PharmaSUG China 2022 - Paper #142

Integrated Process of aCRF with Dual Bookmarking and TOC for SDTM-MSG-V2.0

Weijie Yang, Everest Clinical Research Corporation

Xinran Luo, Everest Clinical Research Corporation

Qiying Wei, Everest Clinical Research Corporation

Shuang Qiu, Everest Clinical Research Corporation

Abstract

As a required document in FDA submission package, annotated Case Report Form (aCRF) is helpful in data transcription from source documents. According to Version 2.0 of the Study Data Tabulation Model Metadata Submission Guidelines: Human Clinical Trials (SDTM-MSG-V2.0) published in April 2021, a new request for Table of Contents(TOC) for submitted aCRF has been raised. Over the past couple of years, though the time-consuming process led to the exploration of automation tools in the pharmaceutical industry, few papers have talked about the whole process of aCRF production from SDTM mapping to TOC, including automatically updating the page numbers in define specifications after the insertion of TOC. Specifically, in studies with multiple database modifications, the integrated process assures consistency and high quality of submission documents in that it links up the raw data, define specifications, CRF annotation and TOC together, and reduces manual issues through semi-automation process. In this paper, we will show the integrated process of semi-automatically generating an annotated CRF, with dual bookmarking and TOC that meets the SDTM-MSG-V2.0 standard, which involves commonly used software/tools in most pharmaceutical companies, mainly SAS, with the support of Visual Basic for Applications (VBA) and Adobe Acrobat.

Introduction

An annotated Case Report Form (aCRF) is a PDF document with annotations that maps each of the items in the CRF to the related variables in the SDTM dataset. It provides the linkage between the questions stored in CRF and the collected clinical data (Figure 1). aCRF is also a required document of study data submissions to FDA and it needs to follow the industry guidelines. Therefore, it is crucial to standardize the high quality aCRF with high efficiency that meets the CDISC standards and company standards. In order to optimize the process and reduce manual issues, we have created an innovative method to generate an aCRF with dual bookmarking and TOC that meets the SDTM-MSG-V2.0 standard semi-automatically.

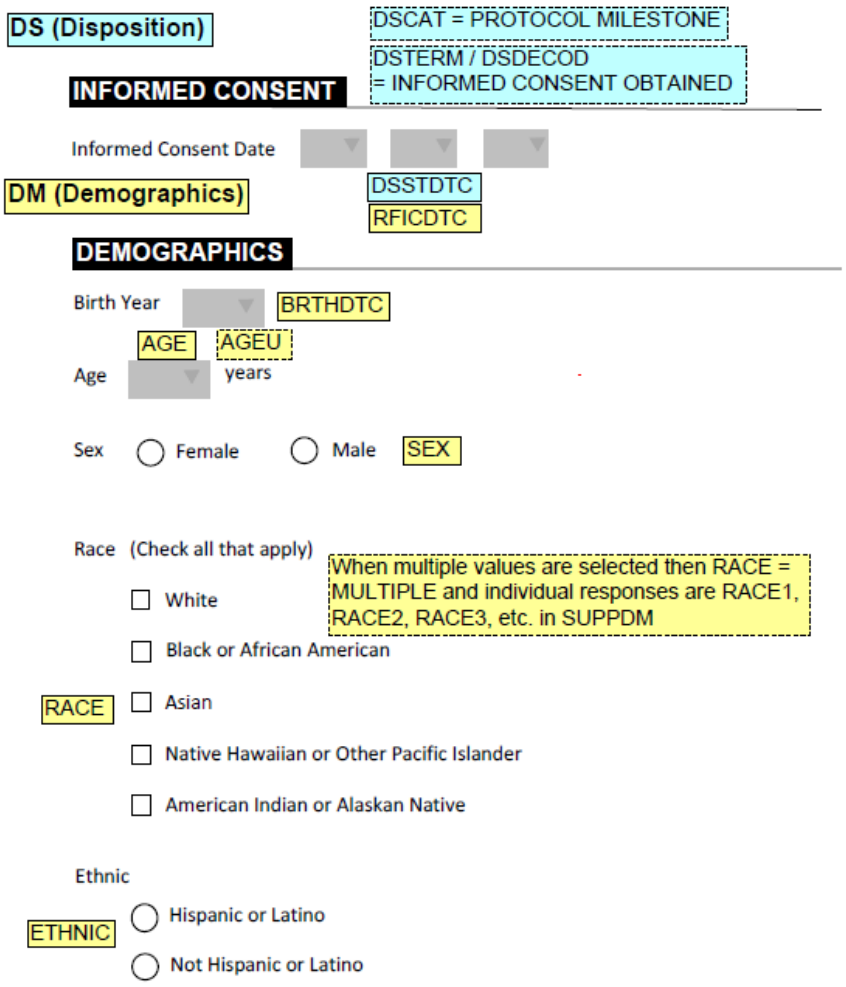


Figure 1 Linkage Between Information Stored in CRF and the SDTM Dataset

Before introducing the integrated process of generating aCRF, dual bookmarking will be described first. Dual bookmarking refers to the presentation of PDF bookmarks in the acrf.pdf file by time point (e.g., Visit 1) and by topic/domain (e.g., ECG). Bookmarks by time point should be ordered chronologically according to the study Time and Events (T&E) Schedule with study-level bookmarks (e.g., Adverse Events) presented last. Within each time point, topic bookmarks should appear in the order that they appear in the aCRF. Unscheduled assessment pages should be grouped under visit "Unscheduled." Bookmarks by topics should be ordered alphabetically. Within each topic, all applicable time points should be ordered chronologically according to the T&E schedule outlined in the Protocol. (Figure 2)

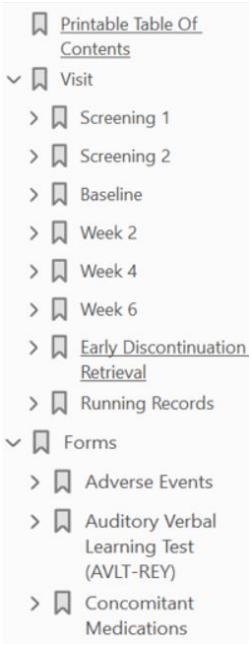


Figure 2 Example of Dual Bookmarking

In order to show the process more clearly, we have created a flow chart shown below. The integrated process of generating an aCRF with dual bookmarking and TOC that meets the SDTM-MSG-V2.0 standard is divided into four steps: 1. Preparing SDTM CRF Mapping; 2. Semi-automated annotation; 3. Creating Dual Bookmarking; 4. Creating Table of Contents. (Figure 3)

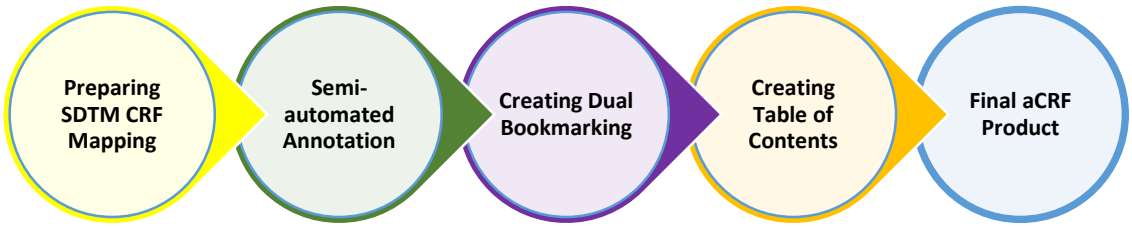


Figure 3 The Integrated Process of Generating an Annotated CRF with Dual Bookmarking and TOC

The advantages of this method are: 1. Maintaining the uniformity of color, text and text across the whole document. 2. Keeping consistent with define specifications for CRF page numbers. 3. Reducing manual issues and improve efficiency.

