

# Vorstellung des CDISC - Standards

---

CDISC – TMF-Workshop  
Berlin, 1.7.05

W. Kuchinke, KKS Düsseldorf

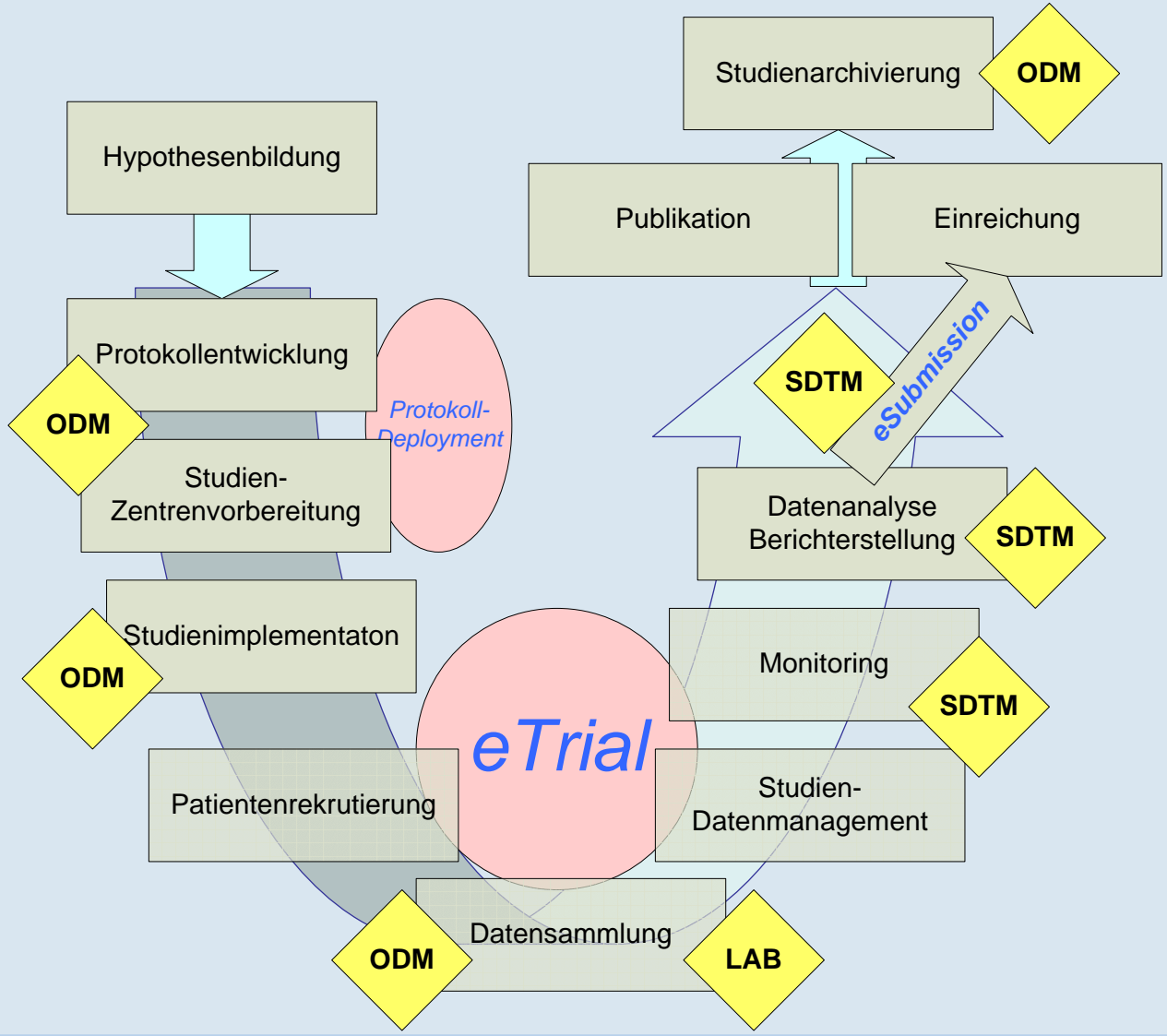
- XML als „Key Enabling Software Technology“
- neue Möglichkeiten zur Integration von Daten und Metadaten
- neue Möglichkeiten zur Vereinheitlichung
- Unterstützt komplexe Datenmodelle
- Automatisierung von Prozessen (Maschinenlesbarkeit)
- Erweiterte Möglichkeiten zur Harmonisierung



