CDISC Italian User Network 2021

Virtual Event | 3 December 2021

SIMeF
Società Italiana di Medicina Farmaceutica
Italian Society of Pharmaceutical Medicine

IBIG
Italian Biostatistics Group

sas

cdisc

ADaM traceability in a respiratory trial

Paola Vaghi

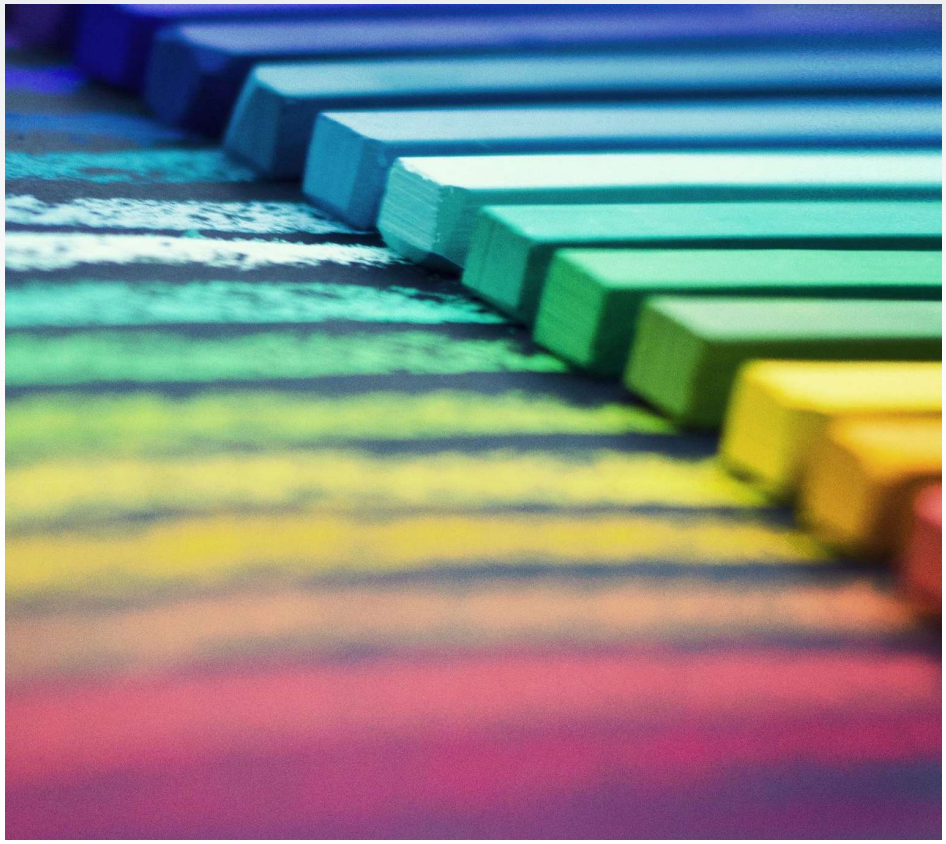
Head of Statistical Programming

Chiesi Farmaceutici S.p.A.

3 December 2021

Agenda

- Why traceability
- Reference Guidelines
- When traceability
- Traceability examples
- Take home messages



Why traceability

- An important component of regulatory review is an understanding of the provenance of the data
- Traceability permits an understanding of the relationships between the analysis results (tables, listings and figures in the study report), analysis datasets, tabulation datasets, and source data
- If the reviewer is unable to trace study data from data collection to the analyses, the regulatory review of a submission may be compromised

