Data standards in clinical research

Present Scenario, challenges and future directions

Mayank Anand
Director INC Research

Agenda

1- Data Standards

2- Current Process Flow

3- Challenges

4- CDISC

5- Road Ahead

6- Q&A

Data Standards

- **Standard**

  - Established by authority, custom or general consent as a model or example.
  - To Optimize the collection, transport and storage of data and simplify the submission of data to the regulatory bodies.

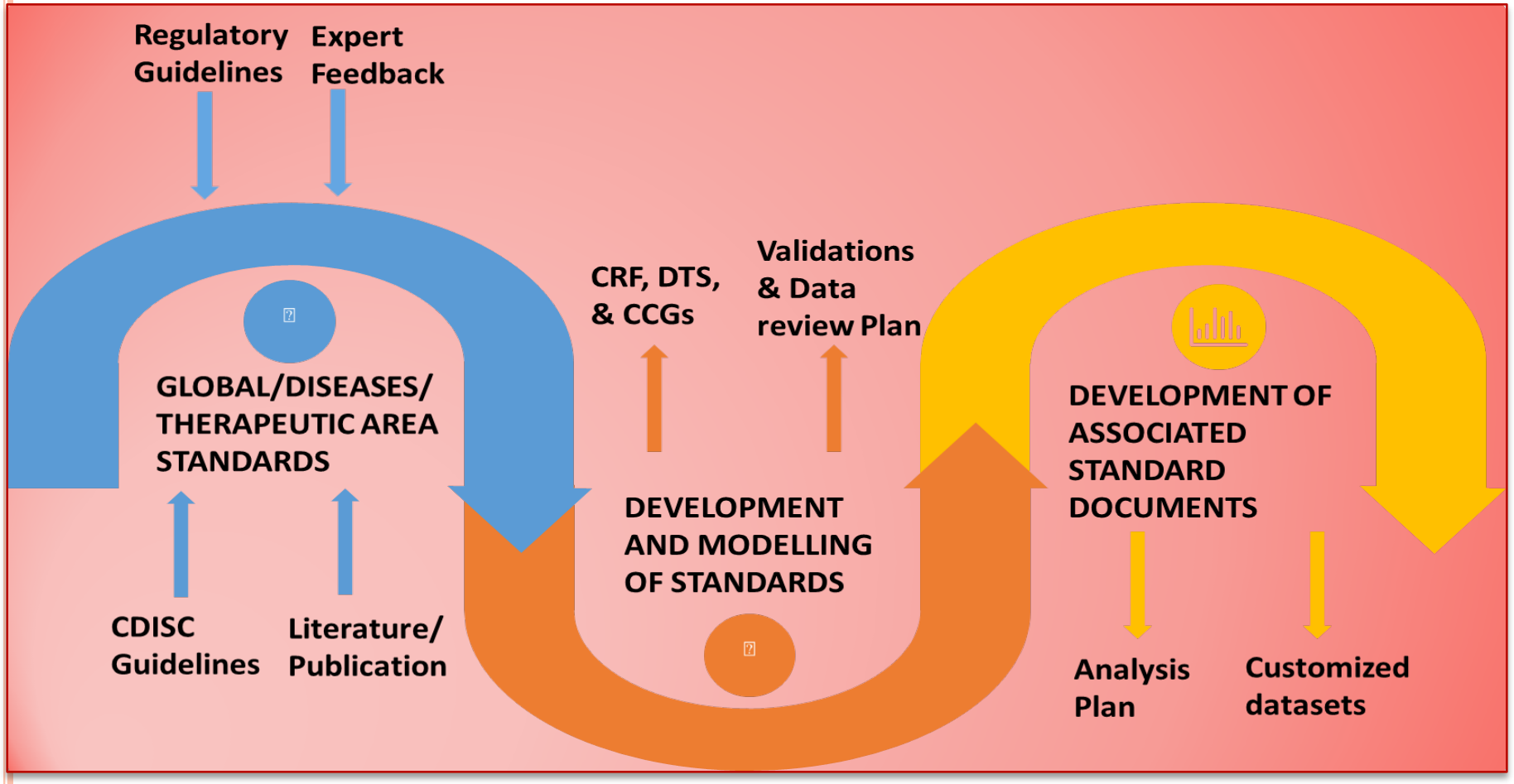
- **Clinical Data Interchange Standards Consortium (CDISC)**

  - Different standards: Foundational, Therapeutic Area, Semantics, Healthcare Link etc.
  - Clinical Data Acquisitions Standards Harmonization (CDASH)

- **Health Level-7 (HL7)**

  - Set of international standards for transfer of clinical and administrative data between software applications used by various healthcare providers.

