

AS05

Improve programming & traceability with datasets for statistical analysis results

Jeppe Rømer Juul, jxju@novonordisk.com

Henning P. Föh, hpf@novonordisk.com

Principal statistical programmers at Novo Nordisk A/S



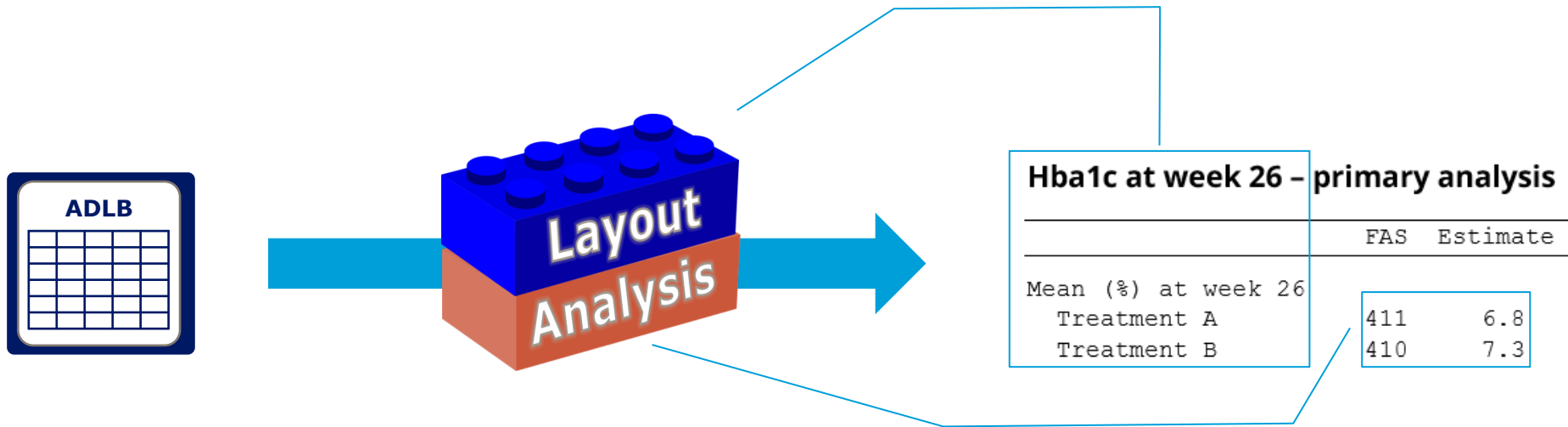
Outline

- What are analysis results datasets?
- What are the benefits?
- Dataset structure
- How to generate analysis results datasets?
- Traceability and documentation
- Authority feedback



What are *Analysis Results Datasets*?