Now, we will start to introduce the detailed methodology of this innovative method to generate an aCRF with dual bookmarking and TOC that meets the SDTM-MSG-V2.0 standard semi-automatically.

Methodology

1. PREPARE SDTM CRF MAPPING

An aCRF is also a tool to document and track SDTM mapping. The below process extracts page numbers and form names from the pdf version of blank CRF and save the information in a .csv file. We can also review the basic information (labels, formats, number of observations) and the structure of the raw data. Specifically, the process introduced can be applied to RAVE database. To ensure the completion of mapping, this process can also reconcile between the raw datasets and CRF pages since some raw data may not be captured from the CRF while some CRF information does not have a corresponding raw dataset, which needs to query to the data management department.

The summary of checking datasets with the mapping forms:

- Run proc contents on all raw datasets
- Run pdftotext -raw -layout acrf.pdf acrf_to_text.csv
- Load acrf_to_text.csv, extract form and page
- Merge datapagename from the datasets with the form extracted from blank CRF (or acrf.pdf)
- Create final output with Raw dataset, CRF form name, Page number

Please see the detailed instructions below.

Step 1: Run proc contents to retrieve raw datasets contents

SAS:

File Edit View Tools Data Solutions Help

	MEMNAME	NAME	TYPE	LENGTH	VARNUM	LABEL	FORMAT	INFORMAT	NOBS
10	AE	SITENUMBER	2	50	10	SiteNumber	$	$	542
11	AE	SITEGROUP	2	40	11	SiteGroup	$	$	542
12	AE	INSTANCEID	1	8	12	Internal id for the instance			542
13	AE	INSTANCENAME	2	255	13	Folder instance name	$	$	542
14	AE	INSTANCEREPEATNUMBER	1	8	14	InstanceRepeatNumber			542
15	AE	FOLDERID	1	8	15	Internal id for the folder			542
16	AE	FOLDER	2	50	16	Folder OID	$	$	542
17	AE	FOLDERNAME	2	255	17	Folder name	$	$	542
18	AE	FOLDERSEQ	1	8	18	Folder sequence number			542
19	AE	TARGETDAYS	1	8	19	Target days from study start			542
20	AE	DATAPAGEID	1	8	20	Internal id for data page			542
21	AE	DATAPAGENAME	2	255	21	eCRF page name	$	$	542
22	AE	PAGEREPEATNUMBER	1	8	22	Sequence number of eCRF page in folder			542
23	AE	RECORDDATE	1	8	23	Clinical date of record (ex: visit date)	DATETIME	DATETIME	542

Figure 4 Layout of Dataset "rawdata"

From RAVE database, DATAPAGENAME contains eCRF page name which should match CRF form name printed on PDF.

Step 2: Extract CRF form and page number from PDF

**(1) use unix command to transfer pdf text into an .csv file; Confidential information is masked. The same color indicates the same information in pdf versus csv.

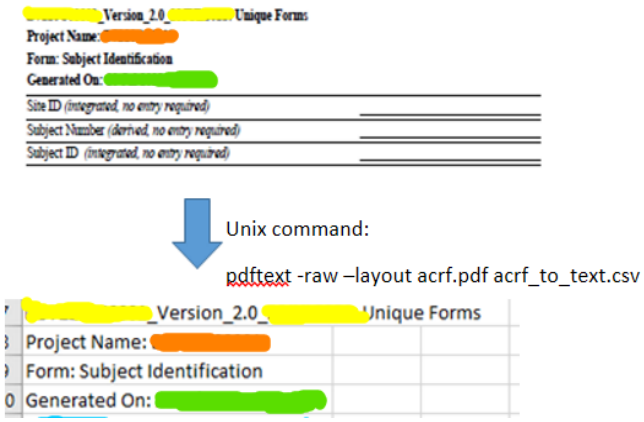


Figure 5 Relationship between acrf.pdf and acrf_to_text.csv

```
**(2) load acrf_to_text.csv using proc import (detailed steps are omitted);

**(3) identify form with the page number;
** this is RAVE form design pattern;
 if upcase(scan(var1, 1,
   ':')) ='FORM' then do;
   form =
   strip(substr(var1, 6));
 end;
 if prxmatch("/\d+ of
   \d+/", var1) then do;
   page =
   input(scan(var1, 1),
   best.); output;
end
```

Step 3: Merge datasets between raw data and CRF extraction

```
**(1) identify dataset from CRF form (detailed steps are omitted);
** RAVE study has variable DataPageName;

**(2) Count number of raw data from CRF form (detailed steps are omitted);

**(3) Retrieve form name from each raw dataset;

%macro get_form(din=);
 %do i=1 %to &ns;
  proc sql noprint;
   select datapagename into :form&i
```

```
      from raw.&&set&i;
   quit;
 %end;

data allforms;
  length dset form $100;
   %do i=1 %to &ns;
      dset = "&&set&i";
      form = "&&form&i";
      output;
   %end;
 run;
%mend get_form;
%get_form;
```

dset	form
AE	Adverse Events
AESI	AESI 1 ALTERED SENSATION BOTH HA 16Sep2021
CM	Concomitant Medications
CONSWD	Withdrawal of Informed Consent
DD	Death Details

Figure 6 Layout of Dataset "allforms"

Step 4: Create final form dataset with raw dataset, CRF form name and page number

```
proc sql noprint;
 create table final as
  select a.dset label ='Raw Data',
        coalesce(a.form, b.form) as form label = 'CRF Form',
         b.page label = 'Page Number'
   from all as a full join forms as b
    on a.form=b.form
     order by a.dset, form;
quit;
```

Step 5: After merging form from raw datasets and CRF form, index file rawinfo.xlsx is generated

	A	B	C	D
1	Raw Data	CRF Form	Page Number	SDTM Domain
2		Overdose Report	35	
3	A	Serious Adverse Events	21	
4		Serious Adverse Events	20	
5	AE	Adverse Events	18	
6	AE	Adverse Events	19	
7	AESI	Adverse Event of Special Interest	22	
8	CM	Concomitant Medications	25	
9	CM	Concomitant Medications	27	
10	CM	Concomitant Medications	26	
11	CM	Concomitant Medications	24	
12	CM	Concomitant Medications	28	
13	CONSWD	Withdrawal of Informed Consent	45	
14	CONSWD	Withdrawal of Informed Consent	46	
15	COVANCE_IMMUNO	B		
16	COVANCE_SAMPLES			
17	DD	Death Details	23	

Figure 7 Layout of "rawinfo.xlsx"

After complete SDTM mapping, all metadata information should be displayed in this index file. Now you can create another column as “SDTM Domain”, the advantages are

- Include all raw datasets
- Include all CRF form
- Some raw datasets can be from external data without CRF form, see case B
- Some CRF form may be missing corresponding raw dataset, see case A – need to confirm with DM

2. Semi-automated Annotation

Due to the homogeneity of CRF production process, a semi-automated annotation process can standardize the CRF generation procedure across different studies, eliminate manual issues and provide high efficiency. Additionally, the semi-automated annotation process also allows the user to map corresponding fields directly from SDTM specification, to keep the consistency between annotation CRF and SDTM specification to a high level.