- Adverse Event Reporting (E2B-Standard)
  - FDA – AERS-Adverse Event Reporting System
- Common Technical Document
  - ICH - eCTD “XML Backbone”
- Structured Product labelling
  - SPL (New standard for Product Labeling content applications)
- Clinical Data
  - CDISC – ODM/SDM (Operational/Submission Data Model)

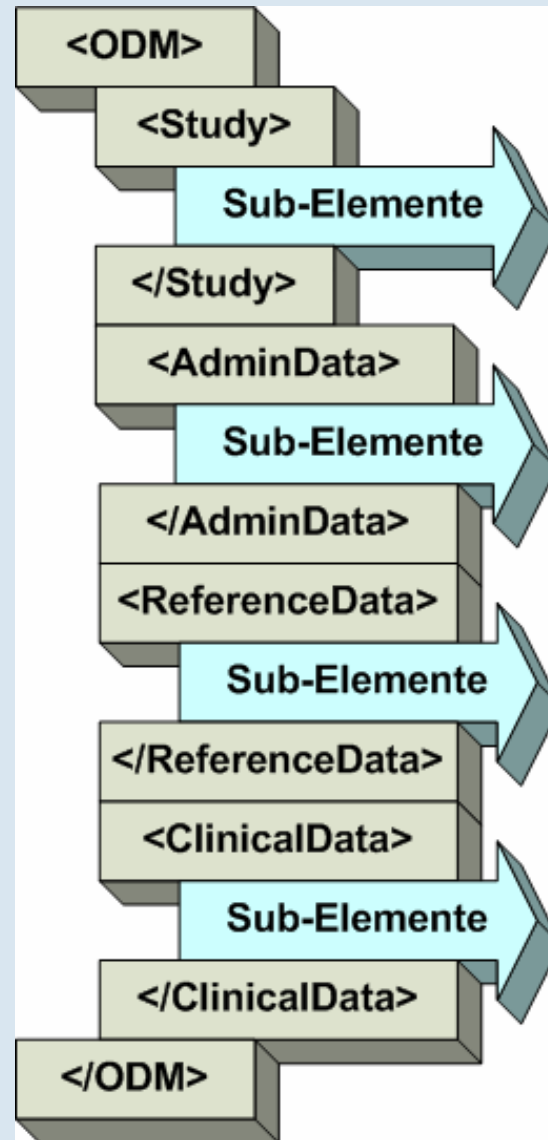


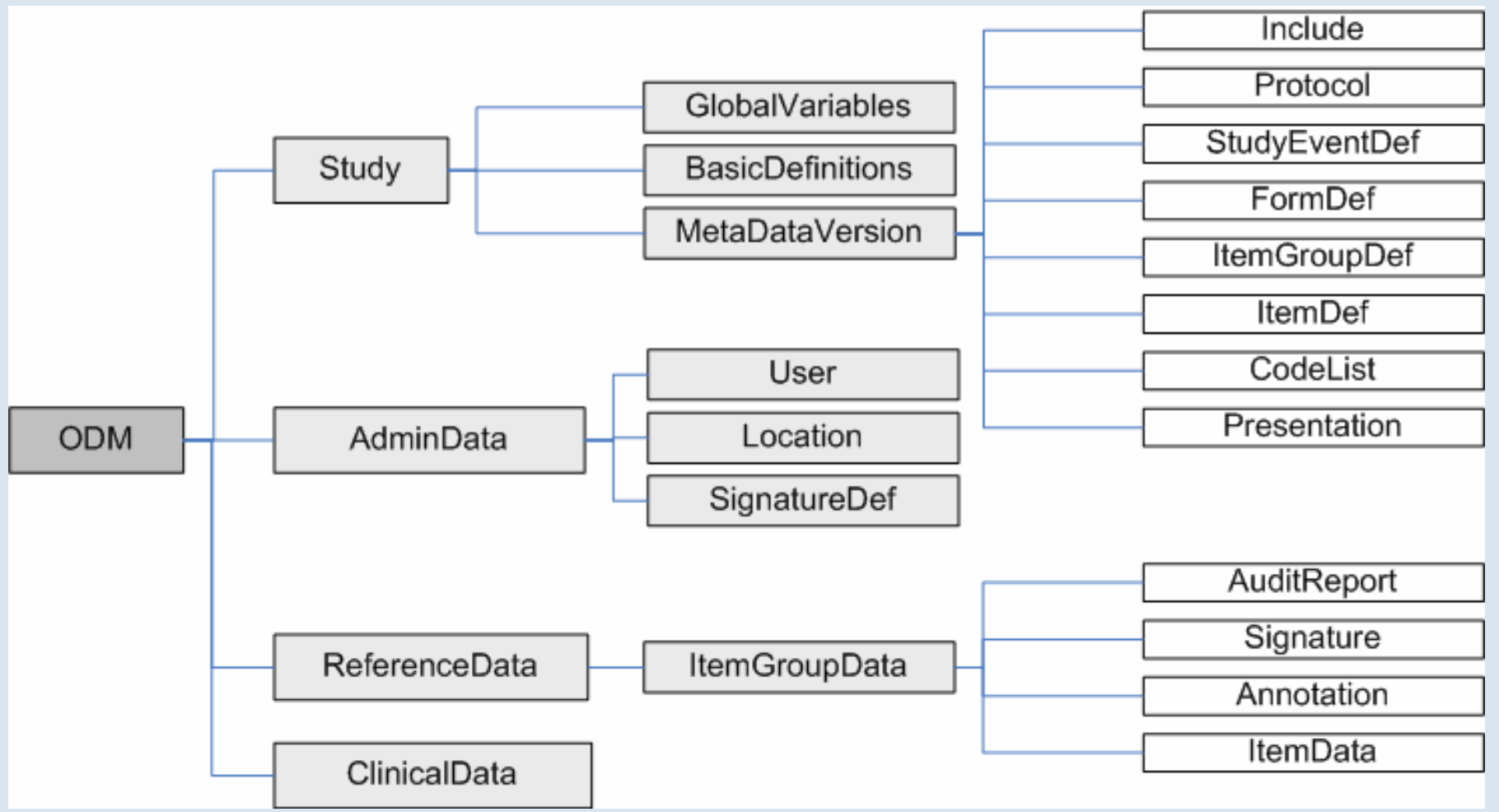
- ODM (Operational Data Model)
- SDS (Submission Domain Standard)
  - SDTM
  - SEND
- LAB (Laboratory Data Standard)
- ADaM



- vollständig XML-basiert
- Bildet die gesamte klinische Studie ab
- Hauptmodule:
  - Studienmetadaten
  - administrative Daten
  - Referenzdaten
  - klinische Daten
- Repräsentiert die Daten die zwischen unterschiedlichen Softwaresystemen in klinischen Studien ausgetauscht werden müssen

Zur Wohlgeformtheit gehört, daß jedes XML-Dokument ein Wurzelement (root tag) hat, welches das gesamte Dokument umschließt







```

<?xml version="1.0" encoding="ISO-8859-1" ?>
<!DOCTYPE ODM (View Source for full doctype...)>
- <ODM TransferType="Snapshot" Granularity="Study" TransferID="987-654-321" PriorTransferID="987-654-320"
  CreationDateTime="2001-10-16T13:27:45">
- <Study ID="123-456-789">
- <GlobalVariables>
  <StudyName>CDISC Test Study II</StudyName>
  <StudyDescription>This file contains test data for the CDISC Connect-A-Thon event scheduled
    for the DIA 11th Annual European Workshop on Clinical Data
    Management.</StudyDescription>
  <ProtocolName>CDISC-123-456-789</ProtocolName>
</GlobalVariables>
- <BasicDefinitions>
- <MeasurementUnit ID="MU.KG" Name="Kilogram">
  - <Symbol>
    <TranslatedText xml:lang="en">English: Kilogram</TranslatedText>
    <TranslatedText xml:lang="fr">French: Kilogram</TranslatedText>
  </Symbol>
</MeasurementUnit>