Standard approach

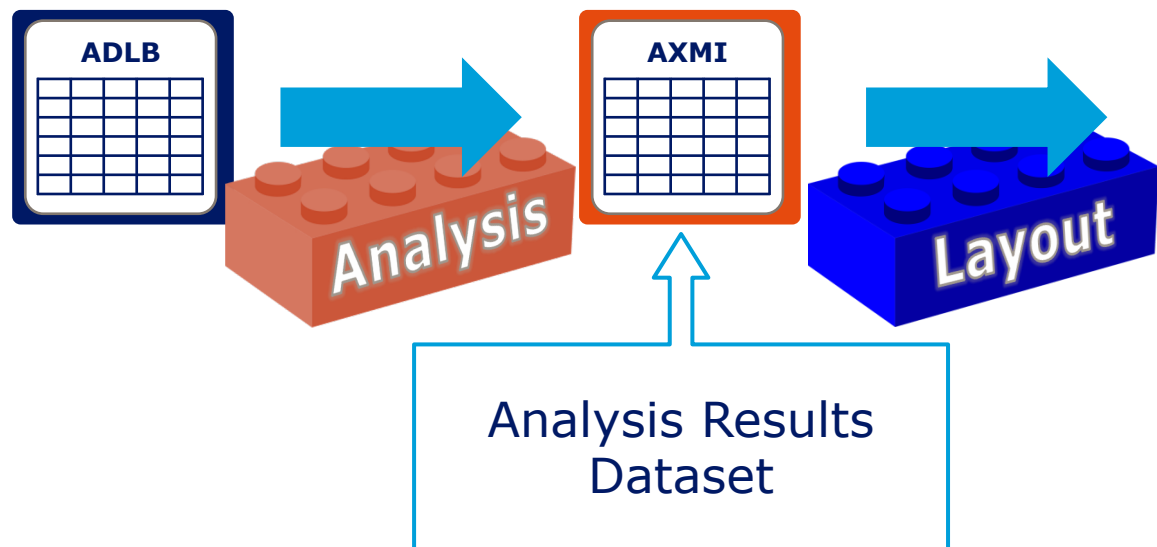


What are *Analysis Results Datasets*?

Suggested approach

Not an ADaM dataset

- *Part of output generation*

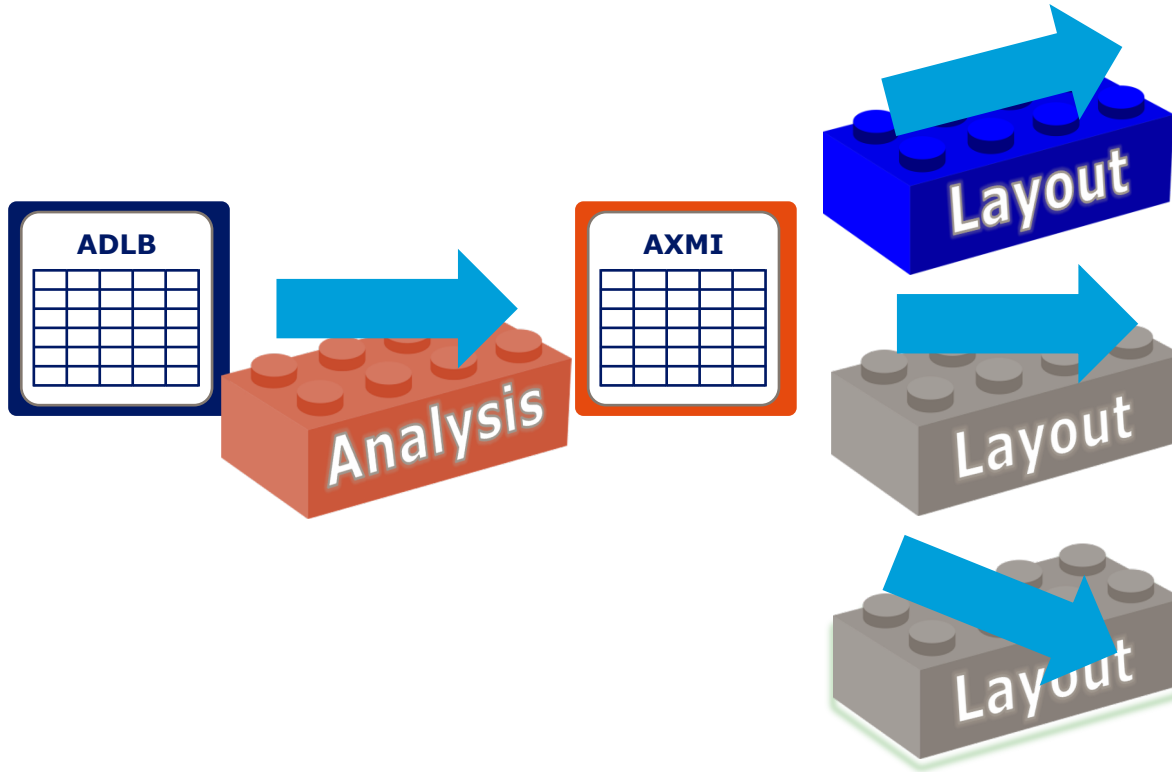


Hba1c at week 26 – primary analysis

	FAS	Estimate
Mean (%) at week 26		
Treatment A	411	6.8
Treatment B	410	7.3

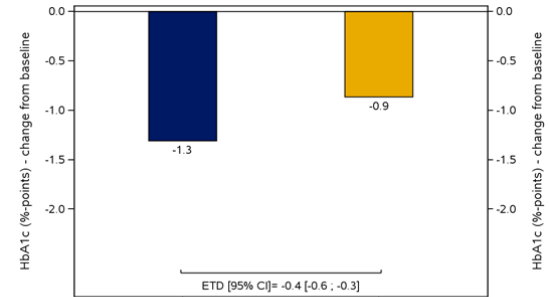
What are the benefits?

