



Creating ADaM Friendly Analysis Data from SDTM Using Meta-data

by Erik Brun & Rico Schiller
(CD10 - 2011)



Agenda

The challenges

The solution

Conclusion

Abbreviations used:

SADs 4 – HLu Statistical Analysis DataSets v.4

DCD – HLu Meta Data Dictionary

CDR – Clinical Data Repository

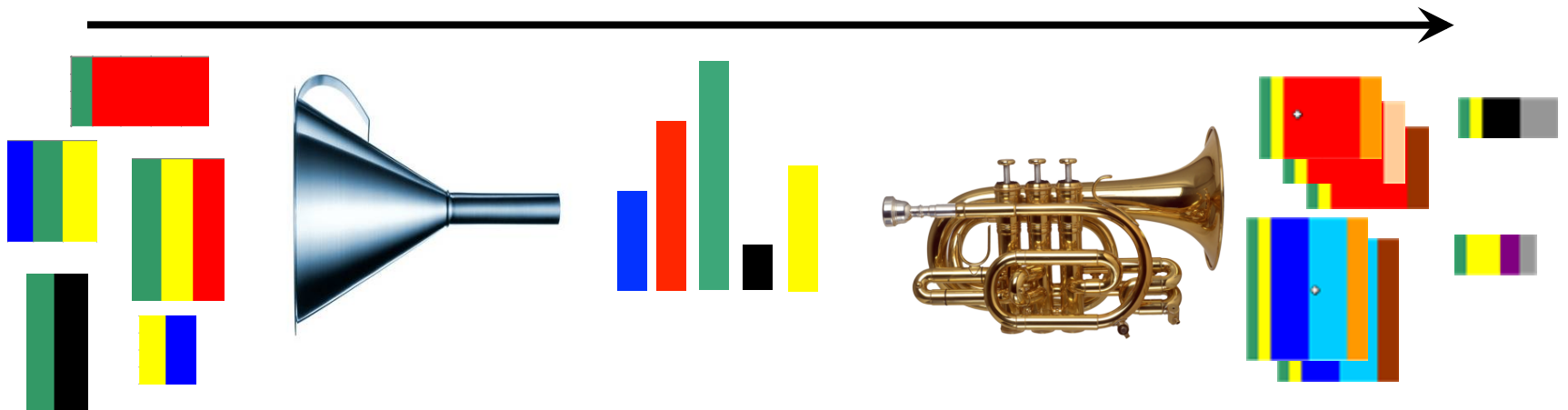
The Challenges

The funnel and the trumpet

SDTM data: Take data from a variety of sources and funnel it into a standard format

Analysis data: Take data from a standard format and expand it into a variety of formats depending on study design (and the statisticians)

Data Flow



The Challenges

Lundbeck challenges with SADs v.3

- Time resolution was date not date-time

- Data model embedded in the code

- Peculiar error and warning messages -
Including reports on data issues

- Only one central lab was assumed used per
study

- Very steep learning curve for new programmers

- Person dependent

- Insufficient for new study designs

The Solution – SADs 4 Requirements

Create the basis upon which the automated and validated production of consistent and standardised statistical analysis reports and listings for safety and efficacy data is possible.