Step 1. Populate the CRF Page where each corresponding mapped SDTM field is collected in the Origin field of the study SDTM specifications:

Prior to the import, we need to complete the Origin page field in the SDTM specifications. For example, if AE Outcome is collected on page 48, the Origin field for AEOUT in the SDTM specifications should say “CRF Page 48”. If the variable should be annotated on multiple unique pages, list the pages separated by a comma, like “CRF Pages 48, 51”. If a variable is collected as electronic data, as well as from CRF page 48 and 51, the Origin field should say “eDT CRF Pages 48, 51”.

CRF fields that have a 1:1 mapping to an SDTM variable should have the CRF page listed in the Origin field of the corresponding SDTM Domain (e.g., the "AE" sheet of the SDTM specifications). If the CRF field cannot be distinguished by the variable name alone, such as when multiple CRF fields are mapped to findings domain variables such as --ORRES or SUPPQUAL datasets using the QVAL variable, the CRF page is listed in the Origin field of the Value Level Metadata sheet of the SDTM specifications.

Step 2. Export the SDTM specifications to three CSV files.

Create three CSV files named "DATASET_METADATA.csv", "VARIABLE_METADATA.csv", and "VALUE_METADATA.csv". These files are similar to the structure of define.xml data specification, "Datasets", "Variables" and "ValueLevel" tab.

DATASET_METADATA file contains the data information, such as Dataset, Description, Class, Structure, Purpose, Key Variables, Repeating, and Reference Data.

VARIABLE_METADATA file contains variable information such as Dataset, Variable, Label, Data Type, Length, Significant Digits, Format, Mandatory, Codelist, Origin, Pages, Method, Predecessor, Role, and Comment.

VALUE_METADATA file contains variable names and corresponding value characteristics, such as Dataset, Variable, Where Clause, Description, Data Type, Length, Significant Digits, Format, Codelist, Origin, Pages, Method and Predecessor.

Step 3. Import the SDTM metadata CSV files into SAS:

Run the following SAS program to read in the CSV files created. The program assumes that the order of the columns in the SDTM specifications is the same as the SDTM specifications. Update the "pgpath" to choose the right location of output.

```
**(1) Set up the libraries and parameters at the beginning;

%let rotat=0; /*set to 0 for Portrait and 90 for Landscape orientation*/
%let tocpage=0;  /*first page number in CRF*/
%let pgpath=%str(./); /*output save location*/
%let ACROBATVERSION=11.0;

**(2) Import Dataset-Level Metadata;
data dataset_meta;
 length spec data $200;
          infile "DATASET_METADATA.csv"  delimiter = ',' MISSOVER DSD lrecl=32767 firstobs=2;
          length Domain Dataset $8 Description Class Structure Purpose Keys Repeating IsReferenceData
         Comment $200.;
          input Dataset Description Class Structure Purpose Keys Repeating IsReferenceData Comment $;
          domain=dataset;
         run;

**(3) Import Variable-Level Metadata (detailed steps are omitted);
```

```
**(4) Find all dataset and domain;

proc sql;
 create table domain
 as
 select distinct domain length=8,  description
 from dataset_meta where domain^='' order by domain;
quit;

proc sort data=specs;
 by domain;
run;

data specs2;
 length domain $8;
 merge specs(in=a) domain;
 by domain;
 if a;
run;

**(5) Keep only records with CRF page references in Origin;

data crfindomain;
 length domain variable $8 cat description origin annotate $1000;
 set specs2;
 if index(ORIGIN, 'CRF');
 origin=compress(origin||pages, ', ', 'kd');
 origin=strip(compbl(translate(origin, ' ',',')));
 cat='variable level';
 annotate=variable;
 if compress(origin)='' then delete;
 keep domain variable origin description cat annotate;
run;

**(6) Import Value-Level Metadata (detailed steps are omitted);

**(7) Get domain description;
data crfinvalue;
 length domain $8;
          merge crfinvalue(in=a) domain(in=b);
          by domain;
          if a;
         run;

**(8) If variable is already defined in Value Level, then it should be excluded from the Variable Level
(detailed steps are omitted);

**(9) Count demo in one page, set up format and color of annotations (detailed steps are omitted);
```

At the end of the program we need to call a macro (named %mwxfdt here), which produces a file called anno.xfdf containing CRF annotation formatting.
The XFDF file produced by the %mwxfdt macro automatically formats the annotations using the conventions.

```
/*write xfdf for annotation*/
%macro mwxfdt(indsn=,fpath=, outfile=);

filename crfanno "&fpath\&outfile..xfdf";

data _null_;
set &indsn end=last;
file crfanno notitles;
/*common PDF attrib*/
creationdate = "D:"||compress(put(datetime(), IS8601DT.),,'kd')||"-04'00'";
 defaultfont=compress(font_size, '.' , 'kd');
fringe='fringe="0.500150,0.500150,0.500150,0.500150" subject="TextBox"';
body='body xmlns="http://www.w3.org/1999/xhtml" xmlns:xfa="http://www.xfa.org/schema/xfa-
data/1.0/" ';
acrobatversion="&acrobatversion";
xfa='xfa:spec="2.0.2"';
if _n_=1 then do;
  put "<?xml version='1.0' encoding='UTF-8'?>";
  put "<xfdf xmlns='http://ns.adobe.com/xfdf/' xml:space='preserve'";
  put "><annots";
end;
put '><freetext color="' color +(-1) '" date="' creationdate +(-1) '" rotation="' "&rotat" '" page="' page +(-
1) '" flags="print" rect="' rect +(-1) '" ' fringe /
     '><contents-richtext' /
     '><' body +(-1) ' xfa:APIVersion="Acrobat:' acrobatversion +(-1) '" ' xfa +(-1) ' style="font-size:'
font_size +(-1) ';font-weight:bold; font-family:Arial"' /
              '><p dir="ltr"' /
              '>' text +(-1) '</p' /
              '></body' /
              '></contents-richtext' /
              '><defaultappearance' /
              ">0 0 0 rg /Arial" defaultfont +(-1) 'Tf</defaultappearance' /
              '><defaultstyle' /
              '>font: Arial' font_size   +(-1) '; text-align:left; color:' colorappear +(-1) '</defaultstyle' /

              '></freetext>';
if last then do;
  put  "></annots>";
  put "></xfdf>";
end;
run;

%mend;
```