```

Studieninfo

Definitionen



# Verzeichnis einiger Elemente in ODM

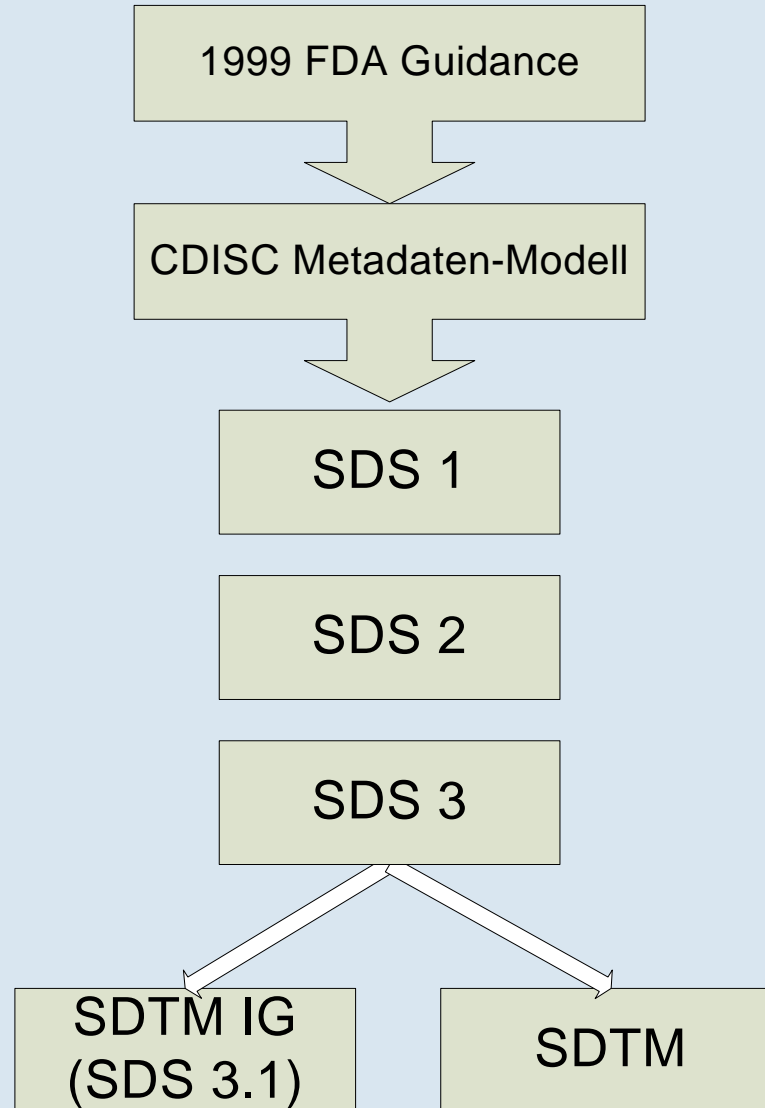
Address  
AdminData  
Annotation  
ArchiveLayout  
ArchiveLayoutRef  
AuditRecord  
BasicDefinitions  
Certificate  
CheckValue  
City  
ClinicalData  
ProtocolName  
Question  
RangeCheck  
ReasonForChange

