CDISC 2019 UK Network Meeting

London | 03 September

Disclaimer and Disclosures

- *The views and opinions expressed in this presentation are those of the author(s) and do not necessarily reflect the official policy or position of CDISC.*

- *The author(s) have no real or apparent conflicts of interest to report.*

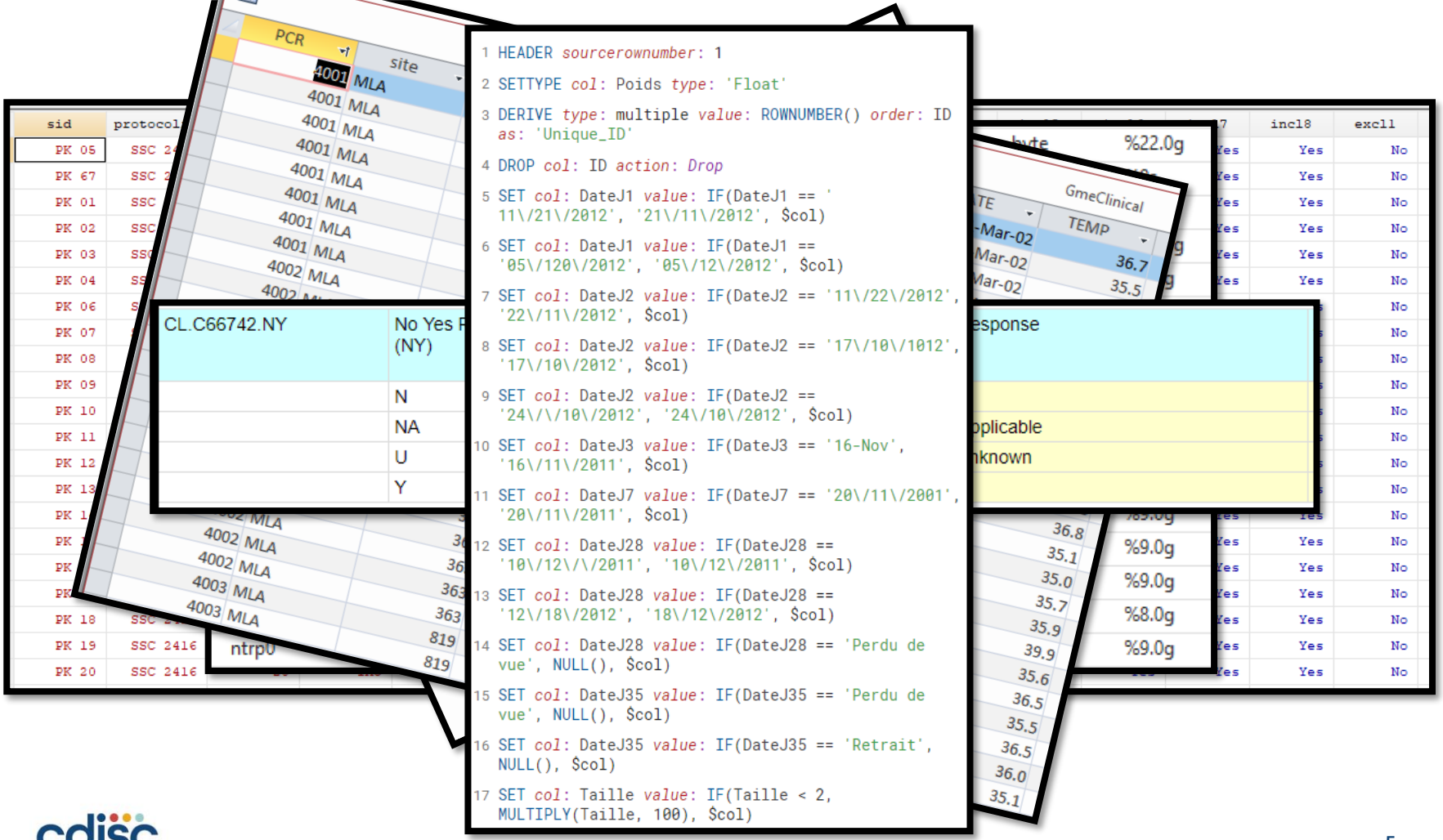
Novel SDTM Implementation to Maximise Benefits of Sharing Legacy Data

