

Migration von Klinischen Studien nach CDISC-SDTM

**KSFE- Fachtagung
Hamburg, 23. Februar 2006**

Claudia Meurer

**Senior Statistical Programmer
Accovion GmbH, Eschborn/Frankfurt**



Inhalt

- ✓ CDISC - allgemein
- ✓ CDISC - SDS
- ✓ CDISC - SDTM
- ✓ Implementierung - SDTM
- ✓ Implementierung - ADM Macro
- ✓ Resümee

CDISC - allgemein

- **C** linical
- **D** ata
- **I** nterchange
- **S** tandards
- **C** onsortium

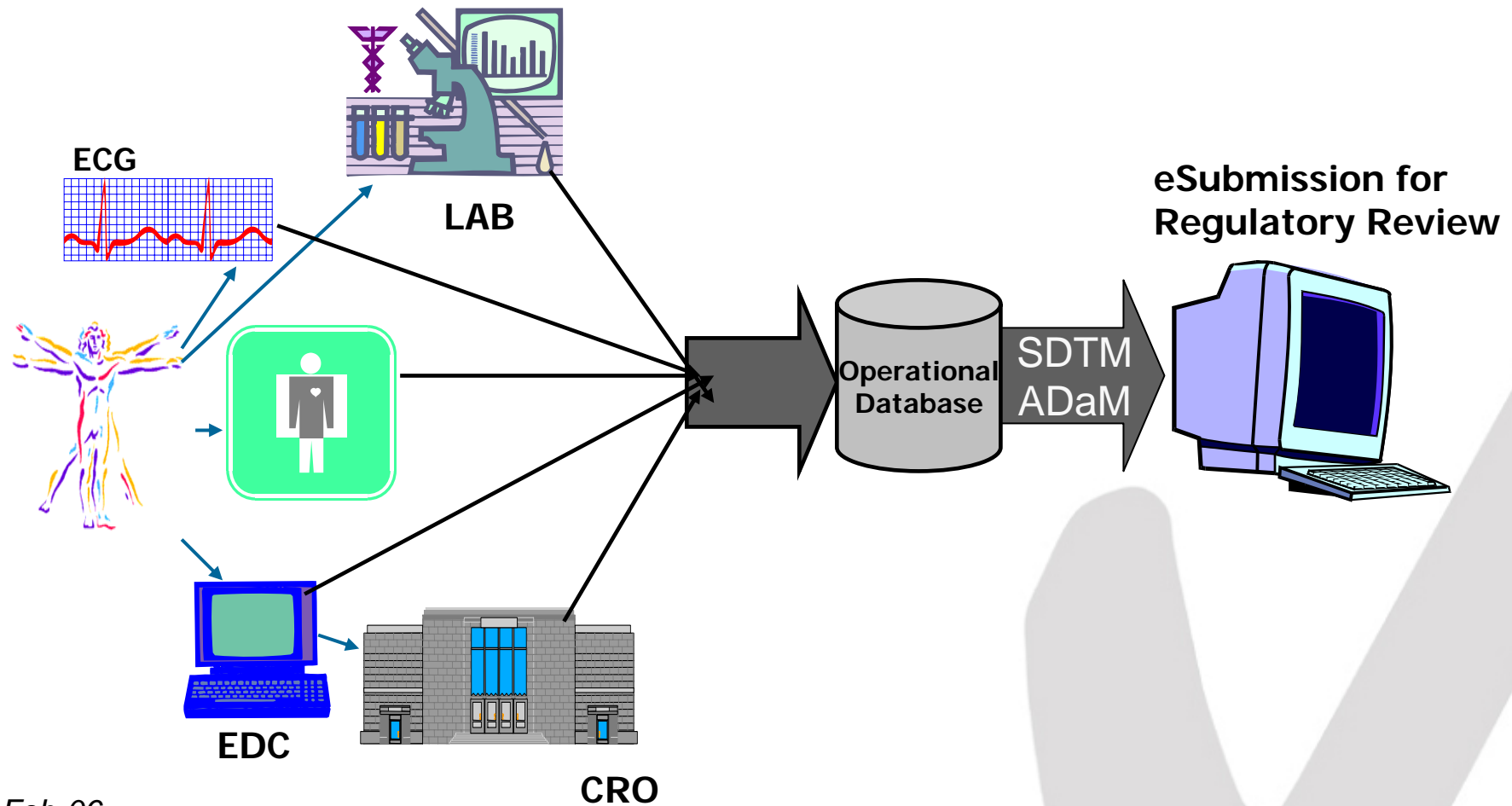
CDISC- allgemein

- ✓ globale, multidisziplinäre Gruppe mit Teilnehmern aus verschiedensten Arbeitsgebieten
- ✓ unabhängig von Behörden, Organisationen oder Firmen
- ✓ Zusammenarbeit mit Behörden
- ✓ Verschiedene Arbeitsgruppen (u.a. für SDTM, ADaM, LAB)

CDISC – allgemein: Ziele

- ✓ Einführung eines klinischen Datenstandards
- ✓ Anbieten von Schulungen
- ✓ einfachere Einreichung klinischer Studien bei Behörden
- ✓ Inhalt und Struktur der Daten plattform-unabhängig

CDISC – allgemein: Datentransfer



Inhalt

- ✓ CDISC - allgemein
- ✓ CDISC - SDS
- ✓ CDISC - SDTM
- ✓ Implementierung - SDTM
- ✓ Implementierung - ADM Macro
- ✓ Resümee

CDISC – SDS (I)

- ✓ Submission Data Standards, Version 3.1
- ✓ CDISC-Arbeitsgruppe, die Organisation, Struktur und Formatierung der Studiendaten für SDTM (Study Data Tabulation Model)- und ADaM (Analysis Data Model)- Datensätze genau definiert hat

CDISC – SDS (II): Domain

- ✓ Daten sind inhaltlich in einer sogenannten „Domain“ zusammengefasst (üblicherweise analog zum Case Record Form)
- ✓ entspricht Datensatz/Tabelle
- ✓ generelle Domain-Kategorien „Interventions“, „Findings“, „Events“
