PharmaSUG 2024 - Paper DV- 278

Standardization of the Patient Narrative Using a Metadata-driven Approach

Kuldeep Sen, Pfizer Inc.

ABSTRACT

Patient narratives are important components of any Clinical Study Report (CSR) and are attached as "Attachment I" in the CSR. They are part of the FDA's and the ICH's requirements to provide information about safety of a study - specifically the events of death, serious adverse events, and other adverse events of clinical importance at patient level. The patient narratives have become one of the statistical programming deliverables, and creating high-quality narratives can be a challenging task due to several reasons such as unconventional nature of the narrative template, dynamic nature of data at subject level, and limited knowledge of its scope within a statistical programming team. This paper presents a metadata-driven end-to-end approach, which begins with standardization of the narrative template and employs metadata-driven programming automation. This approach has been shown to greatly enhance the consistency and efficiency in creating the patient narratives.

INTRODUCTION

Patient narratives are one of the ICH-E3 requirements for the safety evaluation of clinical studies and are an important component of the CSR. Patient narratives are listed in section 12.3.2. According to ICH-E3, analysis of safety-related data is considered at three levels:

1. Extent of Exposure
2. Adverse Events
3. Deaths, Other Serious Events, and Other Significant Adverse Events.

While the CSR tables, listings, and figures (TLFs) can support the analysis of the safety-related data for these categories at the aggregate level, a brief narrative is required for each subject who experienced the event of death, serious adverse events, adverse events leading to study treatment discontinuation, and other events of clinical importance. In addition, the following information should also be included in the narratives:

- Patient identifier
- Age and sex of patient; general clinical condition of patient, if appropriate
- Disease being treated (if the same for all patients this is not required) with duration (of current episode) of illness
- Relevant concomitant/previous illnesses with details of occurrence/duration
- Relevant concomitant/previous medication with details of dosage
- Test drug/investigational product administered, drug dose if this varied among patients, and length of time administered.

Although patient narratives are an important component of any CSR, many statistical programmers are not familiar with them. They may confuse these with Patient Profiles which have a distinctively different purpose than the patient narrative. One of the reasons for this unfamiliarity is that traditionally, patient narratives were handled by groups outside of statistical programming, usually medical writing, or outsourced to a third party. However, in more and more companies, patient narratives are now handled by statistical programming to improve efficiency, accuracy, and consistency. In this setting, the standardization of the patient narrative template and an efficient end-to-end process to create them can be highly beneficial to the company. In this paper, we explore the standardization process of the narrative template and the metadata-driven end-to-end process of creating the narratives themselves.

STANDARDIZATION OF THE PATIENT NARRATIVE TEMPLATE

Patient narratives are not a typical deliverable that statistical programmers are accustomed to and producing them can be a challenging task if not organized and planned thoroughly. Each patient narrative contains a diverse range of information, is very dynamic at the data level, and cannot be fitted into simple tables or listings. It is a mixture of subject-level identifier and safety information in both tabular and paragraph styles. Hence, the first step is the standardization of the patient narrative template at the departmental level while allowing flexibility at the product and study level. In our experience, the standardization of the template is a cross-functional collaboration between Medical Writing, Drug Safety, and Biometrics. After several rounds of discussion, we finalized the department-level template for patient narratives with two parts which provide a robust structure while allowing study specific customization.

PART 1: SUMMARY-LEVEL TEMPLATE

This template provides the list of subjects that have experienced at least one of the qualifying narrative events. It contains the reason(s) for writing a narrative and high-level patient identifier information as shown in Figure 1.

		Reason for Narrative			
Patient Number	Treatment Group	Death	Other SAE	AE Leading to Treatment Discontinuation	Other Event of Interest
SUBJID 1	DRUG A	X	X	X	X
SUBJID 2	DRUG A	X	X	X	X
SUBJID 3	DRUG B	X			X
SUBJID n	DRUG A		X		

Figure 1: Summary-Level Narrative Template

The highlighted events are of critical importance for all studies. The narrative must be written when a subject experienced at least one of these events. The final column in Figure 1 is an additional event of clinical interest that can be study-specific and may not be required based on the pre-defined narrative plan for a given study.