1. Analysis run once and centrally → several outputs



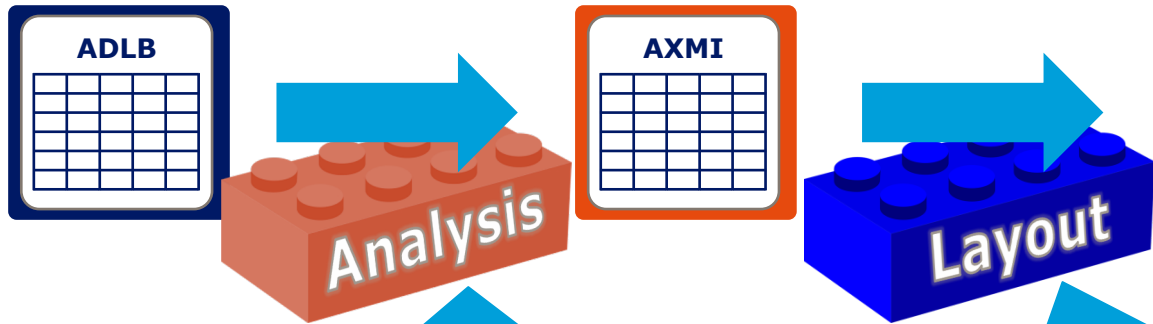
Hba1c at week 26 – primary analysis

	FAS	Estimate
Mean (%) at week 26		
Treatment A	411	6.8
Treatment B	410	7.3



What are the benefits?

2. A focused analysis program and a focused output program



Analysis part:

- Specialized to statisticians
- Important for authorities
- Parallel programmed

Layout part:

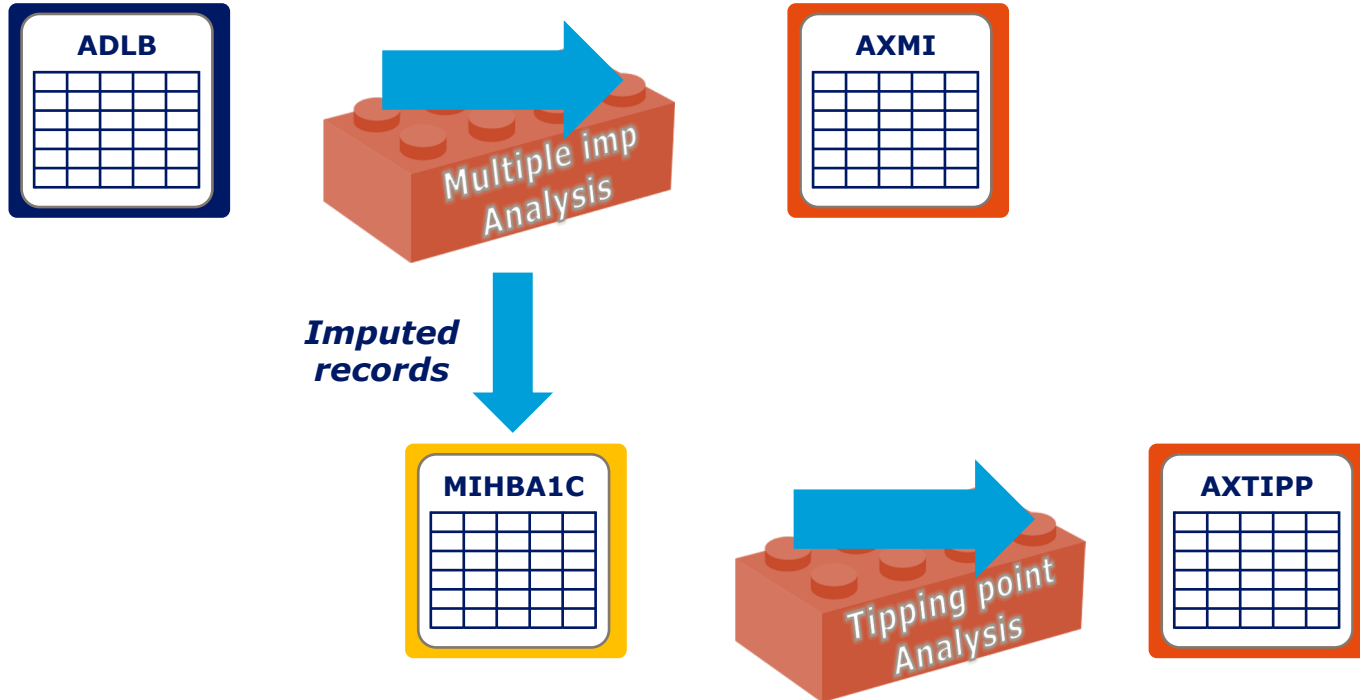
- Specialized to programmers
- Not important for authorities
- Less restrict review

