Metadata Management Tools for Clinical Trials

Seminar "Big Data in Personalised Medicine"

Tom-Hendrik Hülsmann
WS 2017/18

- Introduction
- Clinical Trial Metadata
- Standards (CDISC)
- Management Tools
- Conclusion

Introduction

Clinical Trials

- Key part of the medical research and development circle.
- Used to validate scientific assumptions and prove safety of new drugs.
- Required by government agencies.
- Generate a large amount of data, that has to be stored and organized.

Clinical Trial Metadata

„Metadata is data about data.“

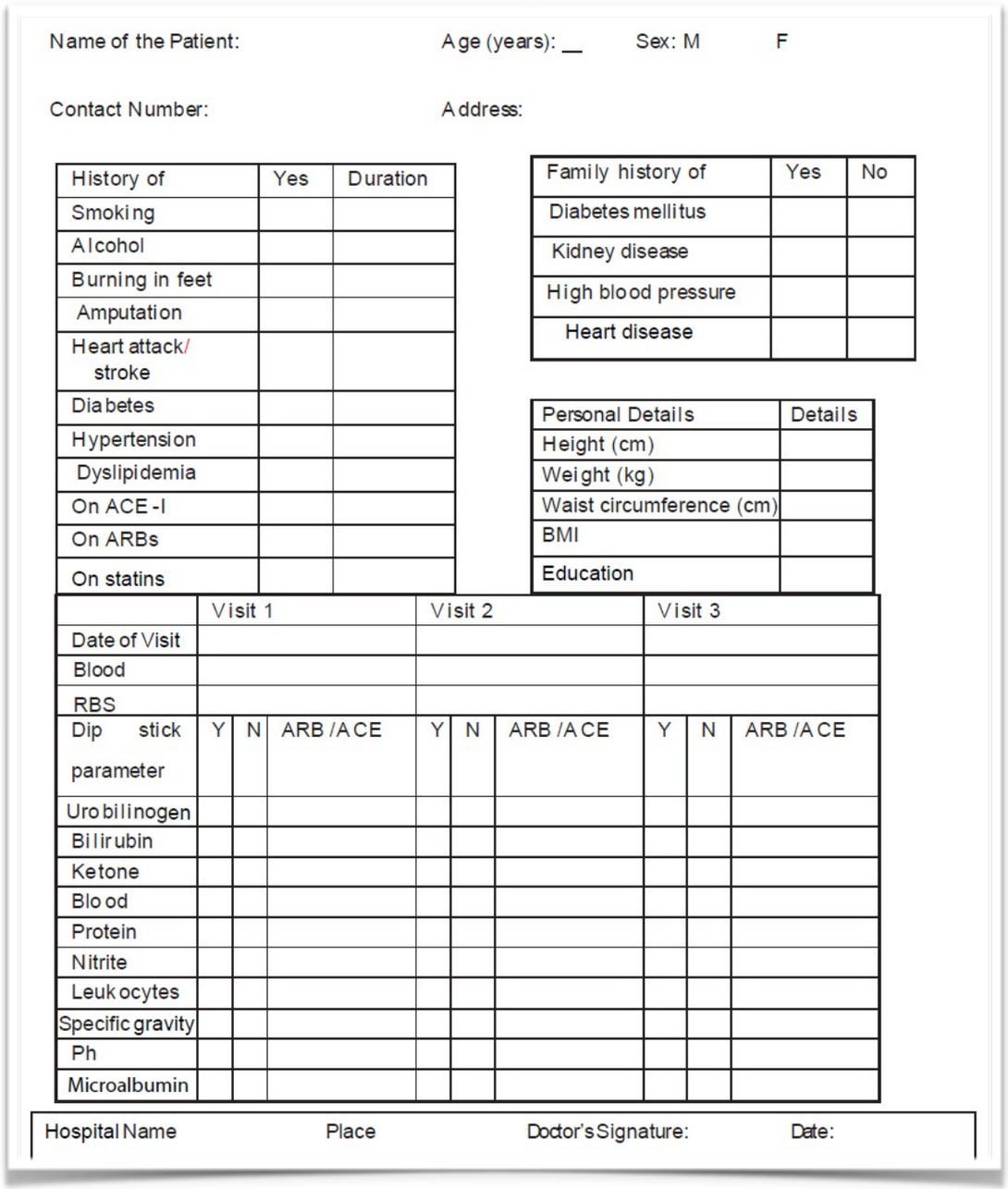
–Everyone

Example

- 38.5
- The datapoint has no meaning without metadata.
- Different metadata types include:
  - **Type:** *Temperature*
  - **Unit:** *Celsius*
  - **Subject-ID:** *404*
  - **Time of measurement:** *31st January 2018*

