



Effortless Table Generation: A SAS Macro For Streamlining Your Workflow

Zhen (Laura) Li

Statistical Programming Associate Director

AstraZeneca

28FEB2024



The opinions expressed in this presentation are those of the author and do not necessarily reflect the views of her company.



Introduction

1

Receive Requests

2

Set Up Programs

3

Run Programs

4

Deliver Outputs



Introduction – Continued

1

Receive Requests

A request requires:

- ☐ Eleven tables (initially)
- ☐ Monthly delivery for three years
- ☐ Non-company-standard templates
- ☐ SDTM data is unvalidated
- ☐ ADaM data is unavailable



Introduction – Continued

1

Receive Requests

Table 3: Baseline Characteristics and Treatment

		Total pts N (%)	
Pts randomized		xx	
Demographics	Age(yrs)	Median	xx
		Q1	xx
		Q3	xx
	Female	xx (xx.xx%)	
	BMI(Kg/m ²)	Median	xx.xx
		Q1	xx.xx
		Q3	xx.xx
	White	xx (xx.xx%)	
Qualifying ACS Event	NSTEMI	xx (xx.xx%)	
	STEMI	xx (xx.xx%)	
Extent of disease at start of angiography	Single vessel disease (>=70%)	xx (xx.xx%)	
	Multivessel disease (>=70%)	xx (xx.xx%)	
	No significant disease in any vessel (<70%)	xx (xx.xx%)	
	Not specified	xx (xx.xx%)	
Procedure type	PCI	xx (xx.xx%)	
	CABG	xx (xx.xx%)	
	No procedure	xx (xx.xx%)	
Time from Qualifying Event to Randomization (median/IQR) (days)		Median	xx
		Q1	xx
		Q3	xx
Past Medical History	Dyslipidaemia	xx (xx.xx%)	
	Hypertension	xx (xx.xx%)	
	Peripheral arterial occlusive disease	xx (xx.xx%)	
	Prior Myocardial Infarction (other than qualifying event)	xx (xx.xx%)	
	Percutaneous coronary intervention (other than qualifying event)	xx (xx.xx%)	
	Coronary artery bypass (other than qualifying event)	xx (xx.xx%)	
	Heart failure	xx (xx.xx%)	
	Stroke	xx (xx.xx%)	
	Transient Ischaemic Attack	xx (xx.xx%)	
	Diabetes Mellitus	xx (xx.xx%)	
Inclusion Criteria Metrics	hsCRP mg/L	Median	xx.xx
		Q1	xx.xx
		Q3	xx.xx
	NT proBNP pg/ml	Median	xx
		Q1	xx
		Q3	xx
Medication	Use of any statin at randomization	xx (xx.xx%)	
	Use of high intensity statin regimen at randomization	xx (xx.xx%)	

Table 6: Events through Follow up

	Last report date:	This report date:
pts randomized	xx	xx
Deaths	xx (xx.xx%)	xx (xx.xx%)
CV death	xx (xx.xx%)	xx (xx.xx%)
MI (STEMI or NSTEMI)	xx (xx.xx%)	xx (xx.xx%)
Unstable Angina	xx (xx.xx%)	xx (xx.xx%)
Stroke (any)	xx (xx.xx%)	xx (xx.xx%)
Ischemic Stroke	xx (xx.xx%)	xx (xx.xx%)
Hemorrhagic Stroke	xx (xx.xx%)	xx (xx.xx%)
Hospitalization for heart failure	xx (xx.xx%)	xx (xx.xx%)





How to efficiently handle this request?



Develop a macro.

Because this can reduce overall workload in the long run and improve programming quality.



Agenda



The Table Generation Macro Introduction



Advantages



Potential Enhancements & Applications



Code Structure Design & Challenges



Programming Practice



Thoughts About Automation

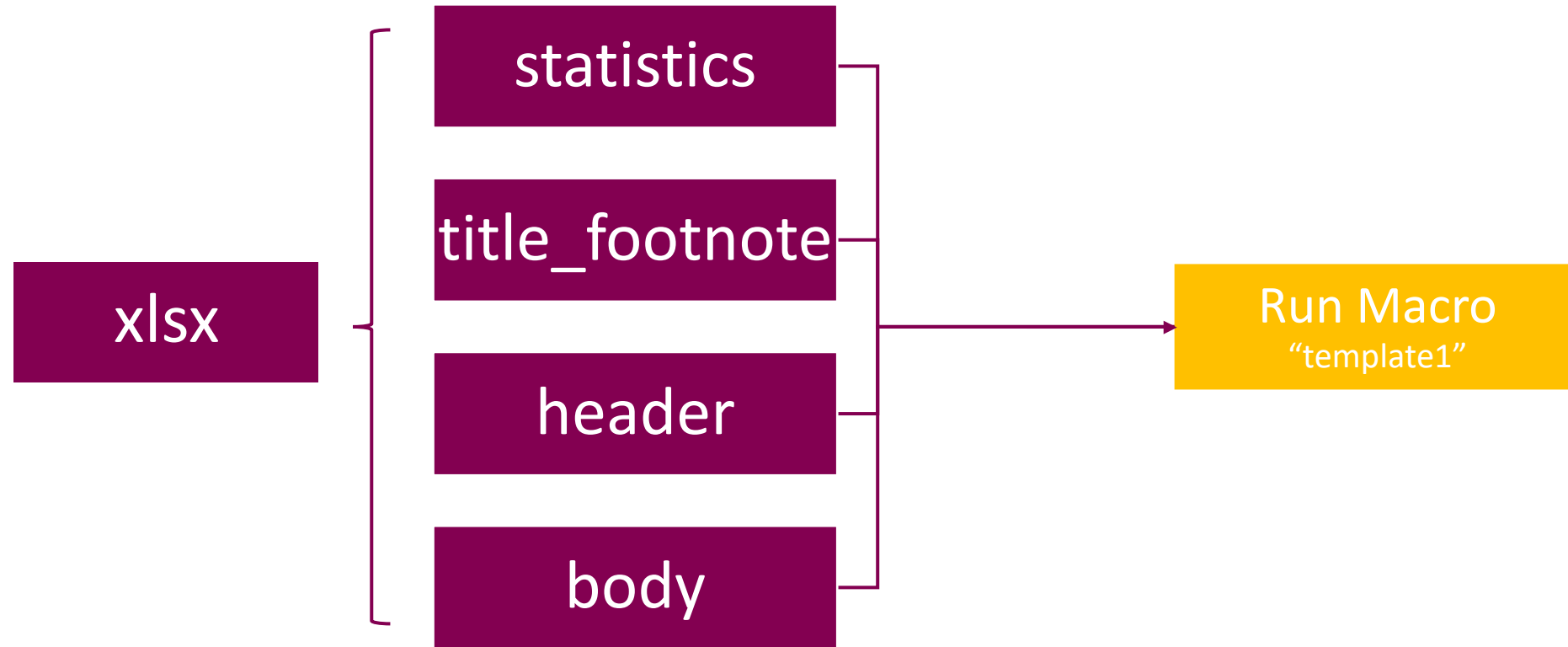


1

The Table Generation Macro Introduction



General Execution Process



Example Template

Table 3: Baseline Characteristics and Treatment			
		Total pts N (%)	
Pts randomized			xx
Demographics	Age(yrs)	Median	xx
		Q1	xx
		Q3	xx
			xx (xx.xx%)
	Female	Median	xx.xx
		Q1	xx.xx
		Q3	xx.xx
			xx (xx.xx%)
	BMI(Kg/m ²)	Median	xx.xx
		Q1	xx.xx
		Q3	xx.xx
Qualifying ACS Event	White		xx (xx.xx%)
	NSTEMI		xx (xx.xx%)
	STEMI		xx (xx.xx%)
Extent of disease at start of angiography	Single vessel disease (>=70%)		xx (xx.xx%)
	Multivessel disease (>=70%)		xx (xx.xx%)
	No significant disease in any vessel (<70%)		xx (xx.xx%)
	Not specified		xx (xx.xx%)
			xx (xx.xx%)
Procedure type	PCI		xx (xx.xx%)
	CABG		xx (xx.xx%)
	No procedure		xx (xx.xx%)
			xx (xx.xx%)
Time from Qualifying Event to Randomization (median/IQR) (days)		Median	xx
		Q1	xx
		Q3	xx
			xx (xx.xx%)
Past Medical History	Dyslipidaemia		xx (xx.xx%)
	Hypertension		xx (xx.xx%)
	Peripheral arterial occlusive disease		xx (xx.xx%)
	Prior Myocardial Infarction (other than qualifying event)		xx (xx.xx%)
	Percutaneous coronary intervention (other than qualifying event)		xx (xx.xx%)
	Coronary artery bypass (other than qualifying event)		xx (xx.xx%)
	Heart failure		xx (xx.xx%)
	Stroke		xx (xx.xx%)
	Transient Ischaemic Attack		xx (xx.xx%)
	Diabetes Mellitus		xx (xx.xx%)
			xx (xx.xx%)
	hsCRP mg/L	Median	xx.xx
		Q1	xx.xx
		Q3	xx.xx
Inclusion Criteria Metrics	NT proBNP pg/ml	Median	xx
		Q1	xx
		Q3	xx
		Median	xx
		Q1	xx
Medication	Use of any statin at randomization		xx (xx.xx%)
	Use of high intensity statin regimen at randomization		xx (xx.xx%)
			xx (xx.xx%)