Hba1c at week 26 – primary analysis

	FAS	Estimate
Mean (%) at week 26		
Treatment A	411	6.8
Treatment B	410	7.3

What are the benefits?

3. Reusability of analysis for sensitivity or new analysis



What are the benefits? - Summary

- **Separation** of *analysis* and *layouts*
- **Storage** of statistical analysis results into **structured** datasets
- **Reusability** of statistical analysis results for
 - several outputs
 - additional analysis
- Work and processing in parallel **specialized units** (stat and prog)

Dataset structure

- BDS-like: key includes PARAMCD and AVISITN but not USUBJID
- Horizontal structure: Several results per line (not one AVAL)

Data selection				Statistical model		Treatment groups and timing			Analysis results			
PARAM	FASFL	ANL01FL	ANL03FL	METHOD	CLASSVAR	TRT30P	CTRT30P	AVISIT	NOBS	EST	ESTSE	PROBT
Parameter	Full Analysis Set Popula...	Analysis Flag 01	Analysis Flag 03	Imputation method	Class. vars. used in analysis model	Planned Treatment for Period 30	Comparator, Planned Treatment	Analysis Visit	Number of subjects in analysis	Mean estimate	Mean estimate - standard error	P-value
HbA1c (%)	Y		Y	Missing at random	STRAT1V	Trial drug A		Visit 2 (week 0)	290	8.2253032929	.	.
HbA1c (%)	Y		Y	Missing at random	STRAT1V	Trial drug A	Trial drug B	Visit 7 (week 16)	290	6.6503219872	0.0419042074	1.9609973E-7
HbA1c (%)	Y		Y	Missing at random	STRAT1V	Trial drug A	Trial drug B	Visit 9 (week 30)	290	6.5044692278	0.0472182671	3.507689E-24
HbA1c (%)	Y		Y	Missing at random	STRAT1V	Trial drug B		Visit 2 (week 0)	287	8.2253032929	.	.
HbA1c (%)	Y		Y	Missing at random	STRAT1V	Trial drug B		Visit 7 (week 16)	287	6.9615276652	0.0427045614	.
HbA1c (%)	Y		Y	Missing at random	STRAT1V	Trial drug B		Visit 9 (week 30)	287	7.1935762647	0.0489427098	.
HbA1c (%)	Y	Y		Treatment status	STRAT1V	Trial drug A		Visit 2 (week 0)	290	8.2253032929	.	.
HbA1c (%)	Y	Y		Treatment status	STRAT1V	Trial drug A	Trial drug B	Visit 7 (week 16)	290	6.6964760637	0.0433465762	2.7640825E-7
HbA1c (%)	Y	Y		Treatment status	STRAT1V	Trial drug A	Trial drug B	Visit 9 (week 30)	290	6.6870276335	0.0514322166	4.834904E-15
HbA1c (%)	Y	Y		Treatment status	STRAT1V	Trial drug B		Visit 2 (week 0)	287	8.2253032929	.	.
HbA1c (%)	Y	Y		Treatment status	STRAT1V	Trial drug B		Visit 7 (week 16)	287	7.015036798	0.0442264295	.
HbA1c (%)	Y	Y		Treatment status	STRAT1V	Trial drug B		Visit 9 (week 30)	287	7.2663651999	0.053055141	.