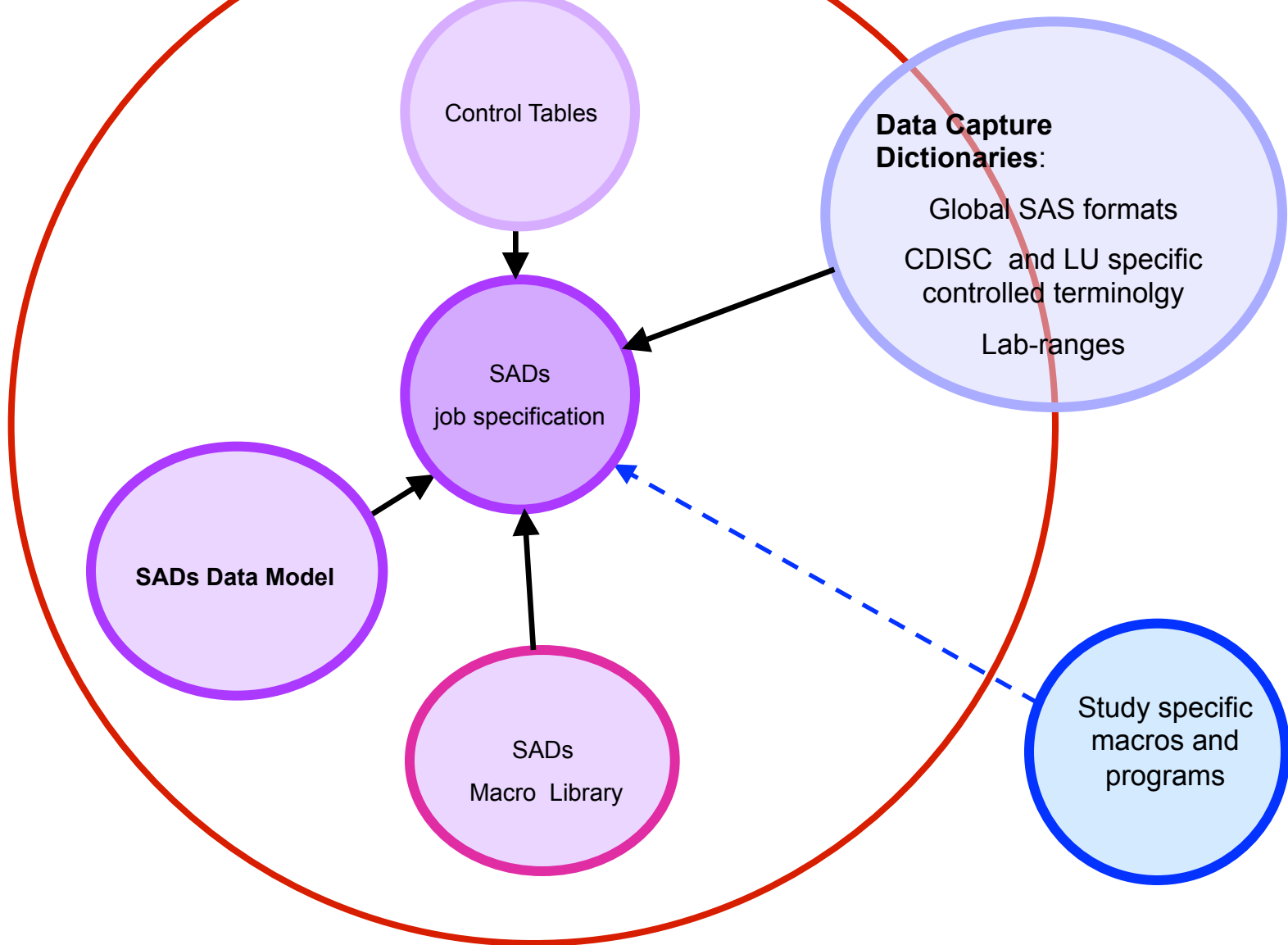
The system should allow for clear documentation of the configuration settings applied in a single study.

The system should be easy to understand and operate and yet flexible to handle a wide range of study designs.

