	ACITM02	PARAM		Parameter	C	16				Value: ADAS-Cog Item 02
11	16	ADQSADAS	ACITM03	PARAM		Parameter	C	16				Value: ADAS-Cog Item 03
12	17	ADQSADAS	ACITM04	PARAM		Parameter	C	16				Value: ADAS-Cog Item 04
13	18	ADQSADAS	ACITM05	PARAM		Parameter	C	16				Value: ADAS-Cog Item 05
14	19	ADQSADAS	ACITM06	PARAM		Parameter	C	16				Value: ADAS-Cog Item 06
15	20	ADQSADAS	ACITM07	PARAM		Parameter	C	16				Value: ADAS-Cog Item 07
16	21	ADQSADAS	ACITM08	PARAM		Parameter	C	16				Value: ADAS-Cog Item 08
17	22	ADQSADAS	ACITM09	PARAM		Parameter	C	16				Value: ADAS-Cog Item 09
18	23	ADQSADAS	ACITM10	PARAM		Parameter	C	16				Value: ADAS-Cog Item 10
19	24	ADQSADAS	ACITM11	PARAM		Parameter	C	16				Value: ADAS-Cog Item 11
20	25	ADQSADAS	ACITM12	PARAM		Parameter	C	16				Value: ADAS-Cog Item 12
21	26	ADQSADAS	ACITM13	PARAM		Parameter	C	16				Value: ADAS-Cog Item 13
22	27	ADQSADAS	ACITM14	PARAM		Parameter	C	16				Value: ADAS-Cog Item 14
23	28	ADQSADAS	ACTOT11	PARAM		Parameter	C	16				Value: ADAS-Cog11 Total Score
24	29	ADQSADAS	*GENERAL*	AVAL		Analysis Value	N	8		4.1		
25	22	ADQSADAS	*DEFAULT*	AVAL		Analysis Value	N	8		3.1		When ADQSADAS.PARAMCD indicates an item score (rather than a total score), AVAL is the corresponding value (for subject and visit) of QS.QSSTRESN when QS.QSTESTCD = ADQSADAS.PARAMCD
26	22	ADQSADAS	ACTOT11	AVAL		Analysis Value	N	8		4.1		Sum of ADAS scores for items 1, 2, 4, 5, 6, 7, 8, 11, 12, 13, and 14, see [I]SAP section 14.2[/I] for details on adjusting for missing values

NOTES | ADQSADAS | CT-VALUES | EXT-DICT

Ready

WE2: CDISC ADaM Sample

- Selected variables from an ADaM questionnaire dataset (adqsadas.xpt of CDISC SDTM/ADaM Pilot)

Unique Subject Identifier (USUBJID)	Visit Name (VISIT)	Analysis Visit Description (AVISITC)	Intent to Treat Visit Flag (ITTV)	Imputation Type (ITYPE)	Analysis Parameter Short Name (PARAMCD)	Numeric value of PARAM (VAL)	Baseline value of VAL (BASE)	Change from baseline (VAL - BASE) (CHG)	ADaM Description of Planned Arm (TRTP)	Age in AGEU at RFSTDTC (AGE)	Sex (SEX)	Safety Population Flag (SAFETY)	Intent to Treat Population Flag (ITT)
01-701-10	BASELINE	BASELINE	Y		ACTOT	13	13	0	Placebo	63	F	Y	Y
01-701-10	WEEK 8	WEEK 8	Y		ACTOT	8	13	-5	Placebo	63	F	Y	Y
01-701-10	WEEK 16	WEEK 16	Y		ACTOT	11	13	-2	Placebo	63	F	Y	Y
01-701-10	WEEK 24	WEEK 24	Y		ACTOT	8	13	-5	Placebo	63	F	Y	Y
01-701-10	BASELINE	BASELINE	Y		ACTOT	13	13	0	Placebo	64	M	Y	Y
01-701-10	WEEK 4	WEEK 4	N		ACTOT	8	13	-5	Placebo	64	M	Y	Y
01-701-10	WEEK 4	WEEK 8	Y		ACTOT	8	13	-5	Placebo	64	M	Y	Y
01-701-10	WEEK 4	WEEK 16	Y	LOCF	ACTOT	8	13	-5	Placebo	64	M	Y	Y