Case Report Forms

- Important part of every trial.
- Filled out by medical doctors while visiting a subject during a clinical trial.
- Contain data with different types of metadata attached.
- Used to evaluate the performance of e.g. a new drug.



Name of the Patient: Age (years): __ Sex: M F

Contact Number: Address:

History of	Yes	Duration
Smoking		
Alcohol		
Burning in feet		
Amputation		
Heart attack/ stroke		
Diabetes		
Hypertension		
Dyslipidemia		
On ACE-I		
On ARBs		
On statins		

Family history of	Yes	No
Diabetes mellitus		
Kidney disease		
High blood pressure		
Heart disease		

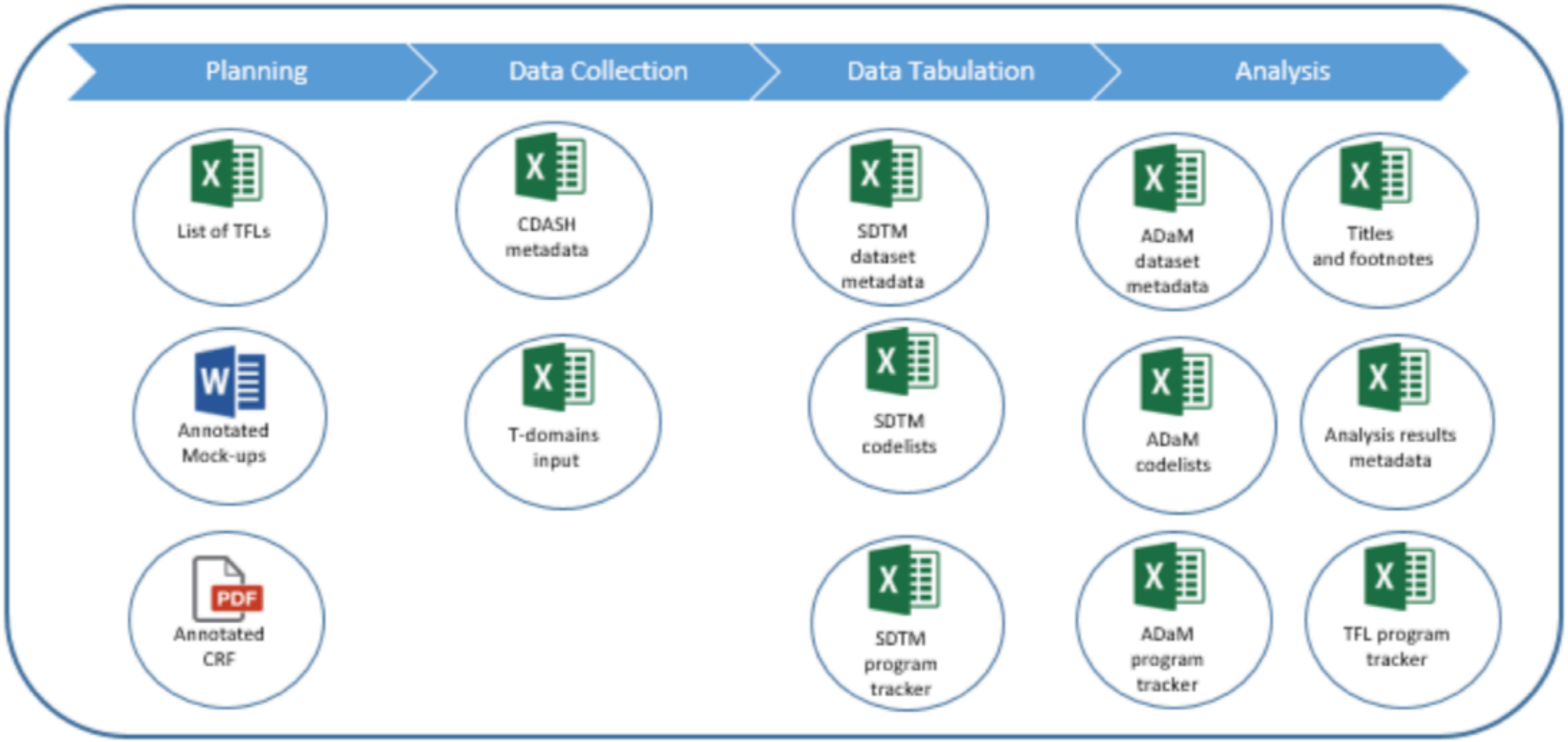
Personal Details	Details
Height (cm)	
Weight (kg)	
Waist circumference (cm)	
BMI	
Education	