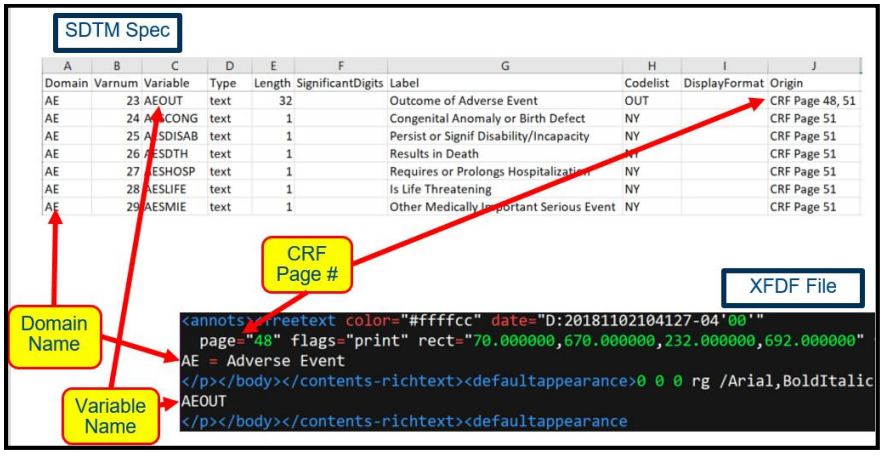


Figure 8 SDTM Spec to XPDF Output

Import the XFDF file into the CRF PDF file in Adobe Acrobat:
In Adobe Acrobat, open the blank CRF file and click Comments Import → Comments → Select XFDF file produced in the last step.

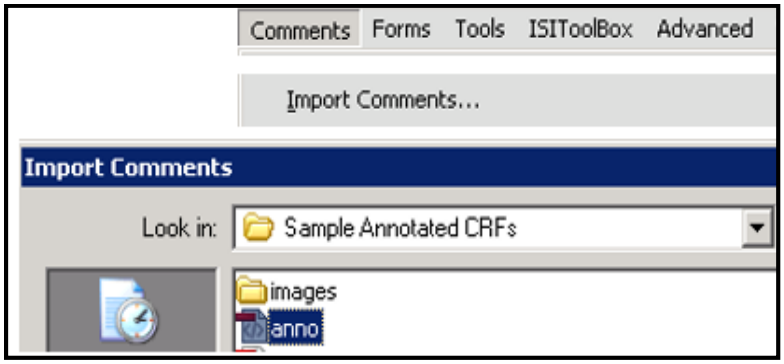


Figure 9 Import the XFDF file into the CRF PDF file in Adobe Acrobat

Reposition the comment annotation boxes to align with CRF fields. Repositioning is required since the XY coordinates of the annotations are not available in the SDTM specifications.

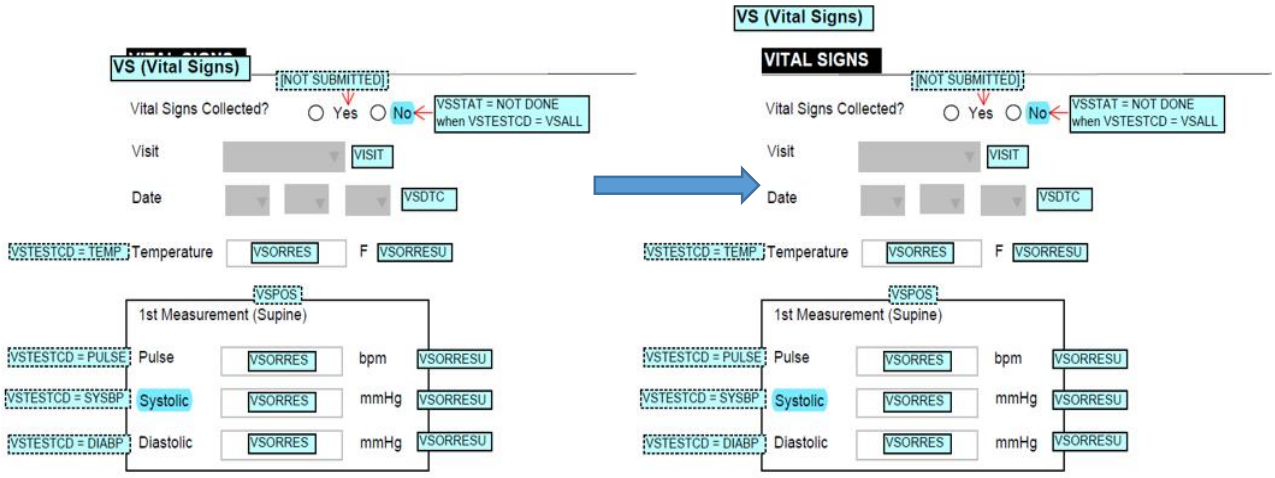


Figure 10 Comments Reposition

3. Dual bookmarking

Dual bookmarking refers to the presentation of PDF bookmarks in the acrf.pdf file by timepoint (e.g., Visit 1) and by topic/domain (e.g., ECG).

Bookmarks by timepoints should be ordered chronologically according to the study Time and Events (T&E) Schedule with study-level bookmarks (e.g., Adverse Events) presented last. Within each timepoint, topic bookmarks should appear in the order that they appear in the aCRF. Unscheduled assessment pages should be grouped under visit "Unscheduled."

Bookmarks by topics need to be ordered alphabetically. Within each topic, all applicable timepoints should be ordered chronologically according to the T&E schedule outlined in the Protocol.

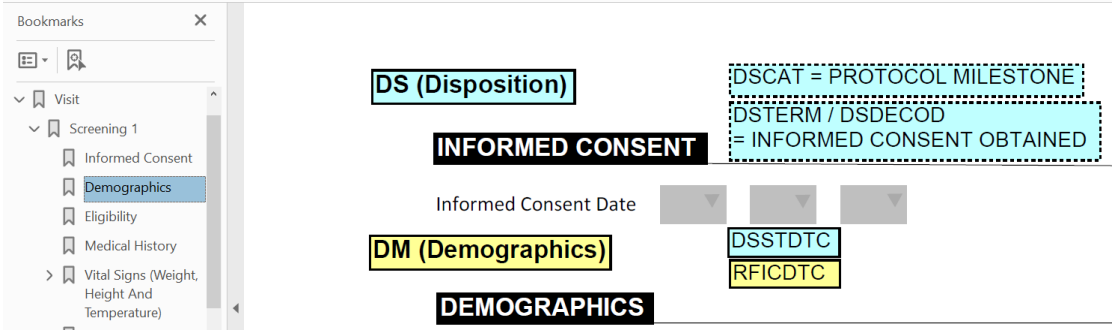


Figure 11 Bookmark by Visit

Step 1: Prepare bookmark csv file

Method 1: From part 1 step 2 in the paper, get dataset forms which include form name and page number

Method 2: Follow SoA from protocol to manually create the template including form name and page number

Follow SoA to add Visits and the selected forms to be done on that visit, add "Running Records" at the end