CURRENT PROCESS FLOW



CHALLENGES

- **Varied requirements led to multidimensional way of data collection**
- **Each study had their own unique set of data standards.**
- **Notion that there is no way we could reuse data or data structures standards.**
- **Multi level expectations from each stake holders.**
- **Hampering the Milestones, Deliverables, Timelines…**
- **Inconsistencies in Clinical Data Collection due to non standard codes and variables. E.g.**

Study 1: Variable : Gender, Values: M,F

Study 2: Variable : SEX, Values: Male, Female

Study 3: Variable: Sexual Category, Values: 1,2

Study 4: Variable: Sexual Characteristics, Values: 0,1

CDISC

The Clinical Data Interchange Standards Consortium (CDISC) is a global, open, multidisciplinary, non-profit organization that has established standards to support the acquisition, exchange, submission and archive of clinical research data and metadata.

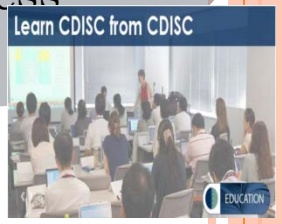
CDISC

Clinical Data Interchange Standards Consortium (CDISC)

Standards Developing Organization (SDO)

- CDISC Standards are open and free

Consistent implementation is supported by authorized training that is built into and supports the standards development process



