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
1 HEADER sourcedocumentnumber: 1
2 SETTYPE col: Poids type: 'Float'
3 DERIVE type: multiple value: ROWNUMBER() order: ID as: 'Unique ID'
4 DROP col: ID action: Drop
5 SET col: DateJ1 value: IF(DateJ1 == '11/21/2012', '21/11/2012', $col)
6 SET col: DateJ1 value: IF(DateJ1 == '05/12/2012', '05/12/2012', $col)
7 SET col: DateJ2 value: IF(DateJ2 == '11/22/2012', '22/11/2012', $col)
8 SET col: DateJ2 value: IF(DateJ2 == '17/10/1012', '17/10/1012', $col)
9 SET col: DateJ2 value: IF(DateJ2 == '24/10/1012', '24/10/1012', $col)
10 SET col: DateJ3 value: IF(DateJ3 == '16-Nov', '16/11/2011', $col)
11 SET col: DateJ7 value: IF(DateJ7 == '20/11/2001', '20/11/2001', $col)
12 SET col: DateJ28 value: IF(DateJ28 == '10/12/2011', '10/12/2011', $col)
13 SET col: DateJ28 value: IF(DateJ28 == '12/18/2012', '18/12/2012', $col)
14 SET col: DateJ28 value: IF(DateJ28 == 'Perdu de vue', NULL(), $col)
15 SET col: DateJ35 value: IF(DateJ35 == 'Perdu de vue', NULL(), $col)
16 SET col: DateJ35 value: IF(DateJ35 == 'Retrait', NULL(), $col)
17 SET col: Taille value: IF(Taille < 2, MULTIPLY(Taille, 109), $col)
Individual patient data (any format)

Clean
Standardise
Map

Repository
Standardised Data Format

Publication
Treatment guidelines
Policy

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

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 anthelmintics 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  
@iddo.org

@IDDOnews

iddo.org
HOW STANDARDS PROLIFERATE:
(SEE: A/C CHARGERS, CHARACTER ENCODINGS, INSTANT MESSAGING, ETC.)

SITUATION: THERE ARE 14 COMPETING STANDARDS.

14?! RIDICULOUS! WE NEED TO DEVELOP ONE UNIVERSAL STANDARD THAT COVERS EVERYONE’S USE CASES. YEAH!

SITUATION: THERE ARE 15 COMPETING STANDARDS.

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
- to the VL Standard
- to the Ebola Standard
- to the Chagas Standard
Study Data Tabulation Model Implementation Guide: Human Clinical Trials

Version 3.3 (Final)

Prepared by the
CDISC Submission Data Standards Team
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:
  - Where can I put this data?
  - Does it make sense to put it there?
  - How might future data change our interpretations?
  - Will it still tell the whole story?
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
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

Go to Jail.
Go directly to Jail.
Do not pass GO.
Do not collect $200.
Inextensible Controlled Terminology

<table>
<thead>
<tr>
<th>Severity</th>
<th>SEVERITY/INTENSITY</th>
<th>CHAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: Mild</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2: Moderate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3: Severe</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4: Life-Threatening</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9: Missing/Blank</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

How to code “Life-Threatening” for AESEV??

AESLIFE

Is Life Threatening | Char | (NY)
### Mismatched Controlled Terminology

<table>
<thead>
<tr>
<th>Action Taken</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0:</td>
<td>None</td>
</tr>
<tr>
<td>1:</td>
<td>Concomitant Medication</td>
</tr>
<tr>
<td>2:</td>
<td>Withdrawn from Study</td>
</tr>
<tr>
<td>3:</td>
<td>Other, Specify</td>
</tr>
<tr>
<td>9:</td>
<td>Missing/Blank</td>
</tr>
</tbody>
</table>

**Can we assume 'DOSE NOT CHANGED'?**

**Can we assume 'DOSE NOT CHANGED'?**

**Can we assume 'DRUG WITHDRAWN'?**
• 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!