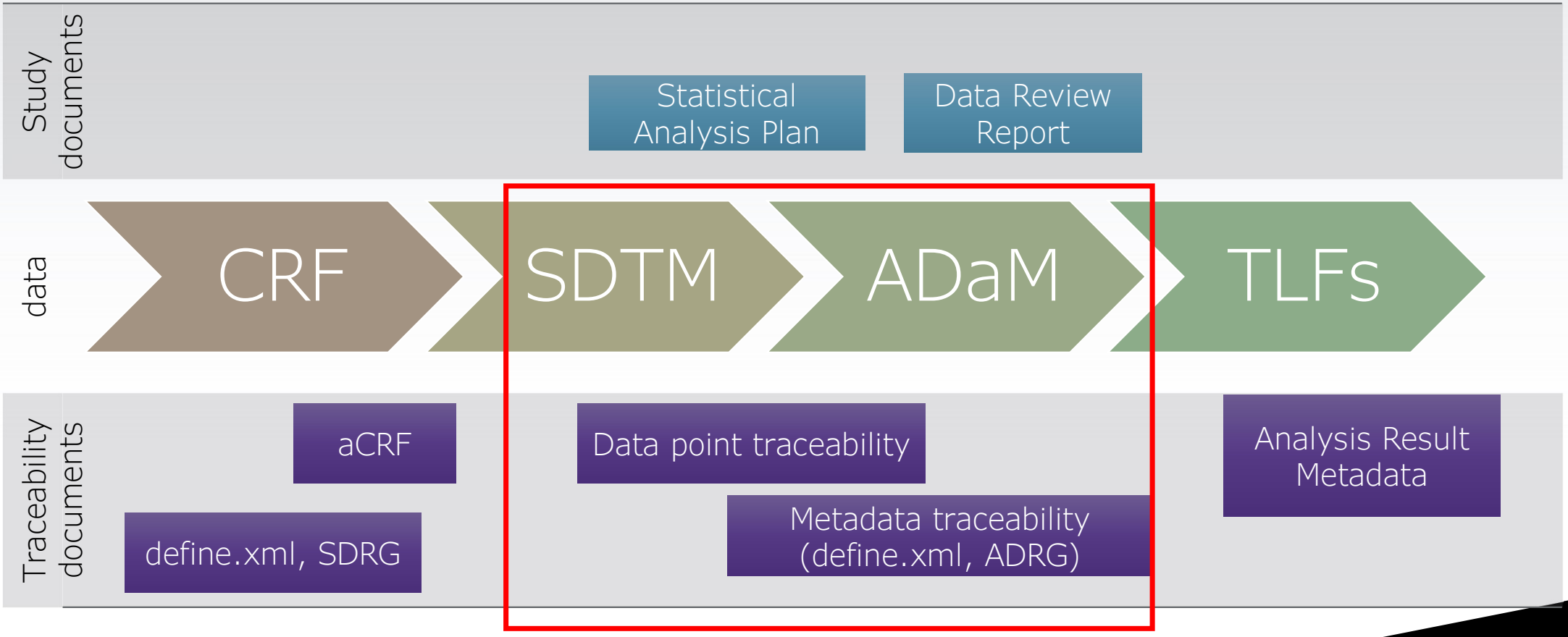
FDA U.S. FOOD & DRUG ADMINISTRATION

STUDY DATA
TECHNICAL CONFORMANCE GUIDE

Respiratory trials Traceability – Reference Guidelines

CDISC ADaM Guidelines	• ADaM IG v1.1 package • ADaM IG v1.2 package can be followed for new studies
CDISC TAUG	• COPD TAUG v1.0 • Asthma TAUG v1.0
Regulatory Authorities	• **FDA** Study Data Technical Conformance Guide v4.8.1 (October 2021) • **China** Guideline on the Submission of Clinical Trial Data (July 2020) • **Pmda** Revision of Basic Principles on Electronic Submission of Study Data for New Drug Applications (March 18, 2020)

Traceability of the results back to CRF data



Traceability in ADaM-When?