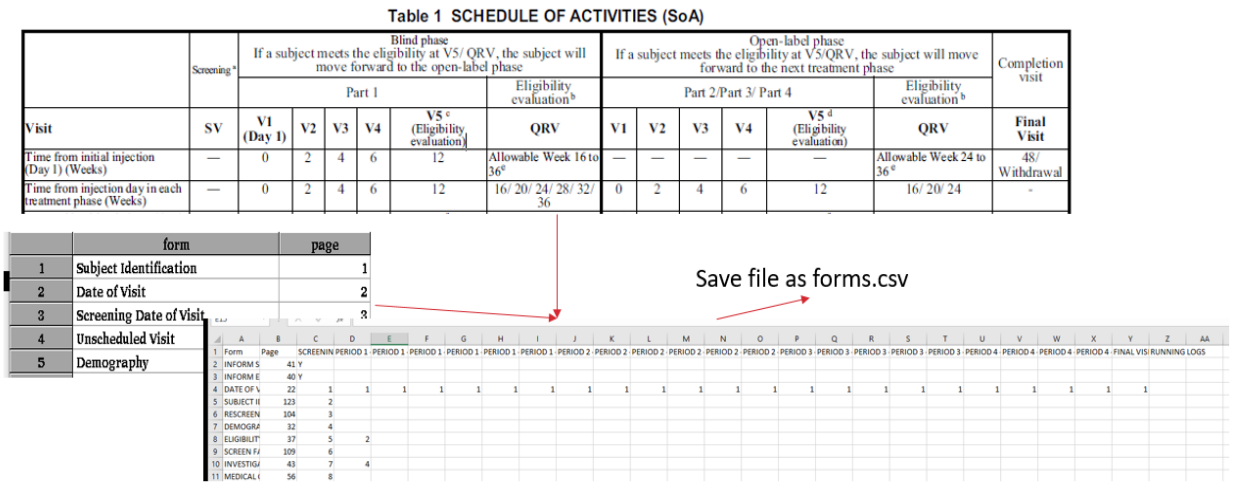


Figure 12 Two Methods to Get forms.csv

Step 2: Run makebookmark.sas on Unix

Copy makebookmark.sas, acrf.csv and CRF to Unix working directory. *%by_visit* and *%by_form* are the 2 main macros in makebookmark.sas. They have similar logic. Below are the key steps of *%by_visit*.

```
%macro by_visit;
  %do i = 1 %to &visitn;
          %put &&vname&i;
          proc freq data=soa(firstobs=2) noprint;
                  tables &&vname&i*var2*var1/out=visit&i(drop=percent count);
                          where not missing(&&vname&i);
                run;
                data visit&i(drop=&&vname&i);
                 set visit&i;
                 length text form $100;
                 level2 = &i;
     text = strip("&&vlabel&i");
                 page = input(var2, best.);
                 form = strip(var1);
    run;

                ** within each timepoint, sort form by page;
    proc sort data=visit&i;
                  by page;
                run;
                data visit&i;
                  set visit&i;
                          level3 = _n_;
                run;
       %end;
  data visits_temp(keep=form page text level:);
```

```
                set   %do i=1 %to &visitn;
                                visit&i
                        %end;
                        ;
        run;
          proc sql noprint;
          create table visit0 as
            select distinct level2, text, min(page) as page
                    from visits_Temp
                            group by level2, text;
        quit;
  data allvisits;
          set visit0
             visits_temp;
          level1 = 1;
          length toc $100 index $10;
          if missing(level3) then toc = text;
          else toc = form;
    if nmiss(level1, level2, level3)=0
      then index = catx('.', catx('.', put(level1, best.), put(level2, best.)), put(level3, best.));
          else if nmiss(level1, level2)= 0
            then index=catx('.', put(level1, best.), put(level2, best.));
        run;

        proc sort data=allvisits(keep=toc index level1 level2 level3 page);
          by level1 level2 level3;
        run;
%mend by_visit;
```

Step 3: Load bookmark acrf.csv

In Adobe Acrobat, open annotation CRF

After import the bookmark, need to review and test the links

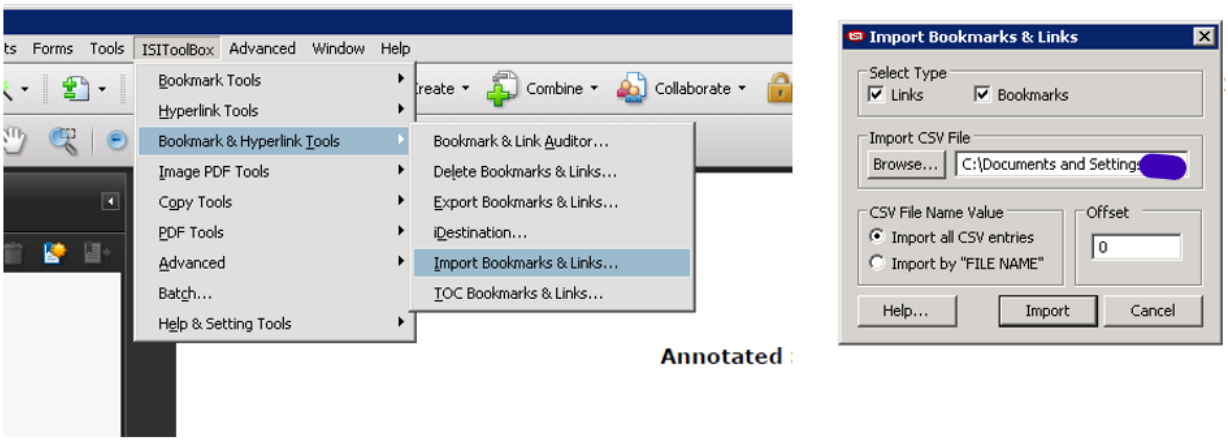


Figure 13 Two Methods to Get acrf.csv

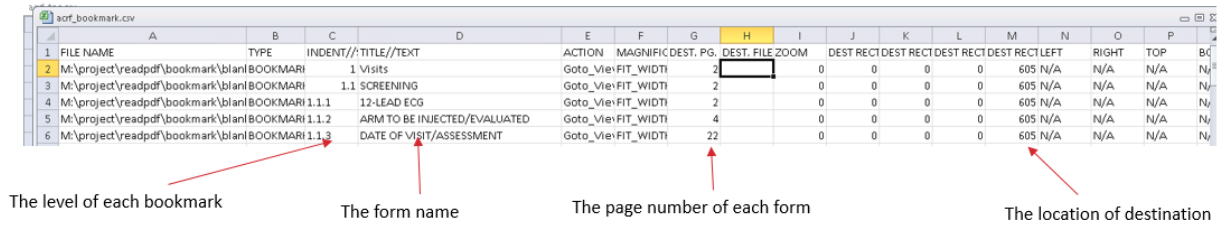


Figure 14 Layout of acrf.csv

4. CREATE TABLE OF CONTENTS

From SDTM-MSG-V2.0, a printable TOC is requested to be included at the beginning of the annotated CRF. The entries in the TOC should be hyperlinked to the respective CRF page, as is done with the corresponding bookmarks. After appending the TOC at the beginning of aCRF, since the page numbers change, this process updates the CRF page in define spec.