Step 1: Fill Out Excel Sheets

xlsx

statistics

title_footnote

header

body

title_footnote

body

			Total pts N (%)
Pts randomized			xx
Demographics	Age(yrs)	Median	xx
		Q1	xx
		Q3	xx
	Female		xx (xx.xx%)
	BMI(Kg/m ²)	Median	xx.xx
		Q1	xx.xx
		Q3	xx.xx
	White		xx (xx.xx%)
Qualifying ACS Event	NSTEMI		xx (xx.xx%)
	STEMI		xx (xx.xx%)
Extent of disease at start of angiography	Single vessel disease (>=70%)		xx (xx.xx%)
	Multivessel disease (>=70%)		xx (xx.xx%)
	No significant disease in any vessel (<70%)		xx (xx.xx%)
	Not specified		xx (xx.xx%)
Procedure type	PCI		xx (xx.xx%)
	CABG		xx (xx.xx%)
	No procedure		xx (xx.xx%)
Time from Qualifying Event to Randomization (median/IQR) (days)		Median	xx
		Q1	xx
		Q3	xx
Past Medical History	Dyslipidaemia		xx (xx.xx%)
	Hypertension		xx (xx.xx%)
	Peripheral arterial occlusive disease		xx (xx.xx%)
	Prior Myocardial Infarction (other than qualifying event)		xx (xx.xx%)
	Percutaneous coronary intervention (other than qualifying event)		xx (xx.xx%)
	Coronary artery bypass (other than qualifying event)		xx (xx.xx%)
	Heart failure		xx (xx.xx%)
	Stroke		xx (xx.xx%)
	Transient Ischaemic Attack		xx (xx.xx%)
	Diabetes Mellitus		xx (xx.xx%)
Inclusion Criteria Metrics	hsCRP mg/L	Median	xx.xx
		Q1	xx.xx
		Q3	xx.xx
	NT proBNP pg/ml	Median	xx
		Q1	xx
		Q3	xx
Medication	Use of any statin at randomization		xx (xx.xx%)
	Use of high intensity statin regimen at randomization		xx (xx.xx%)



The Statistics Sheet

xlsx

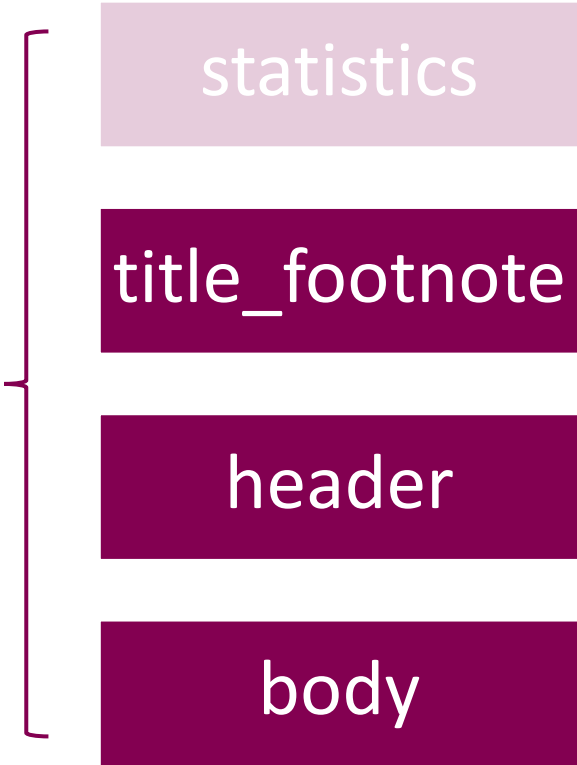


Table 3: Baseline Characteristics and Treatment

			Total pts N (%)
Pts randomized	Demographics	Age(yrs)	xx
		Median	xx
		Q1	xx
		Q3	xx
	Female		xx (xx.xx%)
		BMI(Kg/m²)	xx.xx
		Median	xx.xx
		Q1	xx.xx
	Qualifying ACS Event		xx (xx.xx%)
		STEMI	xx (xx.xx%)
		NSTEMI	xx (xx.xx%)
		White	xx (xx.xx%)
	Extent of disease at start of angiography		xx (xx.xx%)
		Single vessel disease (>=70%)	xx (xx.xx%)
		Multivessel disease (>=70%)	xx (xx.xx%)
		No significant disease in any vessel (<70%)	xx (xx.xx%)
	Procedure type		xx (xx.xx%)
		PCI	xx (xx.xx%)
		CABG	xx (xx.xx%)
		No procedure	xx (xx.xx%)
	Time from Qualifying Event to Randomization (median/IQR) (days)		xx
		Median	xx
		Q1	xx
		Q3	xx
	Past Medical History		xx (xx.xx%)
		Dyslipidaemia	xx (xx.xx%)
		Hypertension	xx (xx.xx%)
		Peripheral arterial occlusive disease	xx (xx.xx%)
		Prior Myocardial Infarction (other than qualifying event)	xx (xx.xx%)
		Percutaneous coronary intervention (other than qualifying event)	xx (xx.xx%)
		Coronary artery bypass (other than qualifying event)	xx (xx.xx%)
		Heart failure	xx (xx.xx%)
		Stroke	xx (xx.xx%)
		Transient Ischaemic Attack	xx (xx.xx%)
		Diabetes Mellitus	xx (xx.xx%)
		hsCRP mg/L	xx.xx
	Inclusion Criteria Metrics		xx.xx
		Median	xx.xx
		Q1	xx.xx
		Q3	xx.xx
			xx
		NT proBNP pg/ml	xx
		Median	xx
		Q1	xx
			xx (xx.xx%)
		Use of any statin at randomization	xx (xx.xx%)
		Use of high intensity statin regimen at randomization	xx (xx.xx%)
		Medication	xx (xx.xx%)

statistics



The Statistics Sheet – Continued

Statistics_name	Statistics_order	Statistics
stat1	1	n
stat1	2	Mean
stat1	3	SD
stat1	4	Median
stat1	5	(Min, Max)
stat3	1	Median
stat3	2	n
stat3	3	Mean
stat3		SD
		Median
		(Min, Max)
		Q1
		Q3

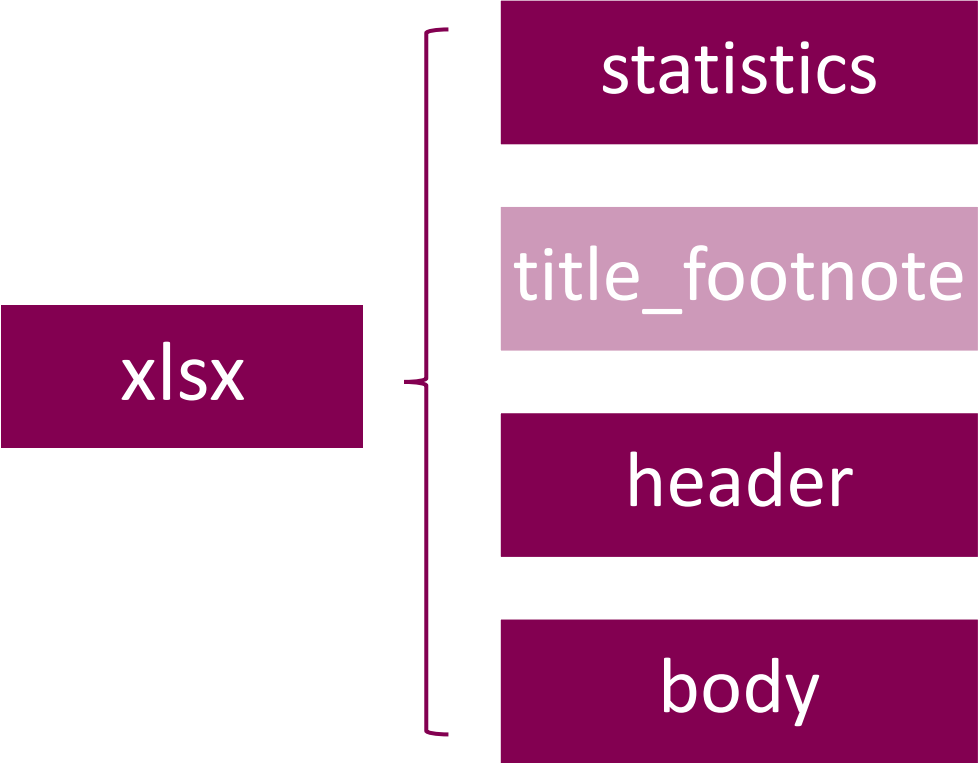
- Statistics_name: User-defined descriptive statistic groups, e.g., stat1, stat3
- Statistics_order: Order of each descriptive statistic
- Statistics: Desired descriptive statistics, e.g., n, Mean, SD, Median, (Min, Max), Q1, Q3