Presented by Kalynn Kennon
Infectious Diseases Data Observatory (IDDO)

09.03.2019

CDISC 2019 UK Network Meeting
London | 03 September

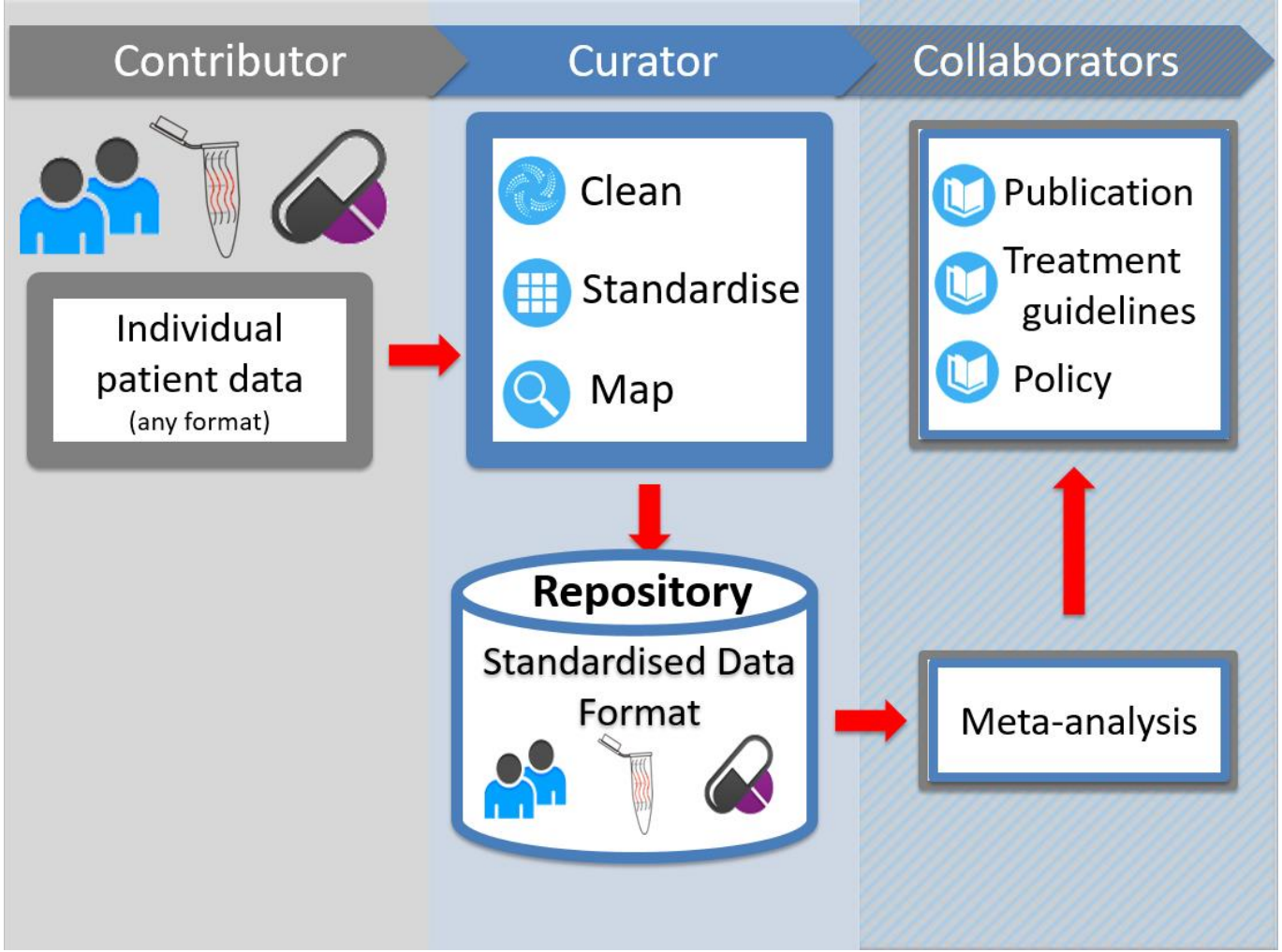
cdisc


HEADER sourcerownumber: 1
SETTYPE col: Poids type: 'Float'
DERIVE type: multiple value: ROWNUMBER() order: ID as: 'Unique_ID'
DROP col: ID action: Drop
SET col: DateJ1 value: IF(DateJ1 == '11\/21\/2012', '21\/11\/2012', $col)
SET col: DateJ1 value: IF(DateJ1 == '05\/120\/2012', '05\/12\/2012', $col)
SET col: DateJ2 value: IF(DateJ2 == '11\/22\/2012', '22\/11\/2012', $col)
SET col: DateJ2 value: IF(DateJ2 == '17\/10\/1012', '17\/10\/2012', $col)
SET col: DateJ2 value: IF(DateJ2 == '24\/\/10\/2012', '24\/10\/2012', $col)
SET col: DateJ3 value: IF(DateJ3 == '16-Nov', '16\/11\/2011', $col)
SET col: DateJ7 value: IF(DateJ7 == '20\/11\/2001', '20\/11\/2011', $col)
SET col: DateJ28 value: IF(DateJ28 == '10\/12\/\/2011', '10\/12\/2011', $col)
SET col: DateJ28 value: IF(DateJ28 == '12\/18\/2012', '18\/12\/2012', $col)
SET col: DateJ28 value: IF(DateJ28 == 'Perdu de vue', NULL(), $col)
SET col: DateJ35 value: IF(DateJ35 == 'Perdu de vue', NULL(), $col)
SET col: DateJ35 value: IF(DateJ35 == 'Retrait', NULL(), $col)
SET col: Taille value: IF(Taille < 2, MULTIPLY(Taille, 100), $col)
PCR
site
CL.C66742.NY
No Yes (NY)
GmeClinical
TEMP
sid
protocol
incl8
excl1

Credit: James Gathany, CDC

Credit: Katie Holt, Sightsavers

cdisc


Contributor
Curator
Collaborators
Individual patient data (any format)
Clean
Standardise
Map
Publication
Treatment guidelines