	Visit 1			Visit 2			Visit 3		
Date of Visit									
Blood									
RBS									
Dip stick parameter	Y	N	ARB /ACE	Y	N	ARB /ACE	Y	N	ARB /ACE
Urobilinogen									
Bilirubin									
Ketone									
Blood									
Protein									
Nitrite									
Leukocytes									
Specific gravity									
Ph									
Microalbumin									

Hospital Name Place Doctor's Signature: Date:

[A]

Present and Past Management



Problems & Challenges

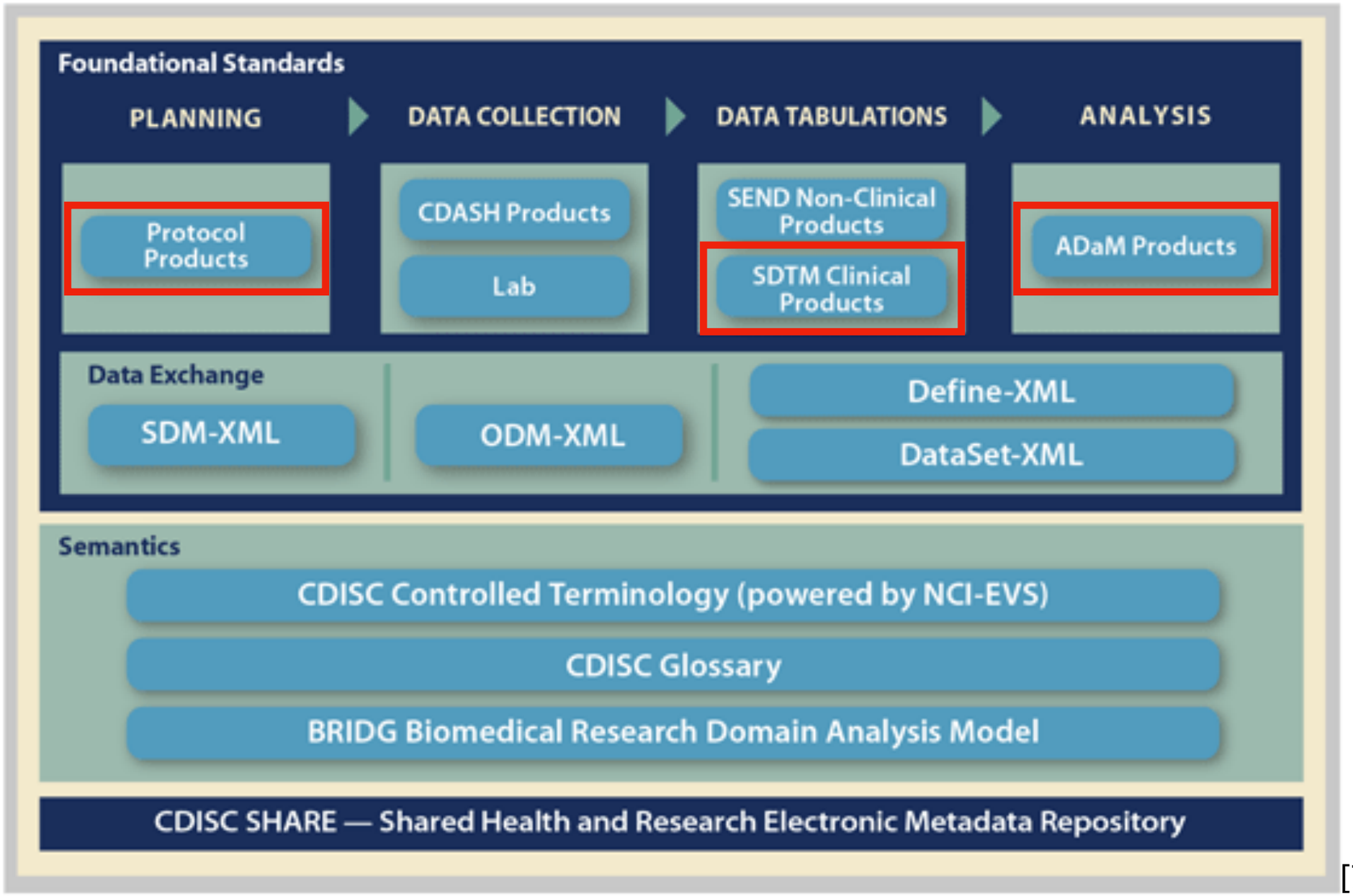
- Metadata is only an afterthought in data centric systems.
- Time is wasted on organizing data manually.
- Manual edits in related files can lead to errors.
- Missing or unused standards hinder collaboration.
- Poor discoverability and therefore reusability of old data.
- Automation is not easily possible.

