



Setting the
Global Standard
for Clinical Data

ODM-SDTM mapping

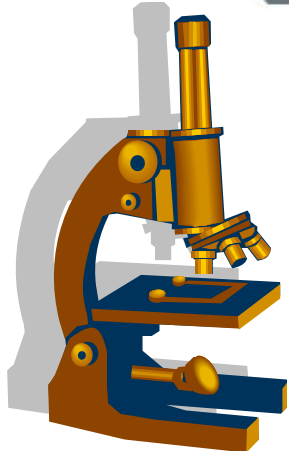
Nicolas de Saint Jorre, XClinical

June 20, 2008
French CDISC User Group
Bagneux/Paris

© CDISC & XClinical, 2008

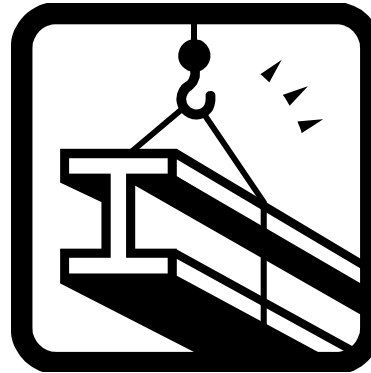
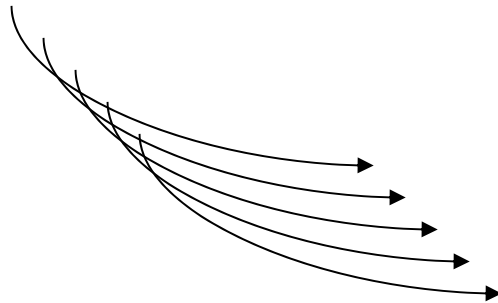
- Founded 2002, fast growing EDC-CDM vendor with offices in Munich, Paris and Cambridge, USA
- Active member of CDISC
- MARVIN, an online, integrated EDC-CDM System
 - GCP system validated, 21 CFR 11 compliant
 - **ODM certified**
- ODM study composer, SDTM tabulator tool
- 50+ studies, 50.000+ patients
- Financially independent and profitable





Terminology

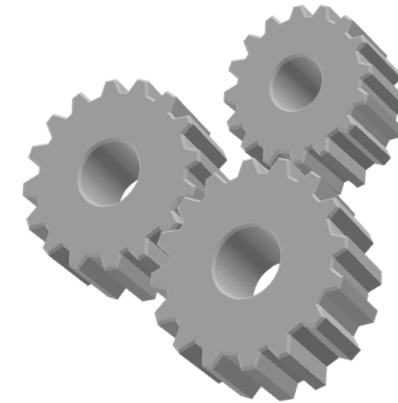
How do we call things?



Structure

Where do we put things?