Policy
Repository
Standardised Data Format
Meta-analysis

MALARIA (WWARN)

Generating innovative resources and reliable evidence on the factors affecting the efficacy of antimalarial medicines.

Estimated cases in 2016: 216 million

MEDICINE QUALITY

Sharing expertise and collating information on the prevalence and impact of substandard, falsified and unregulated medicines.

Estimated number of people affected: Unknown

EBOLA

Facilitating data sharing to improve diagnostics and treatments, optimise outbreak response, and reduce the impact of future epidemics.

Reported cases in 2014-16 outbreak: 28,712

CHAGAS

Scoping the availability of data and opportunity to instigate pooled individual patient data for this neglected tropical disease.

Estimated number of people infected: 6-7 million

IDDO

INFECTIOUS DISEASES DATA OBSERVATORY

NON-MALARIAL FEBRILE ILLNESS

Mapping causes of acute febrile illness in malaria endemic regions of the world.

Estimated cases: Unknown

SOIL-TRANSMITTED HELMINTH INFECTIONS

Collating studies on responses to a variety of antihelmintics to help eliminate morbidity for this neglected tropical disease.

Estimated number of people infected: 1.5 billion

SCHISTOSOMIASIS

Creating a standardised database to answer questions relating to treatment efficacy and optimum dosing.

Estimated cases in 2015: 218 million

VISCERAL LEISHMANIASIS

Pooling individual patient data to guide treatment options and help design better therapies and control strategies to support the world's poorest.