These three columns:

- Equipped with dropdown lists
- Not table-specific, allowing for reuse across different tables



The Title_footnote Sheet



title_footnote

Table 3: Baseline Characteristics and Treatment

		Total pts N (%)	
Pts randomized		Median	xx
Demographics	Age(yrs)	Q1	xx
		Q3	xx
	Female		xx (xx.xx%)
	BMI(Kg/m ²)	Median	xx.xx
		Q1	xx.xx
		Q3	xx.xx
Qualifying ACS Event	White		xx (xx.xx%)
	NSTEMI		xx (xx.xx%)
	STEMI		xx (xx.xx%)
Extent of disease at start of angiography	Single vessel disease (>=70%)		xx (xx.xx%)
	Multivessel disease (>=70%)		xx (xx.xx%)
	No significant disease in any vessel (<70%)		xx (xx.xx%)
	Not specified		xx (xx.xx%)
Procedure type	PCI		xx (xx.xx%)
	CABG		xx (xx.xx%)
	No procedure		xx (xx.xx%)
Time from Qualifying Event to Randomization (median/IQR) (days)		Median	xx
		Q1	xx
		Q3	xx
Past Medical History	Dyslipidaemia		xx (xx.xx%)
	Hypertension		xx (xx.xx%)
	Peripheral arterial occlusive disease		xx (xx.xx%)
	Prior Myocardial Infarction (other than qualifying event)		xx (xx.xx%)
	Percutaneous coronary intervention (other than qualifying event)		xx (xx.xx%)
	Coronary artery bypass (other than qualifying event)		xx (xx.xx%)
	Heart failure		xx (xx.xx%)
	Stroke		xx (xx.xx%)
	Transient Ischaemic Attack		xx (xx.xx%)
	Diabetes Mellitus		xx (xx.xx%)
Inclusion Criteria Metrics	hsCRP mg/L	Median	xx.xx
		Q1	xx.xx
		Q3	xx.xx
	NT proBNP pg/ml	Median	xx
		Q1	xx
		Q3	xx
Medication	Use of any statin at randomization		xx (xx.xx%)
	Use of high intensity statin regimen at randomization		xx (xx.xx%)



The Title_footnote Sheet – Continued

Fill out information
table by table

Output_name	Title1	Location_Title1	Title2	Location_Title2	Footnote1	Footnote2	Footnote3	Output_Generate
Table3	Table 3: Baseline Characteristics and Treatment	C						
Table3	XXXXXX	L	port Date: &dt	L				
		C						
		R						

- Output_name: Table's name
- Title1, Title2: Table's titles and headers
- Location_Title1, Location_Title2: Location of each title or header
 - With dropdown lists ('L', 'C', 'R')
- Footnote1, Footnote2, Footnote3: Table's footnotes
 - Multiple footnotes can be added to a single cell and separated using '^'



The Title_footnote Sheet - Continued

Fill out information
table by table

Output_name	Title1	Location_Title1	Title2	Location_Title2	Footnote1	Footnote2	Footnote3	Output_Generate
Table3	Table 3: Baseline Characteristics and Treatment	C						
Table3	XXXXXX	L	port Date: &dt	L				
		L						
		C						
		R						

- Output_Generate: Specifies whether to generate the table in .rtf format
- Default is to generate it
 - Users can suppress it by specifying 'N'
 - Output dataset is generated for all tables by default



The Header Sheet

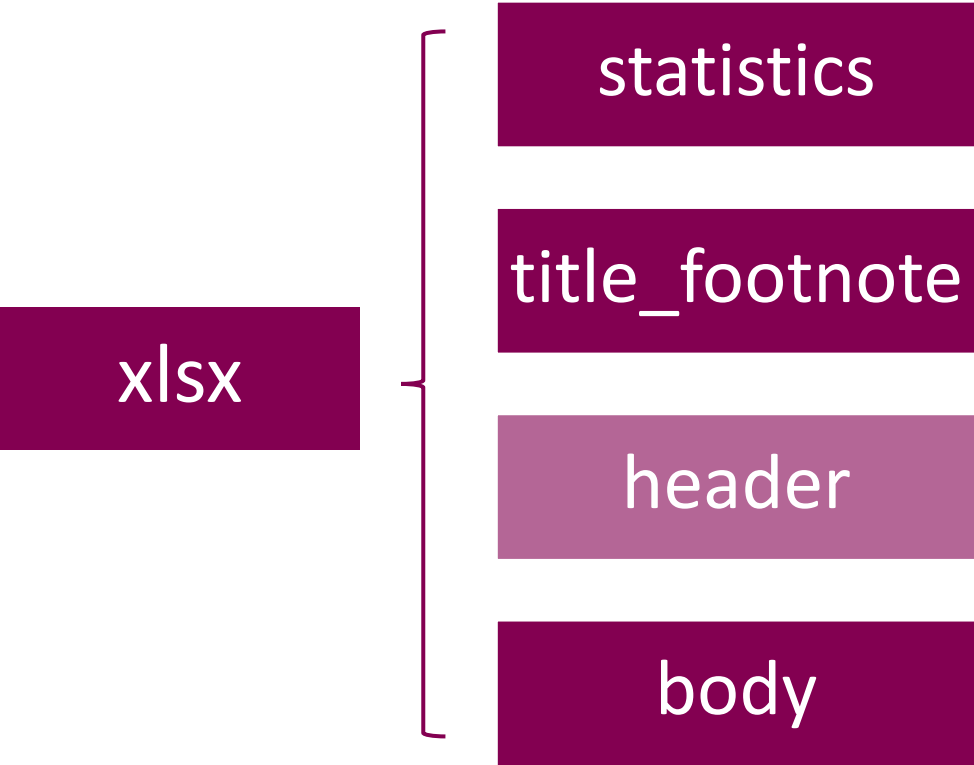


Table 3: Baseline Characteristics and Treatment

		Total pts N (%)	
Pts randomized	Demographics	Age(yrs)	Median
			Q1
			Q3
	Female		xx (xx.xx%)
		BMI(Kg/m ²)	Median
			Q1
	White		xx (xx.xx%)
		NSTEMI	xx (xx.xx%)
		STEMI	xx (xx.xx%)
Extent of disease at start of angiography	Single vessel disease (>=70%)		xx (xx.xx%)
		Multivessel disease (>=70%)	xx (xx.xx%)
		No significant disease in any vessel (<70%)	xx (xx.xx%)
		Not specified	xx (xx.xx%)
Procedure type	PCI		xx (xx.xx%)
		CABG	xx (xx.xx%)
		No procedure	xx (xx.xx%)
Time from Qualifying Event to Randomization (median/IQR) (days)			Median
			Q1
			Q3
Past Medical History	Dyslipidaemia		xx (xx.xx%)
		Hypertension	xx (xx.xx%)
		Peripheral arterial occlusive disease	xx (xx.xx%)
		Prior Myocardial Infarction (other than qualifying event)	xx (xx.xx%)
		Percutaneous coronary intervention (other than qualifying event)	xx (xx.xx%)
		Coronary artery bypass (other than qualifying event)	xx (xx.xx%)
		Heart failure	xx (xx.xx%)
		Stroke	xx (xx.xx%)
		Transient Ischaemic Attack	xx (xx.xx%)
		Diabetes Mellitus	xx (xx.xx%)
Inclusion Criteria Metrics	hsCRP mg/L		Median
			Q1
			Q3
			Median
			Q1
Medication	Use of any statin at randomization		xx (xx.xx%)
			xx (xx.xx%)
			xx (xx.xx%)
			xx (xx.xx%)
			xx (xx.xx%)

header



The Header Sheet – Continued

Fill out information
column by column

Output_name	Column_number	Column_label	Indata	Key_variable	Condition
Table3	1	Total pts N (%)	in.elig2	subject	not missing(iepatien)
Table3					
Table3					
Table6					
Table6					

- Output_name: Populate from the Title_footnote sheet
- Column_number: Order of each column
- Column_label: Label of each column
- Indata: Source data of each column
- Key_variable: Source variable of each column
- Condition: Programming condition(s) of each column