WE2: Analysis Results Sample

Protocol: CDISCPILOT01 Page 1 of 1
Population: Efficacy

Table 14-3.01
Primary Endpoint Analysis: ADAS Cog (11) - Change from Baseline to Week 24 - LOCF

Display

	Placebo (N=79)	Xanomeline Low Dose (N=81)	Xanomeline High Dose (N=74)
Baseline			
n	79	81	74
Mean (SD)	24.1 (12.19)	24.4 (12.92)	21.3 (11.74)
Median (Range)	21.0 (5;61)	21.0 (5;57)	18.0 (3;57)
Week 24			
n	79	81	74
Mean (SD)	26.7 (13.79)	26.4 (13.18)	22.8 (12.48)
Median (Range)	24.0 (5;62)	25.0 (6;62)	20.0 (3;62)
Change from Baseline			
n	79	81	74
Mean (SD)	2.5 (5.80)	2.0 (5.55)	1.5 (4.26)
Median (Range)	2.0 (-11;16)	2.0 (-11;17)	1.0 (-7;13)
p-value(Dose Response) [1][2]			0.245
p-value(Xan - Placebo) [1][3]		0.569	0.233
Diff of LS Means (SE)		-0.5 (0.82)	-1.0 (0.84)
95% CI		(-2.1;1.1)	(-2.7;0.7)
p-value(Xan High - Xan Low)[1][3]			0.520
Diff of LS Means (SE)			-0.5 (0.84)
95% CI			(-2.2;1.1)

Result 1

[1] Based on Analysis of covariance (ANCOVA) model with treatment and site group as factors and baseline value as a covariate.
[2] Test for a non-zero coefficient for treatment (dose) as a continuous variable.
[3] Pairwise comparison with treatment as a categorical variable: p-values without adjustment for multiple comparisons.
Source: C:\cdisc_pilot\PROGRAMS\DRAFT\TFLs\rtf_eff1.sas 21:05 Monday, June 26, 2006

ACCOVION

WE2: Analysis Results Metadata

Metadata Field	*Definition of field*	Metadata
DISPLAY IDENTIFIER	*Unique identifier for the specific analysis display*	Table 14-3.01
DISPLAY NAME	*Title of display*	Primary Endpoint Analysis: ADAS Cog (11) - Change from Baseline to Week 24 – LOCF
RESULT IDENTIFIER	*Identifies the specific analysis result within a display*	Analysis of dose response
PARAM	*Parameter*	ADAS-Cog (11) Total Score
PARAMCD	*Parameter code*	ACTOT11
ANALYSIS VARIABLE	*Analysis variable being analyzed*	CHG
REASON	*Rationale for performing this analysis*	Primary efficacy analysis as pre-specified in protocol
DATASET	*Dataset(s) used in the analysis.*	ADQSADAS
SELECTION CRITERIA	*Specific and sufficient selection criteria for analysis subset and / or numerator*	ITTFL='Y' and AVISIT='Week 24' and PARAMCD='ACTOT11'
DOCUMENTATION	*Textual description of the analysis performed*	SAP Section 10.1.1. Linear model analysis of dose response for the ADAS-Cog(11) total score change from baseline at Week 24 - missing values imputed using LOCF, Efficacy population. Used PROC GLM in SAS to produce p-value (from Type III SS for treatment dose); Independent terms in model are TRTDOSE (randomized dose: 0 for placebo; 54 for low dose; 81 for high dose) SITEGR1 (site group, as a class variable) and BASE (baseline ADAS-Cog score).
PROGRAMMING STATEMENTS	*The analysis syntax used to perform the analysis.*	

Other interesting tasks

- Programming of integrated databases, e.g. ISS, ISE, ...
- Development and validation of generic macros, tools, …
- CDISC conversion consultancy
- Optional: presenting at conferences

Profile of a Statistical Programmer

- Programming skills and enjoy programming
- Some statistical knowledge and / or medical knowledge
- Ability to deal with data
- Communication skills
- English skills
- Office skills
- Team player

Questions?

Nicola Tambascia

Senior Statistical Programmer

www.accovion.com

Accovion GmbH
Helfmann-Park 10
D-65760 Eschborn, Germany
Tel. +49 6196 7709-283
nicola.tambascia@accovion.com

Sven Greiner

Senior Statistical Programmer

www.accovion.com

Accovion GmbH
Helfmann-Park 10
D-65760 Eschborn, Germany
Tel. +49 6196 7709-288
sven.greiner@accovion.com