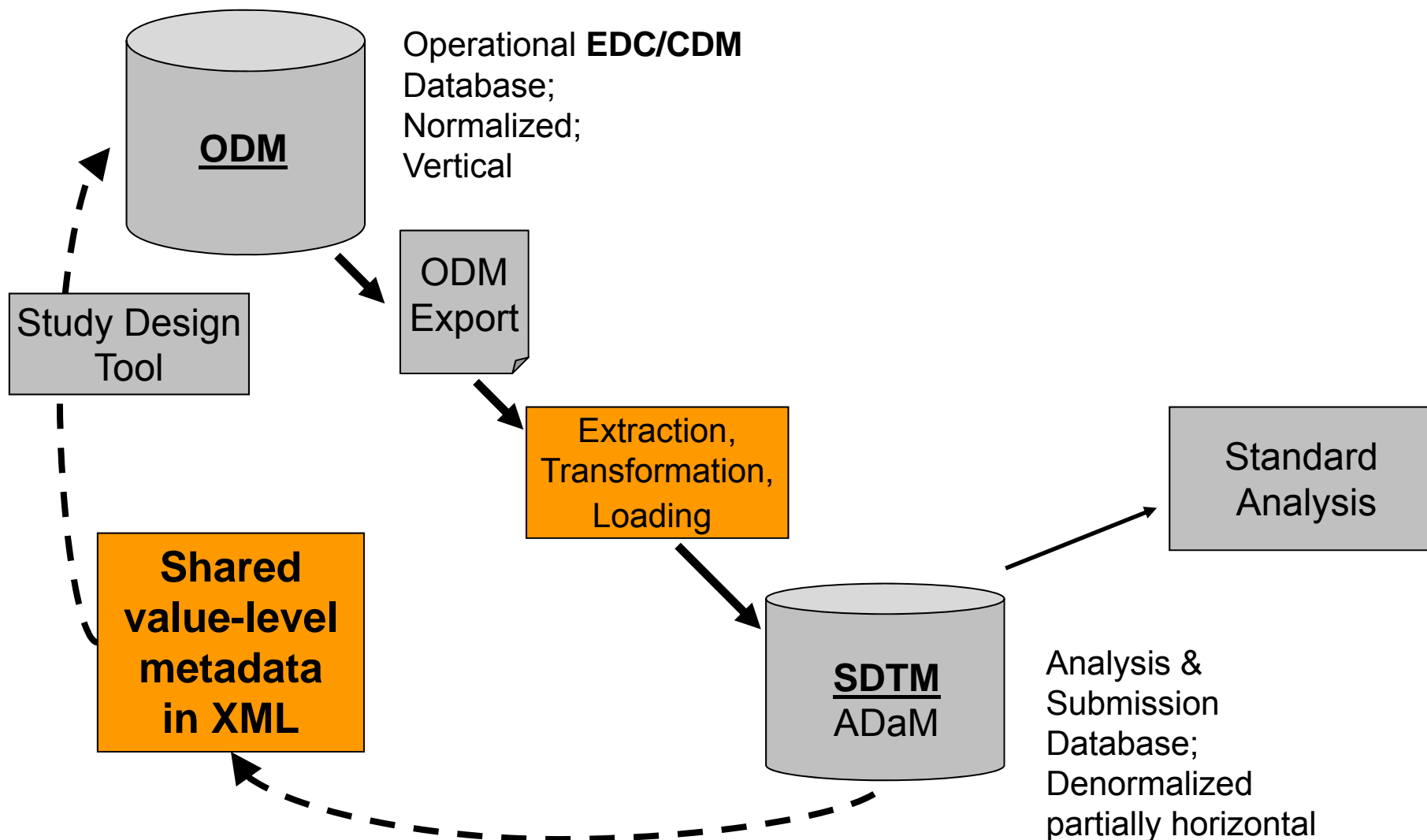
- Protocol (PRG)
- CRF pages
(**ODM**, CDASH)
- Lab forms (LAB)
- Data tables
(SDTM, **define.xml**)
- Analysis tables (ADaM)

