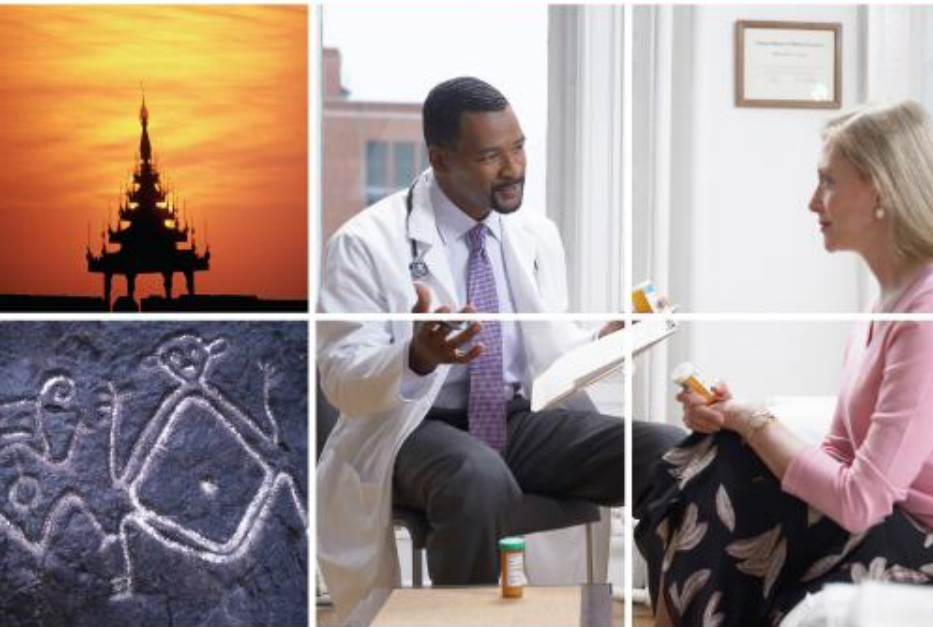




ICON Clinical Research

SAS Clinical DI – will we need new MetaProgrammers ?
Simon Wilcock

Wednesday 21st October 2009

- Introduction to Authors
- Introduction to the Paper
- Main Areas
 - Definitions
 - Why Metadata ?
 - Some examples of using Clinical DI
 - Programmer Reactions to DI
 - Some thoughts
- Conclusions
- Questions



- **Simon Wilcock**
 - 18 years experience in the Pharmaceutical Industry
 - Worked in Big Pharma, Biotech sector and now CRO
 - Over 11 years experience in SAS
 - 2 years experience in CDISC
 - 5 years experience in Enterprise Data Management (Database Administration, Data Modelling, EDI, ETL)
- **Sireesha Potula**
 - 5 + years working in Pharmaceutical Industry
 - 6 years working as a SAS Programmer for Pharma and CRO
 - 2 years within CDISC framework

- There is a lot of talk in the industry, at the moment, around Metadata driven processes
- This paper introduces one tool, SAS Clinical DI, and considers what impact this might have on the traditional programming role
 - Will programmers need different skillsets?
 - What will this mean for roles and responsibilities within an organisation?
- The purpose of the paper is not, necessarily, to answer these questions – but to open a dialogue.



- **Metadata**
 - Data about data.
 - For example: the data that describes things like dataset structures, data types, library names
- **SAS DI Studio**
 - ETL tool that uses Metadata driven processes to map from one data format from another
- **SAS Clinical DI**
 - The Clinical implementation of DI Studio with added



- The continued quest for reusable code
 - to speed up the Clinical Reporting Process
 - to remove the need for constant recoding
- Macros can become unwieldy
 - Programmers have to set multiple macro calls correctly
 - Code needs constant maintenance as new types of data emerge
- Metadata offers the solution
 - *disclaimer...* Metadata offers a solution.