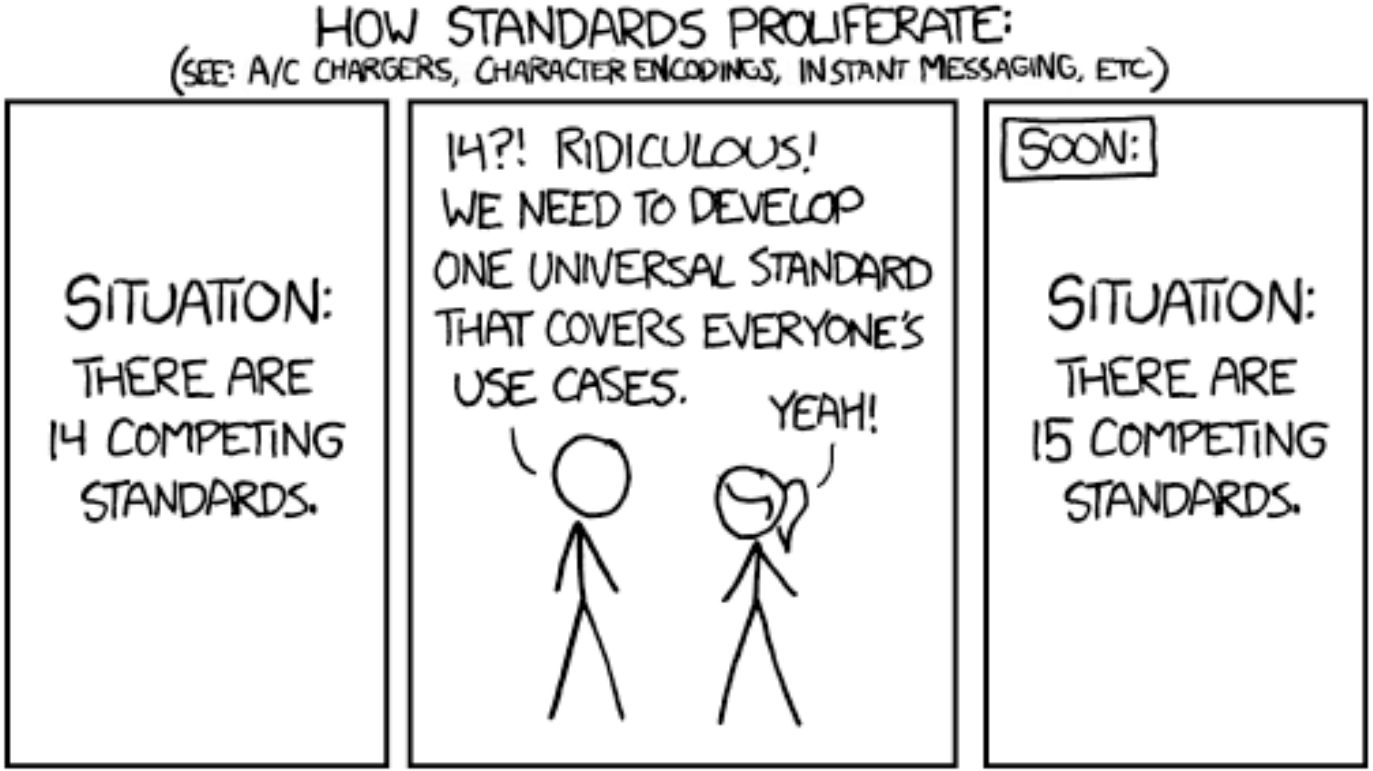
Estimated new cases each year: 50,000-90,000

Research themes: active building

iddo.org

@IDDOnews

cdisc



PERMANENT LINK TO THIS COMIC: HTTPS://XKCD.COM/927/

Data Contributors Provide:

- **Malaria** Individual Patient Data
- **Visceral Leishmaniasis** Individual Patient Data
- **Ebola** Individual Patient Data
- **Chagas** Individual Patient Data

IDDO Data Curators Clean, Standardize, and Map:

- to the Malaria Standard → **Malaria** Repository
- to the VL Standard → **Visceral Leishmaniasis** Repository
- to the Ebola Standard → **Ebola** Repository
- to the Chagas Standard → **Chagas** Repository