- ✓ spezielle Domains: u.a. Demografie (DM), Kommentare (CO)

CDISC – SDS (III): Domain-Kategorien

✓ Interventions

- alle Medikamente, Therapien, Operationen eines Patienten während der Studie
- Domains: EX, ES, CM

✓ Events

- Ereignisse, die vor oder während der Studie auftreten
- Domains: AE, DS, MH

CDISC – SDS (IV): Domain-Kategorien

✓ Findings

- Beobachtungen, die an geplanten Untersuchungsterminen gemacht werden
- Domains (u.a.) : LB, EG, VS

CDISC – SDS (V): Variablen

- ✓ Klassifizierung von Variablen anhand
 - Ursprung (Quelle, berechnet, zugewiesen)
 - Eigenschaften (Label, Typ, Code)
 - Funktion (Schlüssel, Focus des Datensatzes, Selektion, Beginn/Ende/Dauer, Ergebnis)

CDISC – SDS (VI): Variablen

Variable Name	Variable Label	Type	Controlled Terms or Format	Origin	Role	CDISC Notes
STUDYID	Study Identifier	Char		CRF	Identifier	Unique identifier for a study within the submission
DOMAIN	Domain Abbreviation	Char		Derived	Identifier	Two-characters abbreviation for the domain most relevant to the observation
USUBJID	Unique Subject Identifier	Char		Sponsor defined	Identifier	Unique identifier for a study within the submission
AETERM	Reported Term for the Adverse Event	Char	'NOT DONE' 'NONE' or AE.AETERM	CRF	Topic	Verbatim name of the event
AECAT	Category of Adverse Event	Char		Sponsor defined	Grouping Qualifier	Used to define a related category of related records
AECODE	Adverse Event Code	Num		Derived	Synonym Qualifier	
AESOCCD	Body System or Organ Class Code	Num	As per dictionary	Derived	Synonym Qualifier	

...

