



PRAHEALTHSCIENCES

**The TFL Workbench: Tools to
Standardize and Accelerate TFL
Programming**

Iain Humphreys, PRA Health Sciences
Hansjörg Frenzel, PRA Health Sciences



- Issues with traditional TFL programming, and how the TFL Workbench aims to overcome some of them
- Components of the TFL Workbench
- Configuring TFL metadata
- Validated SAS macros
- SAS Code generation
- Structure of output datasets
- Consequences for QC
- Experience to date & future measurements
- Conclusions



Introduction

Overcoming the problems of traditional TFL programming approaches

Long-standing issues

- Traditional approach delivers,but at a cost
- Inefficient, resource intense, inconsistent, error prone

Previous solutions

- Many attempts to automate and standardise TFLs
- Barriers to overcome

Standards ?

- “The key to end-to-end clinical data processing”
- CDASH → SDTM → ADaM [---> TFLs?]
- Challenges for a CRO

Need something new

- Examine what the TFL programmer needs and does
- Capture and exploit TFL metadata

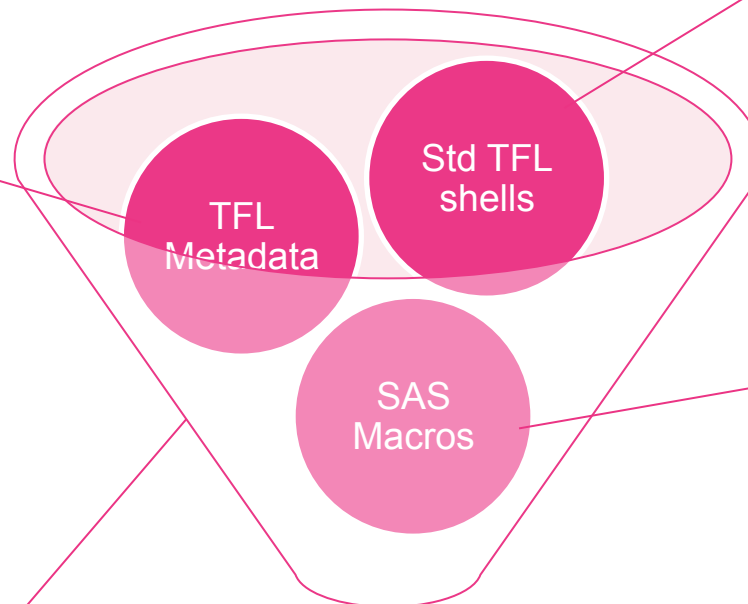


TFL Workbench components

Configuration (in Excel): table templates containing machine-readable TFL metadata:

- Disposition, demography, analysis sets, enrolment
- Occurrence (AE, CM, MH)
- Labs
- Vitals
- ECG
- Exposure

Code-generation (Excel/VB): reads TFL metadata and writes files containing appropriate calls to standard macros to build a TFL program



Specification: PRA or sponsor standard output layouts

Reporting: suite of modular, validated SAS “building block” macros:

- Initialisation
- Treatment columns & Population counts
- Fetch data from ADaM
- Summary statistics
- Frequency counts
- Table dataset
- [Reporting]

Auto-generated TFL program

SAS
ADaM
Database

TFL program run against
ADaM data to produce
output.