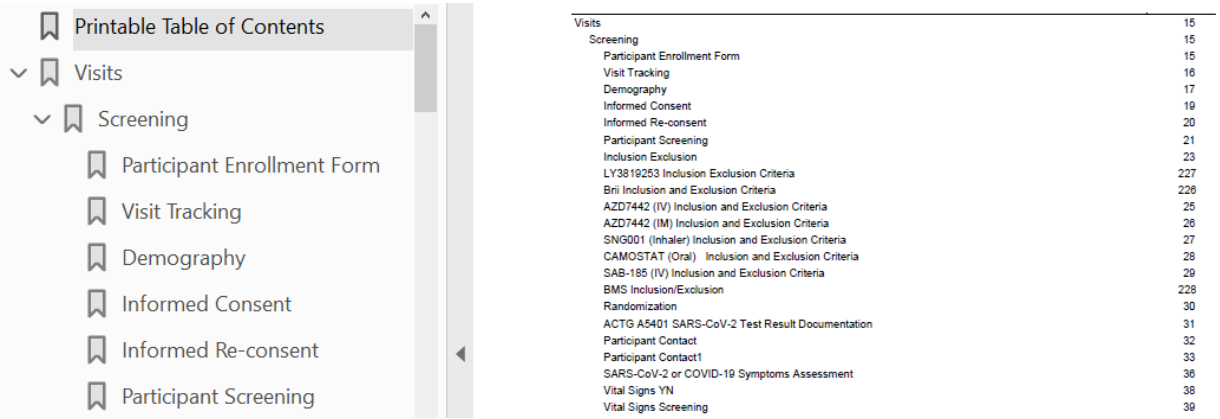


Figure 15 aCRF Bookmark and the Corresponding TOC

The summary of the TOC procedure:

- Generate acrf.csv from dual bookmarking process and save to the study Unix location
- Copy makeacrftoc.sas (please see the key steps below) from to Unix working directory
- This program generates 2 files
  - ➢ toc.doc, the table of contents in word format
  - ➢ acrf_toc.csv, the updated dual bookmarking csv file
- Update file path in the programs

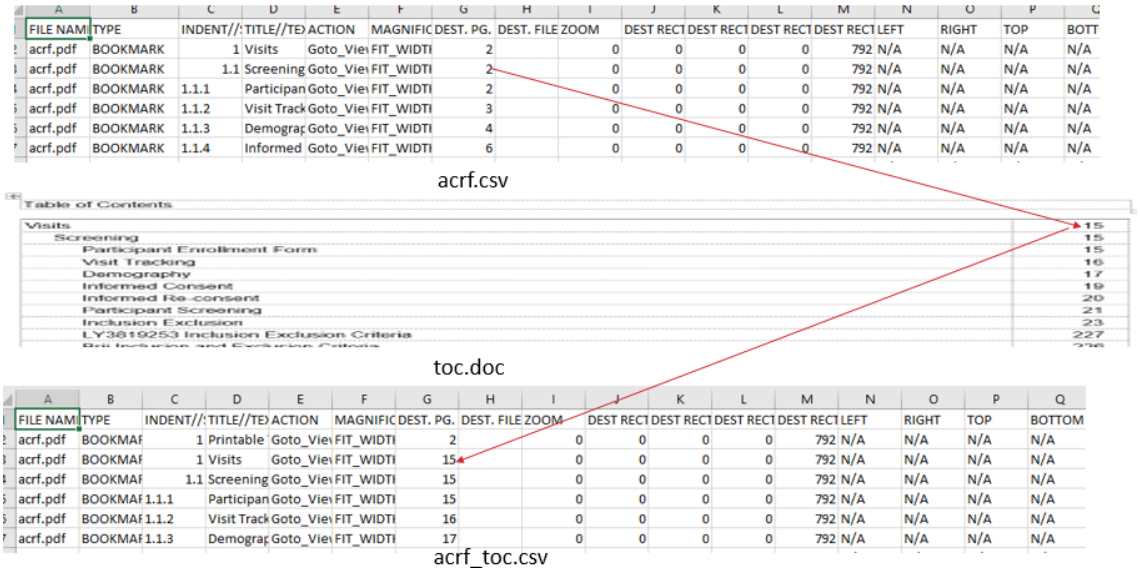
Please see the detailed instructions below.

Step 1: Prepare toc.doc and update dual bookmark acrf.csv page number
makeacrftoc.sas:

```
** define input file name, must be csv;
%let infile = acrf.csv;
```

```
** define output file name, must be doc;
%let outfile = toc.doc;

** define output toc csv file name, must be csv;
%let tocfile = toc_update.csv;

%let ps = %eval(55);

** define output template using proc template (detailed steps are omitted);

/*******Step 1: load acrf.csv******/
** select toc contents;
data toc_1;
 set toc;
 if var2 = 'BOOKMARK';
run;

** keep var3 is for indent, var4 is for content text, var7 is for page number;
data toc_1;
 set toc_1;
 keep var3 var4 var7;
run;
data _null_;
 set toc_1 end=eof;
 if eof then do;
    pn =ceil(_n_/(&ps-5))-1;
   call symputx('tpageno', pn);
 end;
run;
** udpate content based on indentation;
data toc_2;
 set toc_1;
 length text $90;
 text = strip(var4);
 c = countc(var3, '.');
 if countc(var3, '.') >= 1 then text ='   '||trim(text);
 if countc(var3, '.') >= 2 then text = '   '||trim(text);
 if countc(var3, '.') >= 3 then text = '   '||trim(text);
 if countc(var3, '.') >= 4 then text = '   '||trim(text);
 if countc(var3, '.') >= 5 then text = '   '||trim(text);
 pageno = input(var7, best.)+&tpageno;
run;
** use proc report to output toc.doc (detailed steps are omitted);

/*******Step 2: generate acrf_toc.csv******/

data toc_update(drop=pageno);
 set toc;
```

```
 if _n_>1 then do;
   pageno = input(var7, best.)+&tpageno;
   var7 = strip(put(pageno, best.));
 end;
run;
```



acrf.csv

FILE NAM	TYPE	INDENT//	TITLE//TE	ACTION	MAGNIFIC	DEST. PG.	DEST. FILE	ZOOM	DEST RECT	DEST RECT	DEST RECT	DEST RECT	LEFT	RIGHT	TOP	BOTT
acrf.pdf	BOOKMARK	1	Visits	Goto_Vie	FIT_WIDTI	2		0	0	0	0	792	N/A	N/A	N/A	N/A
acrf.pdf	BOOKMARK	1.1	Screening	Goto_Vie	FIT_WIDTI	2		0	0	0	0	792	N/A	N/A	N/A	N/A
acrf.pdf	BOOKMARK	1.1.1	Participan	Goto_Vie	FIT_WIDTI	2		0	0	0	0	792	N/A	N/A	N/A	N/A
acrf.pdf	BOOKMARK	1.1.2	Visit Track	Goto_Vie	FIT_WIDTI	3		0	0	0	0	792	N/A	N/A	N/A	N/A
acrf.pdf	BOOKMARK	1.1.3	Demograp	Goto_Vie	FIT_WIDTI	4		0	0	0	0	792	N/A	N/A	N/A	N/A
acrf.pdf	BOOKMARK	1.1.4	Informed	Goto_Vie	FIT_WIDTI	6		0	0	0	0	792	N/A	N/A	N/A	N/A

toc.doc

Table of Contents

Visits 15
Screening 15
Participant Enrollment Form 15
Visit Tracking 16
Demography 17
Informed Consent 19
Informed Re-consent 20
Participant Screening 21
Inclusion Exclusion 23
LY3819253 Inclusion Exclusion Criteria 227