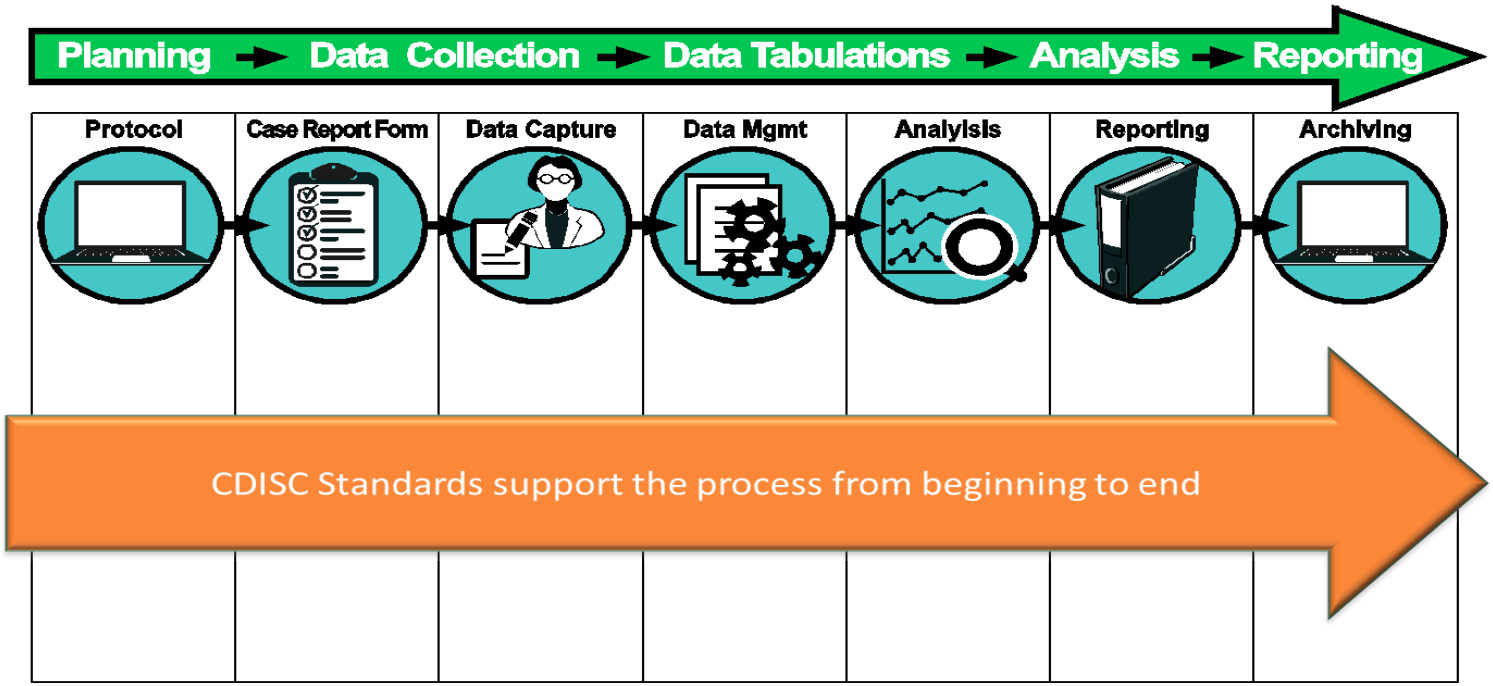
501(c)(3) non-profit organization

- Funded by organizational members, educational events, grants, donations

Standards Developed by Volunteers

- Public Reviews - critical to standards development process

CDISC Standards Support Research…



Purpose of CDASH:

1- Standardize CRF questions
2- Harmonize with SDTM for traceability
3- Use CDISC controlled terminology
4- Provide best practices and implementation guidelines

CDASH is Published as Metadata (not CRFs)

CDASH V

	Question Text	Prompt	SDTM or CDASH Variable Name	BRIDG	Definition	CRF Completion Instructions	Information for Sponsors
3	What is the adverse event term?	Adverse Event	AETERM	PerformedObservation Result.value*	Verbatim (i.e., investigator-reported term) description of the adverse event.	Record only one diagnosis, sign or symptom per line (e.g., nausea and vomiting should not be recorded in the same entry, but as two separate entries). Using accepted medical terminology, enter the diagnosis (if known); otherwise enter a sign or symptom. If a diagnosis subsequently becomes available, then this diagnosis should be entered on the AE form, replacing the original entries, where appropriate. Death should not be recorded as an event but should be recorded as the outcome of the event. The condition that resulted in the death should be recorded as the event. Do not use abbreviations.	In most cases, the verbatim term (i.e., investigator-reported term) will be coded to a standard medical dictionary such as MedDRA, WHO ART, after the data have been collected on the CRF. The coded data will be stored in field(s) not defined by CDASH. *See the BRIDG model for complete path.

Standard data collection metadata for commonly used CRFs:

- Which questions should be asked on each CRF
- How the questions should be asked
- Value lists used for responses
- Variables that are harmonized with SDTM

Value of CDASH

Reusability

- Case Report Forms
- Edit Checks
- Database builds
- Output files
- Downstream programming

Consistency

- For all users - CDM, CRA, Site
- Across partner organizations

Transparency and Traceability

- Built-in harmonization with SDTM and Controlled Terminology
- Supports beginning to end standardization

CDASH Version History

- CDASH "Initiative" started in 2006
- CDASH V1.0 published in 2008
- CDASH V1.1 published in 2011
  - CDASH User Guide
  - CDASH CRF Library (example CRFs0
- CDASH Model started to be developed ~2012/2013
- CDASH Implementation Guide V2.0 and CDASH Model V1.0 currently being finalized following public review period

Comparison of CDASH V1.1 and CDASH V2.0

CDASH V1.1	CDASHIG V2.0
16 Domains (+2 "Common" fields tables)	24 Domains (Header fields integrated)
Metadata tables in PDF	Metadata tables published from SHARE (XLS, ODM, RDF, CSV)
No Model	Model Metadata published from SHARE (XLS, ODM, RDF, CSV)
Domains published alphabetically	Domains organized alphabetically within "Class"
No CRF examples included	CRF examples and Data examples for many domains are included in the CDASHIG
Metadata table had 8 components (some overloaded)	Metadata includes 17 components, including SDTM mappings

EXAMPLES OF DOMAINS

Events

Adverse Events (AE)

Medical History (MH)

Clinical Events (CE)

Disposition (DS)

Protocol Deviations (DV)

Findings

ECG (EG)

Laboratory (LB)

Questionnaire (QS)

Vital Signs (VS)

Inclusion/Exclusion Criteria (IE)

Interventions

Concomitant Medication (CM)

Substance use (SU)

Exposure (EX)

Special Domains

Demography (DM)

Subject Element (SE)

Subject Visits (SV)

Comments (CO)

Trial Design Domains

Trial Arm (TA)

Trial Visit (TV)