Inhalt

- ✓ CDISC - allgemein
- ✓ CDISC - SDS
- ✓ CDISC - SDTM
- ✓ Implementierung - SDTM
- ✓ Implementierung - ADM Macro
- ✓ Resümee

CDISDC – SDTM (I)

- ✓ Study Data Tabulation Model
- ✓ Datensätze, die die fest vorgebene Datenstruktur des CDISC-SDS benutzen
- ✓ spiegelt die Patientendaten aus dem CRF (Case Record Form) wieder
- ✓ soll dem Medizinischen Reviewer/Auditor einen Einblick in die Patientendaten geben

CDISDC – SDTM (II)

- ✓ Erstellung von SDTM Dateien – mögliche Lösungsansätze:
 1. SDTM wird bereits im CDMS (Clinical Data Management System, z.B. CLINTRIAL™ , OracleClinical™) definiert
 2. SDTM wird nach dem Einlesen der Daten in die SAS™ (Statistical Analysis Software) – Umgebung erstellt
 3. “hybrid”- Lösung

CDISDC – SDTM (III)

✓ 1. SDTM innerhalb CDMS

Vorteile:

- Rohdaten in der klinischen Datenbank entsprechen SDTM-Daten, einfache Kommunikation zwischen Datenmanager und Reviewer
- sobald die klinische Datenbank definiert ist, sind auch die SDTM-Datenstrukturen verfügbar

CDISDC – SDTM (IV)

✓ 1. SDTM innerhalb CDMS

Nachteile:

- aufwändig, da **alle** Daten in die CDISC-Struktur überführt werden müssen
- abhängig vom CDMS kann diese Strategie sehr umständlich sein und zu Qualitätsverlust führen

CDISDC – SDTM (V)

✓ 2. SDTM in der SAS™ -Umgebung

Vorteile:

- Änderungen an SDTM möglich, ohne den Datenmanagement-Prozess zu stören
- CDMS macht nur das, was es am besten kann: Dateneingabe, -management
- große Flexibilität von SAS™ erlaubt jede Transformation von CDMS-Daten in SDTM-Struktur

CDISDC – SDTM (VI)

✓ 2. SDTM in der SAS™-Umgebung

Nachteile:

- Verzögerung des Prozesses, erst nach der Erstellung der klinischen Datenbank können die SDTM-Datensätze erstellt werden
- Kommunikation zwischen Datenmanager und Reviewer durch unterschiedliche Datenstrukturen erschwert

CDISDC – SDTM (VII)

- ✓ 3. „hybrid“-Lösung:
SDTM im CDMS teilweise umgesetzt,
„Feintuning“ später in der SASTM-Umgebung

Vorteile:

- einfache Änderungen (Benennung von Datensätzen und Variablen, einfache Berechnungen) im CDMS leicht zu implementieren
- Anpassungen/Umrechnungen von SDTM-Variablen können mittels SASTM automatisch durchgeführt werden

CDISDC – SDTM (VIII)

✓ 3. „hybrid“-Lösung

Nachteile:

- zusätzliche Validierung der in der SASTM-Umgebung durchgeführten Transformationen notwendig
- einige Zeitverzögerung des Prozesses

Inhalt

- ✓ CDISC - allgemein
- ✓ CDISC - SDS
- ✓ CDISC - SDTM
- ✓ Implementierung - SDTM
- ✓ Implementierung - ADM Macro
- ✓ Resümee

Implementierung - SDTM (I)

✓ „hybrid“-Lösung

- Nachteile der beiden anderen Lösungsstrategien werden vermieden
- Pre-Processing innerhalb der klinischen Datenbank (CLINTRIAL™ bzw. OracleClinical™):
 - Anforderungen aus SDTM soweit wie möglich umgesetzt

Implementierung - SDTM (II)

✓ „hybrid“-Lösung

- Datenmanager, die die klinische Datenbank aufbauen, sind mit den SDTM-Strukturen vertraut (Kommunikation mit Reviewer vereinfacht)
- zusätzliche “Komfortvariablen” (z.B. numerische Datum-Variablen, numerische Codes statt Textfeldern) werden generiert, die später vom SAS™-Programmierer bei der Auswertung benötigt werden

Implementierung - SDTM (III)