Standards

CDISC

- ***C**linical **D**ata **I**nterchange **S**tandards **C**onsortium*
- A non-profit organization founded in 1997.
- Develops data standards that enable collaboration, to improve medical research.
- Different standards, based on the *Foundational Standards*.
- Describe how to represent data, not what should be collected, or how.

CDISC


Foundational Standards
PLANNING
DATA COLLECTION
DATA TABULATIONS
ANALYSIS
Protocol Products
CDASH Products
Lab
SEND Non-Clinical Products
SDTM Clinical Products
ADaM Products
Data Exchange
SDM-XML
ODM-XML
Define-XML
DataSet-XML
Semantics
CDISC Controlled Terminology (powered by NCI-EVS)
CDISC Glossary
BRIDG Biomedical Research Domain Analysis Model
CDISC SHARE — Shared Health and Research Electronic Metadata Repository

[7]

PRM

- ***P**rotocol **R**epresentation **M**odel*
- A CDISC Foundational standard released in 2010.
- Used to formally represent a *protocol*, defines the basic structure of a trial.
- Contains basic information about the trial:
  - Title
  - Drug / Procedure Being evaluated
  - Type
  - Goals

PRM

- Traditionally stored as an human readable text document.
- Contains information that can be used during database design and the definition of needed metadata.
- Can be used when registering with regulatory agencies.
- Least used CDISC standard. (Not required)

SDTM

- ***S**tudy **D**ata **T**abulation **M**odel*

- A CDISC Foundational standard released in 2004.

- Regulates the storing of data that was collected through observations of subjects during a trial.

- SDTM describes the model, the SDTM **I**mplementation **G**uide describes the implementation.

SDTM

- Key concepts:
  - Observations
  - Domains
  - Metadata Definitions
  - ...
- *"Subject 404 experienced raised temperature in week 3 of the trial XY42"*
-

STUDYID	USUBJID	VSEVINTX	VSTEST	VSORRES	VSORRESU
XY42	404	WEEK 3	temperature	38	C

SDTM

- Frequently changing, six revisions in fifteen years.
- Already improved the submission process.
- Widely used in the industry today. (Required)