PART 2: PATIENT NARRATIVE TEMPLATE

The main patient narrative template, which is followed and populated for each subject listed in Part 1, contains all the information from section 12.3.2 of ICH-E3 which was briefly summarized in the introduction section of this paper.

Since a patient narrative packs a lot of information, a balance between readability, programming efficiency, and study-specific customization capability becomes paramount. With that in mind, we developed a template as shown in Figure 2 which contains annotations to explain different sections of the narrative template.

Company **XXX** ***Protocol Name/Number***

Patient Narrative

❶

StudyID/Subject Number	Treatment Group	Narrative Category (Code)
xx-xxx	xxx	[Death (D), AE Leading to Treatment Discontinuation (DC), Other SAE (S), Other Event of interest (O)]

Events Requiring Narrative: ❷

Preferred Term (Narrative Code)	Verbatim Term	Onset Date (Study Day)	Grade	Causality by Investigator	Action Taken
AEDECOD 1 (D, SAE)	AETERM 1	ddmmmyy(xx)	x	Related	Dose not changed
AEDECOD 2 (S, DC)	AETERM 2	ddmmmyy(xx)	x	Unrelated	Drug interrupted

All adverse events are treatment emergent.
MedDRA: vXX.x

Other Reasons for Narrative *(Product/Study Specific if applicable):* ❸

[Table shell for Product/Study specific criteria for narratives, if applicable]

Subject Demographics: ❹

Subject [XXX] was a [XX]-year-old [Race] [Ethnicity] [Gender] [enrolled/randomized] in the [Study]. The subject was [enrolled/randomized] to the [Dose] cohort and received their first dose of [dose level / unit] of [treatment group] on [treatment start date (study day)].
(Text/Information in this section can be added/modified per program basis)

Disease Background: ❺

Subject [XXX] was diagnosed with [Disease Diagnostics] on [DD date]. The primary disease site of origin was the [Origin].
(*Other disease specific text to be added on a per program basis)*

Dosing/Disposition Information: ❻

Subject [XXX] received a total of [number] of doses of [treatment group/study drug] and was on study from [treatment start date (study day)] to [treatment end date (study day)] or current [data-cut-off date]. The subject was discontinued from the treatment on [treatment discontinuation date (study day)] due to [discontinuation reason]. *Or* As of [data-cut-off date] the subject remains on treatment. The subject was then followed in long-term follow-up. The subject died on [death date (study day)]. *or* As of [data cut-off date] the subject remains alive in long-term follow-up.
(Text/Information in this section can be added/modified per program basis)

Relevant Medical History: ❼

Reported Term	Preferred Term	Onset Date (Study Day)	Ongoing
		ddmmmyy(xx)	

Previous and Concomitant Medication: ❽

Preferred WHO Term	Reported Term	Indication	Started Prior to first dose of Study Drug?	Start Date (Study Day)/Ongoing?
			Yes No	

❾
Sequence of Events:

Investigator Causality Statement:

Pharmacovigilance Causality Statement:

Sponsor's Interpretation and Comment:

Figure 2: Standard Patient Narrative Template

The following list corresponds to the numbered sections annotated in Figure 2. Note: values inside square brackets in Figure 2 represent information from the input data set, details of which are provided in the following section.

❶ The subject ID and treatment group are listed here. The final column in this section lists the distinct narrative-triggering events a subject experienced in the trial.

❷ Qualified events for narrative writing from the ADAE ADaM data set. The value in parenthesis is the abbreviation of narrative criteria identified in ❶. Study team has the flexibility to add/drop additional columns from ADAE per study needs.

❸ Other study-specific criteria, if any, are listed here.