✓ „hybrid“ -Verfahren“

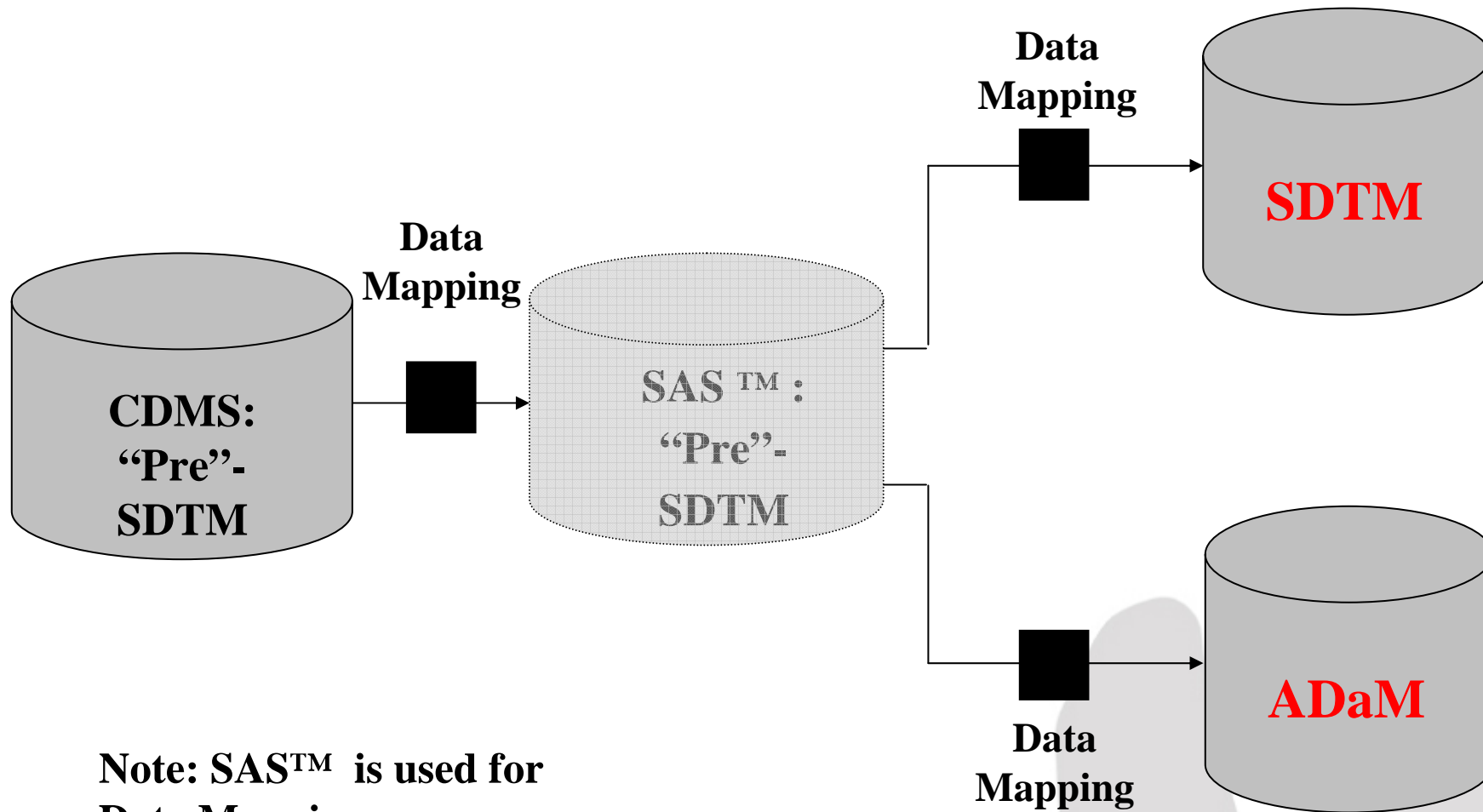
- Post-Processing in der SASTM-Umgebung:
 - zusätzliche Variablen aus anderen Datenbanken ergänzen (z.B. bei der Codierung von Nebenwirkungen)
 - Definition von Auswertungskollektiven für ADaM (Analysis Data Model)-Dateien