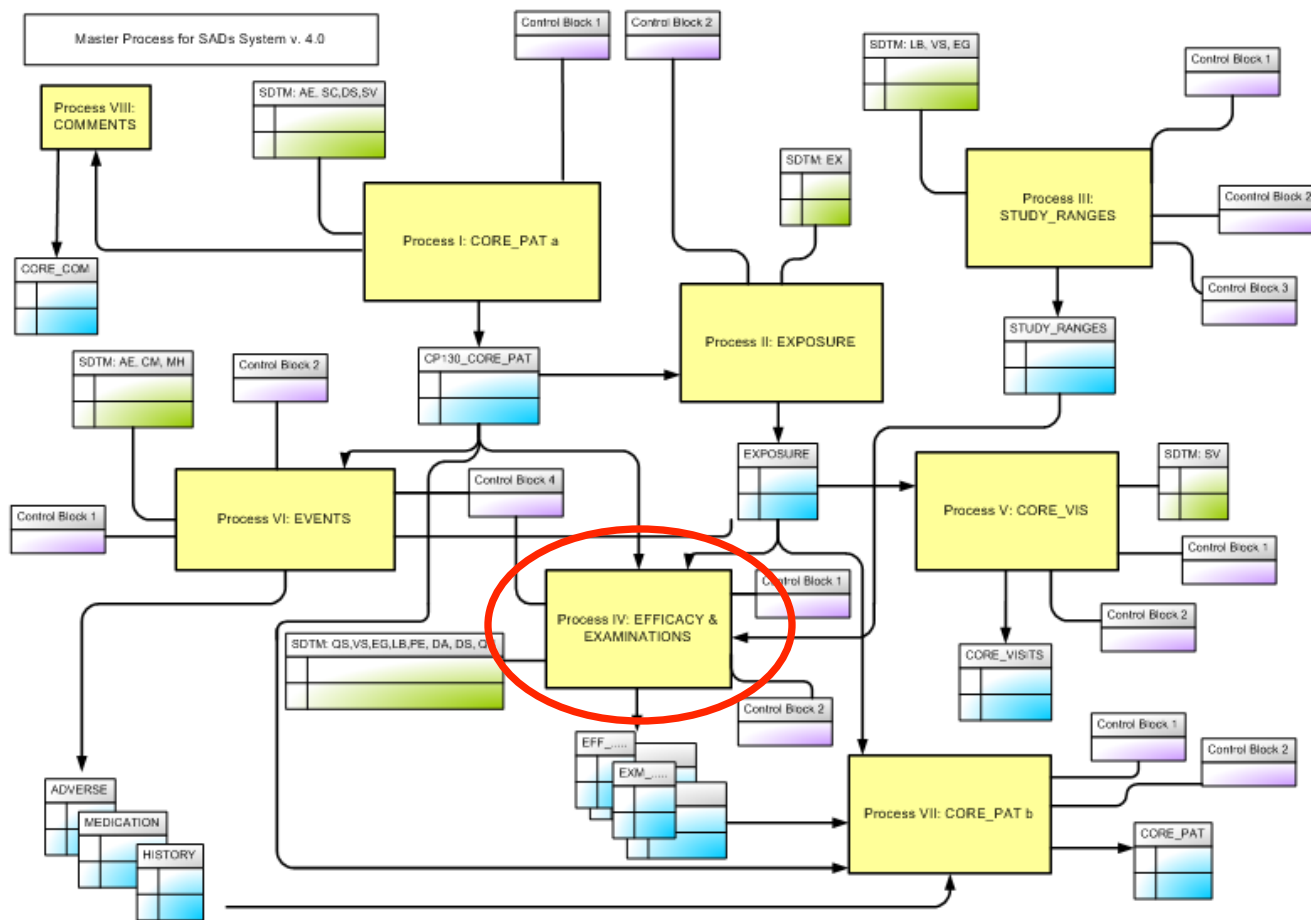
The system should be as CDISC-compliant as possible. Lundbeck pursues a strategy of applying CDISC standards, terminology, and concepts in all scientific data models.

Provide together with CDR a validated and controlled environment for the collection and integration of clinical data across studies within a drug project.