ALWAYS

- Key dates
- When data are excluded from the analyses
  - According to rules defined in the SAP
  - Due to deviations identified during the Data Review Meeting
- On the efficacy parameters, especially when
  - the efficacy assessments are to be derived from observed data
  - Multiple merging of SDTM datasets is required
  - In case of estimands
- When the analysis is focused on a selection of cases (for example a subset of TEAEs of special interest)

Traceability in ADSL

- Common ADSL variables are copied from SDTM
- Other ADSL variables are derived within the ADSL dataset

For many ADSL variables metadata traceability is sufficient

Metadata traceability should clearly identify the source variables and should describe the steps followed to populate the variable

Traceability in ADSL– An Example on key dates

Pay attention: Unclear specifications are not good traceability

Source variables should be clearly identified

Variable	Label	Origin / Source / Method / Comment
RANDDT	Date of Randomization	Derived: Randomisation date

Copy&paste of the SAS code should be avoided

Variable	Label	Origin / Source / Method / Comment

RANDDT	Date of Randomization	data rand;
		set sdtm.ds;
		where dsdecod eq 'RANDOMIZED';
		
		if length(dsstdtc)=10 then
		randdt=input(dsstdtc,is8601da.);
		else if length(dsstdtc) gt 10 then
		randdt=datepart(input(dsstdtc,is8601dt.));
		run;

Provide the source and detailed steps involved in the variable derivation (english language)

Variable	Label	Origin / Source / Method / Comment
RANDDT	Date of Randomization	Derived: Numeric version of DS.DSSTDTC when DS.DSDECOD="RANDOMIZED". If any part of the date is missing do not impute

Traceability in efficacy endpoints (BDS structure)

Efficacy Variable: Pre-dose morning FEV1 values at each visit will be summarised by treatment group using descriptive statistics

Data point traceability: each BDS record can retain SDTM variables to identify the source SDTM records and to help verifying records

AVISIT	PARAMCD	AVAL	BASE	CHG	RESTRESN	RESEQ
Visit 1 (Week -4)	FEV1	0.739	0.541	0.198	0.739	1
Visit 1 (Week -4)	FEV1	0.834	0.541	0.293	0.834	2
Visit 2 (Week 0)	FEV1	0.541	0.541	0	0.541	3
Visit 3 (Week 4)	FEV1	0.617	0.541	0.076	0.617	4
Visit 4 (Week 12)	FEV1	0.619	0.541	0.078	0.619	5
Visit 5 (Week 18)	FEV1	0.604	0.541	0.063	0.604	6
Visit 6 (Week 24)	FEV1	0.529	0.541	0.002	0.529	7

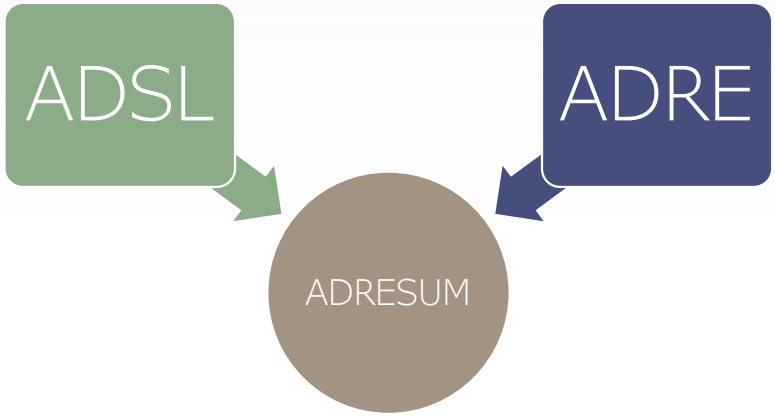
RESTRESN		Numeric Result/Finding in Standard Units	float	BEST12		Predecessor: RE.RESTRESN
BTRGR		BTR Grade after OverRead	text	13		Predecessor: RE.BTRGR
RESEQ		Sequence Number	integer	8		Predecessor: RE.RESEQ

Traceability when multiple datasets are merged

Efficacy Endpoint: FEV1 response at Visit 6:

- *Responder: change from baseline in pre-dose morning FEV1>=100mL;*
- *Non-responder: change from baseline in pre-dose morning FEV1<100mL;*
- *Non-responder: missing FEV1 at Visit 6*

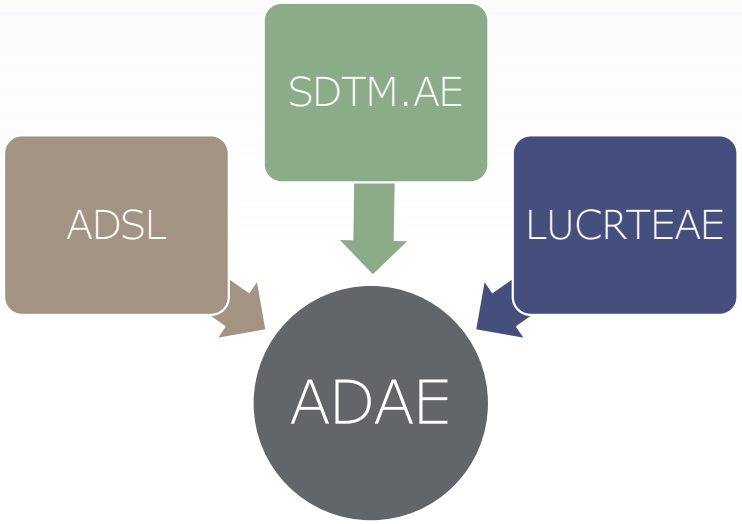
Data point traceability: as ADaM datasets are used as input for another ADaM dataset SRCDOM, SRCVAR, SRCSEQ should be used to guarantee traceability



SUBJID	AVALC	AVALCAT1	SRCDOM	SRCVAR	SRCSEQ
[redacted]2002	Non-Responder	Non-Responder	ADRE	CHGCAT1	14
[redacted]2003	Non-Responder	Non-Responder	ADRE	CHGCAT1	14
[redacted]2005	Non-Responder	Non-Responder	ADRE	CHGCAT1	14
[redacted]2006	Responder	Responder	ADRE	CHGCAT1	18
[redacted]2007	Missing	Non-Responder	ADSL		.
[redacted]2009	Missing	Non-Responder	ADSL		.

Traceability when using a look-up table

- *Class-related TEAEs: The number of class-related TEAEs and the number and percentage of subjects who experienced at least one class-related TEAE will be presented by treatment*
- A look-up table is created with the list of Class Related AEDECOD avoiding a long list of «if-then-else» programming statement



- The look-up table is mentioned in the define.xml

CQ01NAM	Populated by merging SDTM.AE with the look-up table LUCRTEAE. Set to "CLASS RELATED TEAE" if the preferred term in one of those listed in the look-up table; null otherwise

- The look-up table is described in ADRG

1.4 Source Data Used for Analysis Dataset Creation

The ADaM datasets were derived from SDTM version 1.4. The datasets were derived from the final locked database.

In addition to the clinical database, the source data include:

- a lookup file that was used to include the medication common names associated with medication preferred names to be used in the tables.
- a lookup file that was used to identify patients with a pacemaker or with atrial fibrillation.
- a lookup file that was used to calculate SGRQ total and domain scores.

Details about these lookup tables are in the Appendix. The lookup table spreadsheets were converted to SAS transport files and included with the ADaM datasets.

