

Generation of statistical analysis programs made easy (and transparently)

Adrian Czaban¹; Niels Zeuthen²

¹Novo Nordisk A/S, Søborg, Denmark (adcz@novonordisk.com); ²Novo Nordisk A/S, Søborg, Denmark (nzeu@novonordisk.com);

Goal

- To create submission-ready programs in a fast, transparent and user-friendly way.

Rationale

- Datasets holding the results of statistical analysis simplify the process on TLF generation and make it more consistent.

Not an ADoM dataset
• Part of output generation

- However, it is easy to fall into a trap of creating complicated macro programs that are difficult to review.
- There is a need of a setup where multiple complicated analyses can be made by simple programs.
- Computational challenges exist when running all statistical analyses in a single program. Dividing it into small tasks allows for parallelization.

Core Philosophy – the building blocks approach

A program generator approach. A program can be split into multiple sections. We consider those sections as building blocks and each program is created by combining the building blocks. For every section, we can have multiple blocks that we can use. A separate program is created for each analysis defined in the protocol/SAP, with potentially hundreds of programs generated simultaneously.

Header_end1.sas

DataSel1.sas
 DataSel2.sas

Imp1.sas
 Imp2.sas
 Imp3.sas
 Imp4.sas

Model1.sas
 Model2.sas
 Model3.sas
 Model4.sas

Out1.sas
 Out2.sas
 Out3.sas
 Out4.sas

Header

Data Selection

Imputation

Modelling

Output

Intermediate Program

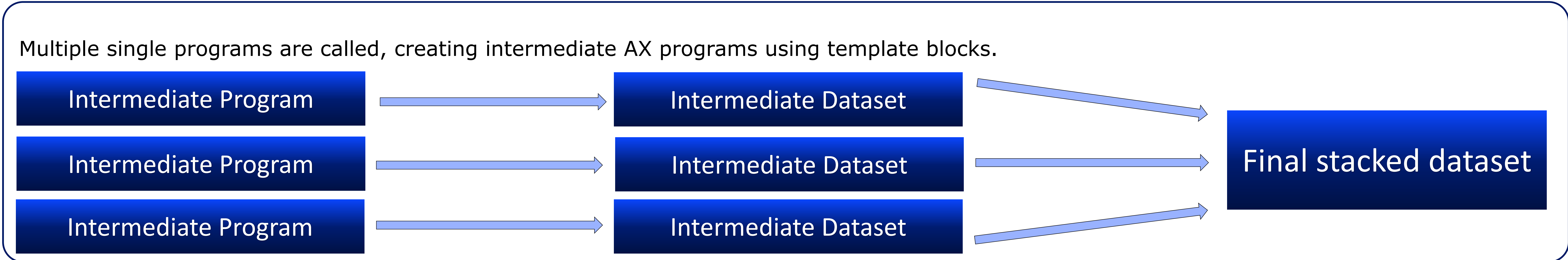
How do we control it?

- 2 levels – done in two excel sheets.
- The entire setup can be done on project level, ensuring alignment between trials.
- 1st level of control – which building blocks do we want to use?