Implementierung - SDTM (IV)

✓ „hybrid“ -Verfahren“

- Post-Processing in der SAS™-Umgebung:
 - Definition von Analysevariablen, die beim Aufsetzen der klinischen Datenbank noch nicht feststehen (z.B. Analysezeitfenster, Baseline-, Endpoint-Definitionen) für SDTM- und ADaM-Dateien
 - Eliminieren der “Komfortvariablen” vor dem Transfer an die Behörden

Implementierung - SDTM (V)



Note: SAS™ is used for Data Mapping

Inhalt

- ✓ CDISC - allgemein
- ✓ CDISC - SDS
- ✓ CDISC - SDTM
- ✓ Implementierung - SDTM
- ✓ Implementierung - ADM Macro
- ✓ Resümee

ADM Macro (I) - allgemein

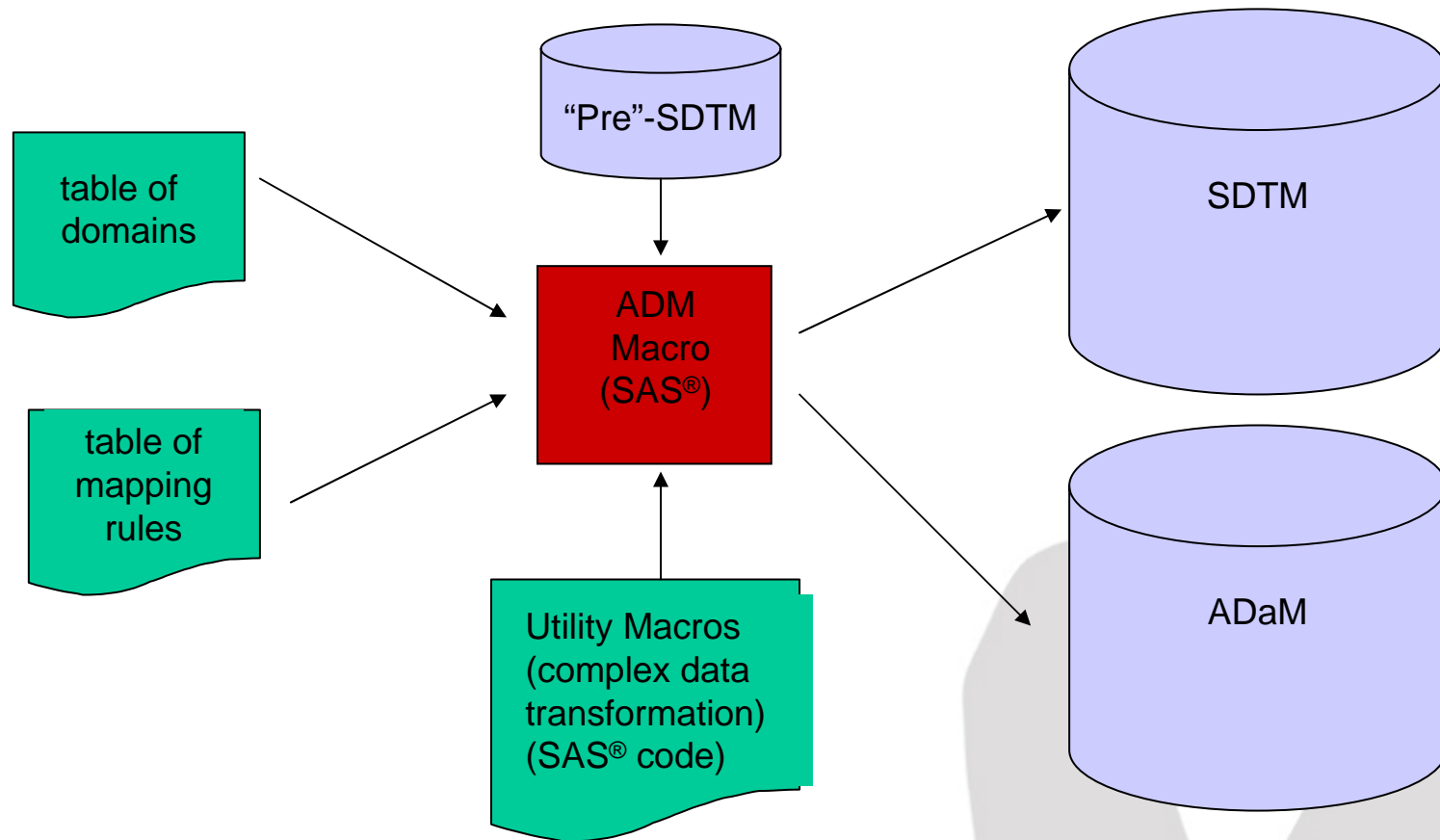
- ✓ ADM = Analysis Dataset Model
- ✓ SASTM-Macro als Tool zur SDTM- und ADaM-Datenerstellung
- ✓ SASTM-Programmcode-Generator, der pro Domain/Datensatz ein lauffähiges SASTM-Programm erstellt