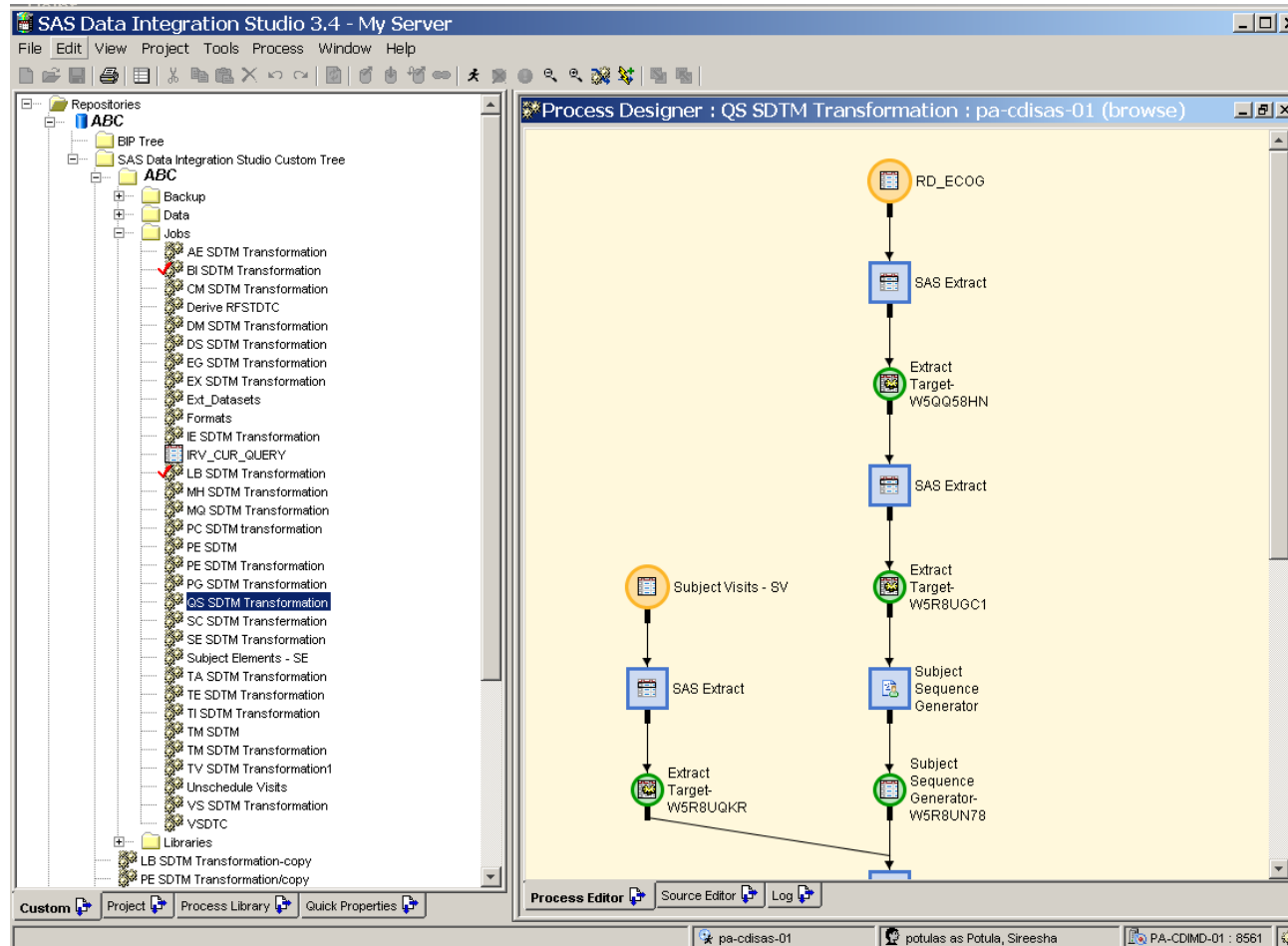


- If Metadata drives the process then the system, in theory, 'knows' the data requirements
 - The structure of input and output datasets
 - The mappings / the macro settings
- We only need to join the dots.
- The potential is there for huge time savings and quality improvements.