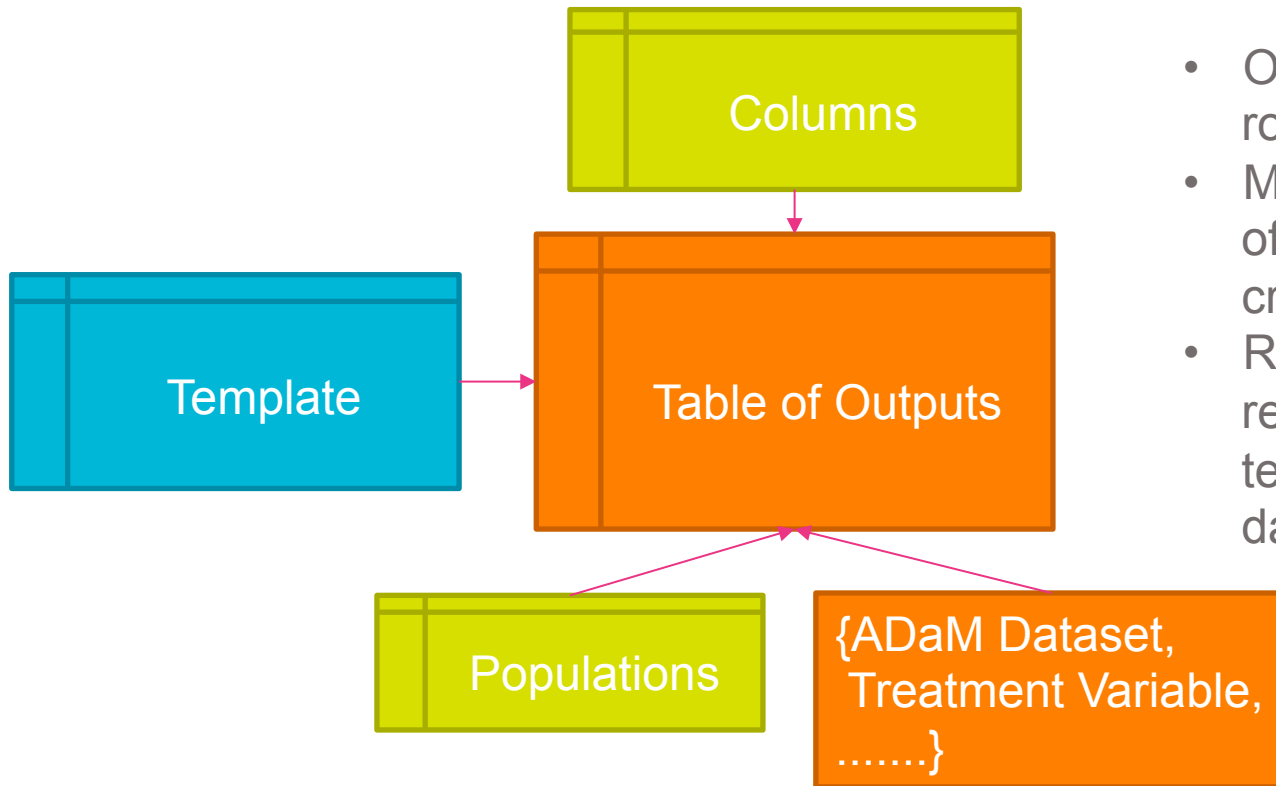


TFL output (RTF)



Configuring of TFL Metadata

Table of Outputs



- One table within each row
- Merges different sets of configurations to create a table
- Repeated tables are recognized by its template and ADaM dataset



Configuring of TFL Metadata

A Table

		Population		Columns		
		Table 14.2.1 Demography by Treatment Safety Analysis Set				
		Drug A, Dose 1 (N=xxx)	Drug A, Dose 2 (N=xxx)	Total Drug A (N=xxx)	Placebo (N=xxx)	Total (N=xxx)
Characteristic						
Categories	Sex [n %]					
	Female	xxx (xx.x)	x (xx.x)	xxx (xx.x)	xxx (xx.x)	xxx (xx.x)
	Male	xxx (xx.x)	x (xx.x)	xxx (xx.x)	xxx (xx.x)	xxx (xx.x)
	Missing	xxx (xx.x)	x (xx.x)	xxx (xx.x)	xxx (xx.x)	xxx (xx.x)
	Race [n %]					
	White	xxx (xx.x)	x (xx.x)	xxx (xx.x)	xxx (xx.x)	xxx (xx.x)
	Non-White ^a	xxx (xx.x)	x (xx.x)	xxx (xx.x)	xxx (xx.x)	xxx (xx.x)
	Black/African American	xxx (xx.x)	x (xx.x)	xxx (xx.x)	xxx (xx.x)	xxx (xx.x)
	Oriental/Asian	xxx (xx.x)	x (xx.x)	xxx (xx.x)	xxx (xx.x)	xxx (xx.x)
	American Indian/Alaskan Native	xxx (xx.x)	x (xx.x)	xxx (xx.x)	xxx (xx.x)	xxx (xx.x)
Statistics	Other	xxx (xx.x)	x (xx.x)	xxx (xx.x)	xxx (xx.x)	xxx (xx.x)
	Missing	xxx (xx.x)	x (xx.x)	xxx (xx.x)	xxx (xx.x)	xxx (xx.x)
	Age (years)					
	n	Xxx		xxx	xxx	xxx
	Mean (SE)	xx.x (x.x)	(x.xx)	xx.x (x.xx)	xx.x (x.xx)	xx.x (x.xx)
	Median	xx.x		xx.x	xx.x	xx.x
	Min, Max	xx, xx	x	xx, xx	xx, xx	xx, xx
	Age Group [n %]					
	<65 years	xxx (xx.x)	x (xx.x)	xxx (xx.x)	xxx (xx.x)	xxx (xx.x)
	≥65 years - <75 years	xxx (xx.x)	x (xx.x)	xxx (xx.x)	xxx (xx.x)	xxx (xx.x)
≥75 years	xxx (xx.x)	x (xx.x)	xxx (xx.x)	xxx (xx.x)	xxx (xx.x)	
Missing	xxx (xx.x)	xxx (xx.x)	xxx (xx.x)	xxx (xx.x)	xxx (xx.x)	