ADM Macro (II) - Input

✓ Was wir benötigen:

- Excel™-sheet mit allen Domains/Datensätzen einer Studie
- je ein Excel™-sheet für jede dieser Domains/Datensätze mit allen zu erstellenden Variablen
- SAS™-Utility-Macros, z.B. zur Berechnung von Standardvariablen, die in mehreren Domains/Datensätzen benutzt werden

ADM Macro (III) - Konzept



ADM Macro (IV) – „table of domains“

✓ Excel™-sheet “table of domains”

- ein Excel™-sheet pro Studie
- legt Reihenfolge der Abarbeitung der Domains/
Datensätze fest
- legt Sortierreihenfolge der Records im späteren
SAS™-Datensatz fest
- wird vom Kunden geprüft, dient gleichzeitig zur
Dokumentation der Kundenvorgaben

ADM Macro (V) - „table of domains“

STUDYID	DOMN	DESCRIPTION	STRCT	SORTORDER
ACCO0001 /1234	EX	Exposure	I	USUBJID PERIOD EXSEQSP
ACCO0001 /1234	DM	Demographics	D	USUBJID
ACCO0001 /1234	AE	Adverse events	E	USUBJID AEBODSYS AECODE AETERM VISITNUM AESTDT AESTTM AEENDT AESEQSP
ACCO0001 /1234	CM	Concomitant medications	I	USUBJID CMCAT CMCODE CMSTDT CMENDT CMSEQSP CMDOSE CMDOSTXT
ACCO0001 /1234	DS	Disposition	E	USUBJID DSCAT DSTERM VISITNUM DSSTDT
ACCO0001 /1234	EG	ECG	F	USUBJID EGTESTCD VISITNUM EGSTDT EGSEQSP