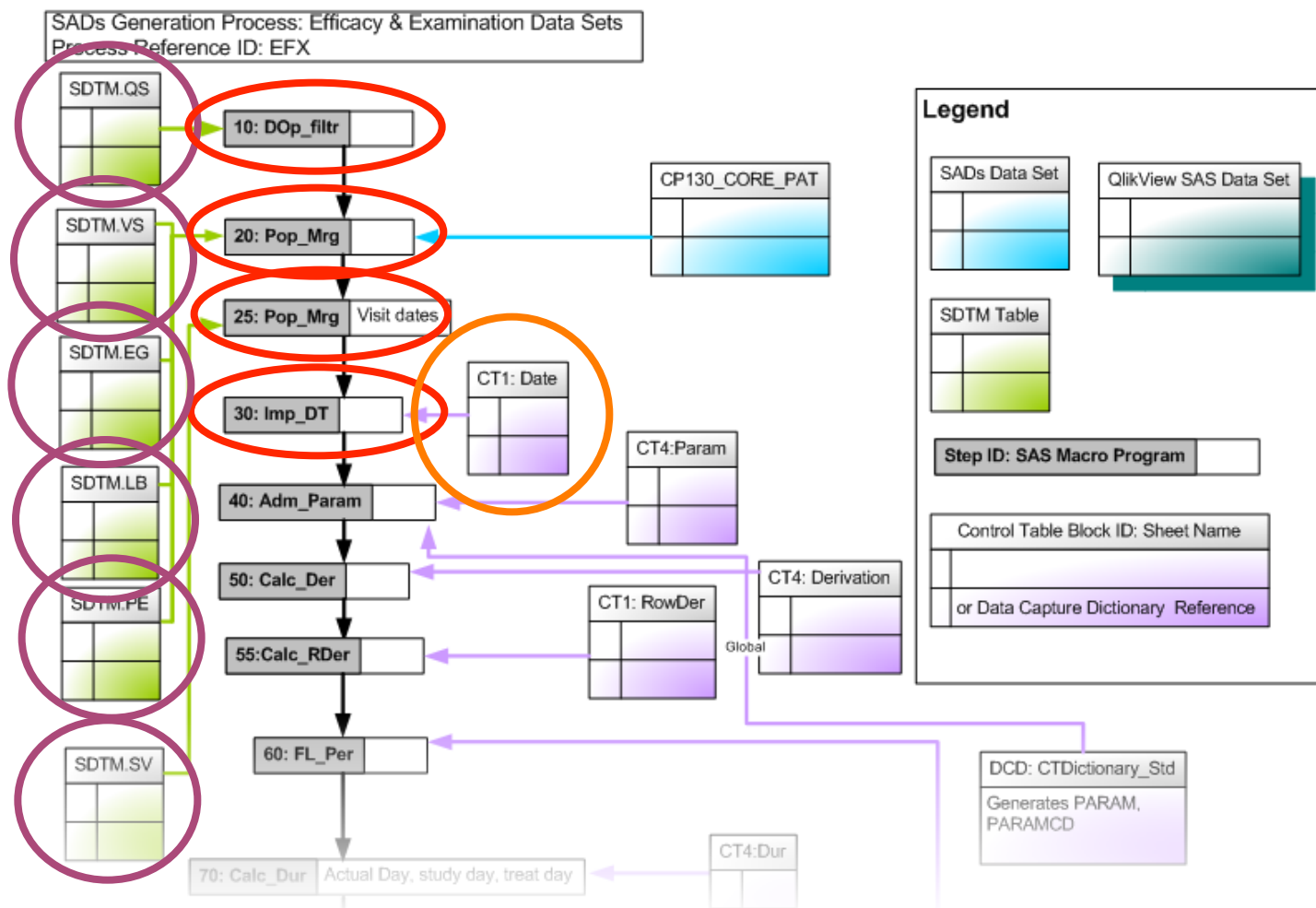
SADs System



SADs 4 – The master process



SADs 4 – Findings process



SADs 4 - Data Model

HISTORY

Position	Name	Label	Type	Format	Origin	Control	Macro	Sponsor Comments	Length
1	USUBJID	Patient	char		SDTM.DM	N/A	Pop_Mrg		40
2	STUDYID	Study Protocol	char		SDTM.DM	N/A	Pop_Mrg		20
3	SITEID	Study Site	char		DERIVED	N/A	Pop_Mrg	Collected from CORE_PATa	8
4	SITEGRP	Pooled Site Group	char		DERIVED	N/A	Pop_Mrg	Collected from CORE_PATa	8
5	COUNTRY	Country	char		DERIVED	N/A	Pop_Mrg	Collected from CORE_PATa	3
6	ARMCD	Arm	char		DERIVED	N/A	Pop_Mrg	Collected from CORE_PATa	8
7	INVTNAM	Investigator Name	char		DERIVED	N/A	Pop_Mrg	Collected from CORE_PATa	40
8	SEX	Sex	char		DERIVED	N/A	Pop_Mrg	Collected from CORE_PATa	8
9	AGE	Age	num	3.	DERIVED	N/A	Pop_Mrg	Collected from CORE_PATa	8
14	ACTUAL_DAY	Days Since Baseline	num		DERIVED	CT4Dur	Calc_Dur		8
15	MHONGO	Diagnosis Ongoing?	char		SDTM.MH	N/A	Pop_Mrg		8
16	STUDY_DAY	Days in Study	num		DERIVED	CT4Dur	Calc_Dur		8
17	TREAT_DAY	Days Since First Dose	num		DERIVED	CT4Dur	Calc_Dur		8
18	MHTERM	Reported Term for the Diagnosis	char		DERIVED	CT4Dur	Calc_Dur		200
19	MHCAT	Category	char	\$DISORDR.	SDTM.MH	N/A	Pop_Mrg		20
20	MHSPID	Sponsor Defined Identifier	char		SDTM.MH	N/A	Pop_Mrg		3
21	MHSEQ	Sequence Number	num	BEST.	SDTM.MH	N/A	Pop_Mrg		8