cdisc

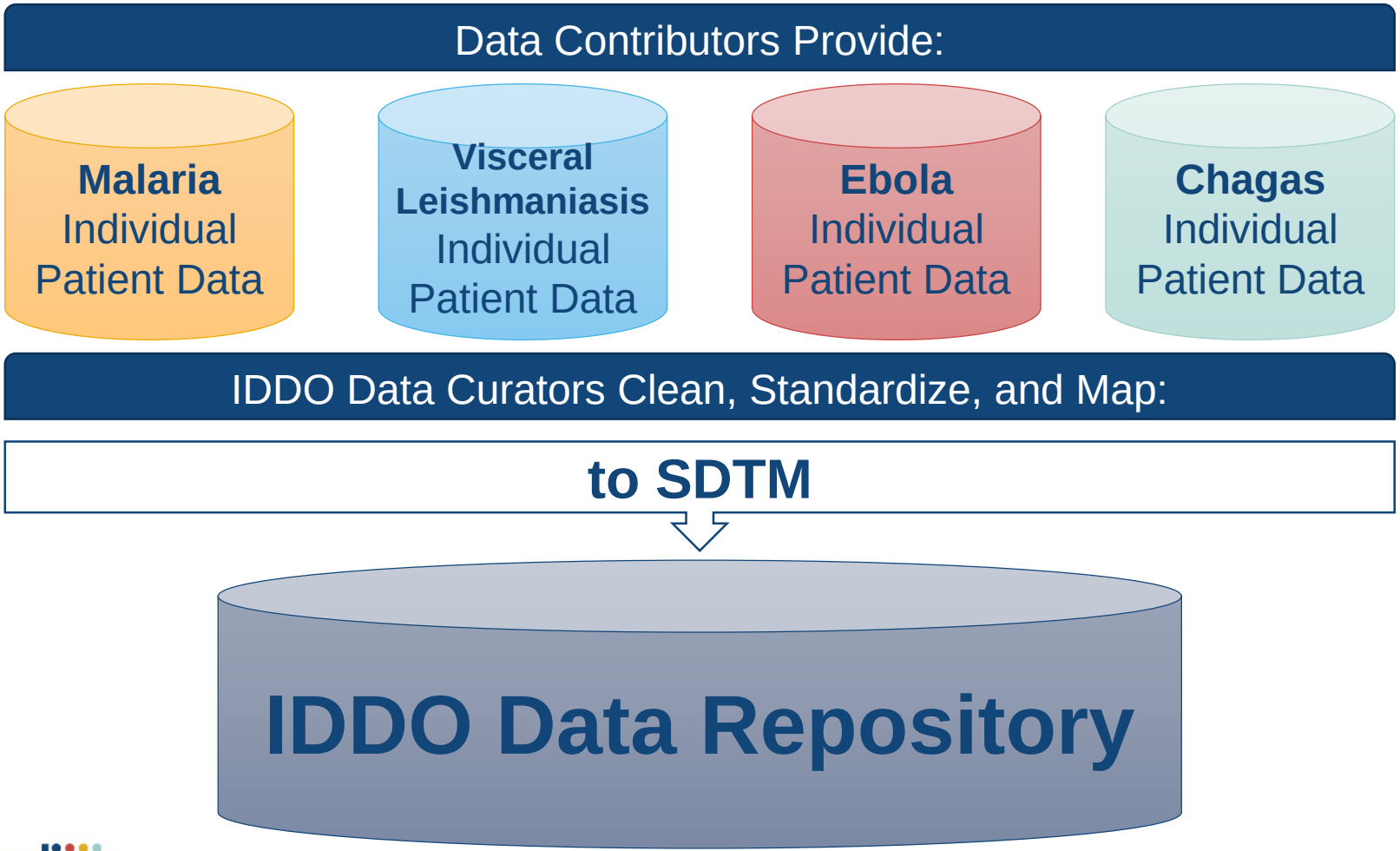
Study Data Tabulation Model Implementation Guide: Human Clinical Trials

Version 3.3 (Final)

Prepared by the

CDISC Submission Data Standards Team

[redacted]


Data Contributors Provide:
Malaria
Individual
Patient Data
Visceral
Leishmaniasis
Individual
Patient Data
Ebola
Individual
Patient Data
Chagas
Individual
Patient Data
IDDO Data Curators Clean, Standardize, and Map:
to SDTM
IDDO Data Repository

The next steps:

Thoughts, challenges, and occasionally solutions

Our Plan

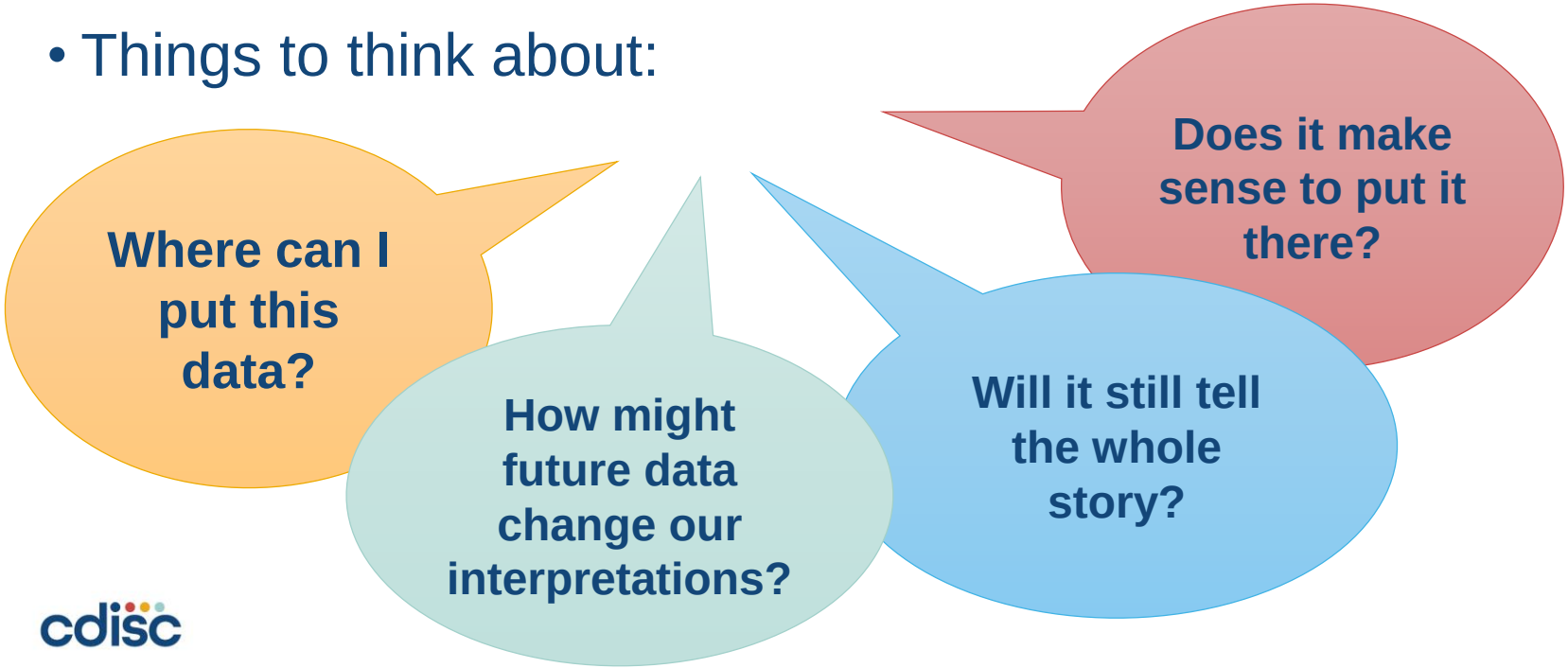
- Tap into the accessibility, user friendliness, and reproducibility of the SDTM standard

- Create a specialist IDDO CDISC Data Dictionary
  - Data and disease agnostic
  - Focusing on aggregation for secondary and meta analysis

- Encourage data re-use

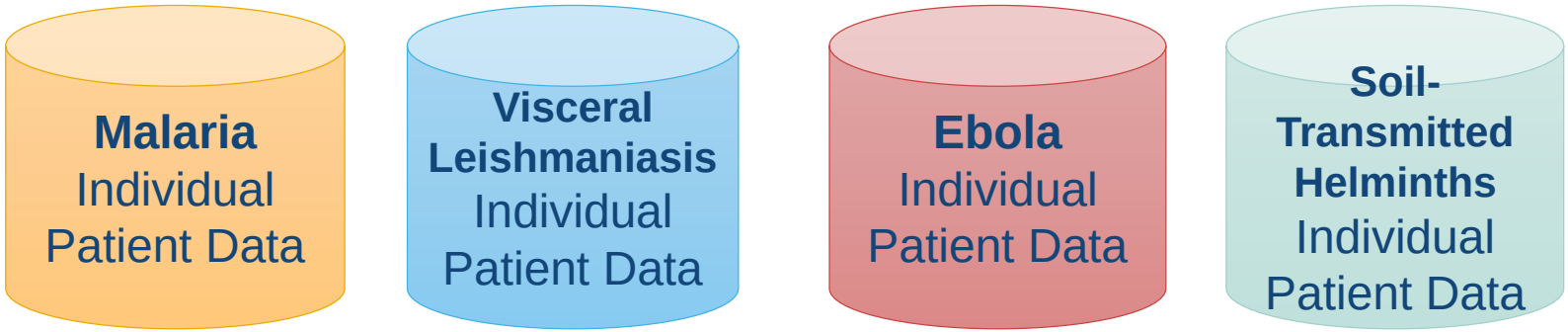
Easier Said Than Done

- Building a disease agnostic CDISC data dictionary is a challenge!
- Things to think about:



What Is On Our Platform?