❹ Subject demographic information is captured here in paragraph style. The information can be tailored to be study specific.

❺ A subject's disease background at high level is described in this section.

❻ A subject's dosing information along with key disposition events are narrated in this section and will be study specific.

❼ This section lists a subject's medical history in tabular format. Columns can be added or dropped based on study needs.

❽ This section lists a subject's previous and concomitant medication in tabular format. Columns can be added or dropped based on study needs.

❾ This is a placeholder for the medical writers to fill in with additional information relevant to the subject. Some of this information may also come from outside of the clinical database.

METADATA-DRIVEN APPROACH TO GENERATE NARRATIVES

Once the narrative templates are finalized for a study, a metadata-driven approach can be implemented to generate them to further enhance efficiency and productivity. Since the narrative template is organized by sections as show in Figure 2, a metadata worksheet can also be built and organized in Excel to mirror the sections in the narrative template. Figure 3 illustrates the input-output flow of this metadata-driven process to generate patient narratives.

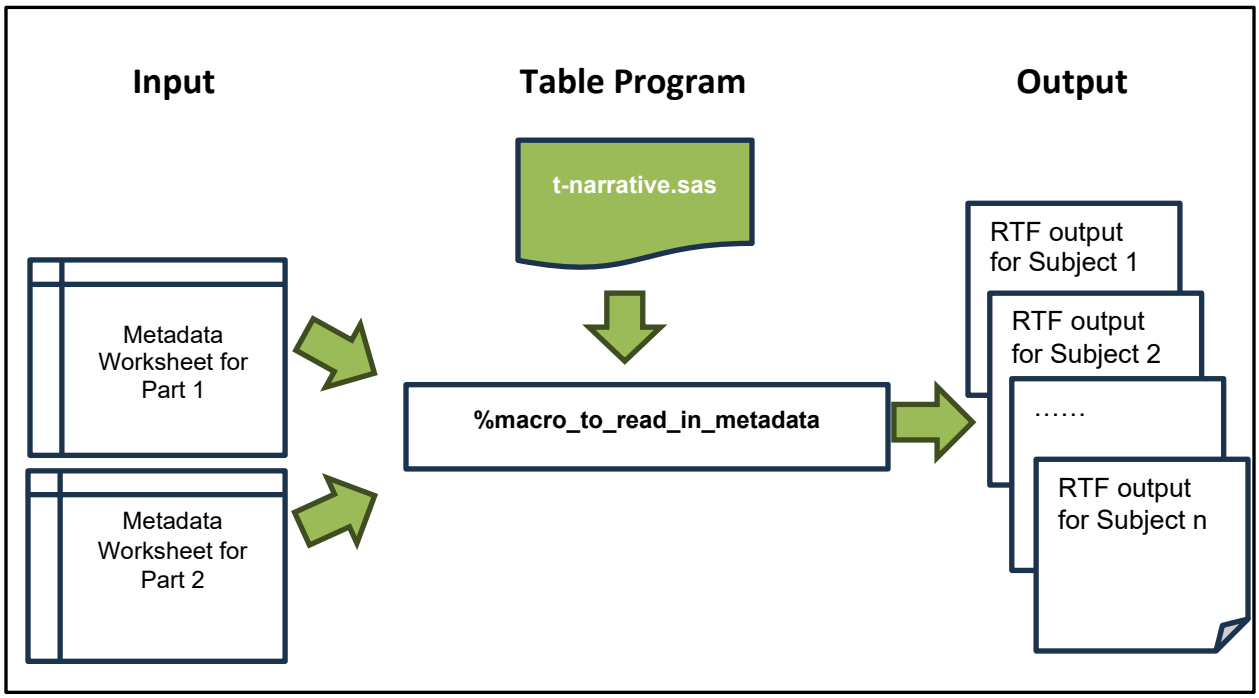

Figure 3: Input and output in patient narrative automation

As shown in Figure 3, the metadata worksheets for Part I (Summary Level Template) and Part 2 (Patient Narrative Template) will serve as an input for SAS table program t-narrative.sas that will generate the final patient narratives. In general, t-narrative.sas will process ADaM data set (ADSL, ADAE, ADCM, and ADMH at minimum, plus other ADaM data set for study specific narrative criteria, if any) to create a temporary data set which is referenced in the column "in_dataset" of metadata worksheet (examples shown in Table 1 and Table 2 below). This temporary input dataset should contain all the information required by the worksheets. Finally, a departmental level macro %macro_to_read_in_metadata can be