The Body Sheet

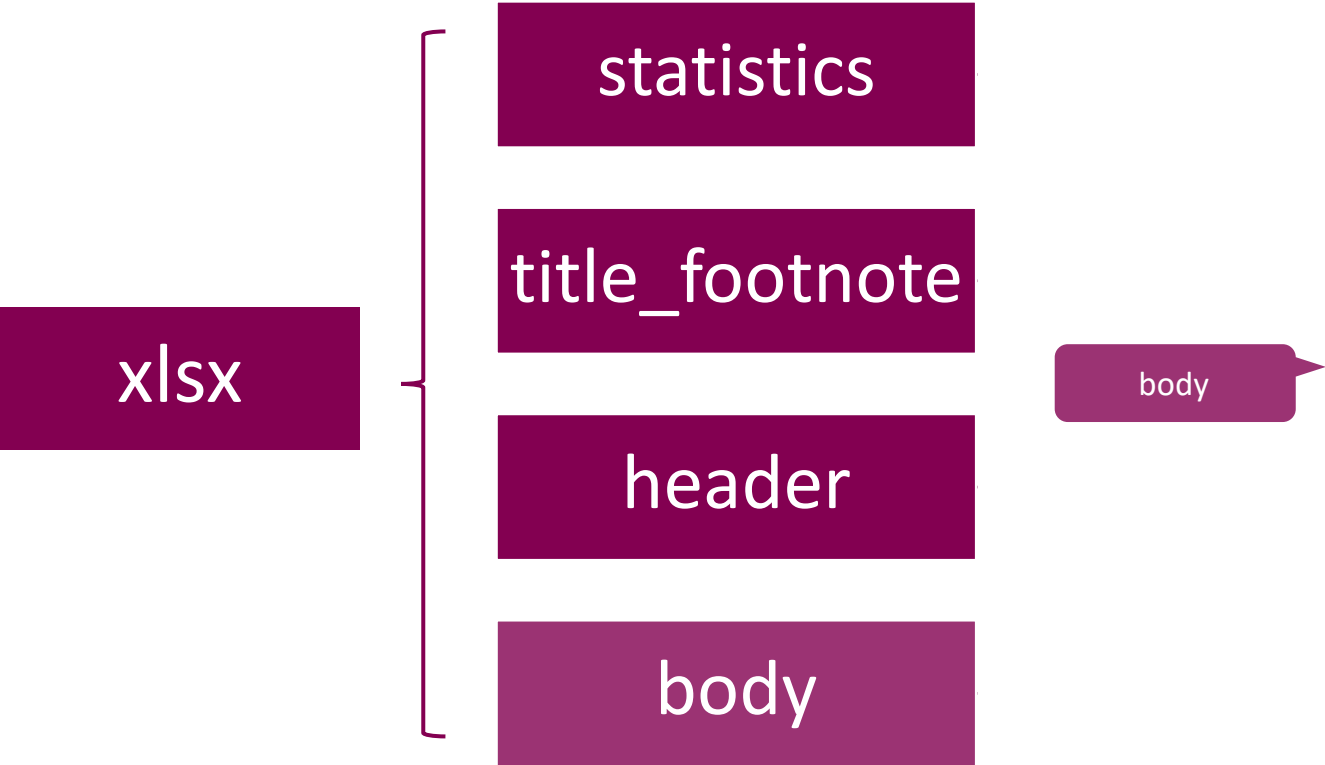


Table 3: Baseline Characteristics and Treatment

			Total pts N (%)
Pts randomized	Demographics	Age(yrs)	xx
		Median	xx
		Q1	xx
		Q3	xx
	Female		xx (xx.xx%)
		BMI(Kg/m ²)	xx.xx
		Median	xx.xx
		Q1	xx.xx
	Qualifying ACS Event		xx (xx.xx%)
		White	xx (xx.xx%)
		NSTEMI	xx (xx.xx%)
		STEMI	xx (xx.xx%)
	Extent of disease at start of angiography		xx (xx.xx%)
		Single vessel disease (>=70%)	xx (xx.xx%)
		Multivessel disease (>=70%)	xx (xx.xx%)
		No significant disease in any vessel (<70%)	xx (xx.xx%)
	Procedure type		xx (xx.xx%)
		Not specified	xx (xx.xx%)
		PCI	xx (xx.xx%)
		CABG	xx (xx.xx%)
	Time from Qualifying Event to Randomization (median/IQR) (days)		xx (xx.xx%)
		No procedure	xx (xx.xx%)
		Median	xx
		Q1	xx
	Past Medical History		xx
		Dyslipidaemia	xx (xx.xx%)
		Hypertension	xx (xx.xx%)
		Peripheral arterial occlusive disease	xx (xx.xx%)
			xx (xx.xx%)
		Prior Myocardial Infarction (other than qualifying event)	xx (xx.xx%)
		Percutaneous coronary intervention (other than qualifying event)	xx (xx.xx%)
		Coronary artery bypass (other than qualifying event)	xx (xx.xx%)
			xx (xx.xx%)
		Heart failure	xx (xx.xx%)
		Stroke	xx (xx.xx%)
		Transient Ischaemic Attack	xx (xx.xx%)
	Inclusion Criteria Metrics		xx (xx.xx%)
		Diabetes Mellitus	xx (xx.xx%)
		hsCRP mg/L	xx.xx
		Median	xx.xx
			xx.xx
		Q1	xx.xx
		Q3	xx.xx
			xx
		NT proBNP pg/ml	xx
		Median	xx
		Q1	xx
	Medication		xx
		Use of any statin at randomization	xx (xx.xx%)
		Use of high intensity statin regimen at randomization	xx (xx.xx%)
			xx (xx.xx%)



The Body Sheet – Continued

Fill out information
row by row

Output_name	Row_number	Label0	Label1	Indata	Id_variable	Summary_method	Variable1	Variable2	Condition
Table3		Is randomized		n.elig2	subject	crk	dummy		not missing(epaten)
Table3		Demographics	Age(yrs)	n.demog	subject	sta3	age		not missing(age)
Table3		Demographics	Female	n.demog	subject	perot	dummy	sexe="Female"	
Table3		Demographics	BMI(Kg/m2)	n.vs_temp	subject	sta3	zbmi		not missing(zbmi)
Table3		Demographics	White	n.demog	subject	perot	dummy	race="white"	
Table3		Qualifying ACS Event	NSTEMI	n.PROCCA	subject	perot	dummy	INDICAT="NSTEMI"	
Table3		Qualifying ACS Event	STEMI	n.PROCCA	subject	perot	dummy	INDICAT="STEMI"	
Table3		Extent of disease at start of angiography	Single vessel disease (>=70%)	n.PROCCA	subject	perot	dummy	CADISEXT_STD="1"	
Table3		Extent of disease at start of angiography	Multivessel disease (>=70%)	n.PROCCA	subject	perot	dummy	CADISEXT_STD="2"	
Table3	1	Extent of disease at start of angiography	No significant disease in any vessel (<70%)	n.PROCCA	subject	perot	dummy	CADISEXT_STD="3"	
Table3		Extent of disease at start of angiography	Not specified	n.PROCCA	subject	perot	dummy	CADISEXT_STD="4"	
Table3	1	Procedure type	PCI	n.PROCCA	subject	perot	dummy	PROCTYPE="PCI"	
Table3	1	Procedure type	CABG	n.PROCCA	subject	perot	dummy	PROCTYPE="CABG"	
Table3	1	Procedure type	No procedure	n.PROCCA	subject	perot	dummy	PROCTYPE="No procedure"	
Table3	1	Time from Qualifying Event to Randomization median(QR) (days)		n.elig2_temp	subject	sta3	qerndur		
Table3	1	Past Medical History	Dyslipidaemia	n.SPECHS	subject	perot	dummy		SPECITEM_STD="1" and spestate="Yes"
Table3	1	Past Medical History	Hypertension	n.SPECHS	subject	perot	dummy		SPECITEM_STD="2" and spestate="Yes"
Table3	1	Past Medical History	Peripheral arterial occlusive disease	n.SPECHS	subject	perot	dummy		SPECITEM_STD="3" and spestate="Yes"
Table3	1	Past Medical History	Prior Myocardial Infarction (other than qualifying event)	n.SPECHS	subject	perot	dummy		SPECITEM_STD="5" and spestate="Yes"
Table3	2	Past Medical History	Percutaneous coronary intervention (other than qualifying event)	n.SPECHS	subject	perot	dummy		SPECITEM_STD="7" and spestate="Yes"
Table3	2	Past Medical History	Coronary artery bypass (other than qualifying event)	n.SPECHS	subject	perot	dummy		SPECITEM_STD="8" and spestate="Yes"
Table3	2	Past Medical History	Heart failure	n.SPECHS	subject	perot	dummy		SPECITEM_STD="9" and spestate="Yes"
Table3	2	Past Medical History	Stroke	n.SPECHS	subject	perot	dummy		SPECITEM_STD="14" and spestate="Yes"
Table3	2	Past Medical History	Transient Ischaemic Attack	n.SPECHS	subject	perot	dummy		SPECITEM_STD="15" and spestate="Yes"
Table3	2	Past Medical History	Diabetes Mellitus	n.SPECHS	subject	perot	dummy		SPECITEM_STD in ("20", "21") and battman = "HS-CRP"
Table3	2	Inclusion Criteria Metrics	hsCRP mg/L	n.bsaf_temp	subject	sta3	crvresn		battman = "HS-CRP"
Table3	2	Inclusion Criteria Metrics	NT-proBNP ng/ml	n.bsaf_temp	subject	sta3	crvresn		battman = "HS-CRP"
Table3	2	Medication	Use of any statin at randomization	n.elig2	subject	perot	dummy		HSYN in ("No", "Yes")
Table3	2	Medication	Use of high intensity statin regimen at randomization	n.elig2	subject	perot	dummy		HSYN="Yes"

