Best practices for annotated CRFs

Amy Garrett, Pinnacle 21

ABSTRACT

There is no doubt that the SDTM annotated CRF (aCRF) is one of the most cumbersome submission documents to create. Once a purely manual task, the extreme burden of creating the aCRF has led to several novel methods to automate or partially automate the process. As the industry moves away from manually annotating CRFs and towards automation, it's more important than ever to truly understand the properties of a high-quality aCRF. This paper reviews published guidance from regulatory agencies and provides best practices for CRF annotations. Following these best practices will ensure your aCRF fulfills current regulatory requirements and meets the needs of internal users and programs.

INTRODUCTION

The annotated case report form (aCRF) is a "PDF document that maps the clinical data collection fields used to capture subject data (electronic or paper) to the corresponding variables or discrete variable values contained within the SDTM datasets"[1]. The aCRF should be created very early in the process because current FDA guidance recommends submitting the aCRF at the same as the protocol[1]. During review, the aCRF is an important reference document when and may be viewed independently document or together with the define.xml. The aCRF can also be used by the study team as a reference document during the course of the study. For example, the aCRF can be used as a resource when creating ADaM specs. It can also be referenced by data management to identify forms containing validation issues from P21, which is useful when issuing site queries. Figure 1 below shows an example of an aCRF page.

Figure 1. Sample aCRF form for a Randomization page

TECHNICAL REQUIREMENTS

FDA ACRF REQUIREMENTS

Any process for creating the aCRF must consider and conform to the FDA technical requirements. The agency doesn't offer robust guidance on the aCRF but existing guidance can be found in section 4.1.4.6 of the Technical Conformance Guide[1]. The guidance is summarized below.

- The aCRF should be submitted preferably at the time a protocol is submitted.

- It should be provided as a PDF with the file name "acrf.pdf".
- The aCRF should include treatment assignment forms, when applicable, and should map each variable on the CRF to the corresponding variables in the datasets.
- The aCRF should include the variable names and coding for each CRF item.
- When data are recorded on the CRF but are not submitted, the CRF should be annotated with the text 'NOT SUBMITTED.' There should be an explanation in the relevant Reviewers Guide stating why these data have not been submitted

FDA PDF REQUIREMENTS

Because the aCRF is a PDF file, it must conform to FDA's *Portable Document PDF Specifications*. Briefly, the key points are summarized below. Please read the entire document for complete specifications[2].

- Acceptable versions of PDF are 1.4 through 1.7, PDF/A-1, PDF/A-2
- Submitted PDF files should be readable by Adobe Acrobat X without the need of any plug-ins or additional software
- Text should be searchable
- Any hyperlinks in the documents should be active
- Submitted PDF should not be password protected
- If possible, use a standard font. If you can't use a standard font, fully embed non-standard fonts into the documents. Standard fonts are: Times New Roman, Arial, Courier, Symbol, and Zapf Dingbats.
- Font sizes should be in a range of 9- 12 points.
- Font color should be black for most text. Hyperlinks can be blue.
- The left margin (top margin for landscape) should be at least 3/4 inch. All other margins should be at least 3/8 inch.
- Include a Table of Contents with hyperlinks to bookmarked pages. Zoom settings should persist when navigating the document.

SUGGESTED GUIDENCE

Finalized in December of 2011, the Guidelines for Annotating and Bookmarking CRFs section of the **Metadata Submission Guidelines (MSG)** provide some additional guidance for sponsors. This document is quite old and newer CDISC Guidelines are currently being developed. In Sept 2019, the **German Speaking CDISC User Group** published a draft version of SDTM aCRF Guidelines which consolidated the technical requirements, recommendations, formats and technical prerequisites from various organizations. This guide has many of the same recommendations as the MSG plus some additional recommendations. Below are some of the key suggestions in these guides[3,4].

FORMATTING RECCOMENDATIONS

- The aCRF should be bookmarked in two ways – by visit and by form
- If more than one domain exists on a page, each domain annotation, and all of its variables, should be color-coded.
- All text in the annotations that represent variable and domain names should be capitalized. If possible, the annotations should not obstruct any text on the CRF page. A sponsor may choose to resize the domain annotation based on the CRF layout.
- Each domain that is represented on a CRF page should have its own annotation on the upper left side of the CRF page with the 2-letter domain code and domain name. To distinguish the domain level annotations from the variable annotations a slightly larger font can be used for the domain annotations.