Analysis Variables
=> Dataset Name

Formatting

^a Counts for categories non-white is the sum of all subjects of all race categories except category 'White'

Programming Note: Display "Missing" rows only in case of actually missing observations for any of the categorical variables



Configuring of TFL Metadata Template

METADATA	SECTION	ROW	DATA TYPE	SUPPRESS IF	ROWLABEL1	INDENT1	TREATMENTS	STATISTIC	PRECISION	ANALYSIS VARS	WHERE	DENOM
NONE												
blank												
row label hdr					Characteristic	0						
label	4	1	CON	N	Age (years)	0						
statistic	4	2	CON	N	n	2	xx	n		0 AGE		
statistic	4	3	CON	N	Mean (std)	2	xx.x (xx.xx)	mean#std		0 AGE		
statistic	4	4	CON	N	Median	2	xx.x	median		0 AGE		
statistic	4	5	CON	N	Q1, Q3	2	xx.x, xx.x	q1#q3		0 AGE		
statistic	4	6	CON	N	Min, Max	2	xx, xx	min#max		0 AGE		
blank												
label	9	1	CAT	N	Cigarette smoking (%)	0						
statistic	9	2	CAT	N	Never smoked	2	xx (xx.x)	n#pct	1	COUNT(_ALL_)	ASMOKE eq 'NEVER'	_BigN_
statistic	9	3	CAT	N	Past smoker	2	xx (xx.x)	n#pct	1	COUNT(_ALL_)	ASMOKE eq 'PAST'	_BigN_
statistic	9	4	CAT	N	Current smoker	2	xx (xx.x)	n#pct	1	COUNT(_ALL_)	ASMOKE eq 'CURRENT'	_BigN_

Row
Sequence

Data type

Row Label
and Indentation

Statistics, Precision
and Analysis Variable

Where-Clause

Denominator



Configuring of TFL Metadata Populations

Population	Where clause	Data
All Randomized	randfl eq 'Y'	ADSL
Full Analysis Set	fasfl = 'Y'	ADSL
Per Protocol Analysis Set	pprotfl = 'Y'	ADSL
Safety Analysis Set	saffl = 'Y'	ADSL

Population Name
(TOO)

Population
Filter

Subject
Level
Dataset



Configuring TFL Metadata

Table Columns

Header	Tmt Column	Column Label	Tmt Variable	Value	Strata Variable	Strata Value	Total	Where for _BigN_	No Display
	▼	▼	▼	▼	▼	▼	▼	▼	▼
Head_SAF	_c1	Drug A#(N=_BigN_)	TRT01AN	1			N	TRT01AN eq 1	N
Head_SAF	_c2	Placebo#(N=_BigN_)	TRT01AN	2			N	TRT01AN eq 2	N
Head_SAF	_c3	Total#(N=_BigN_)	TRT01AN	1, 2			Y	TRT01AN in(1, 2)	N
Head_SAF_Sex	_c1_1	Drug A#!Female#(N=_BigN_)	TRT01AN	1	SEXN	1	N	TRT01AN = 1 and SEXN = 1	N
Head_SAF_Sex	_c1_2	Drug A#!Male#(N=_BigN_)	TRT01AN	1	SEXN	2	N	TRT01AN = 1 and SEXN = 2	N
Head_SAF_Sex	_c2_1	Placebo#!Female#(N=_BigN_)	TRT01AN	2	SEXN	1	N	TRT01AN = 2 and SEXN = 1	N
Head_SAF_Sex	_c2_2	Placebo#!Male#(N=_BigN_)	TRT01AN	2	SEXN	2	N	TRT01AN = 2 and SEXN = 2	N
Head_Std	_c1	Drug A#(N=_BigN_)	TRT01PN	1			N	TRT01PN eq 1	N
Head_Std	_c2	Placebo#(N=_BigN_)	TRT01PN	2			N	TRT01PN eq 2	N
Head_Std	_c3	Total#(N=_BigN_)	TRT01PN	1, 2			Y	TRT01PN in(1, 2)	N