- Output_name: Populate from the Title_footnote sheet
- Row_number: Order of each row
- Label0: Label of the group (left blank if nonexistent)
- Label1: Label of the subgroup (left blank if nonexistent)
- Indata: Source data of each row
- Id_variable: Key variable used for merging the Indata with the datasets specified in the Header sheet



The Body Sheet – Continued

Fill out information
row by row

Output_name	Row_number	Label0	Label1	Indata	Id_variab	Summary_method	Variable1	Variable2	Condition
Table3	1	1 Pt randomized		in.elig2	subject	cnt	dummy		not missing(epaten)
Table3	2	2 Demographics	Age(yrs)	in.demog	subject	sta3	age		not missing(age)
Table3	3	3 Demographics	Female	in.demog	subject	perot	dummy	sex= "Female"	
Table3	4	4 Demographics	BMI(kg/m2)	in.vs_temp	subject	sta3	bmil		not missing(bmli)
Table3	5	5 Demographics	White	in.demog	subject	perot	dummy	race= "White"	
Table3	6	6 Qualifying ACS Event	STEMI	in.PROCCA	subject	perot	dummy	INDICAT= "STEMI"	
Table3	7	7 Qualifying ACS Event	STEMI	in.PROCCA	subject	perot	dummy	INDICAT= "STEMI"	
Table3	8	8 Extent of disease at start of angiography	Single vessel disease (>=70%)	in.PROCCA	subject	perot	dummy	CADISEXT_STD="1"	
Table3	9	9 Extent of disease at start of angiography	Multivessel disease (>=70%)	in.PROCCA	subject	perot	dummy	CADISEXT_STD="2"	
Table3	10	10 Extent of disease at start of angiography	No significant disease in any vessel (<70%)	in.PROCCA	subject	perot	dummy	CADISEXT_STD="3"	
Table3	11	11 Extent of disease at start of angiography	Not specified	in.PROCCA	subject	perot	dummy	CADISEXT_STD="4"	
Table3	12	12 Procedure type	PCI	in.PROCCA	subject	perot	dummy	PROCTYPE="PCI"	
Table3	13	13 Procedure type	CABG	in.PROCCA	subject	perot	dummy	PROCTYPE="CABG"	
Table3	14	14 Procedure type	No procedure	in.PROCCA	subject	perot	dummy	PROCTYPE="No procedure"	
Table3	15	15 Time from Qualifying Event to Randomization (median(IQR)) (days)		in.elig2_temp	subject	sta3	qerndur		
Table3	16	16 Past Medical History	Dyslipidaemia	in.SPECHS	subject	perot	dummy	SPECITEM_STD="1" and speccate="Yes" and speccate="Yes"	
Table3	17	17 Past Medical History	Hypertension	in.SPECHS	subject	perot	dummy	SPECITEM_STD="2" and speccate="Yes" and speccate="Yes"	
Table3	18	18 Past Medical History	Peripheral arterial occlusive disease	in.SPECHS	subject	perot	dummy	SPECITEM_STD="3" and speccate="Yes" and speccate="Yes"	
Table3	19	19 Past Medical History	Prior Myocardial Infarction (other than qualifying event)	in.SPECHS	subject	perot	dummy	SPECITEM_STD="5" and speccate="Yes" and speccate="Yes"	
Table3	20	20 Past Medical History	Percutaneous coronary intervention (other than qualifying event)	in.SPECHS	subject	perot	dummy	SPECITEM_STD="7" and speccate="Yes" and speccate="Yes"	
Table3	21	21 Past Medical History	Coronary artery bypass (other than qualifying event)	in.SPECHS	subject	perot	dummy	SPECITEM_STD="8" and speccate="Yes" and speccate="Yes"	
Table3	22	22 Past Medical History	Heart failure	in.SPECHS	subject	perot	dummy	SPECITEM_STD="9" and speccate="Yes" and speccate="Yes"	
Table3	23	23 Past Medical History	Stroke	in.SPECHS	subject	perot	dummy	SPECITEM_STD="14" and speccate="Yes" and speccate="Yes"	
Table3	24	24 Past Medical History	Transient Ischaemic Attack	in.SPECHS	subject	perot	dummy	SPECITEM_STD="15" and speccate="Yes" and speccate="Yes"	
Table3	25	25 Past Medical History	Diabetes Mellitus	in.SPECHS	subject	perot	dummy	SPECITEM_STD in ("20", "21") and battman = "HS-CRP"	
Table3	26	26 Inclusion Criteria Metrics	hsCRP mg/L	in.bsaf_temp	subj	sta3	crvresn		battman = "HS-CRP"
Table3	27	27 Inclusion Criteria Metrics	NT proBNP pg/ml	in.bsaf_temp	subj	sta3	crvresn		battman = "HS-CRP"
Table3	28	28 Medication	Use of any statin at randomization	in.elig2	subject	perot	dummy	HSYN in ("No", "Yes")	
Table3	29	29 Medication	Use of high intensity statin regimen at randomization	in.elig2	subject	perot	dummy	HSYN="Yes"	

➤ Summary_method:

- Predefined descriptive statistic groups from the Statistics sheet
- Macro-defined statistics:
 - I. Counts ('cnt')
 - II. Percentages using column totals as denominators ('perct')
 - III. Percentages using the total of all subcategories as denominators ('perctcat')



The Body Sheet – Continued

Fill out information
row by row

Output_name	Row_number	Label0	Label1	Indata	Id_variable	Summary_method	Variable1	Variable2	Condition
Table3	1	Pts randomized		in.elig2	subject	crf	dummy		not missing(epsien)
Table3	2	Demographics	Age(yrs)	in.demog	subject	sta3	age		not missing(age)
Table3	3	Demographics	Female	in.demog	subject	perot	dummy		sex="Female"
Table3	4	Demographics	BMi(kg/m2)	in.vs_temp	subject	sta3	zbmi		not missing(zbmi)
Table3	5	Demographics	White	in.demog	subject	perot	dummy		ace="white"
Table3	6	Qualifying ACS Event	NSTEMI	in.PROCCA	subject	perot	dummy		INDICAT="NSTEMI"
Table3	7	Qualifying ACS Event	STEMI	in.PROCCA	subject	perot	dummy		INDICAT="STEMI"
Table3	8	Extent of disease at start of angiography	Single vessel disease (>=70%)	in.PROCCA	subject	perot	dummy		CADISEXTLSTD="1"
Table3	9	Extent of disease at start of angiography	Multivessel disease (>=70%)	in.PROCCA	subject	perot	dummy		CADISEXTLSTD="2"
Table3	10	Extent of disease at start of angiography	No significant disease in any vessel (<70%)	in.PROCCA	subject	perot	dummy		CADISEXTLSTD="3"
Table3	11	Extent of disease at start of angiography	Not specified	in.PROCCA	subject	perot	dummy		CADISEXTLSTD="4"
Table3	12	Procedure type	PCI	in.PROCCA	subject	perot	dummy		PROCTYPE="PCI"
Table3	13	Procedure type	CABG	in.PROCCA	subject	perot	dummy		PROCTYPE="CABG"
Table3	14	Procedure type	No procedure	in.PROCCA	subject	perot	dummy		PROCTYPE="No procedure"
Table3	15	Time from Qualifying Event to Randomization (median/QQR) (days)		in.elig2_temp	subject	sta3	qemdu		
Table3	16	Past Medical History	Dyslipidaemia	in.SPECHS	subject	perot	dummy		SPECITEMLSTD="1" and specstate="Yes"
Table3	17	Past Medical History	Hypertension	in.SPECHS	subject	perot	dummy		SPECITEMLSTD="2" and specstate="Yes"
Table3	18	Past Medical History	Peripheral arterial occlusive disease	in.SPECHS	subject	perot	dummy		SPECITEMLSTD="3" and specstate="Yes"
Table3	19	Past Medical History	Prior Myocardial Infarction (other than qualifying event)	in.SPECHS	subject	perot	dummy		SPECITEMLSTD="5" and specstate="Yes"
Table3	20	Past Medical History	Percutaneous coronary intervention (other than qualifying event)	in.SPECHS	subject	perot	dummy		SPECITEMLSTD="7" and specstate="Yes"
Table3	21	Past Medical History	Coronary artery bypass (other than qualifying event)	in.SPECHS	subject	perot	dummy		SPECITEMLSTD="8" and specstate="Yes"
Table3	22	Past Medical History	Heart failure	in.SPECHS	subject	perot	dummy		SPECITEMLSTD="9" and specstate="Yes"
Table3	23	Past Medical History	Stroke	in.SPECHS	subject	perot	dummy		SPECITEMLSTD="14" and specstate="Yes"
Table3	24	Past Medical History	Transient Ischaemic Attack	in.SPECHS	subject	perot	dummy		SPECITEMLSTD="15" and specstate="Yes"
Table3	25	Past Medical History	Diabetes Mellitus	in.SPECHS	subject	perot	dummy		SPECITEMLSTD in "20", "21" and
Table3	26	Inclusion Criteria Metrics	hsCRP mg/L	in.bsaf_temp	subj	sta3	crvresn		battman = HS-CRP
Table3	27	Inclusion Criteria Metrics	NT proBNP ng/ml	in.bsaf_temp	subj	sta3	crvresn		battman =
Table3	28	Medication	Use of any statin at randomization	in.elig2	subject	perot	dummy		HSYN in ("No", "Yes"
Table3	29	Medication	Use of high intensity statin regimen at randomization	in.elig2	subject	perot	dummy		HSYN="Yes"