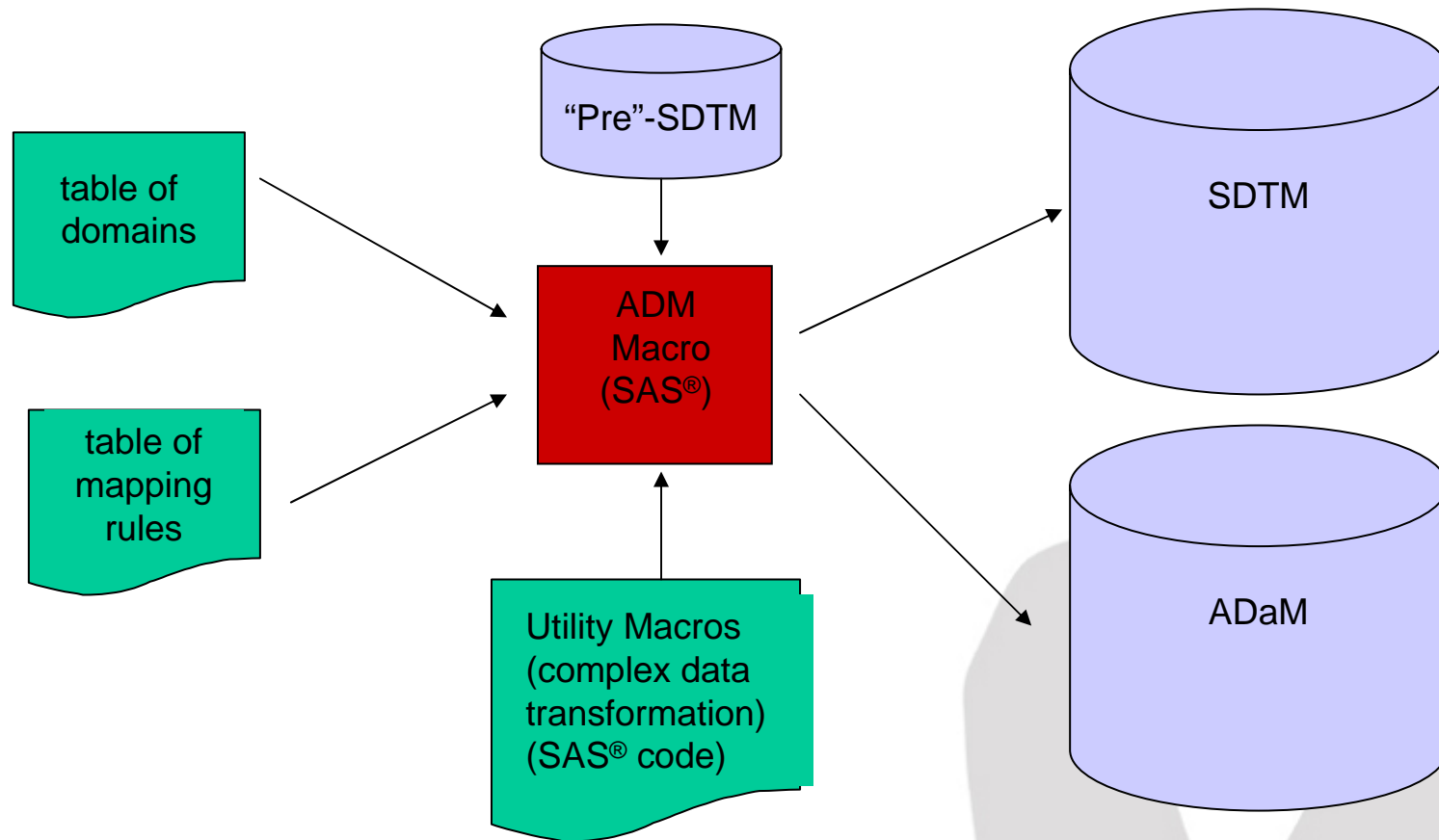
ADM Macro (VI) – „table of mapping rules“

- ✓ Excel™-sheet “table of mapping rules”
 - ein Excel™-sheet pro Domain/Datensatz
 - Definition jeder einzelnen Variablen
 - legt Abarbeitungsreihenfolge der Variablen fest
 - ruft SAS™-Utility-Macros auf
 - wird vom Kunden geprüft, dient gleichzeitig zur Dokumentation der Kundenvorgaben