Variable Name	Variable Label	Type	Role	Description
Topic Variable				
--TERM	Reported Term	Char	Topic	Topic variable for an event observation, which is th pre-specified name of the event.
Qualifier Variables				
--MODIFY	Modified Reported Term	Char	Synonym Qualifier of --TERM	If the value for --TERM is modified for coding pur modified text is placed here.
--LLT	Lowest Level Term	Char	Variable Qualifier of --TERM	MedDRA Lowest Level Term.
--LLTCD	Lowest Level Term Code	Num	Variable Qualifier of --LLT	MedDRA Lowest Level Term code.
--DECOD	Dictionary-Derived Term	Char	Synonym Qualifier of --TERM	Dictionary or sponsor-defined derived text descript variable, --TERM, or the modified topic variable (- applicable. Equivalent to the Preferred Term (PT in
--PTCD	Preferred Term Code	Num	Variable Qualifier of --DECOD	MedDRA Preferred Term code.
--HLT	High Level Term	Char	Variable Qualifier of --TERM	MedDRA High Level Term from the primary path.
--HLTCD	High Level Term Code	Num	Variable Qualifier of --HLT	MedDRA High Level Term code from the primary
--HLGT	High Level Group Term	Char	Variable Qualifier of --TERM	MedDRA High Level Group Term from the primar
--HLGTCD	High Level Group Term Code	Num	Variable Qualifier of --HLGT	MedDRA High Level Group Term code from the p
--CAT	Category	Char	Grouping Qualifier	Used to define a category of topic-variable values.
--SCAT	Subcategory	Char	Grouping Qualifier	Used to define a further categorization of --CAT va
--PRESP	Pre-specified	Char	Variable Qualifier of --TERM	Used to indicate whether the event described by --T specified on a CRF. Value is Y for pre-specified ev spontaneously reported events.
--OCCUR	Occurrence	Char	Record Qualifier	Used to record whether a pre-specified event occur information about the occurrence of a specific even
--STAT	Completion Status	Char	Record Qualifier	Used to indicate when a question about the occurre specified event was not answered. Should be null o of NOT DONE.

[13]

ADaM

- ***A**nalysis **D**at**a** **M**odel*

- A CDISC Foundational standard released in 2006.

- Provides a structure for the data and metadata that is needed during analysis.

- ADaM describes the model, the ADaM **I**mplementation **G**uide describes the implementation.

ADaM

- Build upon two structures: *ADSL* and *BDS*.
- **S**ubject-**L**evel **A**nalysis **D**ataset (ADSL)
  - Structures the dataset around the subject.
- **B**asic **D**ata **S**tructure (BDS)
  - Structured around the parameter that is being analyzed.

ADaM

- Supports the most common types of statistical analysis, but not all.
- Enables to communicate the carried out analysis with third parties.
- Information about underlying assumptions, statistical methods and performed computations.
- Widely used in the industry today. (Required)

Management Tools

Novo Nordisk MMA

- Danish Pharmaceutical company that develops diabetes treatments.

- A *Clinical Data Warehouse* (CDW) was already in use.

- In 2008 decided to add a *Metadata Management Application* (MMA) to the CDW.

Novo Nordisk MMA

Goals

- Increase the capacity of the existing statistical staff.
- Compliance with regulatory agencies.
- Make sharing of clinical data easier.
- Improve the discoverability of historic data.

Novo Nordisk MMA

Implementation

- Designed to make use of metadata throughout the entire process.
- Compatible with SDTM, but uses different standards internally.
- Production of report data is outsourced to different applications.
- Provides a dedicated user interface.

File Edit View Favorites Tools Help Links »

CDW Operations

Metadata Management

eClinical
TOWARDS GLOBALLY NETWORKED R&D

Metadata Trial Manage Logout

Demo/V1R6/KWL

Define Trial Definition Attributes

Define Protocol, Statistical and Cross-Trial Attributes for Trial Definition

Trial ID: BIAsp-1526
Trial Def ID: Clinical Data Migration
Trial Metadata Version: 2
Trial Definition Status: Draft

Protocol Attributes

EUDRACT Number:▪ 0
Trial Phase:▪ Phase III
ICH Category: Therapeutic Confirmatory
CDP Trial:▪ True
Open Label:
Blinding Level: Double Blind
Double Dummy Blinding:
Generic Trial Design Class: Crossover
Trial Title: A Stratified, Randomised, Double-Blind, Two-Period Crossover Trial Comparing the Pharmaco-dynamics and Pharmacokinetics of BIAsp 50 and BIAsp 70 in Non-obese and Obese Subjects with Type 2 Diabetes

Primary Objective Text: To compare the pharmacodynamics (AUCGlu (0-24h)) of thrice daily BIAsp 50 with that of thrice daily BIAsp 70, both after 4 weeks of treatment, in both non-obese (BMI 23-28 Kg/m2) and obese (BMI 30-35 Kg/m2) subjects with type 2 diabetes.

[18]

Clinical Metadata

- English Software Company that is specialized on metadata management.