➤ Variable1: Source variable of each row

- A built-in character-type dummy variable ('dummy') is available

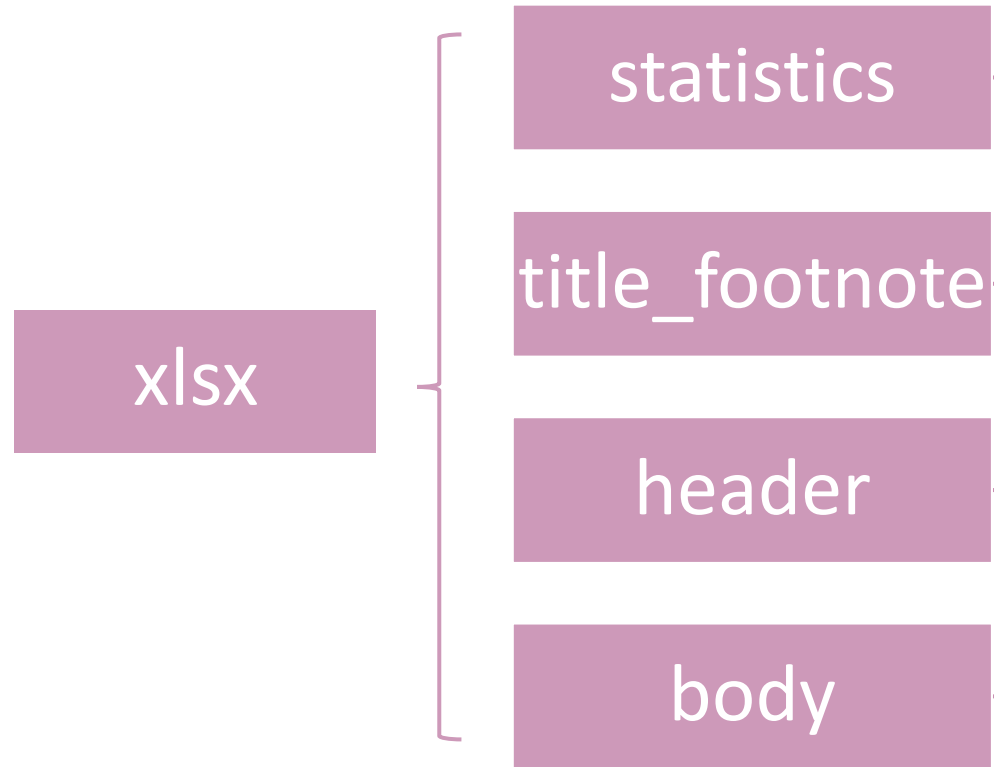
➤ Variable2: Source variable for subgroup analysis of a row (left blank if unnecessary)

- An example will be provided

➤ Condition: Programming condition(s) of each row



Step 2: Run the Macro



Step 2: Run the Macro – Continued

`%template1(Table3);`



XXXXXX
Report Date: 11MAY2022

Table 3: Baseline Characteristics and Treatment

			Total pts N (%)
Pts randomized			228
Demographics	Age(yrs)	Median	60
		Q1	54
		Q3	67
	Female	Median	38 (16.67%)
		Q1	27.78
		Q3	25.37
Qualifying ACS Event	BMI(Kg/m2)	Q3	30.97
		White	193 (84.65%)
		NSTEMI	77 (33.77%)
		STEMI	150 (65.79%)
Extent of disease at start of angiography	Single vessel disease (>=70%)	Multivessel disease (>=70%)	119 (52.19%)
		No significant disease in any vessel (<70%)	108 (47.37%)
		Not specified	4 (1.75%)
		Not specified	2 (0.88%)
Procedure type	PCI	PCI	223 (97.81%)
		CABG	0
		No procedure	4 (1.75%)
		Median	124
Time from Qualifying Event to Randomization (median/IQR) (days)		Q1	86
		Q3	199
Past Medical History	Dyslipidaemia	Dyslipidaemia	177 (77.63%)
		Hypertension	158 (69.3%)
		Peripheral arterial oclusive disease	5 (2.19%)
		Prior Myocardial Infarction (other than qualifying event)	21 (9.21%)
		Percutaneous coronary intervention (other than qualifying event)	33 (14.47%)
		Coronary artery bypass (other than qualifying event)	0
		Heart failure	24 (10.53%)
		Stroke	4 (1.75%)
		Transient Ischaemic Attack	3 (1.32%)
		Diabetes Mellitus	41 (17.98%)
Inclusion Criteria Metrics	hsCRP mg/L	Median	0.26
		Q1	0.16
		Q3	0.42
	NT proBNP pg/ml	Median	361
		Q1	140
		Q3	607
Medication	Use of any statin at randomization	Use of any statin at randomization	228 (100%)
		Use of high intensity statin regimen at randomization	180 (78.95%)



Special Considerations in this Example

Utilize Diverse Source Data

The source data includes both EDC data and TPV data. The patient ID variable is named 'subject' in EDC data, whereas it is 'subjid' in TPV data.

The Header sheet:

Output_name	Column_number	Column_label	Indata	Key_variable	Condition
Table3	1	Total pts N (%)	in.elig2	subject	not missing(iepatien)

The Body sheet:

Output_name	Row_number	Label0	Label1	Indata	Id_variable	Summary_method	Variable1	Variable2	Condition
Table3	26	Inclusion Criteria N	hsCRP mg/L	in.bsaf_temp	subjid	stat3	cnvresn		battrnam = 'HS-CRP'
Table3	27	Inclusion Criteria N	NT proBNP pg	in.bsaf_temp	subjid	stat3	cnvresn		battrnam = 'NTPROBNP'

Require Preliminary Programs

It utilizes **raw, uncleaned** data. A preliminary program may be required prior to the macro execution in certain scenarios:

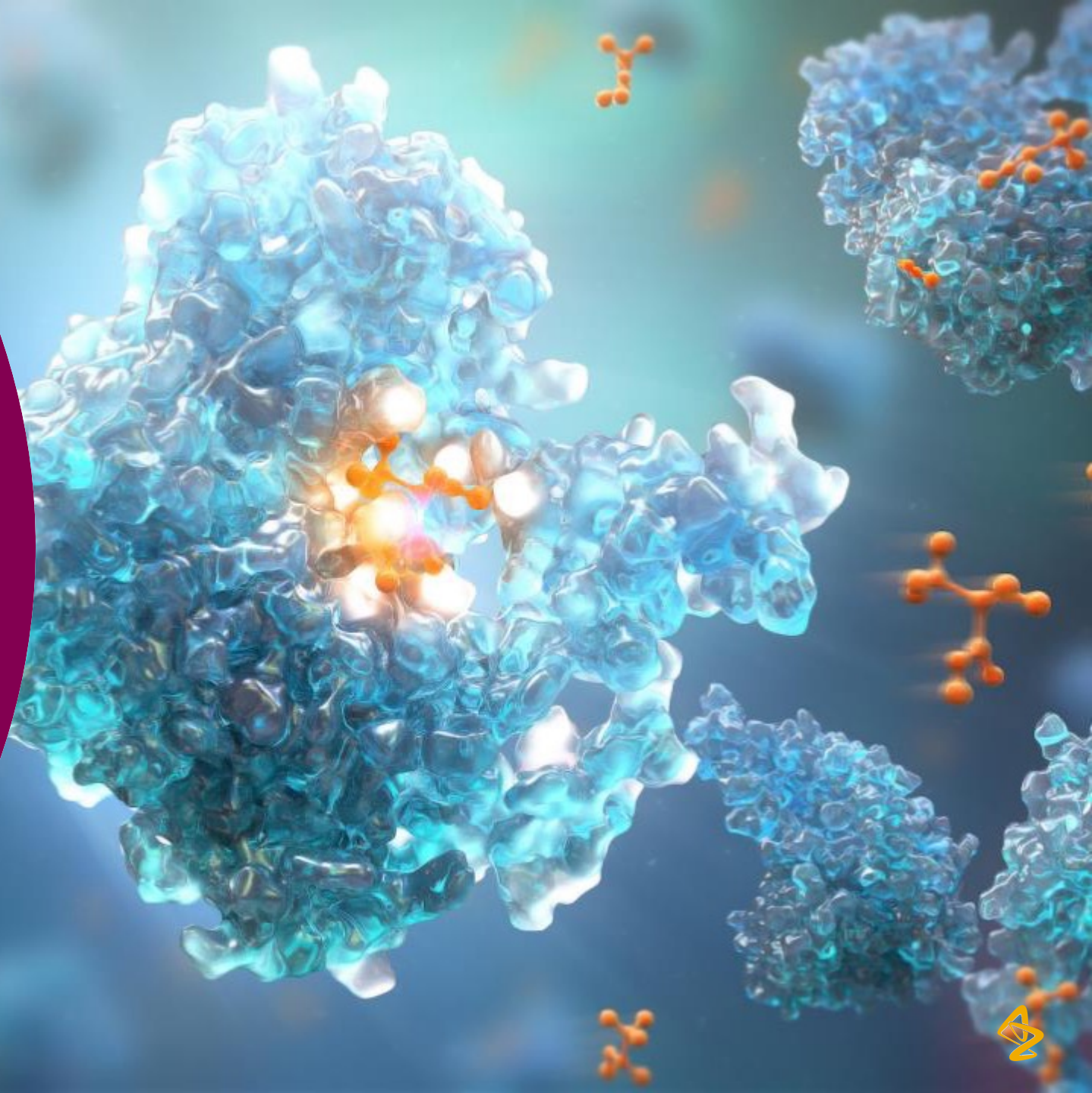
- Data Checking and Cleaning
- Variable Derivation
- Inclusion of a Macro Variable