acrf_toc.csv

FILE NAM	TYPE	INDENT//	TITLE//TE	ACTION	MAGNIFIC	DEST. PG.	DEST. FILE	ZOOM	DEST RECT	DEST RECT	DEST RECT	DEST RECT	LEFT	RIGHT	TOP	BOTTOM
acrf.pdf	BOOKMAI	1	Printable	Goto_Vie	FIT_WIDTI	2		0	0	0	0	792	N/A	N/A	N/A	N/A
acrf.pdf	BOOKMAI	1	Visits	Goto_Vie	FIT_WIDTI	15		0	0	0	0	792	N/A	N/A	N/A	N/A
acrf.pdf	BOOKMAI	1.1	Screening	Goto_Vie	FIT_WIDTI	15		0	0	0	0	792	N/A	N/A	N/A	N/A
acrf.pdf	BOOKMAI	1.1.1	Participan	Goto_Vie	FIT_WIDTI	15		0	0	0	0	792	N/A	N/A	N/A	N/A
acrf.pdf	BOOKMAI	1.1.2	Visit Track	Goto_Vie	FIT_WIDTI	16		0	0	0	0	792	N/A	N/A	N/A	N/A
acrf.pdf	BOOKMAI	1.1.3	Demograp	Goto_Vie	FIT_WIDTI	17		0	0	0	0	792	N/A	N/A	N/A	N/A

Figure 16 Relationship among acrf.csv, toc.doc and acrf_toc.csv

Step 2: Update acrf.pdf

- ❖ Open toc.doc in word, save as toc.pdf
- ❖ Go to remote desktop, in Adobe Acrobat, open annotation CRF, append the toc.pdf at the beginning. If there is cover page of the annotation CRF, then insert toc.pdf after the cover page.

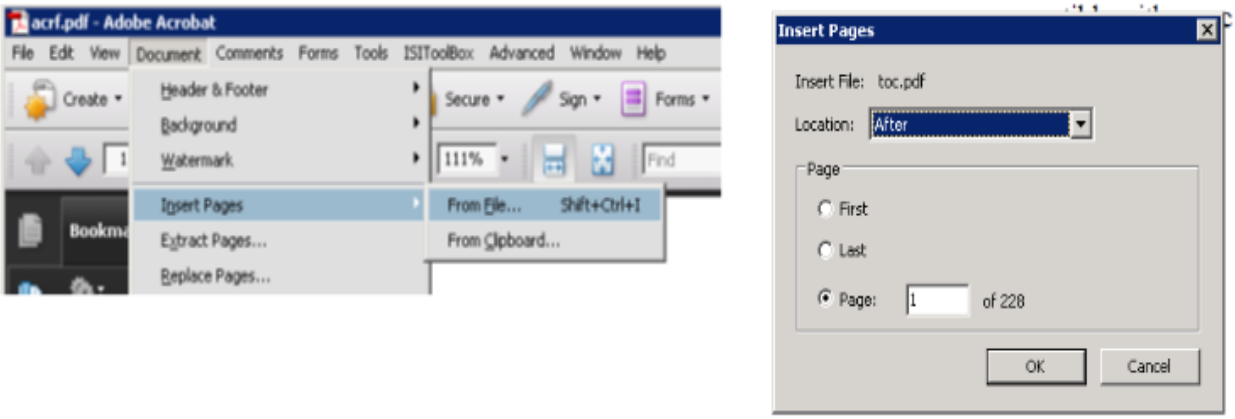


Figure 17 Append the toc.pdf at the Beginning of the Original aCRF

- Set up TOC bookmarks and links
- Go to ISIToolBox - > Bookmark & Hyperlink Tools -> TOC Bookmarks & Links

- Set up the Link Properties, and TOC page ranges, you may need to setup selected area

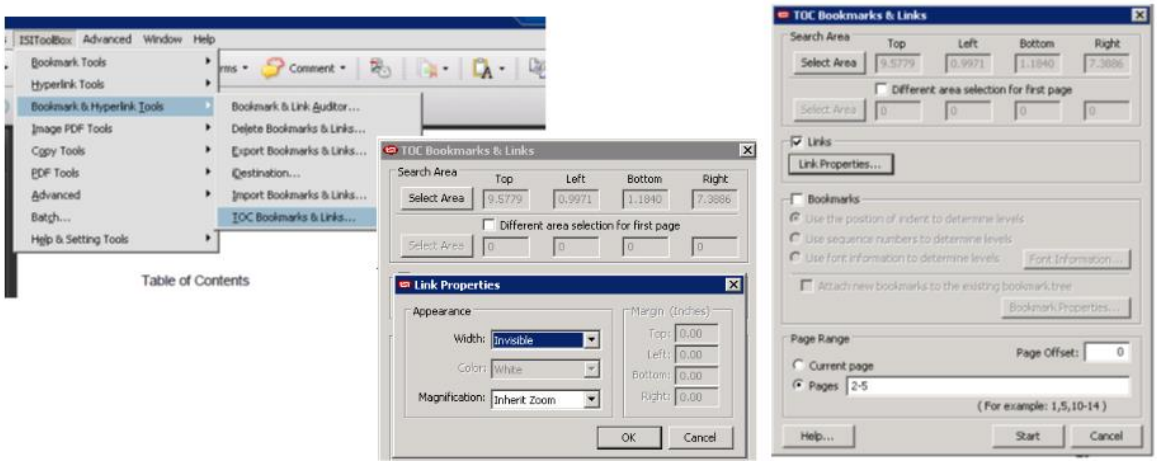


Figure 18 Set up Hyperlink for TOC

Step 3: Update dual bookmark acrf_update.csv

- Since we insert TOC file which makes all page numbers shift, we need to load acrf_update.csv file to update dual bookmarking
- After updating, please check the link from bookmarking and TOC

Step 4: Update acrf page on define spec

- Copy updatedefine.sas (please see the key steps below) to Unix working directory. Please note that when using this program, please pay attention to the change of define spec template. For example, “var 13” in the program may need to be changed to “var 14” if the standard template changes. Please see the corresponding comments in the program when updating.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
	var1	var2	var3	var4	var5	var6	var7	var8	var9	var10	var11	var12	var13	var14	var15
	Order	Dataset	Variable	Label	Data Type	Length	Significant	Format	Mandator	Assigned V	Codelist	Core Varia	Origin	Pages	Method
	1	AE	STUDYID	Study Iden	text	9			Yes				Protocol		
	2	AE	DOMAIN	Domain Ab	text	2			Yes		DOMAIN		Assigned		
	3	AE	USUBJID	Unique Sul	text	17			Yes				Derived		USUBJID
	4	AE	AESEQ	Sequence	integer	8			Yes				Derived		SEQ
	5	AE	AESPID	Sponsor-D	text	2			Yes				Derived		AESPID
	6	AE	AETERM	Reported T	text	43			Yes				CRF	47	