- *Clinical Metadata* tool first released in 2017.

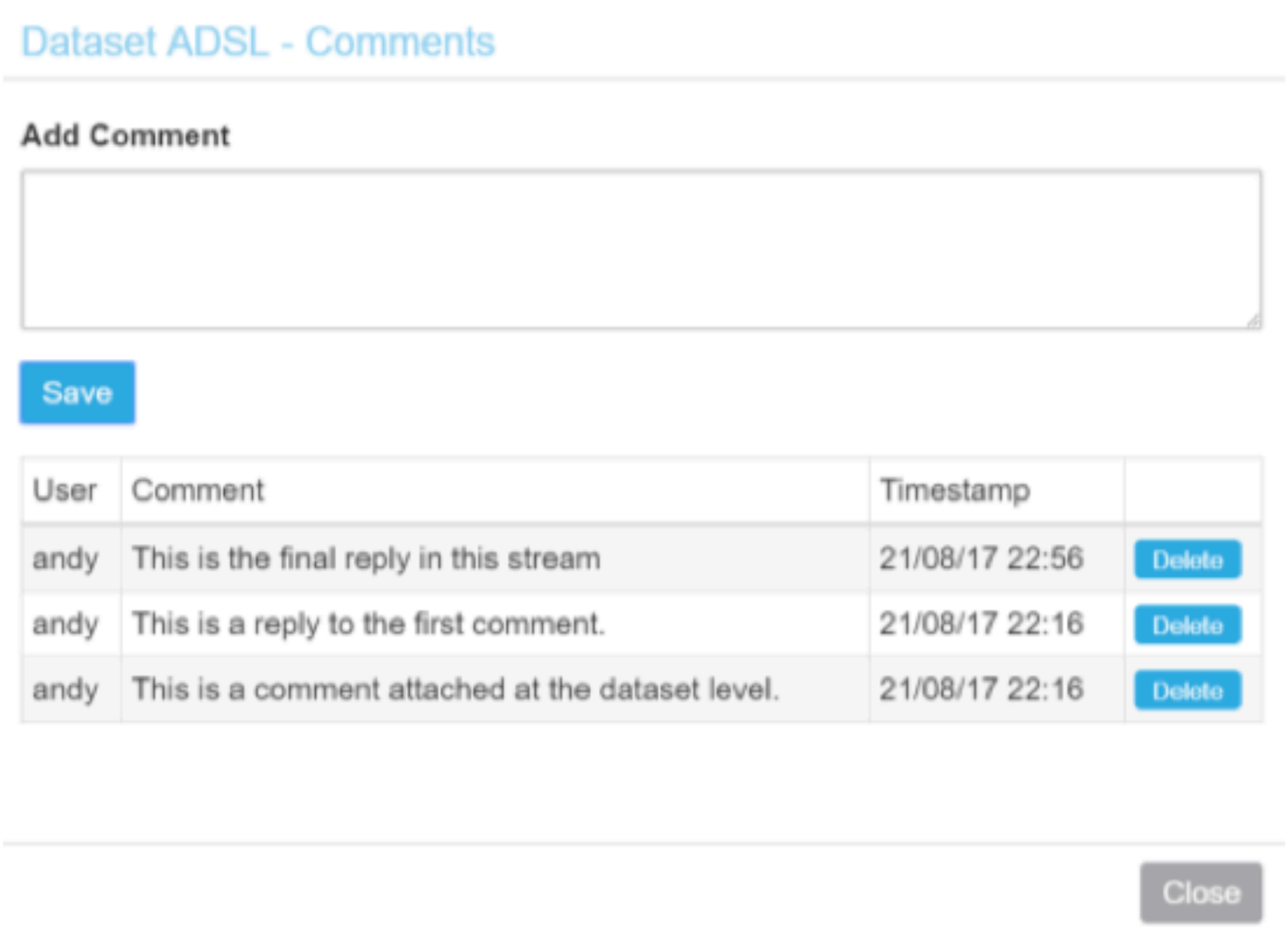
- Designed from the ground up with industry standards in mind.

- Modern, web-based, end-to-end metadata management tool.

- Developed by a software company.