This macro seamlessly handles such name discrepancy scenarios



2

Advantages



Flexible

- ✓ Allows the utilization of various source data (raw data, SDTM data, ADaM data, or the combinations) without worrying about name discrepancies across different datasets
- ✓ Adaptable to various table templates:

XXXXXX
Report Date: 03JAN2024

Table 7: AE by System Organ Class and Preferred Term
Safety Population

	Placebo N=185	IP N=528
Subjects with at least one adverse event	4 (2.16%)	21 (3.98%)
Blood and lymphatic system disorders	1 (0.54%)	0
Lymphadenopathy	1 (0.54%)	0
Cardiac disorders	1 (0.54%)	4 (0.76%)
Myocarditis	0	1 (0.19%)
Atrial fibrillation	0	1 (0.19%)
Cardiac failure	1 (0.54%)	0
Coronary artery disease	0	1 (0.19%)
Sinus bradycardia	0	1 (0.19%)
Tachycardia	0	1 (0.19%)

subgroups

Report Date: 11MAY2022

Table 6: Events through Follow up

	Last report date: 07APR2022	This report date: 11MAY2022
Pts randomized	205	228
Deaths	2 (0.98%)	2 (0.88%)
CV death	0	0
MI (STEMI or NSTEMI)	1 (0.49%)	1 (0.44%)
Unstable Angina	0	0
Stroke (any)	0	0
Ischemic Stroke	0	0
Hemorrhagic Stroke	0	0
Hospitalization for heart failure	1 (0.49%)	1 (0.44%)

from two data transfers

Efficiency

- ✓ Minimize programming efforts
- ✓ Streamline discrepancy investigations and communications
- ✓ Facilitate programming updates

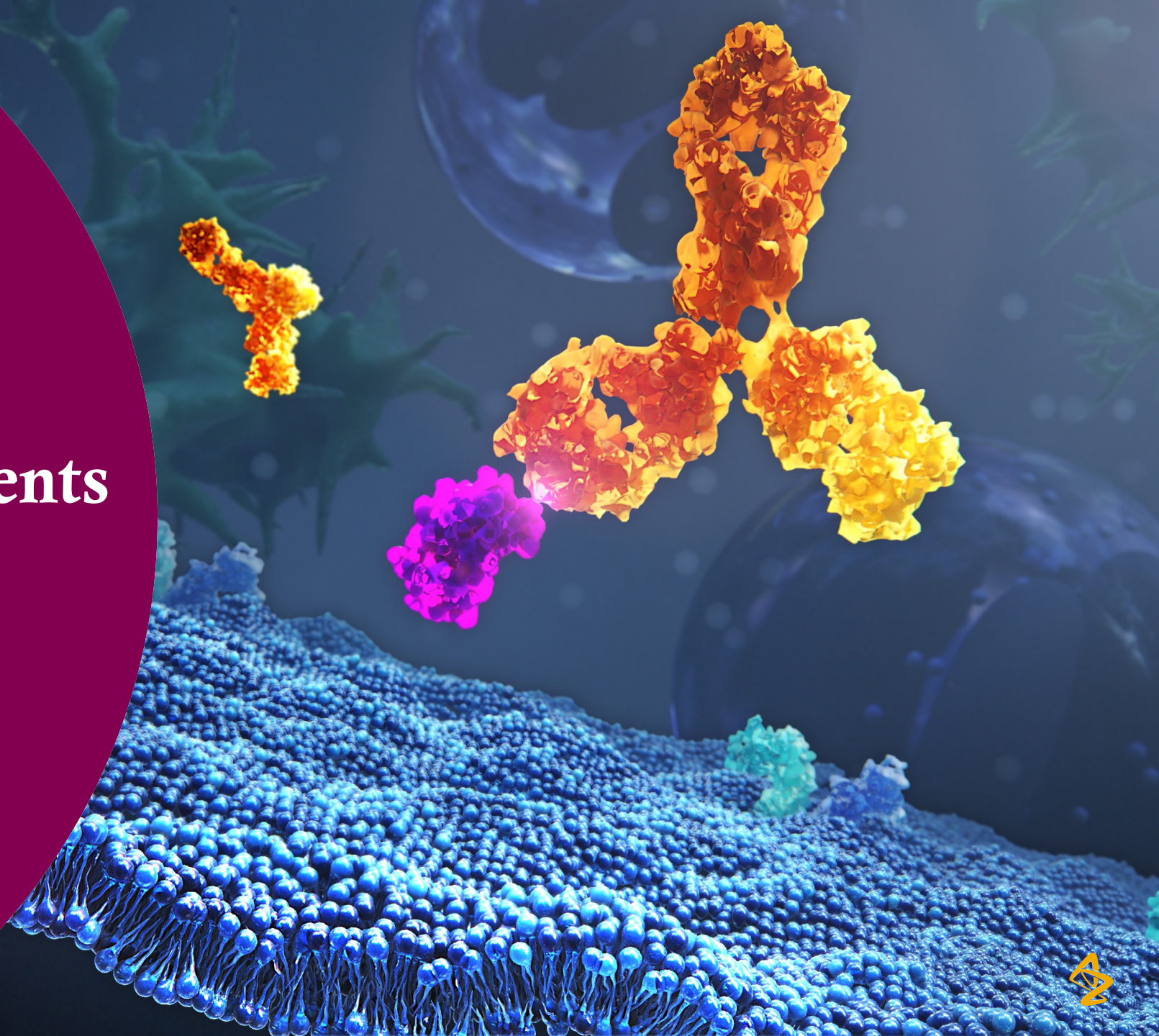
Simplicity

- ✓ Stand-alone macro
- ✓ Not platform-specific
- ✓ User-friendly approach, even for junior programmers and statisticians



3

Potential Enhancements & Applications



Potential Enhancements

Incorporate Macros for Standard Table Formats:

- Integrate macros to generate company-standard table formats

Statistical Analysis Enhancement:

- Incorporate statistical tests like the chi-squared test or t-test
- Introduce some statistical models for a deeper understanding of the data

User-Defined Formats:

- Allow users to define their own formats for variables

Categorical Variables Management:

- Include sorting options for categorical variables to enhance data organization

Missing Category Handling:

- Provide an option to generate a missing category, ensuring comprehensive data representation

Efficiency through Batch Runs:

- Enable batch runs for streamlined execution, facilitating the generation of all or multiple tables at once



Applications

Current:

- Generate tables for ad hoc requests on either the primary or validation side

Potential:

- Generate tables for formal deliverables on either the primary or validation side
- Standardize output datasets to further enhance and simplify the validation process



4

Code Structure Design & Challenges



Code Structure Design



Challenges

Non-programming

- Design columns and sheets of the Excel file, balancing comprehensiveness with user-friendliness
- Define the appropriate scope by considering the balance between the flexibility of using this macro and the complexity of the code
- Concerns regarding the incorporation of information from an Excel file for a formal deliverable:
 - Traceability Gaps: Version control
 - Transparency of the Analytical Process: Risks associated with the information not being directly embedded within the code

Programming

- Issues related to patient ID variable name discrepancies across different source data
 - Resolution: Check whether the Key_variable in the Header sheet matches the Id_variable in the Body sheet; if not, rename the Id_variable in the Body sheet to align with the Key_variable in the Header sheet
- Develop the macro in an updater/reviewer-friendly sequence to enhance robustness and facilitate understanding
- Consider code optimization for efficient execution



5

Programming Practice



Suggested Programming Techniques

Program Structure Design

Follow a modular structure:

- ❑ Group related programming tasks logically
- ❑ Avoid handling mixed cases in one section
- ❑ Ensure ease of reading, understanding, and updating

Improving Computer Processing Resource Efficiency

- ❑ Optimize by selectively reading in and retaining only the necessary datasets, variables, and observations, minimizing CPU and I/O operations. Utilize conditional logic (e.g., WHERE), and control variables using KEEP and DROP statements.
- ❑ Manage space by storing data as character when applicable.
- ❑ In most cases, joining by PROC SQL is more efficient than the merging operation in the data steps.