Figure 19 Layout of variable.csv

updatedefine.sas

```
/*********Step 1: define the total page number need to add from TOC*******/
%let tpageno = 4;

/*********Step 2: load variable sheet csv*******/
data Variables_final;
        set variables;
        length temp tempc $100;
                temp = var14; /***var14 is the page number column***/
if _n_>1 then do;
```

```
    if not missing(var14) then do;
            cn=countc(strip(compbl(var14)), ' ')+1;
            if var13 ne 'CRF' then do; /***var13 is the origin column***/
            put var11 var12;
            var13 = 'CRF';
            end;
            do i = 1 to cn;
            tempc = strip(tempc)||' '||strip(put(input(scan(temp, i, ' '), best.)+&tpageno, best.));
            end;
    end;
end;
run;
/*********Step 3: load valuelevel sheet csv, similarly as step 2*******/
```

- Copy define spec sheet Variables as Variables.csv and sheet ValueLevel as ValueLevel.csv and same to the same location on Unix
- Run updatedefine.sas, it will update CRF page number column for both files.
- Then save these 2 files back to define spec

Conclusion

The integrated process (Figure 20) of generating an annotated CRF with dual bookmarking and TOC that meets the SDTM-MSG-V2.0 standard provides solutions to reduce manual issues and improve efficiency. The mechanism behind this process is easy to implement. Each application utilized in this process to generate annotated CRF is widely used in daily work by most companies. Therefore, this process will provide a more practical and economical way to produce annotated CRF semi-automatically for many companies.

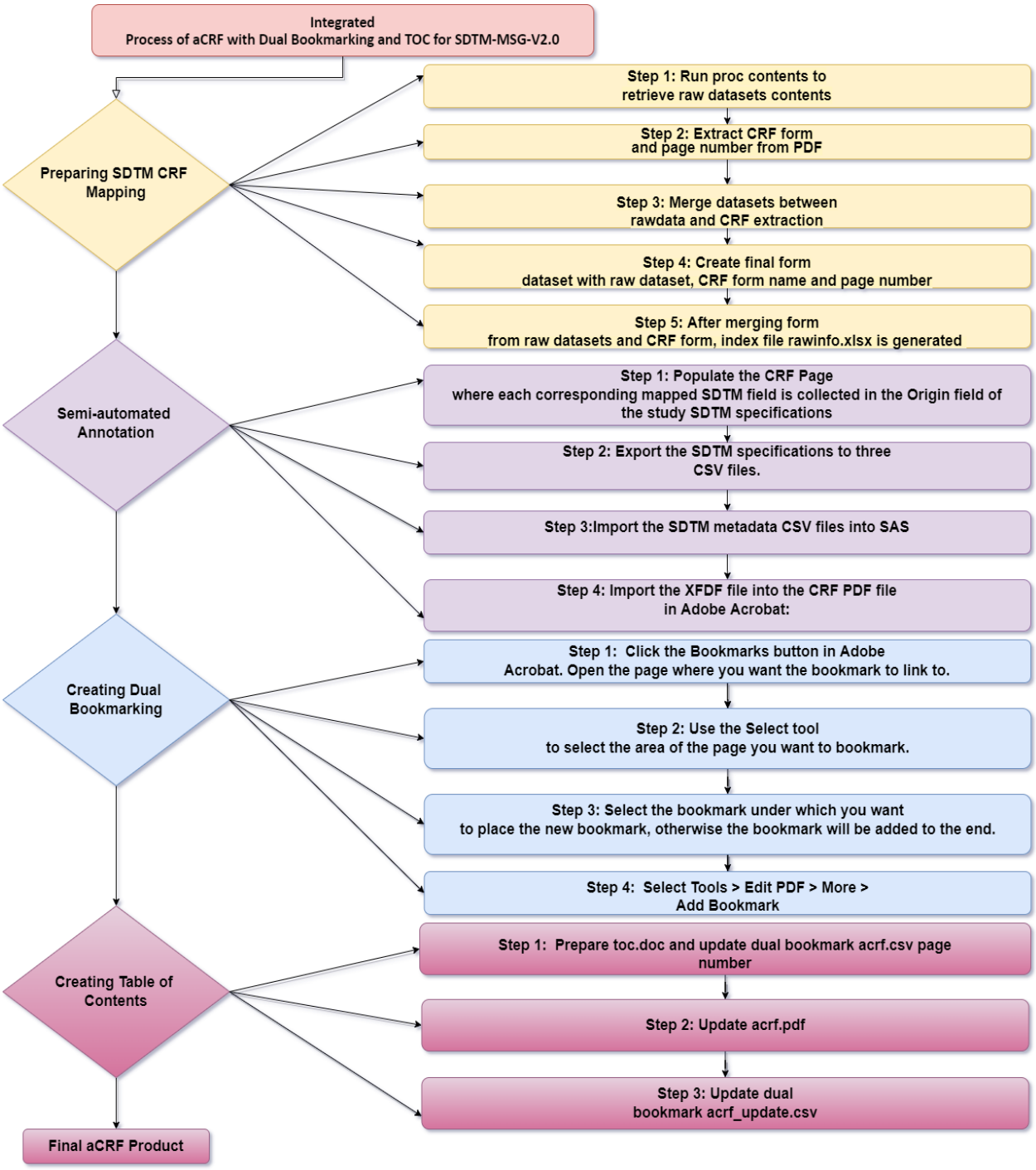


Figure 20 The Integrated Process of Generating an Annotated CRF with Dual Bookmarking and TOC that Meets the SDTM-MSG-V2.0 Standard

References

12 ANNOTATING CRFS Oracle® Clinical Creating a Study Release 4.6.2 E18820-01 Available at: https://docs.oracle.com/cd/E22982_01/index.htm

Developing annotated CRF: SAS, Excel and patience as your friends Ilias Pyrnokokis, OCS Consulting B.V., 's-Hertogenbosch, The Netherlands PhUSE 2015 PP29 Available at: https://www.lexjansen.com/phuse/2015/pp/PP29.pdf

STUDY DATA TECHNICAL CONFORMANCE GUIDE Available at: https://www.fda.gov/media/131872/download

CDISC Study Data Tabulation Model Metadata Submission Guidelines v2.0 (SDTM-MSG) Available at: https://www.cdisc.org/standards/foundational/sdtm/sdtm-metadata-submission-guidelines-v2-0

U.S. Department of Health and Human Services Food and Drug Administration Study Data Technical Conformance Guide Technical Specifications Document Available at: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/study-data-technical-conformance-guide-technical-specifications-document

Portable Document Format (PDF) Specifications, Technical Specifications Document Available at: https://www.fda.gov/media/76797/download

ACKNOWLEDGMENTS

We would like to express my sincere gratitude to our manager Weijie, Yang for the support of our paper.

CONTACT INFORMATION

Your comments and questions are valued and encouraged.

Contact the author at:

Shuang Qiu

Everest Clinical Research Clinical Corporation

+86 15621762998

Miya.Qiu@ecrscorp.com

Xinran, Luo

Everest Clinical Research Clinical Corporation

+86 021.5079.1081

Flora.Luo@ecrscorp.com

Qiying, Wei

Everest Clinical Research Clinical Corporation

+86.021.5079.1036

Katrina.Wei@ecrscorp.com

Weijie, Yang

Everest Clinical Research Clinical Corporation

+1.647.230.3448

Weijie.Yang@ecrscorp.com