Column
Set
(TOO)

Column Name
and Label

Treatment

Strata

Column
Filter Display



Configuring TFL Metadata

Table of Outputs

Output ID	Population	Headers	Treatment as	Template	Row label	Table Filter	Data Source	Treatment
			Columns					Variable
14-01-02-01	Full Analysis Set	STD1_N	Y	PRA-STD-T-006			ADSL	TRT01P
14-01-02-02	Per Protocol Analysis	STD1_N	Y	PRA-STD-T-006			ADSL	TRT01P
14-01-02-03	Safety Analysis Set	SAF1_N	Y	PRA-STD-T-006			ADSL	TRT01A
14-01-02-04	Safety Analysis Set	SAF1	Y	PRA-STD-T-006			ADSL	TRT01A
14-01-02-05	Safety Analysis Set	SAF1_N	Y	PRA-STD-T-006			ADSL	TRT01A

Output
Identifier

Population

Column Set

Template

Row Label

Table Where
Clause

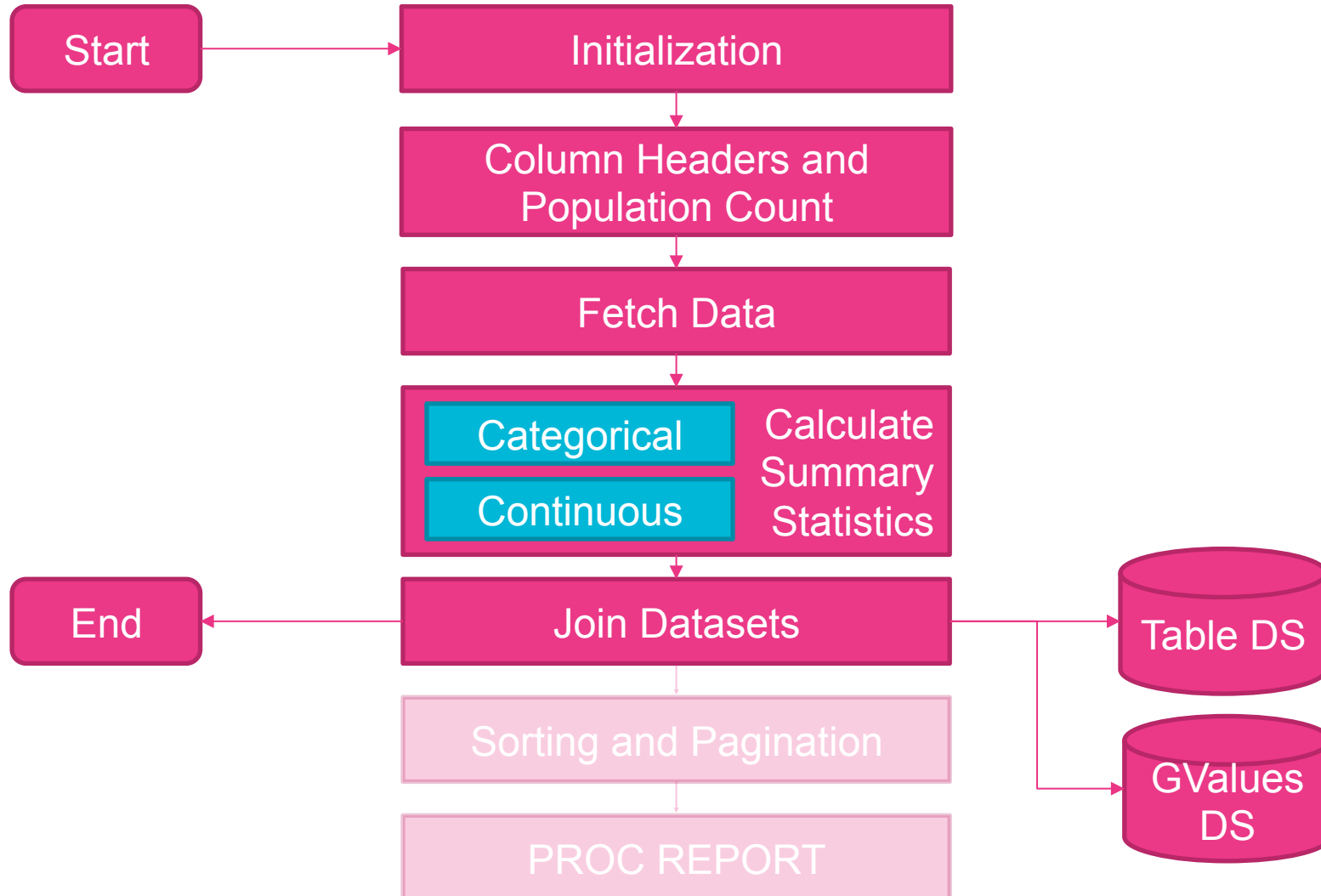
Analysis
Dataset

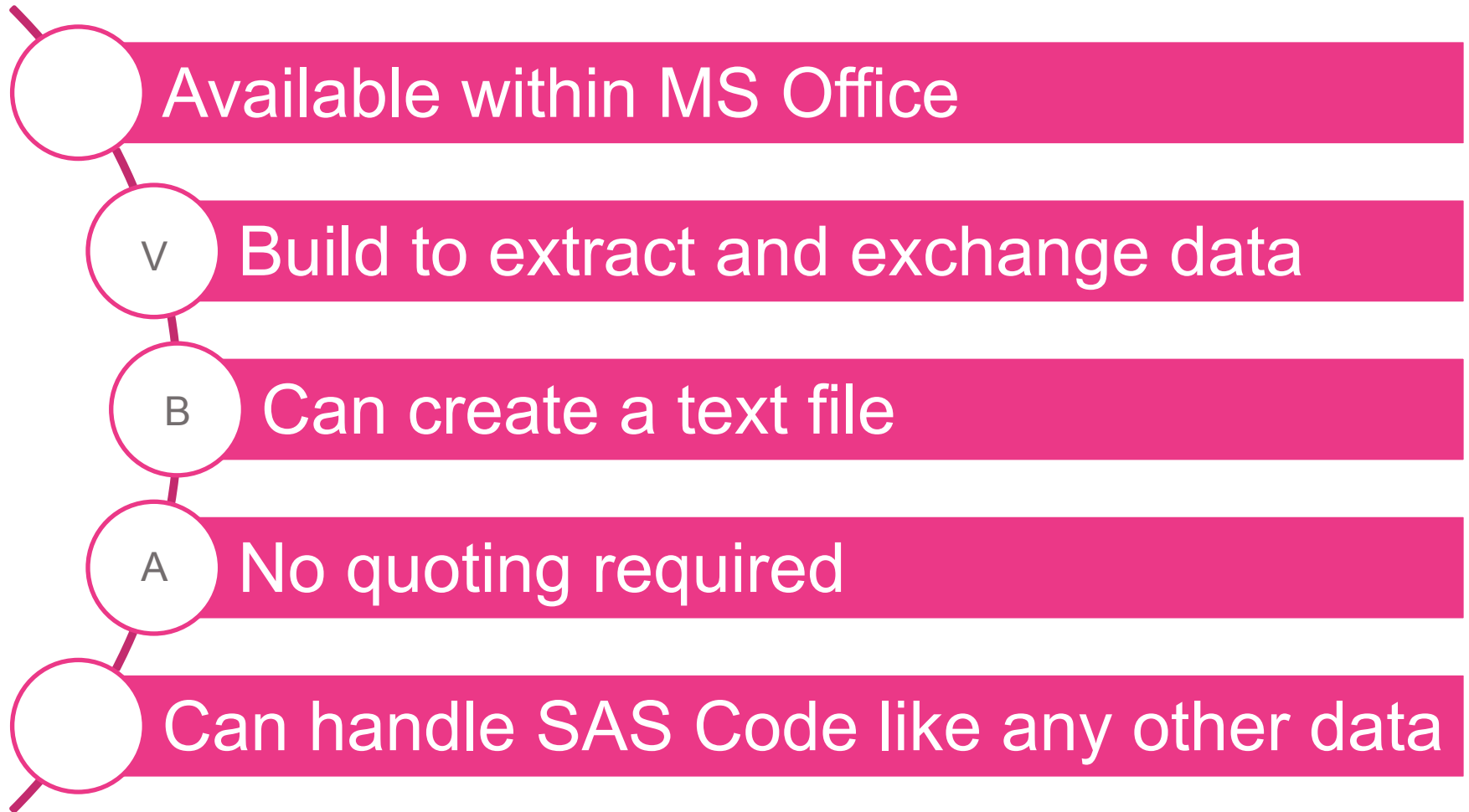
Treatment
Variable



Validated SAS Macros

Elements of a Table Program

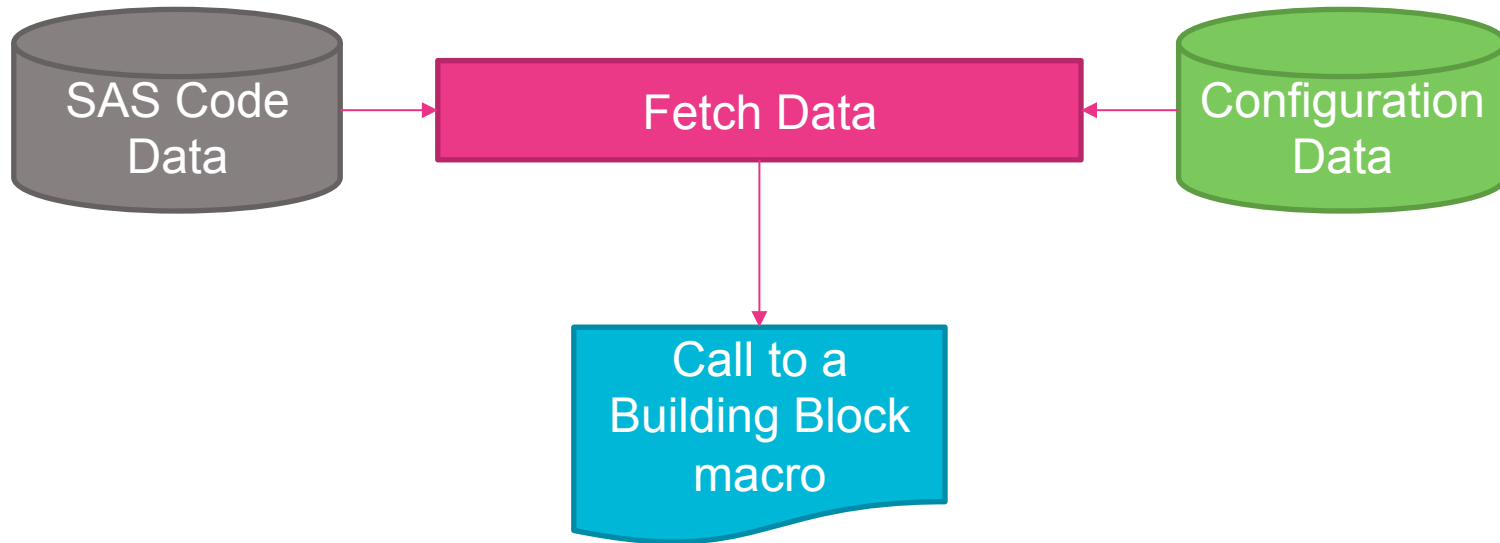






SAS Code Generation

How a Program is Written





SAS Code Generation

SAS Code Storage

Call ID	SAS Code	Parameter Name	Metadata Source	Return Type
FETCHDATA	*** fetch the data being analyzed *** ;			
FETCHDATA	%g_sTFL_fetchData(tfldata = #1	txtDataSource	qryTOO	FIRST
FETCHDATA	, tfltreat = #1	txtTMTVar	qryTOO	LIST
FETCHDATA	, tflwhere = #1	memTableFilter	qryTOO	LIST
FETCHDATA	, out = _INDS1			
FETCHDATA	, treat2col = #1	ynTMTAsCols	qryTOO	FIRST
FETCHDATA	, strat2col = #1	ynStrataAsCols	qryTOO	FIRST
FETCHDATA	, other2col = #1	ynOtherVarsAsCols	qryTOO	FIRST
FETCHDATA	, altODS = #1	ynAlternativeEscChar	qryTOO	FIRST
FETCHDATA	, outname = #1	Output DS	tblGlobalInfo	NFIRST
FETCHDATA	)			
FETCHDATA	%if &_rc_fetchData gt 0 %then %do ;			
FETCHDATA	%put %str(WARN)ING: Failure in g_sTFL_fetchData - EARLY TERMINATION OF _tflFrame ;			
FETCHDATA	%goto exit ;			
FETCHDATA	%end ;			