Clinical Metadata

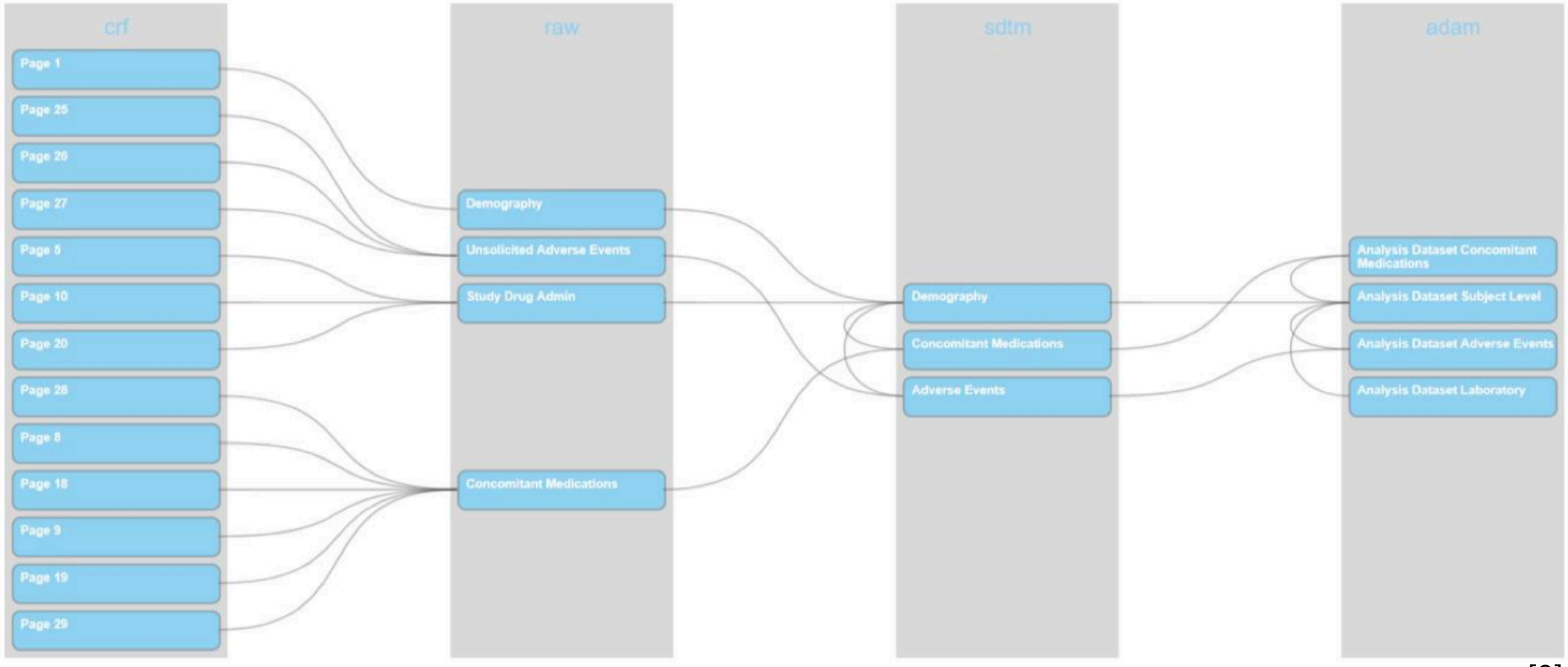
Implementation



[2]

Clinical Metadata

Implementation



[2]

Comparison

Tool	PRM	SDTM	define.xml	ADaM	Code Generation	Visualization	Platform
Novo Nordisk MMA	✘	✔	✔	✘	✘	✘	Java
Clinical Metadata	✘	✔	✔	✔	✔	✔	Web App
Anzo	✘	✔	✔	✔	✔	✘	Excel & Web App

Conclusion

- Metadata often implemented as an afterthought.
- Tools need to be designed specifically for metadata and with standards in mind.
- Standards are needed. Should evolve, but be backward compatible.
- Enables collaboration and increases efficiency.
- Already improved the clinical trial process.
- Has the opportunity to accelerate medical process as a whole.

- [A] http://www.picronline.org/articles/2014/5/4/images/PerspectClinRes_2014_5_4_159_140555_f3.jpg
- [?] All other figures were taken from the corresponding references in the paper.