CONTENT RECCOMENDATIONS

- The aCRF should include additional collection documents if those data were not transcribed into the CRF (i.e. patient diaries).
- The aCRF should not have database annotations on it.
- The aCRF must be complete – it contains all final unique CRF pages/modules. It should not contain any blank pages.
- If additional data sources are used (i.e. patient diaries), they should be appended to the aCRF instead of submitted separately or as part of the SDRG. In this case, the origin should still be eDT. Note, this requirement is only needed if the information in the additional data source is not transcribed into the eCRF.

ADDITIONAL GUIDENCE

An extensive review of the literature revealed additional recommendations not found in any CDISC or FDA publication. When followed, these can improve the process of creating the aCRF.

- Use conventions for annotations[5] – conventions for annotations help automate the process of creating and QC'ing the aCRF. A sample table of conventions is shown in Table 1.
- When applicable, annotate the --CAT. When it printed on the CRF, an annotation --CAT next to the category printed on the CRF is sufficient. If not printed on the CRF, it can be annotated as --CAT = "VALUE"
- Annotate variables by listing them separately. Only combine variables that map to the same value. For example, ORRES/STRESC not ORRES/ORRESU.
- When the variables prefix doesn't match the domain's prefix, include the domain name in the annotation (for example: APDM.SREL)

Type	Annotation Pattern	Examples
Domain	[Domain Short Name] = [Domain Long Name]	(1) AE = Adverse Events
Variable	VARNAME	(1) AETERM (2) DS.DSSTDTC
Variable & Value	VARNAME = VALUE	(1) DSTERM = "RANDOMIZED" (2) DSTERM/DSDECOD = "RANDOMIZED"
Supplemental QVAL	QVAL where QNAM = VALUE	(1) QVAL where QNAM = "RACE1" (2) SUPPDM.QVAL where QNAM = "RACE1"
Value Level (non-SUPP)	VARNAME1 where VARNAME2 = VALUE1 VARNAME1 where VARNAME2 = VALUE1 AND VARNAME3 = VALUE2	(1) VSORRES where VSTESTCD = "SYSBP" (2) VS.VSORRES where VSTESTCD = SYSBP (3) FAORRES where FATESTCD = "ONSETDTC" and FAOBJ = "TRIAL DISEASE"
Unsubmitted data	[NOT SUBMITTED]	[NOT SUBMITTED]

Table 1. Common SDTM annotations construct patterns (adapted from *SDTM Annotations: Automation by implementation a standard process*, 2015, Geo Joy & Andre Couturier)

METHODS TO CREATE ACRF

There are several methods to create the aCRF but all the of them involve manually processing with a PDF editor at some point. The entire aCRF can be created manually in a PDF editor like Adobe Acrobat but this process is time consuming and tedious because it requires the creator to manually create, position, format, and type in hundreds (if not thousands) of annotation boxes. This problem has led to the publication of several semi-automation techniques.

AUTOMATION METHODS

Most automation techniques involve the creation of a forms data format (FDF) file that contains the annotation text and container information (color, size, position, etc.). After the fdf file is created, it can be imported into the PDF document to populate all annotations at once. The methods of creating the FDF file are either SAS® Programming or a combination of Excel and SAS[6,7]. In-house tools have also been developed for creating the FDF file[8]. Figure 2 shows what a sample process with FDF automation might look like. More novel methods that leverage machine learning and MDR systems recently emegerged[9,10]. In addition to automation of annotations, bookmarks can also be created in an automated way[11].

Excel Spec with pages listed → SAS Processing → Output FDF File (annotations.fdf) → Import annotations.fdf into blank crf → Populated aCRF.pdf
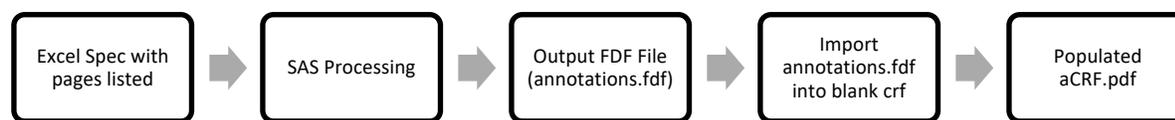
Figure 2. Sample process to populate the aCRF with fdf file.