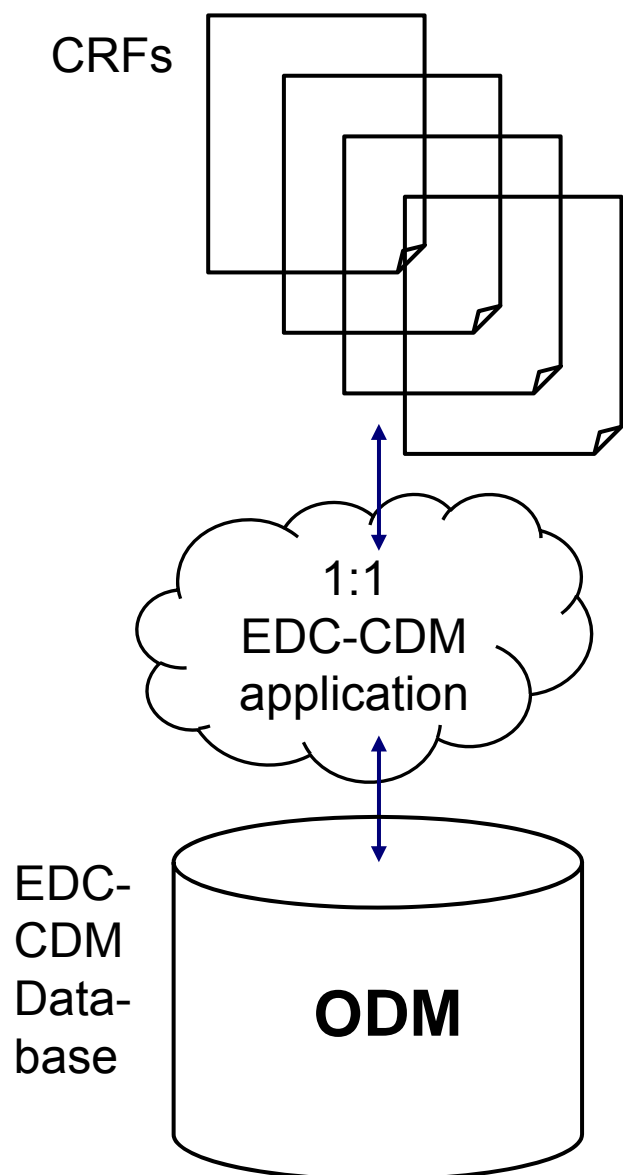


Format

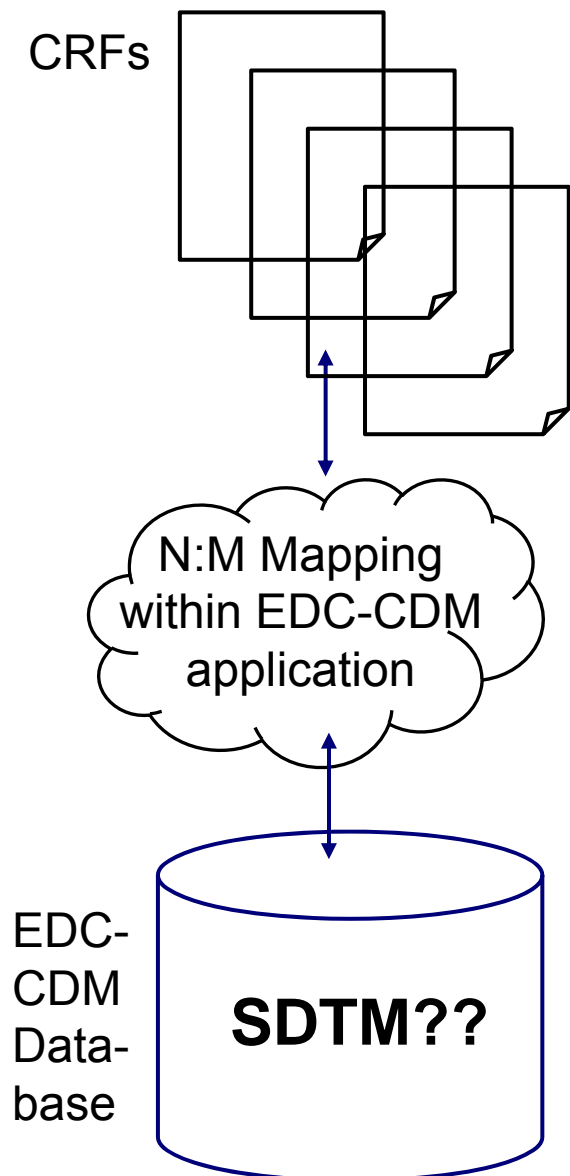
How do computers store and exchange things?