Streamlining Debug/Update/Review Processes

- ❑ Improve code clarity by incorporating comments, documentation, and maintaining well-formatted code.
- ❑ Assign meaningful names and labels to datasets and variables to aid reviewers, code updaters, and validation programmers.
- ❑ Employ DO loops or macro programs to streamline repetitive tasks and minimize the risk of updating programs for new data.
- ❑ Consider the utilization of macro variables when assigning values that may undergo changes.

Enhancing Overall Robustness

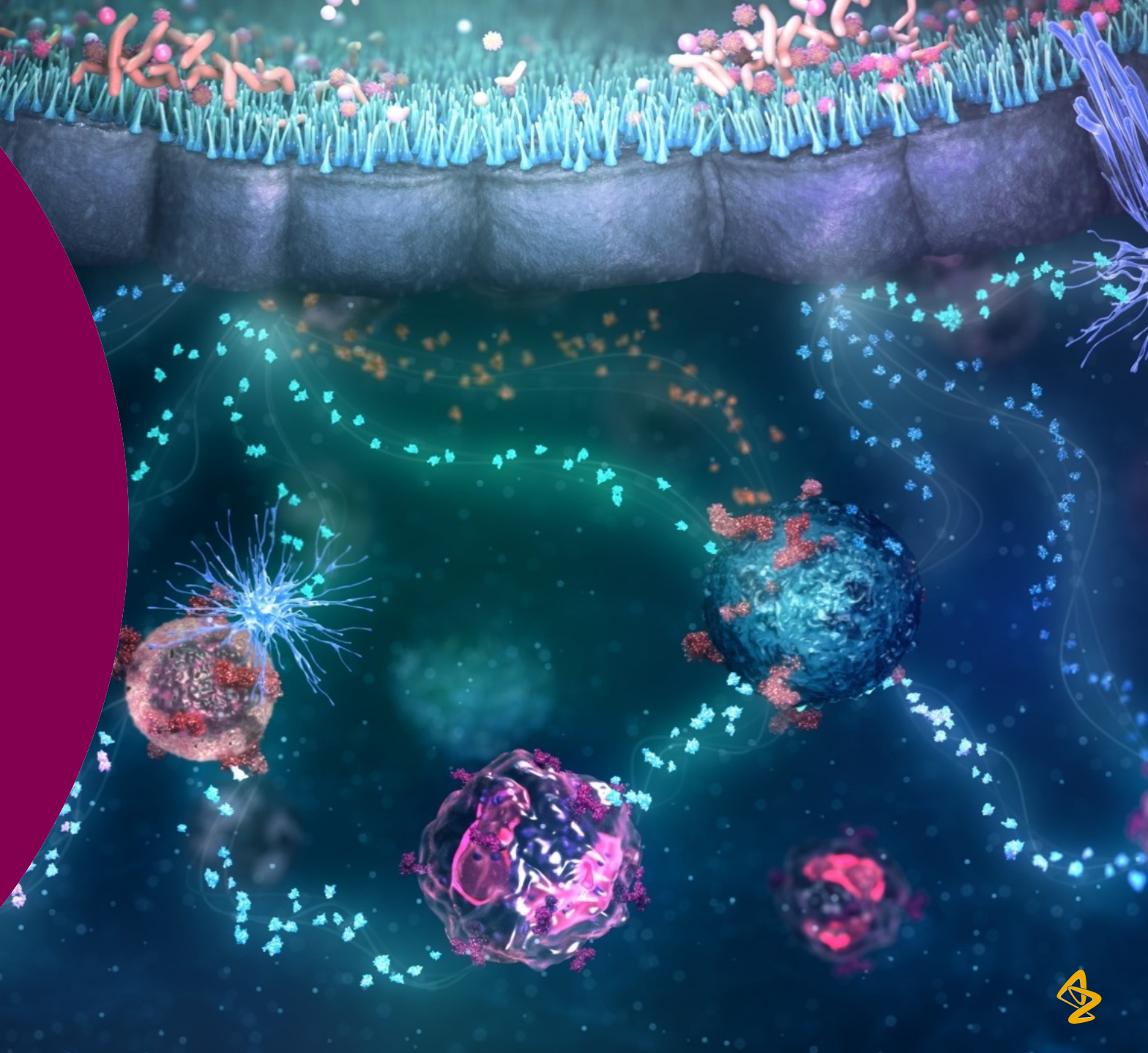
- ❑ Incorporate data checks to identify issues within uncleaned data, such as duplicates, missing values, or outliers. Generate user-defined logs, errors, notes, or employ PUT/%PUT statements for effective debugging.
- ❑ Maintain the length of numeric data to prevent inaccuracies in storage.
- ❑ Read in source data only once in the entire program to reduce the risk of programming mistakes or accidental modifications to source data.
- ❑ Utilize PROC DATASETS at the conclusion of the program to delete unnecessary datasets from the work library.

Note: These examples offer a glimpse into suggested programming techniques. For more detailed insights, readers are encouraged to explore the referenced materials listed in the 'References' section of the paper.



6

Thoughts About Automation



Implementing Macros/Tools

1

Receive Requests

2

Set Up Programs

3

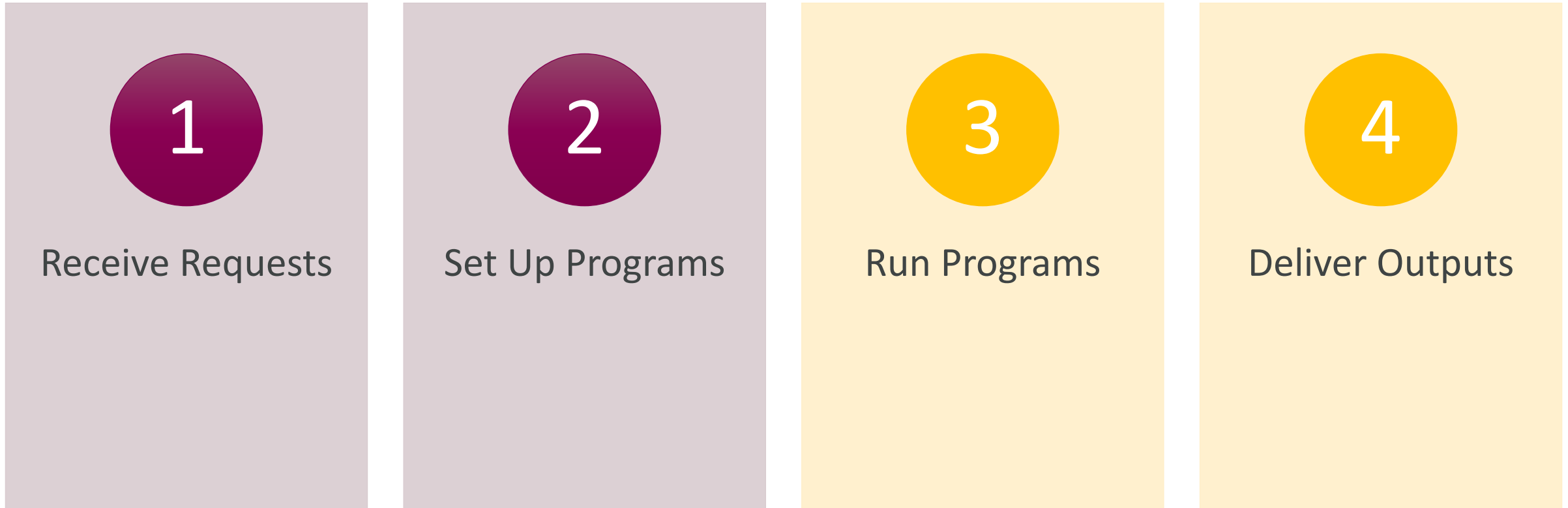
Run Programs

4

Deliver Outputs



Process Innovation



Implementing Macros/Tools + Process Innovation

1

Receive Requests

2

Set Up Programs

3

Run Programs

4

Deliver Outputs



A portion of the
detailed code is
available in the
appendix of the
paper



Contact Information

Your comments and questions are valued and encouraged. Contact the author at:

Zhen (Laura) Li
AstraZeneca
zhen.li@astrazeneca.com



Thank You!