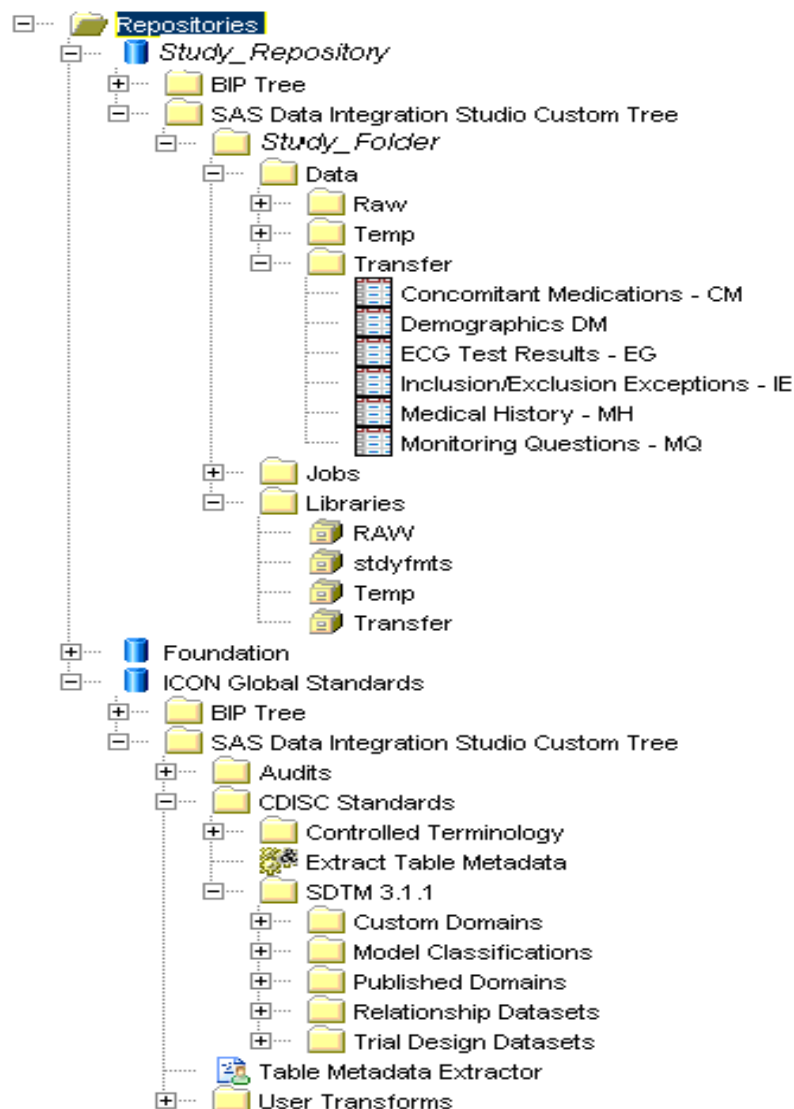


Impact on Programming

- But what is it like to work in a MetaData environment ?