- XML (**ODM**)
- ~~ASCII~~
- ~~SAS~~





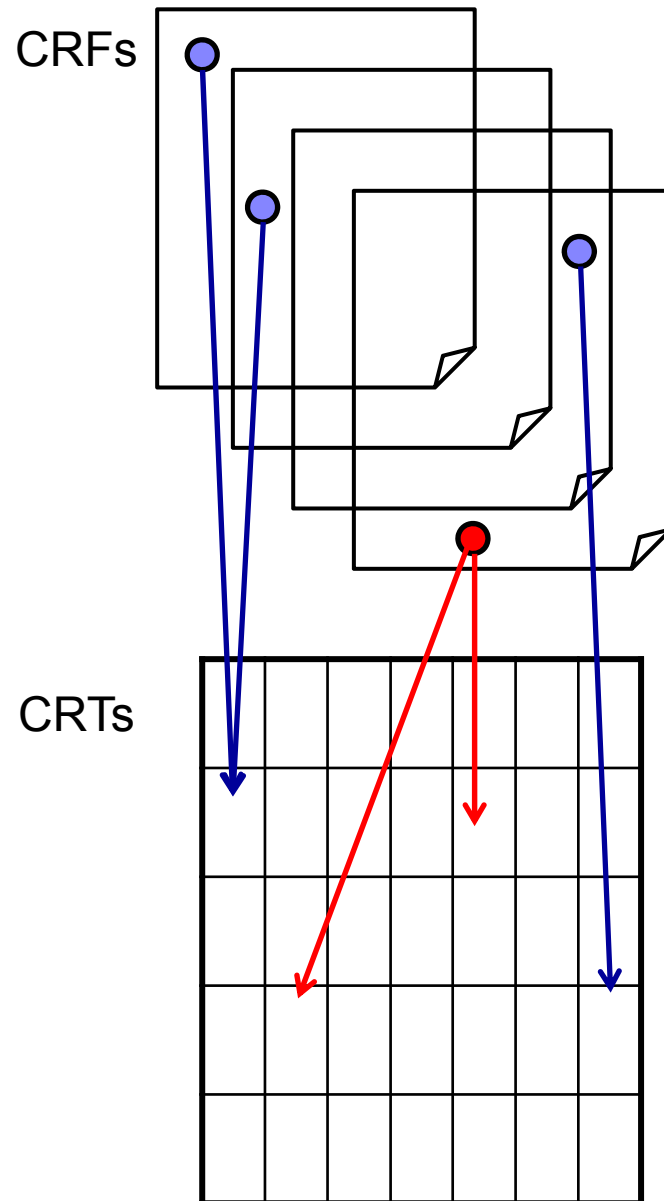
- ODM contains audit trail, signatures, internationalization
- ODM is extremely flexible to adapt to any kind of CRFs
- Mapping between ODM and CRFs is trivial (1:1)
- ODM contains XML-based value-level metadata that can be shared with SDTM
- ODM can integrate the SDTM controlled terminology



- SDTM does not contain an audit trail
- Audit trail, signatures, administrative data would have a proprietary format within the EDC-CDM application
- Mapping between SDTM tables and CRF pages needed for every trial

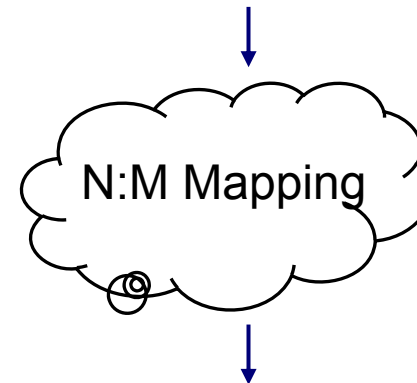
- Data is captured and cleaned (→ audit trail) in an EDC database and is exported in ODM format
 - Single ODM items may/can/should conform to the SDTM standard (value-level metadata)
 - But CRF forms do not fully match SDTM domains
- CRF Data exported from an EDC system needs to be mapped into SDTM domains

- ODM-SDTM mapping is nothing else than what data managers have always done:
pull data out of skinny, vertical, normalized tables into partially denormalized tables
- ODM data are typically stored all in one items table, or they may be stored in itemgroup- or form-level tables, corresponding to the CRF structure
- SDTM domains are technically represented by a different set of database tables, some horizontal (like demographics or adverse events), some vertical (like vital signs, physical examinations, etc.)