22	MHBODSYS	System Organ Class	char		SDTM.MH	N/A	Pop_Mrg		200
23	BODYSYSC	System Organ Class Code	num	BEST.	SDTM.SUPPMH	N/A	Pop_Sup		8
24	MHDBCD	Preferred Term	char		SDTM.MH	N/A	Pop_Mrg		200
25	PTCODE	Preferred Term Code	num		SDTM.SUPPMH	N/A	Pop_Sup		8
26	HLGT	High Level Group Term	char		SDTM.SUPPMH	N/A	Pop_Sup		200
27	HLGTCODE	High Level Group Code	num	BEST.	SDTM.SUPPMH	N/A	Pop_Sup		8
28	HLT	High Level Term	char		SDTM.SUPPMH	N/A	Pop_Sup		200
29	HLTCODE	High Level Code	num	BEST.	SDTM.SUPPMH	N/A	Pop_Sup		8
30	LLT	Low Level Term	char		SDTM.SUPPMH	N/A	Pop_Sup		200
31	LLTCODE	Low Level Code	num	BEST.	SDTM.SUPPMH	N/A	Pop_Sup		8
32	MHGEN	Gender Specific Code Flag	num		SDTM.SUPPMH	N/A	Pop_Sup		3
33	MHSTDTC	Date of Diagnosis	char	ISO8601	SDTM.MH	N/A	Pop_Mrg		20
34	MHSTRTN	Date of Diagnosis	num	DATETIME.	DERIVED	CT1Date	Imp_DT		8

One sheet per data set

Examinations (LB, PE, EG, VS) data sets are normalised

You can add study specific variables, but you cannot remove variables

Generic solution for all scales data sets (SDTM.QS)

SDTM names are kept for unchanged values

SDTM naming fragments are used [SDTMig v3.1.2 appendix D]

ADaM friendly: AVAL

AVISIT/AVISITN

PARAM/PARAMCD

SADs 4 – Control Tables

Assign group centre

Rules for date imputations

Add treatment code

Derivations:

Type casting
Scale totals
etc. Etc.