ItemGroupData  
ItemGroupDef  
ItemGroupRef  
ItemRef  
LastName  
LegalReason  
Location  
LocationRef  
Role  
Signature Role  
Signature  
StreetName  
Study  
StudyDescription  
StudyEventData  
StudyName  
SubjectData

```

+ <Study OID="KKS">
+ <AdminData>
- <ClinicalData StudyOID="KKS" MetaDataVersionOID="MDV.001">
  - <SubjectData SubjectKey="SUBJECT.21" TransactionType="Insert">
    - <Annotation SeqNum="1">
      <Comment>zu löschen</Comment>
    </Annotation>
    <StudyEventData StudyEventOID="VISIT_2" />
    <StudyEventData StudyEventOID="AE" />
  - <StudyEventData StudyEventOID="CN">
    - <FormData FormOID="mnpkkstermination">
      - <AuditRecord>
        <UserRef UserID="USER.21" />
        <LocationRef LocationOID="LOCATION.1" />
        <DateTimeStamp>2004-05-17 17:30:57 +0200</DateTimeStamp>
      </AuditRecord>
    - <ItemGroupData ItemGroupOID="FS" TransactionType="Insert">
      <ItemData ItemOID="FS.0" Value="0" />
      <ItemData ItemOID="FS.1" Value="1" />
      <ItemData ItemOID="FS.2" Value="0" />
      <ItemData ItemOID="FS.3" Value="0" />
      <ItemData ItemOID="FS.4" Value="0" />
      <ItemData ItemOID="FS.5" Value="0" />
      <ItemData ItemOID="FS.6" Value="0" />
      <ItemData ItemOID="FS.7" Value="0" />
    </ItemGroupData>
    - <ItemGroupData ItemGroupOID="FG.448" ItemGroupRepeatKey=""
      TransactionType="Insert">
      <ItemData ItemOID="FF.behandlung_448" Value="0" />
    </ItemGroupData>
  </SubjectData>
</ClinicalData>

```



- SDTM (entspricht SDS V3 Kapitel 2 Specs)
- SDTM IG
- SDS V3.1 Comments Reconciliation
- CRT Data Definition Specifications (define.xml)

- Jede Domäne hat dieselbe Datensatzstruktur: Schlüsselvariable, Selections-Variable, Review-Variable, und Support-Variable
- Schlüsselvariablen werden zur eindeutigen Identifizierung des Datensatzes, Selektionsvariablen werden zur Unterteilung der Daten für die Berichterstellung benutzt
- Die SDM-Attribute sind z.B. Schlüsselvariable, Zweck, Datensatzstruktur,...
- Die Metadaten für die Variablen beinhalten die Basis-SAS Variablenattribute (Variablen-Name, Label, Beschreibung, Datentyp, Format)
- Dazu kommen SDM-spezifische Informationen (Ursprung der Variable, wo die Werte einer Variablen herkommen z.B. CRF, derived), Rolle)

Während dem ODM das CRF-Paradigma hinterlegt ist, zeigt das SDM einen weniger abstrakten Ansatz. Das SDM V2 besteht inzwischen aus 12 klinischen Standard-Domänen (SDTM besteht aus 23 definierten Domänen innerhalb sechs allgemeinerer Kategorien):

- **Demography**
- **AE**
- **Concomitant Medication**
- **ECG**
- **Drug Exposure**
- **Chemistry Labs**
- **Hematology Labs**
- **Urinalysis Labs**
- **Medical History**
- **Physical Examination**
- **Vital Signs**
- **Subject Disposition**

Alle Studienbeobachtungen müssen in eine von drei Klassen fallen:

- Interventionen = Behandlungen oder Prozeduren
- Events = Sachen die passieren (AE, Medizinische Vorgeschichte)
- Findings = allgemeine Patientenbeobachtung, z.B. Fragen, Tests (80% der Daten)



- Während der ODM Standard sowohl das Datenmodell, als auch seinen Darstellungsmechanismus via XML beschreibt, ist das Study Data Tabulation Model (SDTM) ein reiner Dateninhaltstandard
- Data Tabulation Data Sets sind eine Art die human Subject Case Report Tabulation (CRT) und entsprechende Tierdaten so zu präsentieren, daß sie beim FDA eingereicht werden können
- Möglichkeit zu custom-defined Domänen mit Sets von Standard Variable Definitions
- Der Prefix am Anfang aller Variablen ist eine Domänenabkürzung aus zwei Buchstaben
- Die SDTM-Metadaten für Domänen und Variablen werden mit Hilfe der Case Report Tabulation Data Description (CRT DD) Specificationen bereitgestellt

