Leveraging CDISC Standards for AE Narrative Automation

Presented at CDISC Webinar
Carol Vaughn (d-wise Life Sciences Consultant)

April 3, 2019
Let's Get on the Same Page

What is an AE Narrative

Regulatory agency submission deliverable

Summary of select events experienced by particular subjects during a clinical trial with other information about those subjects' participation in the trial.

The criteria for including a subject/events are study specific.

Typical criteria include a serious AE, an AE of Special Interest, or an AE leading to discontinuation from the study.

What is the Goal of Automating

Generate a narrative for any study, based on a standard therapeutic area template, as close to final product as possible with the need for, at most, minimal further post generation manual entries or modifications (limited to medically subjective assessments).
So You Want Your Computer to Write Your Narrative
Intended Take-aways

Why solid sponsor defined SDTM/ADaM standards are critical

Why AE narrative templates need to be air tight

Why enabling study specific input is necessary

Why datapoint metadata is an automation development linchpin

What to look for in an implementation method
Use SDTM or ADaM?

➢ Can be dependent on:

- How fully developed are sponsor defined SDTM versus ADaM standards
- The timing of when the sponsor develops their ADaM datasets relative to when it is desired to be able to generate the narrative
- How much derived data (only present in ADaM) is desired to be included in the narrative

Maybe start with SDTM datasets, then switch the source to ADaM datasets once available to take advantage of derived flag variables
Source Data – Standardization is Critical

ERROR

The data point could not be found!

OK
Examples of Impact of Lack of Standardization

At the time of the AE, the following concomitant medications were ongoing: <<Standardized Medication Name>> for <<Standardized Indication>>.

CM.CMDECOD

CM.CMiINDC

SUPPCM.QNAM = 
INDCPT
REASPT
PREFERI
CODEDIN

???
The subject completed the study on **<<Completion Date>>**.

Or

The subject did not complete the study due to **<<Reason for Non-completion>>**.

**DS.DSDECOD** = 
where **DS.EPOCH** =

<table>
<thead>
<tr>
<th>DSDECOD</th>
<th>EPOCH</th>
</tr>
</thead>
<tbody>
<tr>
<td>COMPLETED</td>
<td>TREATMENT</td>
</tr>
<tr>
<td>COMPLETED</td>
<td>STUDY TREATMENT</td>
</tr>
<tr>
<td>COMPLETED</td>
<td>RUN-IN TREATMENT</td>
</tr>
<tr>
<td>COMPLETED</td>
<td>FOLLOW UP</td>
</tr>
<tr>
<td>COMPLETED RUN IN</td>
<td>TREATMENT</td>
</tr>
<tr>
<td>COMPLETED STUDY ACCORDING TO PROTOCOL</td>
<td>STUDY</td>
</tr>
<tr>
<td>COMPLETED</td>
<td>STUDY WASHOUT</td>
</tr>
</tbody>
</table>
Studies mapped to newer versions of CDISC standards may map to different domains or use different CT. For example:

Older studies may map primary cause of death to SUPPDS whereas studies mapped to SDTMIG 3.2 and higher may map it to the DD domain (automation can look for both).

Sponsor defined standards may similarly evolve. Sponsors should factor in tool dependencies when considering sponsor defined version changes.

As studies use newer standards after the automation is in place, automation adjustments may be needed.
Priority needs to be given to developing standards for any piece of data which will be used by an automated tool. This includes standardizing mapping to:

- Domains
- Variables
- Controlled Terminology
- SUPPQUALs (subset of Controlled Terminology)
Why solid sponsor defined SDTM/ADaM standards are critical
At the time the subject experienced `<<(AEDECOD) AE preferred term>>` he/she was also experiencing `<<(NEARAES1) List of coma separated concurrent AEs sorted by onset date, SDC, and PT - each followed by AE severity in parentheses>>`

{If no concurrent events replace with: At the time the subject experienced `<<(AEDECOD) AE preferred term>>` he/she was not experiencing any other adverse events.}
From Template to Output