Element Name

SAS Code

Configuration

Return Type

Placeholder

Configuration
Storage



SAS Code Generation

The Result

```
*** fetch the data being analyzed *** ;
%g_sTFL_fetchData( tfldata      = ADSL
                  , tfltreat    = TRT01P
                  , tflwhere     =
                  , out          = _INDS
                  , treat2col    = Y
                  , strat2col    = N
                  , other2col    = N
                  , altODS       = N
                  , outname      = T_14_02_01
                  ) ;

%if &_rc_fetchData gt 0 %then %do ;
  %put %str(WARN)ING: Failure in g_sTFL_fetchData - EARLY TERMINATION OF _tflFrame ;
  %goto exit ;
%end ;
```



Output Dataset

	_section	_row	AVISITN	_rtext1	_c1	_c2	_c3	_count1	_count2	_count3
1	1	1	2	Week 1	325 (99.1)	309 (98.4)	634 (98.8)	325	309	634
2	1	1	3	Week 4	302 (92.1)	294 (93.6)	596 (92.8)	302	294	596
3	1	1	4	Week 7	280 (85.4)	276 (87.9)	556 (86.6)	280	276	556
4	1	1	5	Week 10	250 (76.2)	254 (80.9)	504 (78.5)	250	254	504
5	1	1	6	Week 13	243 (74.1)	243 (77.4)	486 (75.7)	243	243	486
6	1	1	7	Week 16	197 (60.1)	209 (66.6)	406 (63.2)	197	209	406
7	1	1	8	End of Treatment [^] (super a)	325 (99.1)	310 (98.7)	635 (98.9)	325	310	635

- Reliable Column Naming
- Tables can be compared
 - Amongst each other
 - Between shipments
- Population counts, column labels, population filters etc. can be compared
- Eyeballing can be automated to a certain point.

	symbol	value	descr
1	_popwhere	ITTFL eq 'Y'	Population where clause
2	_adam	adam	Libname for source data
3	_poptreat	trt01pn	Population treatment variable
4	_stratvar		Stratifying variable
5	_popdata	ADSL	Dataset containing population set
6	_rtext1hdr	Visit	Column heading for first row label column
7	_n_col	3	Number of column variables
8	_col_1	_c1	Column variable name
9	_col_2	_c2	Column variable name
10	_col_3	_c3	Column variable name
11	_n_1	328	Column population count
12	_n_2	314	Column population count
13	_n_3	642	Column population count
14	_label_1	Drug A#(N=328)#n (%)	Column header
15	_label_2	Placebo#(N=314)#n (%)	Column header
16	_label_3	Total#(N=642)#n (%)	Column header
17	_where_1	TRT01PN eq 1	Population count where clause
18	_where_2	TRT01PN eq 2	Population count where clause
19	_where_3	TRT01PN in (1, 2)	Population count where clause



- Tables not covered by the TFL Workbench follow the QC process currently in place (double-programming)
- Amount of deviation from the standard determines the level of QC
 - Cosmetic changes might require a configuration review
 - More complex changes might still require double-programming
- Initial assessment of the Statistician and a compliance review will drive the QC level



‘Retrospective’ pilots:

- Aims: Prove concept of the metadata-driven approach; confirm statistics generated are correct; confirm reduction in effort
- Findings:
 - High proportion of target tables reproducible with TFL Workbench
 - Quick and easy to:
 - make simple study-specific changes
 - select and summarise data

Pilot conclusions

- Significant savings possible, under certain conditions:
 - Established standard is configured in the TFL Workbench
 - Study follows standard to a reasonable high degree
 - Input datasets are ‘analysis-ready’
 - Training and subsequent experience is gained in operation