How to generate analysis results datasets?

Programs should be:

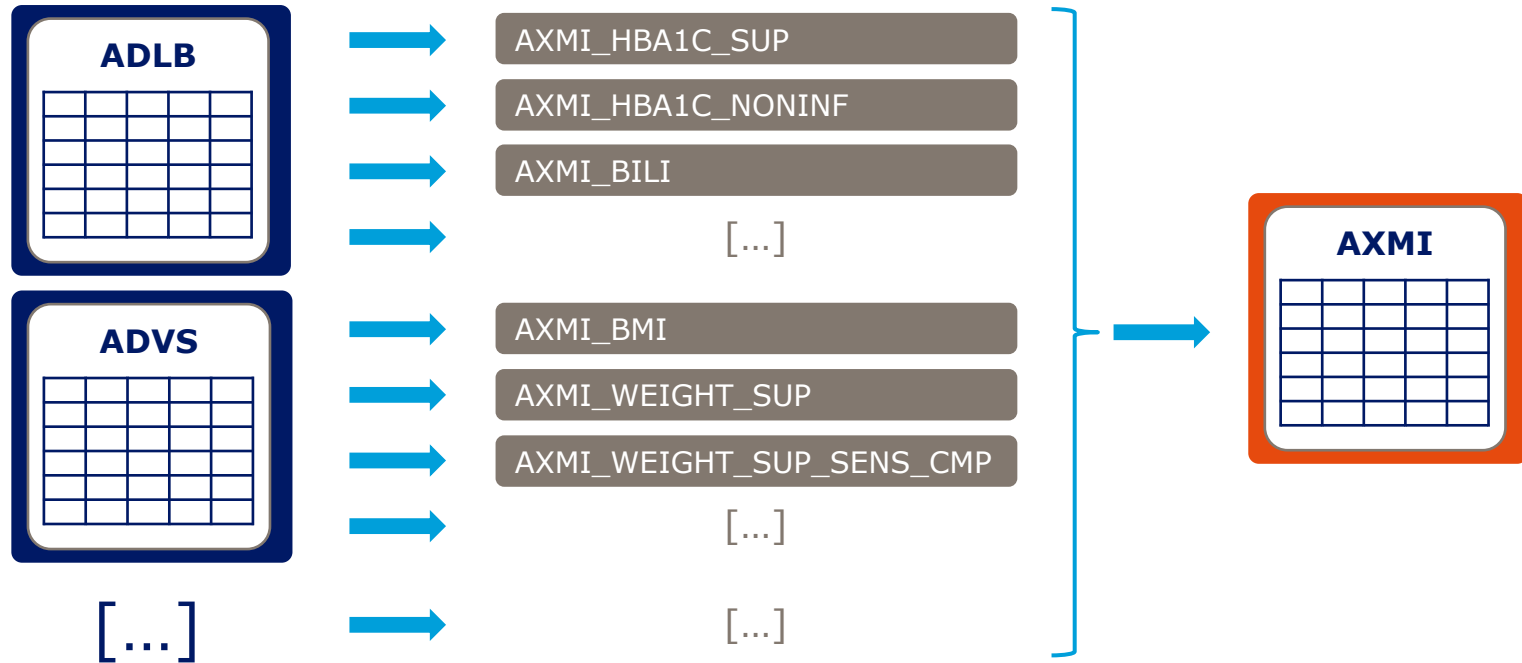
- Short and simple to read
- Fast to execute
- Single or few PARAMs and analyses per program

Datasets should be:

- Few, so you know where to look
- Basis for forest plots where multiple results are summarised
- Multiple PARAMs and analyses per dataset

Solution: Many programs making intermediate analysis results datasets, which are then stacked