		Header HeaderX				Data selection DataSelX				Imputation ImpX				Modelling ModelX				Output OutX			
Model/imputation type	Program part-name	STEP 1	STEP 2	STEP 3	STEP 4	STEP 1	STEP 2	STEP 3	STEP 4	STEP 1	STEP 2	STEP 3	STEP 4	STEP 1	STEP 2	STEP 3	STEP 4	STEP 1	STEP 2	STEP 3	STEP 4
ANCOVA - RD-MI	rdmi	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
LR - RD-MI	lrdmi	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2
ANCOVA - J2R-MI	j2rmi	1	1	1	1	1	1	1	1	2	2	2	2	1	1	1	1	1	1	1	1
LR - J2R-MI	lj2rmi	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2
ANCOVA - S1-SI	s1	1	1	1	1	1	1	1	1	3	3	3	3	3	3	3	3	1	1	1	1
LR - S1-SI	lrs1	1	1	1	1	1	1	1	1	3	3	3	3	4	4	4	4	2	2	2	2
ANCOVA - S2-SI	s2	1	1	1	1	1	1	1	1	3	3	3	3	3	3	3	3	1	1	1	1
LR - S2-SI	lrs2	1	1	1	1	1	1	1	1	3	3	3	3	4	4	4	4	2	2	2	2
ANCOVA - TP-MI	tpmi	1	1	1	1	1	1	1	1	1	1	1	1	5	5	5	5	3	3	3	3
LR - TP-MI	lrtpmi	1	1	1	1	1	1	1	1	1	1	1	1	6	6	6	6	4	4	4	4
MMRM1	mrm1	1	1	1	1	1	1	1	1	NA	NA	NA	NA	7	7	7	7	1	1	1	1
LR - MMRM1	lrmrm1	1	1	1	1	1	1	1	1	4	4	4	4	4	4	4	4	2	2	2	2
MMRM2	mrm2	1	1	1	1	1	1	1	1	NA	NA	NA	NA	7	7	7	7	1	1	1	1
LR - MMRM2	lrmrm2	1	1	1	1	1	1	1	1	4	4	4	4	4	4	4	4	2	2	2	2
LR - NR	lnr	1	1	1	1	1	1	1	1	5	5	5	5	4	4	4	4	2	2	2	2
LR - RD-MI - CO	lrdmico					2				6				8				5			
LR - MMRM2 - CO	lrmrm2co					2				7				9				5			

- 2nd level of control – populating macro variables inside individual blocks. A separate program is created for each analysis defined in the protocol/SAP

model	estimand	param1	dataset1	chdtype1	impfactors	analfactors	popfl	anfl	trvar_adsl	bL_visit_wk	lm_visit_wk
ANCOVA - RD-MI	Treatment policy	Combined LDL cholesterol (mg/dL)	ADLB	LOGR2BA	SEX BMIG3BL			FASFL	ANL01FL m02p	20	68
ANCOVA - RD-MI	Treatment policy	Combined LDL cholesterol (mmol/L)	ADLB	LOGR2BA	SEX BMIG3BL			FASFL	ANL01FL m02p	20	68
ANCOVA - RD-MI	Treatment policy	VLDL Cholesterol (mg/dL)	ADLB	LOGR2BA	SEX BMIG3BL			FASFL	ANL01FL m02p	20	68
ANCOVA - RD-MI	Treatment policy	VLDL Cholesterol (mmol/L)	ADLB	LOGR2BA	SEX BMIG3BL			FASFL	ANL01FL m02p	20	68
ANCOVA - RD-MI	Treatment policy	Free Fatty Acids Serum-Fasting (mg/dL)	ADLB	LOGR2BA	SEX BMIG3BL			FASFL	ANL01FL m02p	20	68
ANCOVA - RD-MI	Treatment policy	Free Fatty Acids Serum-Fasting (mmol/L)	ADLB	LOGR2BA	SEX BMIG3BL			FASFL	ANL01FL m02p	20	68
ANCOVA - RD-MI	Treatment policy	Triglycerides (mg/dL)	ADLB	LOGR2BA	SEX BMIG3BL			FASFL	ANL01FL m02p	20	68
ANCOVA - RD-MI	Treatment policy	Triglycerides (mmol/L)	ADLB	LOGR2BA	SEX BMIG3BL			FASFL	ANL01FL m02p	20	68
ANCOVA - RD-MI	Treatment policy	SF36 V2.0 Acute PRO-CoRE V1 - Role Physical	ADQSSSF36	CHG	SEX BMIG3BL			FASFL	ANL01FL m02p	20	68
ANCOVA - RD-MI	Treatment policy	SF36 V2.0 Acute PRO-CoRE V1 - Bodily Pain	ADQSSSF36	CHG	SEX BMIG3BL			FASFL	ANL01FL m02p	20	68
ANCOVA - RD-MI	Treatment policy	SF36 V2.0 Acute PRO-CoRE V1 - General Health	ADQSSSF36	CHG	SEX BMIG3BL			FASFL	ANL01FL m02p	20	68



Outcome

- A large number of high quality, submission-ready programs have been created.
- No repetitive programming without macro usage
- Reviewing or adding new analyses was easy – once a certain block was validated, an analysis could be added very fast
- Analysis Results Metadata can be generated for each individual analysis, making life easy for the FDA reviewer.
- Please refer to presentation AS05 for more details.