- Following FDA guidance, the look-up table is submitted as a SAS transport file within the submission package

Traceability in case of timepoints exclusion from the analyses

In documents

Efficacy: change from baseline in FEV1 AUC 0-12h

- The following timepoints should be excluded from SAP:
  - in case of more than two missing consecutive timepoints the AUC should be missing
  - In case of more than three timpeoints the AUC should be missing
  - etc
- From Data Review Report:
  - from the PP analysis timepoints collected after 6h from rescue intake will be excluded

In ADaM

- An algorithm is applied to exclude records as per SAP rule
- Timepoints to be excluded as per DRR are mapped into SDTM.DV
- To improve traceability a look-up table can be created from the table in DRR to clearly identify timepoints to be excluded
- SDTM.RE, SDTM.DV and the look-up table are joined to identify both
  - Timepoints to be included in AUC derivations
  - AUC that cannot be derived for more than allowed missing timepoints

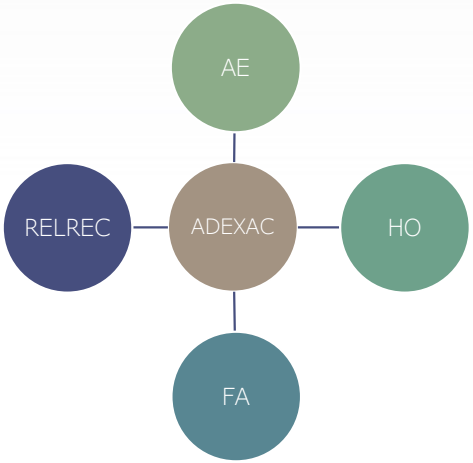
Patient ID	Deviation Term	Deviation Category	Deviation Subcategory	Period	Day	Comment
[redacted]004	Rescue Medication use during serial Spirometry	Major	Local	2	28	Spirometries at 23.5H and 24H are taken after intake of rescue medication. Trough FEV_1 is not considered in the PP analysis.

Complex traceabilty example

- In case the parameter is heavly derived from collected data, multiple approaches should be put in place to guarantee the traceability

Efficacy endpoint in respiratory trials: the rate of Moderate and Severe exacerbations

- Exacerbation data as in the eCRF are far away from exacerbations variables used in the analyses
- Exacerbations in SDTM.AE are combined with multiple SDTMs with related informations -> ADEXAC
- The rate of moderate and severe exacerbations is stored in ADXASUM -> ADXASUM



ADEXAC (OCCDS)

SUBJID	AESEQ	AESPID	AETERM
[redacted] 2002	1	1	COPD Exacerbation
2002	2	2	COPD Exacerbation
2003	1	1	COPD Exacerbation
2003	2	2	COPD Exacerbation
2007	2	1	COPD Exacerbation
3001	1	1	COPD Exacerbation
3005	1	1	COPD Exacerbation

ADXASUM (BDS)

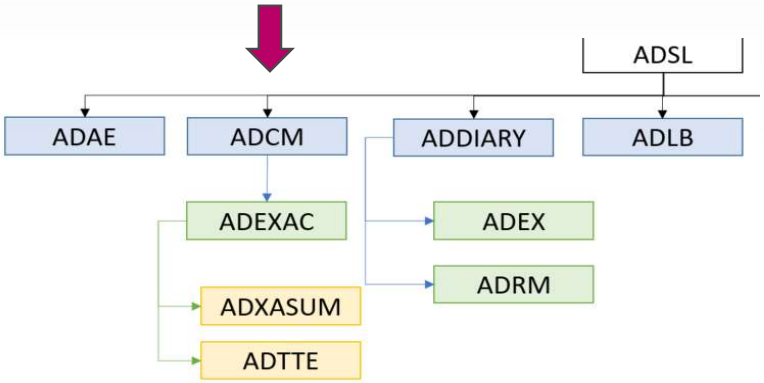
SUBJID	PARAM	PARAMCD
[redacted] 2002	Number of Moderate/Severe Exacerbations	MSEX
2002	Number of Moderate Exacerbations	MODEX
2002	Number of Severe Exacerbations	SEVEX
2002	Number of Exacerbations Leading to Death	DEATHEX
2002	Number of Exacerbations Requiring Hospitalisation	HOSPEX
2002	Moderate/Severe Exacerbations Individual Rate	RMSEX
2002	Moderate Exacerbations Individual Rate	RMODEX
2002	Severe Exacerbations Individual Rate	RSEVEX
2002	Exacerbations Leading to Death Individual Rate	RDEATHEX
2002	Exacerbations Requiring Hospitalisation Individual	RHOSPEX