- Legacy data from external researchers
- Not restricted to trial data – we accept any health data people are willing to share
- We are standardizing to encourage re-use of data



Balancing Standardization With Flexibility

Creating something that is comprehensive, flexible, and dynamic

That is clear, concise, and readily accessible for everyone

Data Categories

- What are our common types of variables?
  - Test data
  - Treatment data
  - Signs and symptoms data
  - Patient information
  - Outcomes

The Easy Ones!

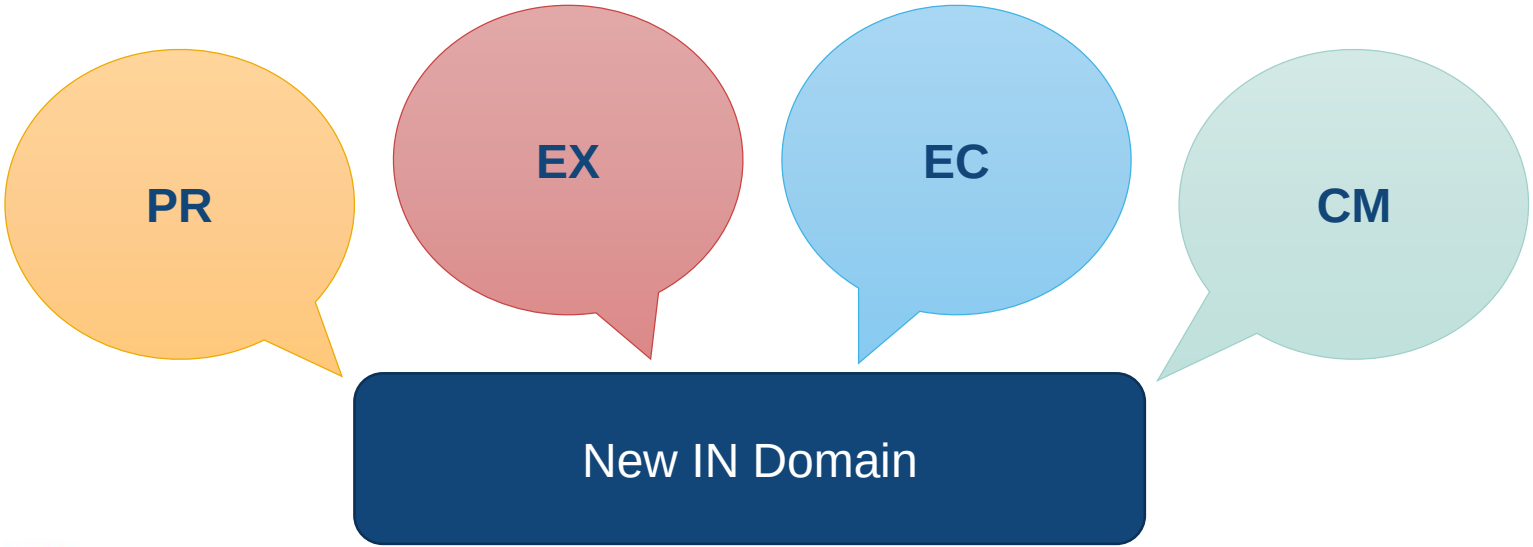
- Testing data
  - Lab
    - Hematology
    - Biochemistry
    - Microbiology
  - In the clinic
    - Vital signs
    - Functional clinical tests
- Not too difficult to identify where some things belong

Harder than it seems!