Metadata Directory Structure

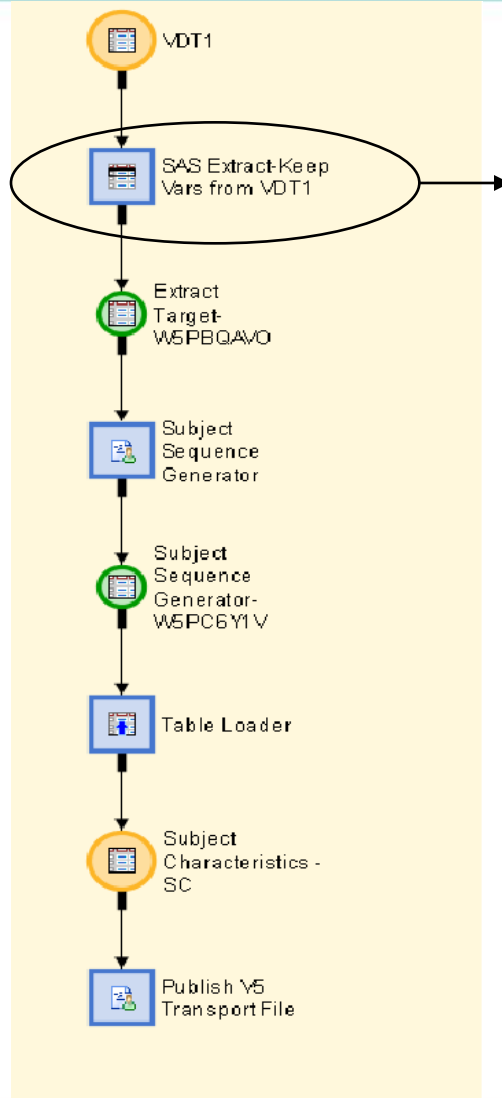


This is not the Concomitant Medication dataset but a dataset that describes the variables *in* the CM dataset

This subfolder lists all the library mappings

The CDISC SDTM structures are all kept together in this Global Repository

Performing a simple mapping



SAS Extract-Keep Vars from VDT1 Properties *

GeneralWhereOrder ByGroup ByProcessMappingOptionsNotesExtended AttributesAdvanced

Source table: VDT1

#	Column
1	STUDY
2	DCMNAME
3	DCMSUBNM
4	SUBSETSN
5	DOCNUM
6	INVSITE
7	INV
8	PT
9	ACCESSTS
10	LOGINTS
11	LSTCHGTS
12	LOCKFLAG
13	CPEVENT
14	DCMDATE
15	DCMTIME
16	REPEATSN
17	ACTEVEN
18	SUBEVENT_NUMBER
19	VISIT_NUMBER
20	QUALIFYING_VALUE
21	QUALIFYING_QUESTION
22	LAB
23	LABRANGE_SUBSET_NUI
24	LAB_ASSIGNMENT_TYPE
25	LAB_ID
26	VISDT
27	OK
28	SUBJINI
29	VISDT_FUL
30	VISDT_EXC

Target table: WSPBQAVO

#	Column	Expression	Column Description	Mapping Type	Type
1	STUDYID	study1	Study Identifier	Derived	Character
2	DOMAIN	"SC"	Domain Abbreviation	Derived	Character
3	USUBJID	"study1" '-' trim(substr(INVSITE...	Unique Subject Identifier	Derived	Character
4	SCSEQ		Sequence Number	(None)	Numeric
5	SCTESTCD	"SUBJINIT"	Subject Characteristic Shor...	Derived	Character
6	SCTEST	"Subject Initials"	Subject Characteristic	Derived	Character
7	SCORRES	SUBJINI	Result or Finding in Original ...	Derived	Character
8	SCSTRESC	SUBJINI	Character Result/Finding in ...	Derived	Character
9	SCSTRESN		Numeric Result/Finding in St...	(None)	Numeric
10	SCDTC	VISDT_FUL	Date/Time of Collection	Derived	Character

A more complex transformation

VIEWTABLE: Work.Pe1

	SUBJECTID	SITEID	VISITID	VISITINDEX	FORMINDEX	PEORRES1	PEORRES2	PEORRES3	PEORRES4	PEORRES5	PEORRES6
1	001	101	1	1	1	ABNORMAL	NORMAL	NORMAL	NORMAL	ABNORMAL	NORMAL