Add population flags

Baseline definitions

Windowing of Visits

Sort order of output datasets

Period definitions

Study specific additions to the data model

... and much more

SADs 4 - Control Tables

Date and Date-Time

Original SDTM value --DTC

Numerical SADs value --DTN (date-time)

Imputation rule applied --DT_CD

Data Set	Variable	Rule	Expected	Limit
ADVERSE_EVENT	AEENDTN	LATE	MINUTE	DOSE_EXDTN
ADVERSE_EVENT	AESTDTN	EARLY	MINUTE	DOSE_STDTN
CORE_PAT	BRTHDTN	EARLY	DAY	
CORE_PAT	DOSE_ENDTN	LATE	DAY	
CORE_PAT	DOSE_STDTN	EARLY	DAY	
CORE_PAT	DTHDTN	EARLY	DAY	
CORE_PAT	INCOLDTN	LATE	MINUTE	
CORE_PAT	INCONDTC	LATE	DAY	
CORE_PAT	RESTDTC	EARLY	DAY	

SADs 4 – Control Tables

Input (SDTM)	Settings	Output
AESTDTC = "2011-08-07"	Rule="EARLY" Expected="DAY"	AESTDTN = 07AUG2011:00:00:00 AESTDT_CD="Expected accuracy"
AESTDTC="2011-08"	Rule="EARLY" Expected="DAY"	AESTDTN = 01AUG2011:00:00:00 AESTDT_CD="Early; Day unknown"
AEENDTC="2011-08"	Rule="LATE" Expected="DAY"	AESTDTN=31AUG2011:00:00:00 AESTDT_CD="Late; Day unknown"
AEENDTC="2011-08-31"	Rule="LATE" Expected="MINUTE"	AESTDTN=31AUG2011:23:59:00 AESTDT_CD="Late; Hour unknown"
AESTDTC="2011-08"	Rule="EARLY" Expected="DAY" Limit=DOSE_STDTN (DOSE_STDTN=07AUG2011)	AESTDTN=07AUG2011:00:00:00 AESTDT_CD="Early; Day unknown"

SADs 4 – Control Tables

Timing

*

Window_model	PERIODN	PERIOD	AVISIT	TRIGGER	NOMINAL_DAY	WFROM	WTO	BASELINE
default	1	SCREEN	SCREEN		-7			
default			BASELINE		0			
default	3	TREATMENT	Visit 3	DOSE_STD TN	7	1	10	PRIMARY
default			Visit 4		14	11	17	
default			Visit 5		21	18	24	
default			Visit 6		28	25	35	
default			COMPLETION		42	36	56	
default	4	WASH OUT	Down Taper		49	0		WASHOUT
default	5	FOLLOW UP	Follow Up	DHMS(DATEPART(EXENDTN+7), HOUR(TIMEPART(EXENDTN)), MINUTE(TIMEPART(EXENDTN)), SECOND(TIMEPART(EXENDTN)))	70	0		

*Columns omitted for simplicity and readability

Conclusions

We have a validated system that works!

It is flexible

SDTM 3.1.x can be used as source

It has been used with success on a wide range of indications and study designs

Easy to use

A junior programmer can make a good draft set-up of a study in 1½ day

Integration of studies made much easier

The SADs data sets work for our standard reporting system

“Real” ADaM data sets can easily be created from SADs 4

Renaming and type casting is all what is needed

Conclusions

A system generating SDTM has since been made applying the same methodologies, both in development and use

SAS-DI can not be recommended as a tool for developing systems like this

It requires not only dedicated and skilled resources to develop such a system. They must also be assigned wholehearted by their managers to the project

The future: Move away from Excel as control tables

CDISC PRM (Protocol Representation Model) , it could reduce and/or simplify the control tables, and the stat.prog. will not have to re-enter a lot of information

SADs 4

???

Contact

Erik Brun, System & Process Specialist
H. Lundbeck A/S
Ottiliavej 9
2500 Valby
Denmark
erik@lundbeck.com

Rico Schiller, Head of Section
H. Lundbeck A/S
Ottiliavej 9
2500 Valby
Denmark
rico@lundbeck.com