- Alle Daten als eine Reihe Beobachtungen gesammelt
- Grundlegende Patienten Informationen sind in „Demographics“
- Drei Beobachtungsklassen: Interventionen, Events, Findings
- Charakterisierung der Beobachtung:
  - Topic: Variable od. Parameter (Fokus der Beobachtung)
  - Timing (Beobachtung relativ zur Zeit)
  - Unique Identifiers (identifiziert Studie oder Subjekt)
  - Qualifiers (zusätzliche Merkmale der Beobachtung)
- Variablen können mehrere Rollen für verschiedene Zwecke haben

Behandlung	Ereignis	Findings	anderers
------------	----------	----------	----------

Exposure	Adverse Events (AE)	Labs	InclExcl	Demog
ConMeds	Disposition	Vitals	SubjChar	RELREC
SubstUse	MedHist	PhysExam	ECG	SuppQual
		QS*	andere	Comments
				TrialDesign

- Define.xml spezifiziert einen Standard für Datenbeschreibungen für Case Report Tabulations, die der Einreichung zulassungsrelevanter Daten an Behörden, wie die FDA, dienen
- Das XML-Schema, das die Struktur dieser XML-Datei definiert, basiert auf einer Erweiterung (Extension) des CDISC ODM-Modells
- Der define.xml-Standard basiert also auf ODM

SEND basiert auf SDTM. SEND definiert eine Reihe von Domänen. Eine Domäne ist eine Sammlung von topic-spezifischen Beobachtungen

## *Interventions*

- **Exposure — EX**

## *Special Purpose Domains*

- **Comments – CO**

## *Findings*

- **Animal Disposition — AD**
- **Body Weights — BW**
- **Clinical Pathology — CP**
- **Clinical Signs, Physical and Ophthalmoscopic Examination — CS**
- **Drug/Metabolite Levels — DL**
- **ECG Test Results — EG**
- **Food and Water Consumption — FW**
- **Group Characteristics — GC**
- **Macroscopic Findings — MA**
- **Microscopic Findings — MI**
- **Organ Weights — OW**
- **Subject Characteristics — SC**
- **Study Design Summary — SS**
- **Vital Signs – VS**

Findings – One record per lab test per subject

Variable Name	Variable Label	Type	Controlled Terms or Format	Origin	Role	Usage Notes
STUDYID	Study Identifier	Char			Identifier	Unique identifier for a study within the submission.
USUBJID	Animal Identifier	Char			Identifier	Animal identifier.
DOMAIN	Domain	Char	CP		Identifier	Two-character code for the domain most relevant to the observation.
CPREFID	Pool Identifier	Char			Identifier	Unique identifier for a pool of samples. Not used if there is no pooling of samples.
CPSEQ	Sequence Number	Num			Identifier	Sequence number given to ensure uniqueness within a dataset for an animal. Can be used to join related records.
CPGRPID	Group ID	Char			Identifier	Used to link together a block of related records for an animal in a single domain. This is not the treatment group number.
CPTESTCD	Lab Test Short Name	Char			Topic	Topic variable for CP. It is the short name for the value in CPTEST, which can be used as a column name when converting the dataset from a vertical format to a horizontal format. The short names can be up to 8 characters.
CPTEST	Lab Test Name	Char			Qualifier	Represents the name of the measurement. Name of the lab test.
CPCAT	Category for Test or Examination	Char	CHEM, HEM, UA, COAG, MN, OTHER		Qualifier	The following categories are used: CHEM = Clinical Chemistry HEM = Hematology UA = Urinalysis COAG = Coagulation MN = Rodent Micronucleus tests OTHER is used for findings that do not fit into any of the categories mentioned above.
CPSCAT	Subcategory for Test or Examination	Char			Qualifier	Contains subcategory for the test, if necessary. For example, Gross Analysis and Chemical Analysis