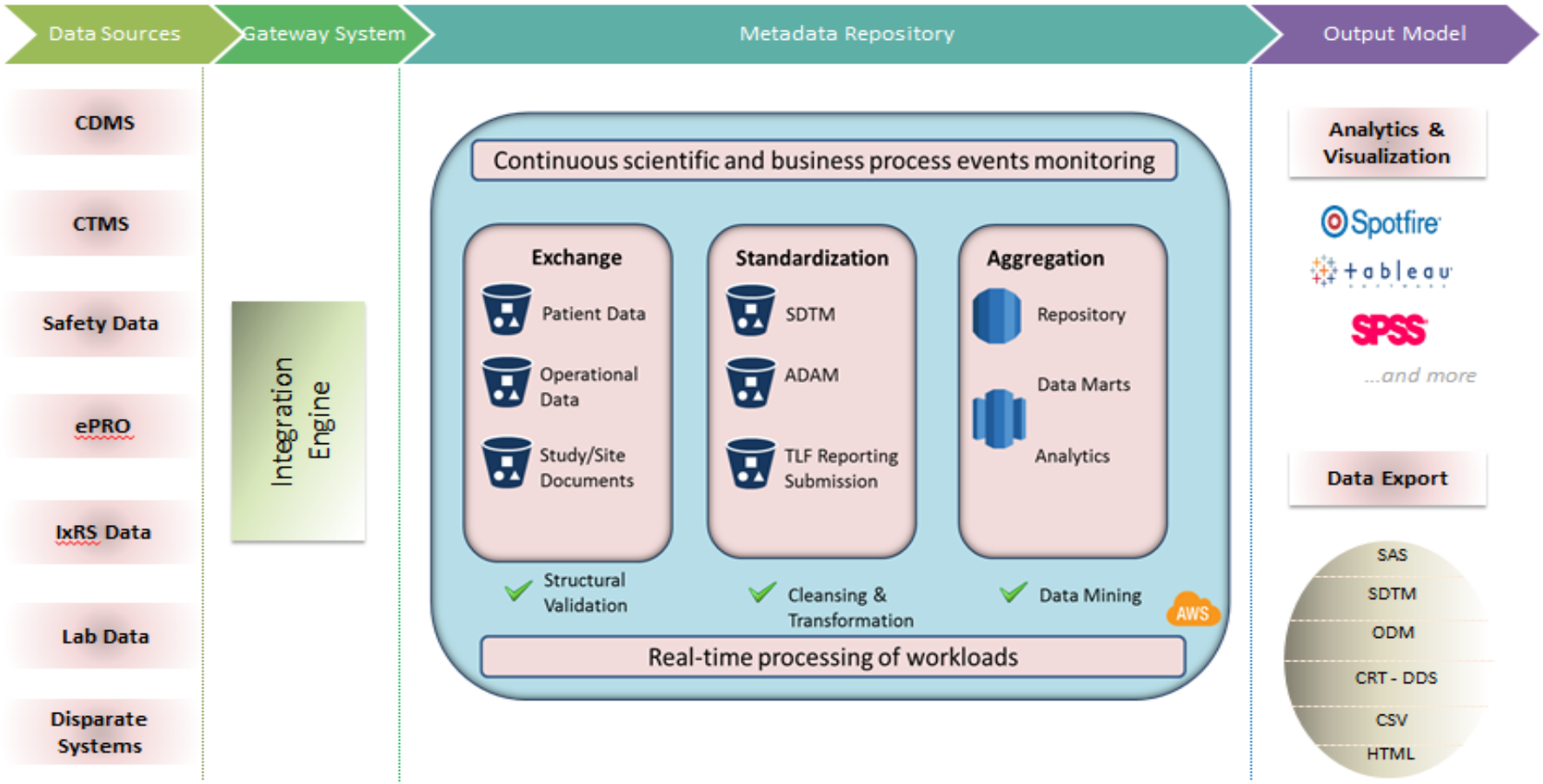
Trial Summary (TS)

Trial Elements (TE)