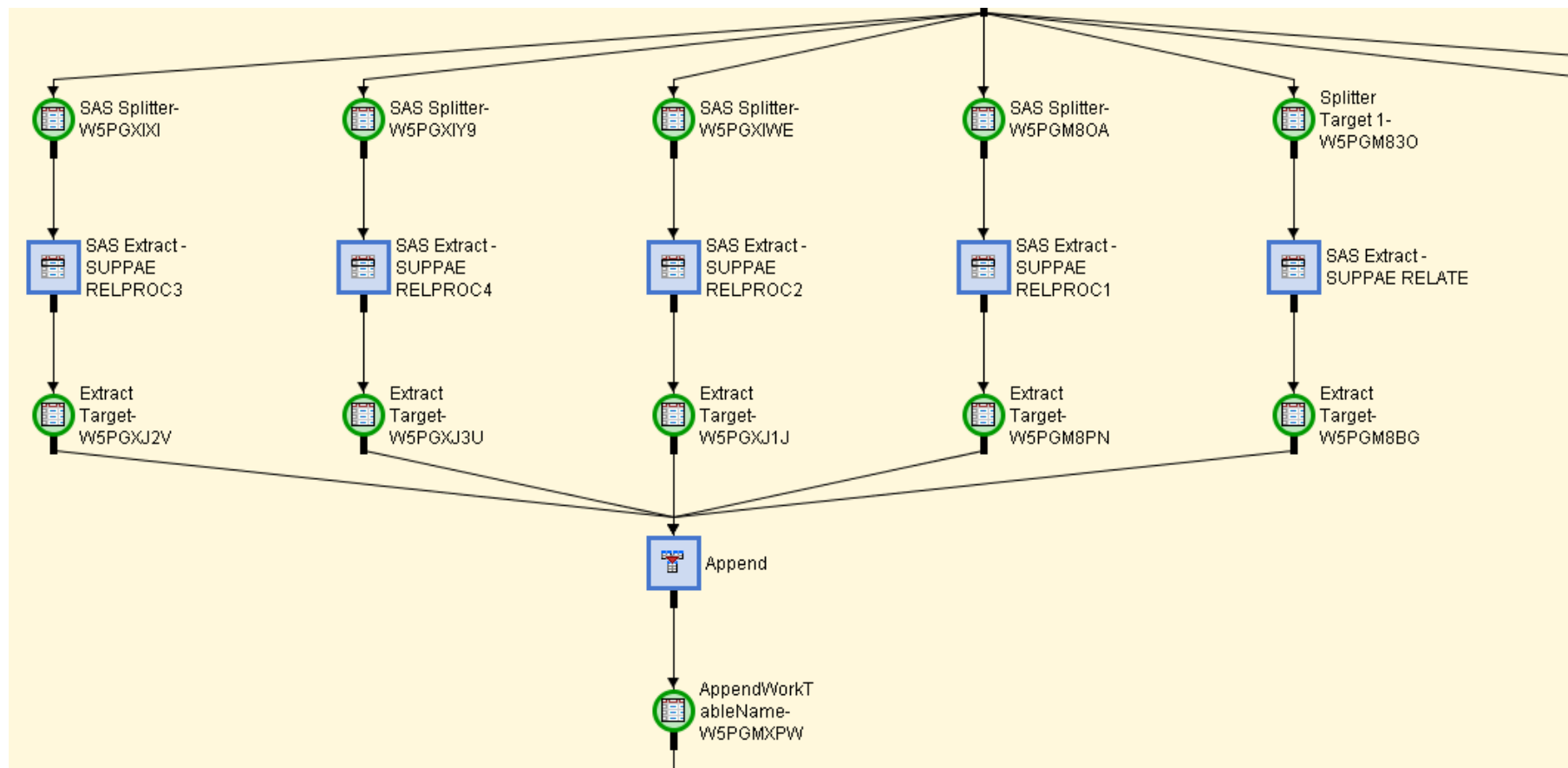
	SUBJECTID	SITEID	VISITID	VISITINDEX	FORMINDEX	DOV	_name_	col1
1	001	101	1	1	1		PEORRES1	ABNORMAL
2	001	101	1	1	1		PEORRES2	NORMAL
3	001	101	1	1	1		PEORRES3	NORMAL
4	001	101	1	1	1		PEORRES4	NORMAL
5	001	101	1	1	1		PEORRES5	ABNORMAL
6	001	101	1	1	1		PEORRES6	NORMAL
7	001	101	1	1	1		PEORRES7	COUGH
8	001	101	1	1	1		PEORRES8	NORMAL
9	001	101	1	1	1		PEORRES9	NORMAL