USUBJID	CPTSTCD	CPTST	CPCAT	CPSCAT	CPORRES	CPORRESU	CPSTAT	CPREASND
4103M	WBC	White Blood Cell Count	HEM		11.71	10E+9/L		
4103M	RBC	Red Blood Cell Count	HEM		10.62	10E+12/L		
4103M	HB	Hemoglobin	HEM		166	g/L		
4103M	pH	pH	UA	Chemical Analysis	8.5			
4103M	PROT	Protein	UA	Chemical Analysis	100	mg/dL		
4103M	COL	Color	UA	Gross Analysis	NORMAL			
4103M	TRAN	Transparency	UA	Gross Analysis	NORMAL			
4104M	WBC	White Blood Cell Count	HEM				NOT DONE	Not Sufficient Quantity
4104M	RBC	Red Blood Cell Count	HEM				NOT DONE	Not Sufficient Quantity
4104M	HB	Hemoglobin	HEM				NOT DONE	Not Sufficient Quantity

Beispiel für einen Datenset. Es zeigt Daten der klinischen Pathologie für ein Tier aus mehreren Tests und Zeitpunkten. ( 'NORMAL' = keine abnormalen Ergebnisse).

- Im Gegensatz zum hierarchischen ODM hat LAB eine flache hierarchische Organisation
- Basiert auf der Definition von „klinischen Labordaten“ durch die Pharmaindustrie
- Im April 2004 wurde die finale Präsentation des XML-Schema für das Models freigegeben
- Die flache Hierarchie beinhaltet z.B. Investigator-ID, Subjekt ID, Sex Value, Initialen, Geburtstag, Visit-ID



1. Good Transmission Practice
2. Study
3. Site
4. Investigator
5. Subject
6. Visit
7. Accession
8. Record Extension Type
9. Base Specimen
10. Base Battery
11. Base Test
12. Base Result

```
<?xml version="1.0" encoding="ISO-8859-1" ?>
```

```
- <!--
```

CDISC Lab Model: Sample Output File

- ```
- <GTP ModelVersion="01-0-01" CreationDateTime="2003-08-07T14:09:44-05:00">
  _ <TransmissionSource ID="A1234" Name="Central Lab ABC" />
  - <Study ID="CDISC Test 1" Name="CDISC Test 1" TransmissionType="C">
    - <Site ID="11">
      - <Investigator ID="11" Name="John Smith, M.D.">
        - <Subject>
          _ <ScreenID>8222</ScreenID>
          _ <Sex Value="M" CodeListID="HL7 V2.5 Gender Vocabulary Domain"
            />
          _ <Confidential Initials="ABC" Birthdate="1968-08-12" />
          - <Visit ID="01" Name="Screen" Type="S">
            - <Accession ID="C434382" LastActiveDateTime="2001-05-
              10T11:34:50-05:00">
              _ <CentralLab ID="C1234" Name="Central Lab ABC" />
              - <BaseSpecimen ID="2">
```

## ADaM

Zwei Modelle herausgegeben (Released for Comment):  
Change from Baseline  
Categorical Data Analysis  
Subject-Level Analysis" Model example (Review)  
ADaM General Considerations Version 1.0 (bestimmt den  
allgemeinen Inhalt, Struktur und Metadaten für Analyse  
Datasets) (Available)

## ODM V1.2.1

zwei neue Schemas zur besseren Unterstützung von  
Extensionen (Available for Implementation)

## **SDTM V1.1 (Published for Use)**

sieben neue Variablen

neuer Trial Summary Dataset

kleinere Änderungen

Case Report Tabulation Data Definition Specification  
(define.xml), Version 1.0 (Available for Implementation)

SDTM Controlled Terminology Package 1 Draft (Available for  
Comment)

## **SEND (Standard for Exchange of Nonclinical Data Models )**

SEND Version 2.1 (Draft)

## **LAB**

Base Model Version 1.0.1

Final XML Schema Version 1.0.1

Microbiology Extension (Review)

ECG Extension (Review)



**Vielen Dank für Ihrer Aufmerksamkeit!**

---

Mehr Informationen:

<http://www.uniklinik-duesseldorf.de/kks>

<http://www.tmf-ev.de/>