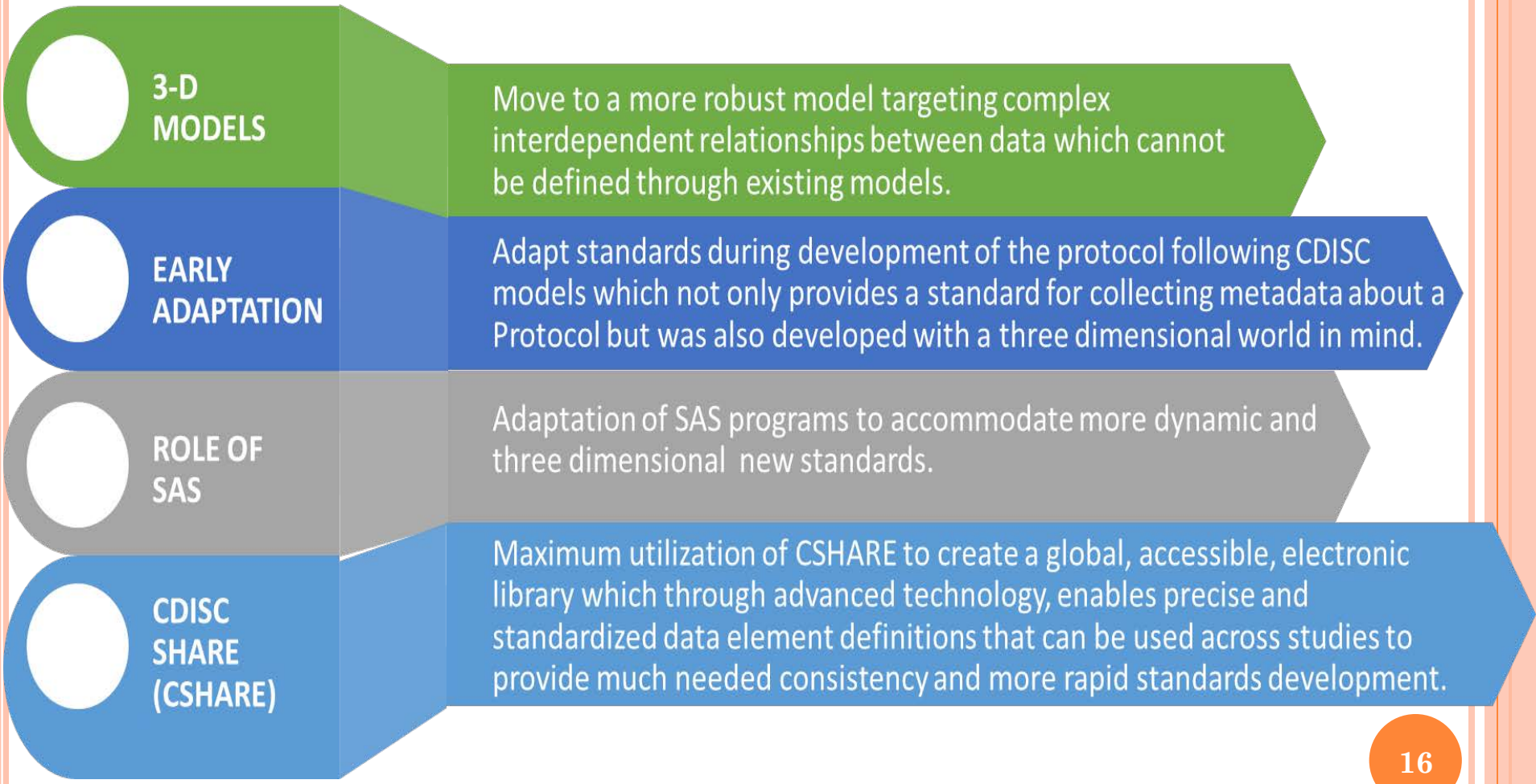
METADATA

- Metadata is the data about data.
- Metadata answers fundamental questions about a bit of information, where in the data flow it is used.
- Metadata maximizes the value of data standards.

SMART Design: Metadata Repository



ROAD AHEAD


3-D MODELS
Move to a more robust model targeting complex interdependent relationships between data which cannot be defined through existing models.
EARLY ADAPTATION
Adapt standards during development of the protocol following CDISC models which not only provides a standard for collecting metadata about a Protocol but was also developed with a three dimensional world in mind.
ROLE OF SAS
Adaptation of SAS programs to accommodate more dynamic and three dimensional new standards.
CDISC SHARE (CSHARE)
Maximum utilization of CSHARE to create a global, accessible, electronic library which through advanced technology, enables precise and standardized data element definitions that can be used across studies to provide much needed consistency and more rapid standards development.

Q & A

THANK YOU