Complex traceabilty example – Respiratory exacerbations

Traceability using and intermediate dataset

Example : The number and percentage of exacerbations treated with systemic corticosteroids and antibiotics are presented by treatment group

- The intermediate dataset ADCM is derived to clearly identify records taken from SDTM.CM that should be joined ot SDTM.AE



ADCM (OCCDS structure)

- Analysis flags to indentify:
  - all the medications related to an exacerbation (ANL01FL)
  - systemic corticosteroids for an exacerbation (ANL02FL)
  - antibiotics for an exacerbations (ANL03FL)

ADEXAC (OCCDS structure)

- ESATRT is populated from the analyses flags derived in ADCM

SUBJID	AESEQ	AETERM	ESATRT
[redacted] 2002	1	COPD Exacerbation	Systemic Corticosteroids and Antibiotics
2002	2	COPD Exacerbation	Systemic Corticosteroids and Antibiotics
2003	1	COPD Exacerbation	Systemic Corticosteroids and Antibiotics
2003	2	COPD Exacerbation	Systemic Corticosteroids and Antibiotics
2007	2	COPD Exacerbation	Antibiotics Only

Complex traceabilty example – Respiratory exacerbations

Traceability in combined episodes

Two COPD exacerbations will be considered as a single episode in the statistical analysis if:

- *the second exacerbation started less than 10 days after the end of the systemic corticosteroids and/or antibiotics intake for the previous exacerbation*
- *the second exacerbation started less than 10 days after the onset of the previous exacerbation*

SUBJID	AESEQ	AESPID	EGRPID	ANL02FL
[redacted]2002	1	1	A	Y
[redacted]2002	2	2	A	

All exacerbations are collected in ADaM.ADEXAC

An analysis flag on the first occurrence within a combined episode allows the right count of exacerbations

EGRPID groups exacerbations that should be a single episode

Complex traceabilty example – Respiratory exacerbations

Traceability with estimands

- *Primary analysis- treatment policy strategy: n. of moderate and severe exacerbations occurring during the planned weeks of treatment*
- *Alternative estimand – hypotethical estimand: n. of moderate and severe exacerbations based only on treatment data*

ADEXAC (OCCDS)

Two flags to identify exacerbation of each estimand

ANL01FL	Flag to identify exacerbation occurred during the overall period: […]
ANL03FL	Flag to identify exacerbation occurred during the on-treatment period: […]

ADXASUM (BDS)

Two different parameters for the count of moderate and severe exacerbation of each estimand

SUBJID	PARAM	PARAMCD	PARAMN	AVAL
[redacted]002	Number of Moderate/Severe Exa (Hypothetical estimand)	MSEXHE	1	1
[redacted]002	Num of Moderate/Severe Exa (Treatment Policy)	MSEXTP	2	2

Traceability in ADRG

A description of the dataset in ADRG can support in complex data derivations

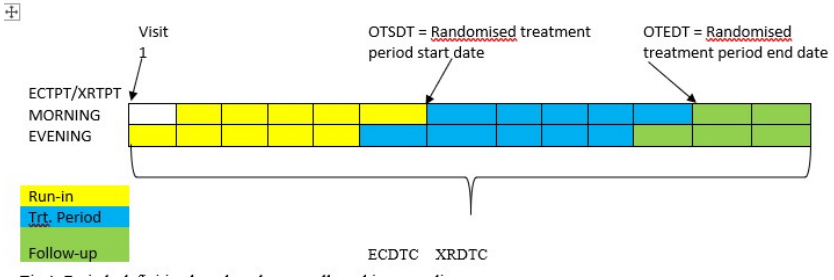
- *Compliance is heavly derived from data collected in the daily diary and the intake at the clinic*
- *A description on how daily intake have been considered for the compliance purposes compared to how they are collected in the diary is provided in ADRG*

date + 1, ECTPT/XRTPT = MORNING correspond to Visit 1 evening session).