1:1

ODM = CRF metadata and data = Visits, Forms, ItemGroups, Items



SDTM = Case Report Tabulations = ***different*** (!) „ItemGroups“, Items

- Decoding:
ODM ClinicalData are coded, SDTM ORRES variables are decoded
- Transcoding:
the CRF may have used different codes than SDTM requires
- Labeling:
some ODM item names may not correspond to SDTM controlled terminology yet
- Measurement Units:
the EDC system may collect data in different measurement units, SDTM requires that the result is also calculated in standard units
- Time durations:
the ODM (= CRF) database typically collects time points, SDTM needs time durations

- 1:1 mapping
 - Date of birth on a CRF page → Column “BRTHDTC” in SDTM DM table (horizontal)
 - Sex on a CRF page → Column “SEX” in SDTM DM table (horizontal)
 - Weight and Height on a CRF page → Weight and Height in the column “VSORRES” of the SDTM VS table (vertical)
- 1:N mapping
 - Visit date on one CRF page → Visit date in many SDTM tables
- M:1 mapping
 - Date of FU visit on a CRF page - Date of baseline Visit on another CRF page → Study day in SDTM




WHAT IS SDTM VALUE-LEVEL METADATA? (1/2)

define.xml

```
<ItemDef OID="VS.VSTESTCD" Name="VSTESTCD" DataType="text"
  Length="8"
  Origin="CRF Page"
  def:Label="VITAL SIGNS TEST SHORT NAME">
  <def:ValueListRef ValueListOID="VL.VSTESTCD"/>
</ItemDef>
```

Actual TESTCD values limited to 8,
OIDs not limited



```
<def:ValueListDef OID="VL.VSTESTCD">
  <ItemRef ItemOID="VS.HEIGHT" OrderNumber="1" Mandatory="Yes"/>
  <ItemRef ItemOID="VS.WEIGHT" OrderNumber="2" Mandatory="Yes"/>
  <ItemRef ItemOID="VS.BMI" OrderNumber="3" Mandatory="Yes"/>
  <ItemRef ItemOID="VS.SYSBP" OrderNumber="4" Mandatory="Yes"/>
  <ItemRef ItemOID="VS.DIABP" OrderNumber="5" Mandatory="Yes"/>
  <ItemRef ItemOID="VS.POS" OrderNumber="6" Mandatory="Yes"/>
  <ItemRef ItemOID="VS.PULSE" OrderNumber="7" Mandatory="Yes"/>
</def:ValueListDef>
```



WHAT IS SDTM VALUE-LEVEL METADATA? (2/2)

```
<ItemDef OID="VS.SYSBP"
```

```
  Name="Systolic Blood Pressure" /
```

```
  def:Label="Systolic Blood Pressure"
```

```
  DataType="float"
```

```
  Length="5" SignificantDigits="2"
```

```
  SASFieldName="SYSBP" SDSVarName="SYSBP">
```

```
  <Question>
```

```
    <TranslatedText xml:lang="en">Systolic Blood Pressure</TranslatedText>
```

```
    <TranslatedText xml:lang="de">Systolischer Blutdruck</TranslatedText>
```

```
  </Question>
```

```
  <MeasurementUnitRef MeasurementUnitOID="MU.mmHg"/>
```

```
</ItemDef>
```

```
  Name="SYSBP"
```

define.xml

ODM.xml

Sharing metadata between SDTM (define.xml) and CRF metadata (ODM.xml) is easy

- SAS
 - programming effort
 - GCP system validation effort
- Programming in PLSQL etc.
- XML based commercial tools