ADM Macro (VII) – „table of mapping rules“

studyid	domain	variable	label	type	length	format	eval_ord	src_lib	src_dsn	src_var	task	transform
ACCO0001/123	DM						0-01				macro0	%mmerge(dsin=ODS.RANDOM ODS.DEMOG, dsout=DEMO, keep1=sunivid subno rndt, byvar=sunivid subno, inkeep=in2)
ACCO0001/123	DM						0-02				macro0	%mmerge(dsin=WORK.DEMO ODS.COMPWI ODS.INFCON, dsout=DEMO, keep2=sunivid subno w d, keep3=sunivid subno ifcndt, byvar=sunivid subno, inkeep=in1)
ACCO0001/123	DM	STUDYID	Study identifier	Char	20			WORK	DEMO		assignment	"ACCO0001/123"
ACCO0001/123	DM	DOMAIN	Domain Abbreviation	Char	2			WORK	DEMO		assignment	'DM'
ACCO0001/123	DM	USUBJID	Unique Subject Identifier	Char	40			WORK	DEMO		assignment	'ACCO0001/123' '1001/' PUT(_SUNVID,Z4.0) '/' PUT(_SUBNO,Z7.0)
ACCO0001/123	DM	BIRTHDT	Date of Birth	Num	8	DATE9.		WORK	DEMO	BIRTHDT	no change	
ACCO0001/123	DM	BIRTHCDT	Date of Birth (Text)	Char	11		1-02	WORK	DEMO		macro1	%date_num_to_char(invar=birthdt, outvar=birthcdt)
ACCO0001/123	DM	AGE	Age in AGEU	Num	8			WORK	DEMO	AGE	no change	
ACCO0001/123	DM	AGEU	Age Units	Char	8			WORK	DEMO		assignment	PUT(4,TIMEU_AVE)
ACCO0001/123	DM	AGEUCD	Age Units Code	Num	8			WORK	DEMO		assignment	4
ACCO0001/123	DM	SEX	Sex	Char	40			WORK	DEMO	SEX	derive	select(__sex); w hen('M') sex='Male'; w hen ('F') sex='Female'; otherw ise; end;
ACCO0001/123	DM	SEXCD	Sex Code	Num	8			WORK	DEMO	SEX	derive	select(__sex); w hen('M') sexcd=1; w hen ('F') sexcd=2; otherw ise; end;

ADM Macro (VIII)



Inhalt

- ✓ CDISC - allgemein
- ✓ CDISC - SDS
- ✓ CDISC - SDTM
- ✓ Implementierung - SDTM
- ✓ Implementierung - ADM Macro
- ✓ Resümee

Resümee (I)

- ✓ Excel™-sheets bedeuten ein erhöhtes Maß an Dokumentation, dies bewirkt aber auch eine hohe Transparenz
- ✓ hohe Flexibilität, studien- und kundenübergreifend
- ✓ jederzeit Updates innerhalb des Prozesses durch Änderungen an den Excel™-sheets möglich

Resümee (II)

✓ schneller und genauer

- schnellere Erstellung von SASTM-Datensätzen durch große Wiederverwertbarkeit von ExcelTM-sheets und SASTM-Utility-Macros
- geringere Fehleranfälligkeit durch das ADM-Macro und die SASTM-Utility-Macros
- geringerer Arbeits- und Trainingsaufwand

Resümee (III)

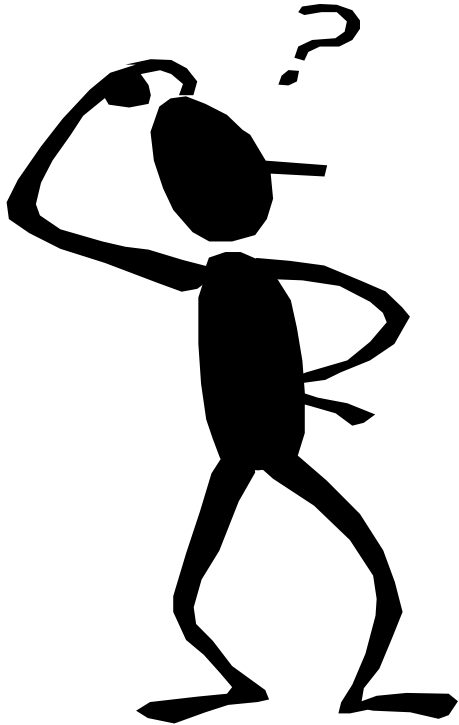
✓ effizienter

- leichte Integration von Studiendaten in bestehende Datenbanken
- durch die Benutzung von Standardmacros weniger Validierungsaufwand

✓ “mehr klinische Forschung für weniger Geld”

Referenzen

- ✓ www.cdisc.org
- ✓ www.cdisc.org/models/sds/V3.1
(Submission Data Standards)
- ✓ www.cdisc.org/models/sdtm/v1.1/index.html
(Study Data Tabulation Model, Version 1.1)
- ✓ www.fda.gov



Claudia Meurer
Statistical Programming

Accovion GmbH

Helfmann-Park 2

65760 Eschborn

Tel.: 06196/7709-387

claudia.meurer@accovion.com

Vielen Dank für Ihre Aufmerksamkeit !