Template Sentence

At the time the subject experienced <<(AEDECOD) AE preferred term>> he/she was also experiencing <<(NEARAES1) List of coma separated concurrent AEs sorted by onset date, SOC, and PT – each followed by AE severity in parentheses>>

{If no concurrent events replace with: At the time the subject experienced <<(AEDECOD) AE preferred term>> he/she was not experiencing any other adverse events.}

Output Sentence

At the time the subject experienced traumatic liver injury she was also experiencing angina pectoris (Grade 2), plural effusion (Grade 2), and pulmonary oedema (Grade 3).
Who Defines the Templates

- Clinical Development: responsible for writing the narratives
- Automation Development: to ensure all specifications have been covered
- Biostatistics: to advise on logic rules
### Rules to be Defined

Examples:

- Date imputation rules

Logic for selecting events/interventions:

- “during”
- “prior”
- “after”
- “at baseline”
- “ongoing”
Factors Determining Contents

**Therapeutic Area**

- **Type of data collected in a study**
  - E.g., hospitalizations, autopsy, substance use, tumor response...

- **Study design**
  - E.g., single dose/multi dose, has follow-up, open label/blinded...

- **Clinical events of interest**
  - E.g., signs and symptoms, vital signs, abnormal labs...

- **Disease under study**
  - E.g., background medications, diagnosis differences...
Why AE narrative
templates need to be air tight
Study Specific Information

- **Not in the study data**
  - e.g., to select subjects discontinuing within x number of days after last dose
  - Tell it: The value of x.

- **Display instead**
  - e.g., to display “last drug prior to”
  - Tell it: When you see EXTRT of MIR99 display instead Miracle Drug 99.

- **How to find in the study data**
  - e.g., to display “all disease under study medications”
  - Tell it: What’s the value of CMCAT to use to find these?

- **Discretionary**
  - e.g., whether to display a statement about the cohort ---or--- which criteria to use to select subjects/events
  - Tell it
The subject was taking the following disease under study medications at the start of the study: albeuteral and fluticasone propionate.

The subject discontinued the study due to AE within 180 days of her last dose of Tizlepan (25 mg).
Why enabling study specific input is necessary
What is Data Point Metadata?

➢ Metadata for the data points around which text is wrapped in the narrative

At the time the subject experienced <<(AEDECOD) AE Preferred Term >>, he/she was also experiencing <<NEARAES1 List of comma separated concurrent AEs sorted by onset date, SOC, and PT, each followed by AE Severity in parentheses>>.

➢ Conceptually similar to CDISC SDTMIG metadata:
# Data Point Metadata Purpose

<table>
<thead>
<tr>
<th>Enable Possible Metadata Driven</th>
<th>Facilitate Troubleshooting, Maintenance, and/or Enhancement of the automation</th>
<th>Provide Traceability back to the source (SDTM/ADaM) data</th>
</tr>
</thead>
<tbody>
<tr>
<td>creation of data points (at least shells)</td>
<td>Facilitate Validation of the data points</td>
<td>Establish Assumptions of the (SDTM/ADaM) source data</td>
</tr>
<tr>
<td>Build in Consistency of conventions across TA</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Why datapoint metadata is an automation development linchpin
Keys to a Successful Automation Method

✓ correctly pulls in the data as intended
✓ can generate multiple versions of narrative templates
✓ can accept study specific user entry to control the data and display
✓ missing data and/or source data which fails to meet standards does not cause malfunction
✓ produces helpful error messages identifying issues with the source data
✓ provides a user-friendly user interface
✓ Can be run by the clinical group responsible for writing the AE Narrative without any programmer intervention
What to look for in an implementation method
Why solid sponsor defined SDTM/ADaM standards are critical

Why AE narrative templates need to be air tight

Why enabling study specific input is necessary

Why datapoint metadata is an automation development linchpin

What to look for in an implementation method