Within Fig.2 below, a visual representation of diary data collection within SDTM datasets has been created. Each cell represents a diary session (ECTPT and XRTPT equal to MORNING/EVENING) and each column represents a diary day (XRSTDTC/ECSTDTC). Based on the SDTM datasets creation the analysis periods have been derived as follow:

'The run-in period for diary data is defined from the morning session of date of Visit 1 to the evening session of the date of start of randomised treatment period - 1 (both inclusive).

The randomised treatment period for diary data is defined from the morning session of date of start of randomised treatment period to the date of end of randomised treatment period.'



Analysis outputs programs can be submitted helping the reviewer in understanding the process by which the variables for the analyses were derived

7. Submission of Programs

All programs for analysis datasets and primary and secondary efficacy results are submitted. They were all created on a SAS platform using version 9.4. The internal reference date used to create dates in ADaM datasets is January 1, 1960.

7.1 ADaM Programs

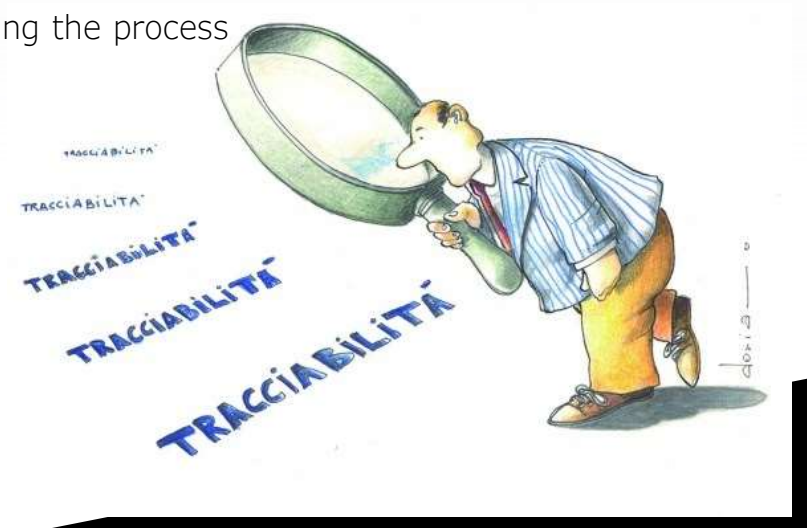
Program Name	Output	Macro Used
Aditt.sas	ADITT	Cleanwork; adsl_trtdates; date; mergesupp; adqs_sgrq; adqs_cat; adsl_phase_period; adamcheck; adamseq; adamspec
Adsl.sas	ADSL	Cleanwork; date; mergesupp; adsl_demo; adsl_pop; adsl_trtdates; adsl_phase_period; adsl_dispo; adsl_smoking; adsl_primhist; adamcheck; adamspec

7.2 Analysis Output Programs

Program Name	Output Number	Title	Input
t_re_fev1_model_itt.sas	Table 14.2.1.1.2	Table: Statistical Analysis of Change from Baseline in Pre-dose Morning FEV1 (L) (Intention-to-Treat set)	ADRE
t_re_fev1_model_pp.sas	Table 14.2.1.2.2	Table: Statistical Analysis of Change from Baseline in Pre-dose Morning FEV1 (L) (Per Protocol set)	ADRE
t_re_fev1resp model_itt.sas	Table 14.2.2.5.2	Table: Statistical Analysis of FEV1 Response at Visit 6 (Week 24) (Intention-to-Treat set)	ADRESUM