- Treatments
  - ‘Trial’ vs ‘no trial’
  - Atypical treatments
  - Cross-dataset comparison
- Consistency is important
- Future proofing for future analysis

Treatments and Interventions (IN) Domain

- Custom Interventions Observation Class domain
- Generality and flexibility of original Interventions model



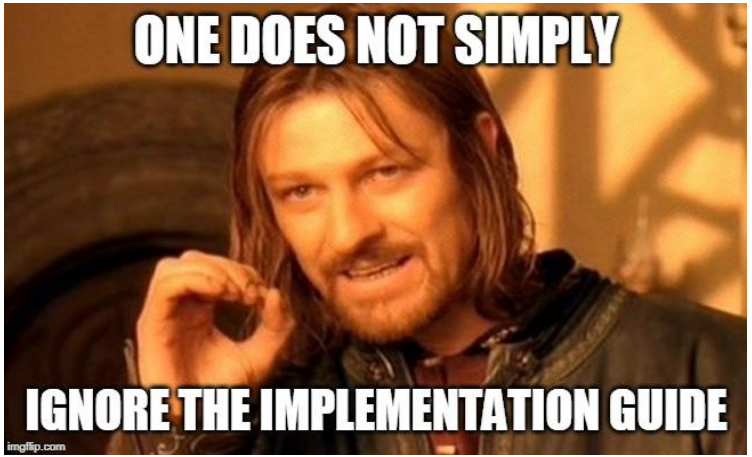
More Flexibility

- Develop our own constraints on content

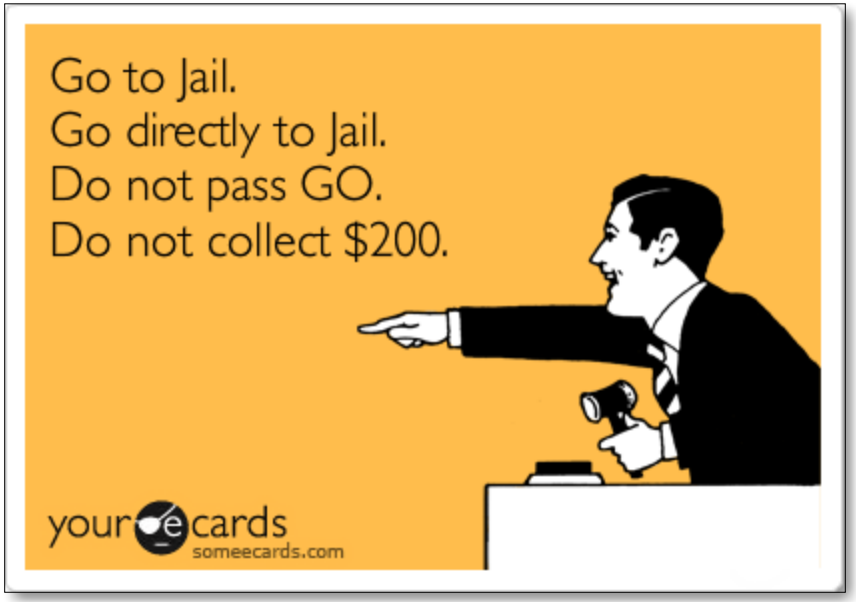
IN
Individual
Patient Data

Slippery Slopes

- How can we avoid them?
- Still utilizing SDTMiG rules as they stand as much as possible
- Investigate other ways to incorporate atypical data



Controlled Terminology



Inextensible Controlled Terminology

Severity	1:	Mild
	2:	Moderate
	3:	Severe
	4:	Life-Threatening
	9:	Missing/Blank

AESEV	Severity/Intensity	Char	(AESEV)

Severity/Intensity Scale for Adverse Events (AESEV)	text Extensible: No
MILD	
MODERATE	
SEVERE	

AESLIFE	Is Life Threatening	Char	(NY)

How to code "Life-Threatening" for AESEV??

Mismatched Controlled Terminology

Action Taken	0:	None
	1:	Concomitant Medication
	2:	Withdrawn from Study
	3:	Other, Specify
	9:	Missing/Blank

Can we assume 'DOSE NOT CHANGED'?

Can we assume 'DOSE NOT CHANGED'?

Can we assume 'DRUG WITHDRAWN'?

Action Taken with Study Treatment (ACN)	text Extensible: No
DOSE INCREASED	
DOSE NOT CHANGED	
DOSE RATE REDUCED	
DOSE REDUCED	
DRUG INTERRUPTED	
DRUG WITHDRAWN	
NOT APPLICABLE	
UNKNOWN	

AEACNOTH	Other Action Taken	Char	

- Challenges have resulted in some implementation rule bending
- SDTM is flexible
- This is a work-in-progress
- SDTM was built by group collaboration so;

We welcome your thoughts!

Thank You!