VERIFICATION/VALIDATION OF ACRF

Automation outputs are only as good as the inputs. Ensuring the quality of the aCRF is incredibly important to helping reviewers properly navigate study data. Below are some best practices for verification and validation of the aCRF.

VERIFICATION – PROGRAMATIC COMPARISONS

One way to check the aCRF for technical conformance is to compare annotation(s) on the page to the actual page number(s) listed in the define.xml. This can be done programmatically by exporting the fdf files and comparing it to the XML where Origin = CRF[12]. If programming specs contain page numbers of form references, a similar approach can be taken to compare programming specs to fdf files. Using annotation conventions will make this process easier because the annotation text file can be parsed in a reliable way.

VALIDATION – USING A CHECKLIST

Unfortunately, programmatic comparisons are not sufficient to fully ensure the quality of the aCRF. It's the best practice to do a manual review as well. When doing manual review, a checklist can be your best friend. A sample checklist is below; the items in this checklist ensure that the aCRF meets official guidance and adheres to company standard practices.

The document meets FDA technical requirements

The document meets MSG requirements

Each domain that is represented on the page should be listed at the top of the page

All annotations follow the pre-defined convention syntax

When all annotations on the page belong to the same domain, the background color should be light blue

When annotations on the page come from multiple domains, the annotations should be color-coded by domain

EVAL is only annotated on the CRF when collected

Annotations do not obstruct any text on the page

It's clear for each annotation which domain it belongs to

All data points are annotated, [NOT SUBMITTED] is used when data are not mapped to SDTM

When applicable, --CAT is annotated

All annotations do not contain typos and are searchable

Annotation text does not contain pseudo-code

CONCLUSION

The aCRF is a critical document for review and following the best practices presented in the paper can help you create a high quality aCRF.

REFERENCES

1. FDA. STUDY DATA TECHNICAL CONFORMANCE GUIDE (2019). Retrieved from https://www.fda.gov/media/131872/download

2. FDA. PORTABLE DOCUMENT FORMAT (PDF) SPECIFICATIONS (2016). Retrieved from https://www.fda.gov/media/76797/download

3. CDISC SDS Metadata Team. Study Data Tabulation Model Metadata Submission Guidelines (SDTM-MSG) (2011). Retrieved from https://www.cdisc.org/standards/foundational/sdtmig

4. CDISC German Speaking User Network. SDTM annotated Case Report Form (aCRF) Guideline (2019). Retrieved from https://wiki.cdisc.org/pages/viewpage.action?pageId=64684119

5. Joy, G., & Couturier, A. (2015). SDTM Annotations: Automation by implementing a standard process. In *Proceedings of the PharmaSUG US Annual Conference*.

6. Zeng, X., & Zhang, S. (2016). Define.xml Content Validation – CRF Page Check. In *Proceedings of the PharmaSUG China Annual Conference*.

7. Pyrnokokis, I. (2015). Developing annotated CRF: SAS, Excel and patience as your friends. In *Proceedings of the PhUSE Annual European Conference.*

8. Zhang, B., & Kelly, T. (2015). A Unique Way to Annotate Case Report Forms (CRFs) in PDF, Using Forms Data Format (FDF) Techniques. In *Proceedings of the PharmaSUG US Annual Conference*.

9. Weising, E., & Ram, S. (2018). Automated CRF Annotation – A Shift from Manual. In *Proceedings of the PhUSE US Connect*.

10. McDade, J. (2019). Automating the blankcrf.pdf and define.xml with your database build. In *Proceedings of the PhUSE Annual European Conference*.

11. Zhou, Y. (2019). Adding Table of Contents (TOC) for the Annotated CRF Automatically. In *Proceedings of the PharmaSUG US Annual Conference*.

12. Pedersen, R., & Both, N. (2016). QC of SDTM and aCRF using SAS. In *Proceedings of the PhUSE Annual European Conference*.

CONTACT INFORMATION

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

Amy Garett
Pinnacle 21
agarrett@pinnacle21.com

www.pinnacle21.com

Any brand and product names are trademarks of their respective companies.