Take home messages

- Traceability is an important component in data submission; poor traceability can compromise the entire submission
- Data point traceability should be applied if feasible as it's the best way to identify the source records
- Use of an intermediate datasets, intermediate parameters and look-up tables can improve traceability. Data dependency and intermediate datasets should be described in the ADRGs; a graphical representation is better
- Metadata traceability should clearly identify the source data and should describe all the steps to be followed to populate derived data
- A description of the datasets in the ADRG can support complex data derivations
- Analysis output programs can be submitted helping the reviewer in understanding the process by which the varaibles were derived



Thank You

Paola Vaghi

P.vaghi@chiesi.com

Back-up slides

Traceability when adding a Row to a BDS structure

Safety: in case of multiple measurements associated to the same timepoint the average value will be considered for SBP and BDP

- Data point traceability: DTYPE variable is used to indicate when a new derived row has been added to the dataset and to define how the analysis values was derived

SUBJID	AVISIT	AVISITN	ADT	ATM	ADTM	PARAMCD	PARAMN	AVAL	DTYPE	ANL01FL	ANL02FL	VISITNUM	VSSEQ
[redacted]9006	Visit 1 (Week -4)	1	09OCT2019	9:30	09OCT19:09:30:00	SYSBP	1	156		Y	Y	1	13
[redacted]9006	Visit 2 (Week 0)	2	04NOV2019	8:42	04NOV19:08:42:00	SYSBP	1	163		Y		2	14
[redacted]9006	Visit 3 (Week 4)	3	04DEC2019	8:30	04DEC19:08:30:00	SYSBP	1	146		Y		3	15
[redacted]9006	Visit 5 (Week 18)	5	12MAR2020	9:08	12MAR20:09:08:00	SYSBP	1	185				5	16
[redacted]9006	Visit 5 (Week 18)	5	12MAR2020	9:24	12MAR20:09:24:00	SYSBP	1	174				5	17
[redacted]9006	Visit 5 (Week 18)	5	12MAR2020	.	.	SYSBP	1	179.5	AVERAGE	Y		.	.

DTYPE=AVERAGE when AVAl is the average of the two-pre-dose assessments

ANL01FL=Y only on records to be analysed

VSSEQ is missing for the average record

Analysis Result Metadata Traceability

- *The change from baseline of pre-dose FEV1 was analyzed by means of a MMRM with treat, visit, treat by visit interaction, […] and base by visit interaction as cov*.

Display	**14.2.1.2 Analysis of Pre-dose Morning FEV1 at Visit 6 (ITT Population)**
Analysis Result	**Analysis of Primary Efficacy Variable : Change from baseline in Pre-dose Morning FEV1 at Visit 6**
Analysis Parameter(s)	PARAMCD"FEV1" (Forced Expiratory Volume in 1 Second)
Analysis Variable(s)	CHG (Change from Baseline)
Analysis Reason	Specified in SAP
Analysis Purpose	Primary Outcome Measure
Data References (incl. Selection Criteria)	ADRE [ITTFL = ""Y" and AVISITN=6 and PARAMCD "FEV1"]
Documentation	The CFB pre-dose FEV1 was analyzed by means of a MMRM with treat, visit, treat by visit interaction, US regions, IVRS ICS dose before study and the base value and base by visit interaction as cov. SAP section 8.2

Programming Statements	[SAS version 9.3]
	`proc mixed data = ADRE NOCLPRINT=100;`
	`class usubjid TRT01Pn AVISIT REGION1 BGDC;`
	`model CHG = TRT01Pn AVISIT REGION1 BGDC TRT01Pn*AVISIT BASE BASE*AVISIT / DDFM=KR;`
	`repeated AVISIT / subject=usubjid type=UN;`
	`lsmeans TRT01Pn*AVISIT / cl OM AT MEANS;`
	`lsmestimate TRT01Pn*AVISIT "Trt A vs Placebo week 8" 0 1 0 0 0 0 0 -1 0 0,`
	`[… removed for this presentation purpose]`
	tanal_prim.sas