created to process worksheets from Table 1 and Table 2 to create the final patient narrative.

An example of the metadata worksheet for the summary-level narrative template (Figure1) is shown in Table 1 below.

in_dataset	narrative_cat	narr_cat_label	narr_cat_code
work.ADAE	DTHFLAG='Y'	Death	D
	AEDISC='Y'	AE Leading to Treatment Discontinuation	DC
	SAEFL='Y'	Other SAE	S
	HYGLYFL='Y' or PNFL='Y'	Other Event of Interest	OAE1

Table 1: Metadata worksheet for patient narrative summary in Figure 1

The input dataset work.ADAE (created in t-narrative.sas in Figure 3) contains flag variables to flag the events qualifying for narrative writing as shown in the 'narrative_cat' column of the example metadata file. The final two columns store the label and annotation for display as shown in the final summary-level template (Figure 1).

The main body of the patient narrative contains several sections. However, they can be categorized into two display types: tabular style and paragraph style. Sections ❶❷❼❽ in Figure 2 are in tabular style while the remaining sections ❸❹❺❻ are written in paragraph style. For illustration purposes, a screenshot of meta-data covering each type is shown below:

	A	B	C	D	E	F	G	H	I
1	SecNum	Type	Section_label	In_dataset	narr_filter	narr_body	narr_table_col_label	col_width	SortVar
2	2	tabular	Events Requiring Narrative:	work.ADAE	TRTEMFL='Y'	SPI_AEDECOD	Preferred Term|(Narrative Code)	20	AESTDTC
3	2	tabular	Events Requiring Narrative:	work.ADAE		AETERM	Verbatim Term	17	
4	2	tabular	Events Requiring Narrative:	work.ADAE		ifC(n(AESTDY),strip(AESTDTC)||' ('||cats(AESTDY)||')',strip(AESTDTC))	Onset Date|(Study Day)	13	
5	2	tabular	Events Requiring Narrative:	work.ADAE		put(AETOXGRN,best.)	Grade	6.5	
6	2	tabular	Events Requiring Narrative:	work.ADAE		propcase(AEREL)	Causality by Investigator	12	
7	2	tabular	Events Requiring Narrative:	work.ADAE		propcase(AEACN)	Action Taken	10	
8	6	text		work.ADEXSUM	NDOSE>.	Subject [SUBJID] received a total of [NDOSE] doses of [TRT01A] from [propcase(put(TRTSDT,date11.))] to [propcase(put(TRTEDT,date11.))]			
9	6	text		work.ADSL	DCTREAS NE ""	The subject was discontinued from the treatment on **[propcase(put(EOTDT,DATE11.))||' (Day '|| strip(put(EOTDY,best.))||')']** due to [propcase(DCTREAS)].			
10	6	text		work.ADSL	DCTREAS EQ ""	As of [put(DCODT, date11.)] the subject remains on treatment.			
11	6	text		work.ADSL	LTFUFL='Y' and DTHDT NE .	The subject was then followed in long-term follow-up. The subject died on [propcase(put(DTHDT,DATE11.))||' (Day '|| strip(put(DTHDY,best.))||')']			
12	6	text		work.ADSL	LTFUFL='Y' and DTHDT EQ .	The subject was then followed in long-term follow-up. The subject died on [propcase(put(DTHDT,DATE11.))||' (Day '|| strip(put(DTHDY,best.))||')']			

Table 2: Metadata worksheet for narrative summary template

Below is a brief description of each column in the metadata worksheet in table 2:

SecNum: Corresponds to section numbers annotated in the standard narrative template in Figure 2. For illustration, only section numbers 2 (tabular style) and 6 (paragraph style) are displayed.