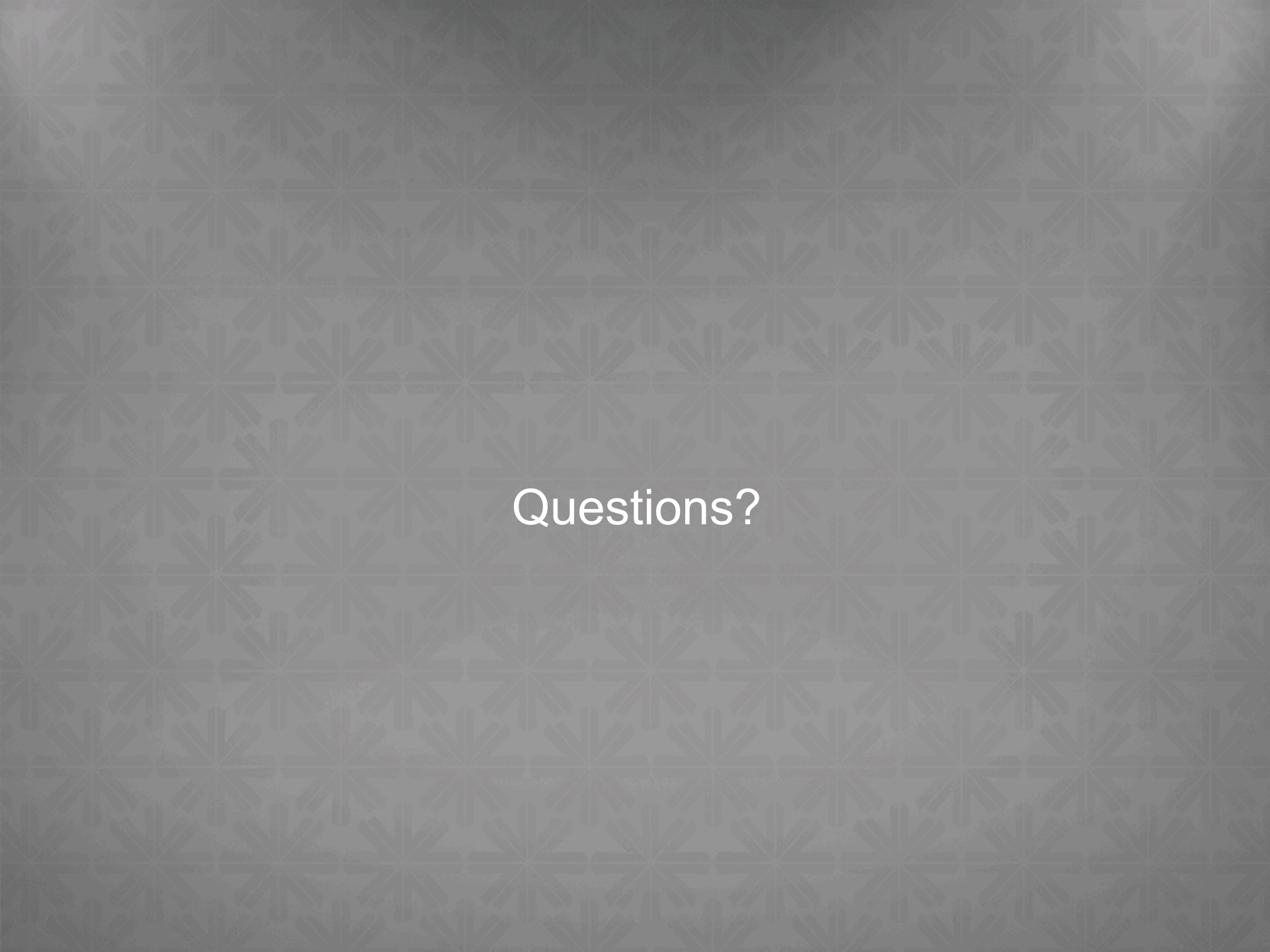
Compliance to a standard

- Comparing study-level instance versus ‘master’ version
 - Understanding the level of change from the standard determines the QC effort needed
 - Quantitative measure – assists in assessing the degree of uptake of the standard & possible changes needed



Conclusions

- Metadata-driven approach shows early indications of success in overcoming some of the problems and barriers in TFL standardisation & automation
- Initial pilots favourable; further 'live' pilots will follow
 - Ease of use, flexibility, reduction in effort all demonstrated
- More work to do, but the journey has begun towards accelerating TFL production, and thereby reducing cost and improving quality



Questions?