- Define Domains and Variables in define.xml format
- Define rules to write records into domains (as an extension of the define.xml schema)
- Refer to ClinicalData in an ODM file using XPath syntax to tell the tool which ODM items should go into which place in the SDTM domains
- Run an XSLT or a Java engine to execute the rules and produce SDTM datasets in XML format (which may in the end also be stored in relational database tables)
- Hide all the XML technology behind a drag-and-drop graphical user interface



Tabulator: XML-based mapping tool



Tabulator - workshop/tabulator.xml - Eclipse SDK

File Edit Navigate Search Project Run Window Help

Validate Tbt

Tabulator

Input ODM

(Subject Protocol)

- EVT00001 - "Screening.Visit 1"
- FRM00001 - "Form: Obtain informed consent"
- FRM00002 - "Form: Assign patient number"
- FRM00003 - "Form: ADAS-Cog(14) questionnaire"
- FRM00004 - "Form: Modified Hachinski Ischemic Score"
- FRM00005 - "Form: Mini-mental State Examination"
- FRM00006 - "Form: CIBIC+ questionnaire"
- FRM00007 - "Form: NPI-X questionnaire"
- FRM00008 - "Form: HbA1c"
- FRM00009 - "Form: UA w/ Dipstick, no Micro"
- FRM00010 - "Form: Disability Assessment for Dementia"
- FRM00011 - "Form: Comprehensive Hist, Phys & Vitals"
- ITG00005 - "Comprehensive Hist, Phys & Vitals"
- ITD00022 - "W(lb)"
- ITD00023 - "H(in)"
- ITD00024 - "TempO"
- ITD00025 - "DOB"
- ITD00026 - "DAP"
- ITD00027 - "Pulse"
- ITD00028 - "Race"
- ITD00029 - "Sex"
- ITD00030 - "Standing DBP"
- ITD00031 - "Standing DBP3"
- ITD00032 - "Standing SBP"
- ITD00033 - "Standing SBP3"
- ITD00034 - "Supine DBP"
- ITD00035 - "Supine SBP"
- FRM00012 - "Form: Habits Questionnaire: Alcohol, Tob"
- FRM00013 - "Form: Chest Xray"
- FRM00014 - "Form: ECG w/ Interpret. & Report"
- FRM00015 - "Form: Placebo test"
- FRM00016 - "Form: Head CT Scan w/ Contrast"
- FRM00017 - "Form: Review Concomitant Medications"
- FRM00018 - "Form: Chemistry Panel"
- FRM00019 - "Form: CBC w/ Plate & Auto Diff"
- FRM00020 - "Form: Thyroid Function Tests"
- FRM00021 - "Form: Serum Folate"
- FRM00022 - "Form: Serum B-12"
- FRM00023 - "Form: VDRL"
- FRM00024 - "Form: Adverse Event Data Capture"
- EVT00002 - "Screening.Visit 2"
- EVT00003 - "Placebo.Visit 3"
- EVT00004 - "Placebo.Visit 4"
- EVT00005 - "Placebo.Visit 5"
- EVT00006 - "Placebo.Visit 7"
- EVT00007 - "Placebo.Visit 8"
- EVT00008 - "Placebo.Visit 9"
- EVT00009 - "Placebo.Visit 10"
- EVT00010 - "Placebo.Visit 11"
- EVT00011 - "Placebo.Visit 12"

Mappings

Domain: VS (2 Selectors)

Selector: Sel.1

Mapping

Selector OID: Sel.1

Context path: /SubjectData/StudyEventData[@OID="EVT00001"]/FormData[@OID="FRM00011"]/ItemGroupData[@OID="ITG00005"]/ItemData[@OID="ITD00026"]

Path Variables:

```
$contextNode := .  
$parentNode := ..  
$SubjectKey := /SubjectData/@SubjectKey
```

Condition: isSet(\$contextNode)

Writers:

STUDYID	C	DEMO-TRIAL
DOMAIN	C	VS
USUBJID	E	concat("DEMO-TRIAL-", \$SubjectKey)
VSSEQ	S	VS.Seq1
VSTESTCD	C	VSSYSBP
VSTEST	C	Systolic Blood Pressure
VSPOS	N	[null]
VSORRES	E	\$contextNode
VSORRESU	C	mmhg
VSTRESN	N	[null]
VSTRESU	N	[null]
VSTAT	N	[null]
VSLOC	N	[null]
VISIT	N	[null]
VISITNUM	N	[null]
VISITDY	N	[null]
VSDTC	N	[null]
VSDY	N	[null]

CRT Variables

OID	Da...	L...	Label
AEACN	text	200	Action Taken with Study Treatment
AEACNOTH	text	200	Other Action Taken
AEBODSYS	text	100	Body System or Organ Class
AECAT	text	200	Category for Adverse Event
AECONTRT	text	10	Concomitant or Additional Trtmtnt Given
AEDECOD	text	200	Dictionary-Derived Term
AEDUR	text	510	Duration of Adverse Event
AEENDTC	text	17	End Date/Time of Adverse Event
AEENDY	int...	8	Study Day of End of Adverse Event
AENRNF	text	20	End Relative to Reference Period
AEGRPID	text	200	Group ID
AELOC	text	200	Location of the Reaction
AEMODIFY	text	250	Modified Reported Term
AEOCCUR	text	200	Adverse Event Occurrence
AEOOUT	text	200	Outcome of Adverse Event Char
AEPATT	text	200	Pattern of Adverse Event
AEREPID	text	100	Reference ID
AEREL	text	200	Causality
AERELNST	text	200	Relationship to Non-Study Treatment
AESCAN	text	100	Involves Cancer
AESCAT	text	200	Subcategory for Adverse Event
AESCONG	text	100	Congenital Anomaly or Birth Defect
AESDISAB	text	100	Persist or Signif Disability/Incapacity
AESDTH	text	100	Results in Death
AESEQ	int...	10	Sequence Number
AESER	text	100	Serious Event
AESEV	text	100	Severity/Intensity
AESHOSP	text	100	Requires or Prolongs Hospitalization
AESLIFE	text	100	Is Life Threatening
AESMIE	text	1	Other Medically Important Serious Even
AESOD	text	100	Occurred with Overdose
AESPID	text	510	Sponsor-Defined Identifier
AESTDTC	text	17	Start Date/Time of Adverse Event
AESTDY	int...	8	Study Day of Start of Adverse Event
AETERM	text	510	Reported Term for the Adverse EventVt
AETOXGR	text	510	Standard Toxicity Grade
AGE	int...	2	Age in AGEU at RFSTDTC
AGEU	text	5	Age Units
ARM	text	50	Description of Element
ARMCD	text	50	Arm Code
BRTHDTC	text	10	Date/Time of Birth
COUNTRY	text	5	Country
DMDTC	text	10	Date/Time of Collection
DMDY	int...	4	Study Day of Collection
DOMAIN	text	2	Domain Abbreviation
IDVAR	text	5	Identifier Variable
IDVARVAL	text	5	Identifier Variable Value
QLABEL	text	250	Qualifier Variable Label
QNAM	text	20	Qualifier Variable Name
QORIG	text	10	Origin

Overview Domains Mappings Source

- Use XML-based metadata as early as possible, enabling automatic generation of e-CRFs, paper CRFs, annotated CRFs, queries, etc.
- Share value-level metadata and controlled terminology between EDC database and SDTM data repository
- Use the freedom of ODM to design whatever CRF pages you and your investigators like; mapping this to SDTM is easy (once you've done it the first time 😊)

Contact:

Nicolas de Saint Jorre

XClinical France

86, Rue du Cherche-Midi

75006 Paris

France

nicolas.desaintjorre@xclinical.com

Tél.: +33 (0) 1 / 45 44 14 25

Fax: +33 (0) 1 / 45 44 14 25