Type: Section display type – either tabular or text (for paragraph style). For type=" tabular," each record in the worksheet stores the metadata for corresponding columns of the sections in the patient narrative. For example, columns 'narr_table_col_label', 'col_width', and 'SortVar' in SecNum=2 stores column label, column width, and sort by variable information respectively for ❷ "Events Requiring Narrative:"

Section_label: Stores the explicit label for each section in the narrative template.

In_dataset: Input dataset for SecNum. For type='tabular', one input data set per SecNum is allowed. For type='text', one input data set per row in the metadata file is allowed.

Narr_filter: Filter (where clause) to be applied to the input data set for a section.

Narr_body: Contents of each section for display will be stored here. For type = "tabular", they can be variables from input dataset, SAS functions or a combination of both. For example, in cell F3 of Table 1, an existing variable from input dataset work.ADAE was used to display "Verbatim Term", whereas in cell F4, a combination of SAS functions and input data set variables was used to display "Onset Date (Study Day)" for "Events Requiring Narrative:" section of the narrative. Similarly, for type= 'text', values enclosed in a bracket, example cell F9, indicate the input data set variable, SAS function, or combination of both. The rest are the text entered per study needs. In addition, for type='text', the contents under each SecNum can be broken down to subsection by entering them as distinct rows in the worksheet. This will not only help simplify the paragraph style reporting, but also allow for dynamic processing of the information. For example, cell E11 and E12 are mutually exclusive conditions, and depending on the subject's status in the study, contents of either E11 or E12 will be displayed in the narrative template.

Narr_table_col_label: Column labels for the tabular section. Applicable for type='tabular' only.

Col_width: Assigned column width for tabular section. Applicable for type='tabular' only.

SortVar: Sort variable for tabular section. Applicable for type='tabular' only.

Finally, a SAS macro (%maro_to_read_in_metadata) was developed to read in the metadata worksheets (Table 1 and Table 2) to automate the generation of the populated summary-level template (Figure 1) and patient narrative (Figure 2) for each qualifying subject.

CONCLUSION

Patient narratives are an important component of any CSR. They are unlike traditional TLFs that statistical programmers are accustomed to. In addition, the dynamic nature of the data at subject level and wide variety of information in different styles can make programming of patient narratives a very difficult task. However, despite its complexity, the requirements for writing these patient narratives are defined very clearly. For these reasons, the standardization of a patient narrative template and a metadata-driven automation to populate them can bring a significant improvement in consistency and efficiency to the statistical programming team. Finally, medical writers can use such patient narratives provided by statistical programming as a starting point, then add or edit information from data sources such as Patient Profiles, CIOMS (Council for International Organizations of Medical Sciences) and TLFs to create the final version to be attached to the CSR.

REFERENCES

ICH Harmonized Tripartite Guideline: Structure and Content of Clinical Study Reports E3. Section 12.3 https://database.ich.org/sites/default/files/E3_Guideline.pdf

Tammisetti R. and Bhavadas K. 2017. "A Guide to Programming Patient Narratives", Proceedings of PharmaSUG 2017, Baltimore, MD

ACKNOWLEDGMENTS

I want to thank Seagen Inc's (now Pfizer) medical writing, biostatistics, and drug safety teams for their participation and valuable feedback in standardizing the narrative template. I also want to thank Seagen's Statistical Programming Infrastructure team, especially Ramesh Inaganti, for developing the tools and macros to automate the creation for the patient narratives. Finally, I want to thank Wei Qian and Michael Hagendoorn for their review, feedback, and constant support and guidance for this project.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Kuldeep Sen
Pfizer Inc.
kuldeep.sen@pfizer.com