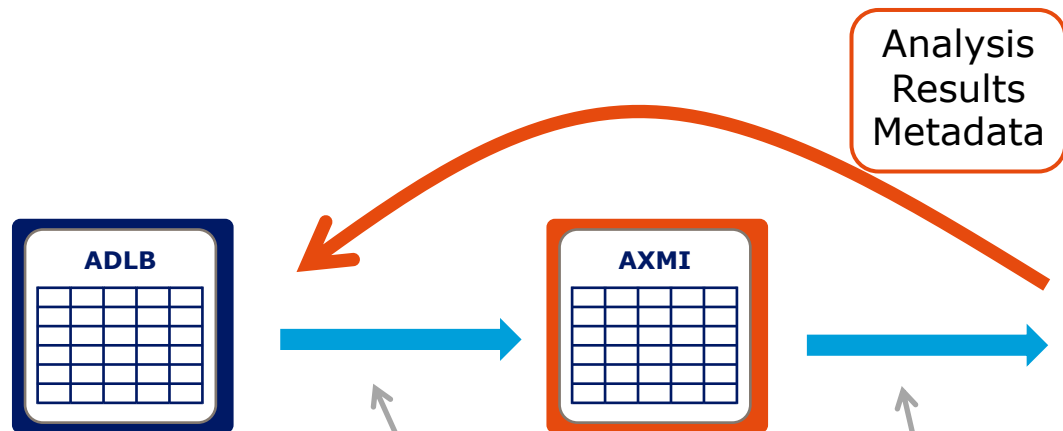
How to generate analysis result datasets?



Traceability and Documentation

- What will be submitted to authorities:
 - ✓ Selected analysis results programs (generating analysis results datasets)
 - ✓ Analysis results metadata linking to these programs
 - ✓ Description of programs in ADRG section 7: "Submitted programs"
- What will not be submitted:
 - ✗ Analysis results dataset as .xpt
 - ✗ Layout programs
 - ✗ Dataset description in define-xml

Traceability and Documentation



Hba1c at week 26 – primary analysis

	FAS	Estimate
Mean (%) at week 26		
Treatment A	411	6.8
Treatment B	410	7.3

Submit:

- ✓ ADaM dataset
- ✓ Analysis results programs

Do not submit:

- ✗ Layout programs
- ✗ Analysis results dataset

Traceability and Documentation

Analysis Results Metadata - Detail

table 14.2.3

Display	HbA1c at week 26 - primary analysis
Analysis Result	
Analysis Parameter(s)	PARAMCD = "HBA1C"
Analysis Variable(s)	ADLB.AVAL (Analysis Value)
Analysis Reason	SPECIFIED IN SAP
Analysis Purpose	PRIMARY OUTCOME MEASURE
Data References (incl. Selection Criteria)	ADLB [PARAMCD = "HBA1C" and ANL02FL = "Y"] .
Programming Statements	[SAS version 9.4] ax-mmrm2-mre-prim.txt 

Authority feedback

- Novo Nordisk has used Analysis Results Datasets in several projects submitted to e.g. FDA and PMDA
- Positive feedback received
- Note: the approach has been refined over time

3.1 Data and Analysis Quality

Data and analysis programs provided by the applicant were consistently well organized.

Source: FDA statistical review and evaluation document
www.accessdata.fda.gov/drugsatfda_docs/nda/2019/213051Orig1s000StatR.pdf

Summary

- Analysis Results Datasets improves performance and consistency of statistical analysis programs
- Analysis results are stored in BDS-like structure
- Analysis programs are short, specialized, and submission-ready
- Analysis results are simple to document in analysis results metadata
- Positive feedback received from FDA and PMDA



For implementation examples and details:

- **PP08** *Generation of statistical analysis programs made easy (and transparently)*
by Adrian Czaban

Acknowledgement

Lars Lehmann Andersson, Gitte Dam Jepsen,
Camilla Mondrup Andreassen, Julie Kjærulff Furberg,
Jan Skowronski, Mads-Jeppe Tarp-Johansen, Jacob Bonde Jacobsen