Base SAS approaches:

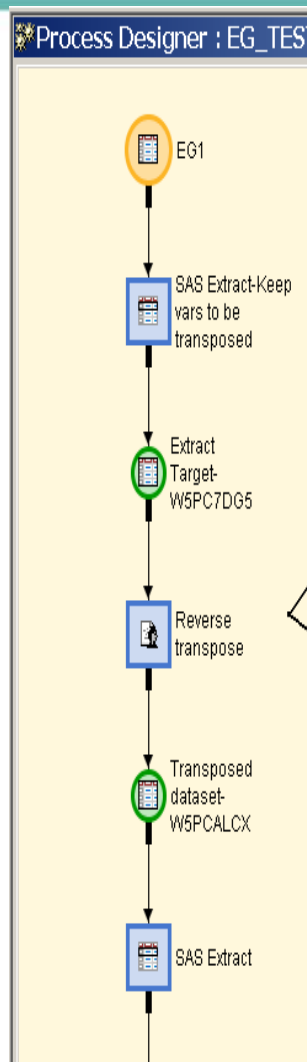
DATA step with OUTPUT statements
PROC SQL
PROC TRANSPOSE



The “output” way



Trying SQL

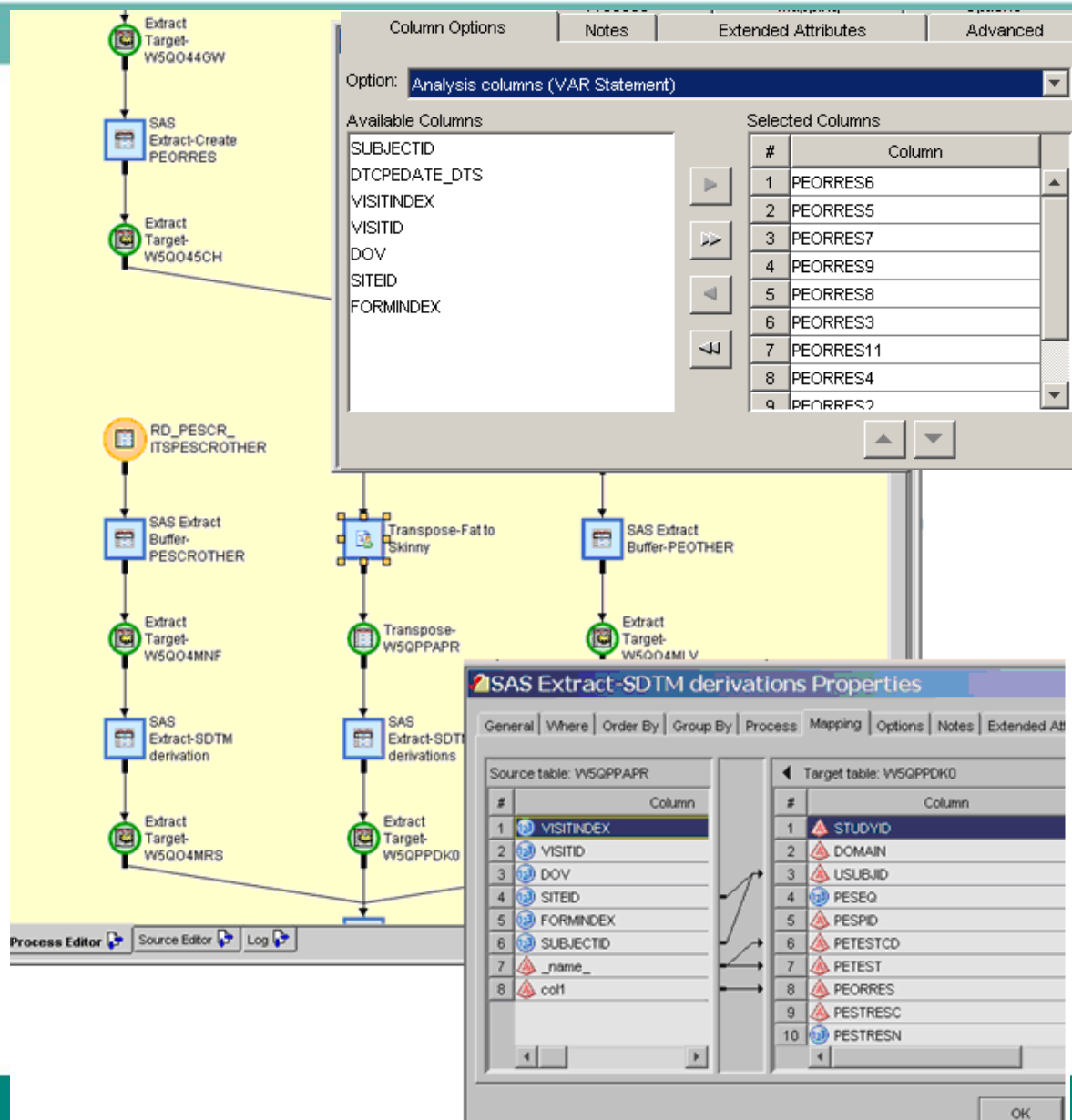


```
Proc SQL;
create table &_output
as
select invsite, pt, cpevent, ecgtim_ful, ECGCHQ,'HR' as EGTESTCD, 'Heart rate' as EGTEST, HR as EGORRES from &syslast
union all
select invsite, pt, cpevent, ecgtim_ful, ECGCHQ,'RR' as EGTESTCD, 'RR Interval' as EGTEST, RR as EGORRES from &syslast
union all
select invsite, pt, cpevent, ecgtim_ful, ECGCHQ,'PR' as EGTESTCD, 'PR Interval' as EGTEST, PR as EGORRES from &syslast
union all
select invsite, pt, cpevent, ecgtim_ful, ECGCHQ,'QRS' as EGTESTCD, 'QRS Interval' as EGTEST, QRS as EGORRES from &syslast
union all
select invsite, pt, cpevent, ecgtim_ful, ECGCHQ,'QT' as EGTESTCD, 'QT Interval' as EGTEST, QT as EGORRES from &syslast
union all
select invsite, pt, cpevent, ecgtim_ful, ECGCHQ,'QTc' as EGTESTCD, 'QTc Interval' as EGTEST, QTC as EGORRES from &syslast
union all
select invsite, pt, cpevent, ecgtim_ful, ECGCHQ,'INTP' as EGTESTCD, 'ECG Interpretation' as EGTEST, ECGINT as EGORRES from
&syslast
order by pt;
quit;
```

/*--- Start of User Written Code ---*/

/*--- End of User Written Code ---*/

PROC TRANSPOSE



Processing within the mapping

Process Designer : MH SDTM Transformation : pa-cdisas-01 *

SAS Extract - prepare to merge with RFSTDTC/ RFENDTC

Extract Target- W5P9U0H2

SQL Join - merge with RFSTDTC

SQL Join-Inner join Properties

General Designer Status Handling Parameters Process Notes Extended Attributes Advanced

Navigate

- SQL Join-Inner join
 - Create
 - SQL Target - W5PD9GT1
 - From
 - Transposed dataset - W5PCALCX
 - Inner
 - Extract Target - W5PD8V50
 - Where

Tables

- Transposed dataset - W5PCALCX
- Extract Target - W5PD8V50
- Select Results

Source tables:

#	Column	Column D
1	EGORRES	Result or F
2	INVSITE	INVSITE
3	PT	PT
4	CPEVENT	CPEVENT
5	EGTEST	ECG Test c
6	EGTESTCD	ECG Test c
7	ECGTIM_FUL	ECGTIM_FI
8	STUDY	STUDY
9	CPEVENT	CPEVENT
10	VISIT_NUMBER	VISIT_NUM
11	ECGTIM	ECGTIM
12	DAY	DAY
13	ECGCHQ	ECGCHQ
14	ECGDT	ECGDT
15	CHGFROM	CHGFROM
16	ECGDT_FUL	ECGDT_FI
17	ECGDT_FUL	ECGDT_FI
18	PT	PT

Target table: W5PD9GT1

#	Column	Expression	Column Description	Mapping Type	Type	Le
1	EGTEST		ECG Test or Examin...	1:1	Character	
2	EGTESTCD		ECG Test or Examin...	1:1	Character	
3	INVSITE		INVSITE	1:1	Character	
4	PT		PT	1:1	Character	
5	CPEVENT		CPEVENT	(None)	Character	
6	VISIT_NUMBER		VISIT_NUMBER	1:1	Numeric	
7	DAY		DAY	1:1	Numeric	
8	ECGCHQ		ECGCHQ	1:1	Character	
9	ECGDT		ECGDT	1:1	Character	
10	CHGFROM		CHGFROM	1:1	Character	
11	ECGDT_FUL		ECGDT_FUL	1:1	Character	
12	Untitled37		(None)		Character	
13	EGORRES	CASE...	Result or Finding in ...	Derived	Character	
14	EGORRESU	CASE...	Original Units	Derived	Character	
15	EGSTRESC	CASE...	Character Result/Fin...	Derived	Character	
16	EGSTRESN	CASE...	Numeric Result/Find...	Derived	Numeric	
17	EGSTRESU	CASE...	Standard Units	Derived	Character	
18	VISIT	CASE...	Visit Name	Derived	Character	

Case Expression

Operand: Column...

#	WHEN Condition	THEN Result
1	W5PD9GT1.EGORRES = '1'	'WITHIN NORMAL LIMITS'
2	W5PD9GT1.EGORRES = '2'	'ABNORMAL, NOT CLINICALLY SI...
3	W5PD9GT1.EGORRES = '3'	'ABNORMAL, CLINICALLY SIGNIFI...

ELSE Result: W5PD9GT1.EGORRES

OK Cancel Help

Display columns in alphabetical order

Select Source Log

- Traditional SAS programmers can be left tearing their hair out with the new tool because:
 - Clinical DI does not program as well as they do
 - There is no FIRST. /LAST. functionality (yet)
 - Adding additional steps, in a job, can be cumbersome
 - The program is never laid out as a single code stream (except in the log)
 - They like to look at the data when programming each step



- Metadata concepts shift the focus away from the actual content to understanding how data is structured and how it needs to be processed.
- Organisations need to consider what impact this different focus may have
- This approach frees up those working in the Data Transformation team to be more 'Data Modeller' than 'Clinical Programmer'
 - CDISC models, aCRFs, Data Definitions, define.xml and more



Conclusion.. or ideas to take away

Clinical Research

- **The potential:**
 - Organisations, recruiting into a data transformation group, will increasingly need skill sets around data modeling rather than traditional SAS programming. The role becomes one of understanding data relationships, and the relevant industry data standards, and *not the language that processes it*.
- **This is a good thing:**
 - it is the data and its interpretation that is really important, not the effort that goes into getting it into shape.
- **The Challenge**
 - Staff need an ETL mindset but still have a strong background



- Your comments and questions are encouraged, please feel free to contact the author:

Simon Wilcock
ICON Clinical Research
2 Globeside
Globeside Business Park
Marlow
Buckinghamshire
SL7 1HZ
England

Email: Simon.Wilcock@iconplc.com

Questions

