

Do It Yourself: Case Report Form (CRF) Annotation

Caro Sluijter, OCS Life Sciences, 's-Hertogenbosch, the Netherlands

ABSTRACT

There is more than one way to skin a cat; and there is more than one way to annotate Case Report Forms (CRF). As regulatory authorities and CDISC have guidelines to deal with metadata submission, we have a road map to create annotated CRFs. However, some guidelines contradict each other, and what do you do when your aCRF cannot adhere to all guidelines? As a CRO we annotate CRFs, but we also work with CRFs annotated by clients and we see differences regarding annotations and bookmarking. These different approaches are not wrong, but to be consistent we try to annotate CRFs following the FDA Technical Specifications Document and CDISC Metadata Submission Guidelines. CDISC guidelines are updated in March 2021, therefore we took a closer look at how we think we can best annotate CRFs. Furthermore, by automating annotation and bookmarking of CRFs we can speed up the process considerably.

INTRODUCTION

After everyone got used to using the Metadata Submission Guideline (MSG) Version 1.01 an update, MSG Version 2.02, was released at the 30th of March 2021. This update was welcomed and clarified some outstanding issues. As the new version has only recently become available, we still see a lot of differences between annotated CRFs from different sponsors.

In the first part of this paper, we will address the issues being resolved in MSG V2.0, and some pointers regarding this new guideline. In addition, the MSG team provided guidance especially on how to annotate a CRF consisting of only unique pages. However, there are still studies where the CRF contains all pages used per visit even when they are repeated pages, and we feel that it is important to think about the linking of those repeated pages. Furthermore, as CRFs are more and more developed with the CDISC SDTMs in mind, it is also possible to re-use annotations in multiple studies. We will show some nice features from PDF editing software that could make our CRF annotation task easier.

For the second part of the paper, we will focus on the other task of preparing a CRF for submission: bookmarking. We want to discuss the updating bookmarking according to MSG V2.0 and give some pointer on dealing with a CRF with non-unique pages. As bookmarking can also be repetitive across different studies, we also want to show some nice features to re-use bookmarkings from one study to the next.

CRF ANNOTATION

ANNOTATION V1.0 VS V2.0: IMPROVEMENTS AND POINTS OF ATTENTION

As a clinical research organisation (CRO) we transform study data to CDISC standards for various clients/sponsors. For most of these studies the CRF annotation is done by our client. We have noticed that although they all adhere to MSG V1.0, they do not look alike or harmonised. For example, we found that the colour of the annotation boxes could differ between clients, and also the content of the annotation boxes was not aligned between clients, while these are both aspects that are covered by the MSG. Where one client would use blue-coloured boxes with a blue bold font, another client could use green boxes, with a red border and with a red bold/italic font. Overall, the look of some CRF pages looked crowded with all the annotations present as suggested by the MSG V1.0 (red, bold and italic).

When looking at the MSG V2.0 we see that the MSG team has corrected for this in the new guidelines. Instead of the style suggested in V1.0 (red, bold and italic), it is now suggested to use black, regular lettering. We agree with these update as it makes the CRF pages feel less crowded and the annotations easier to read. To reduce the space needed for annotation even more, some additional suggestions are made: not using quotation marks is now recommended; the equals sign between the domain abbreviation and the domain full name is replaced by parentheses (see figure 1).

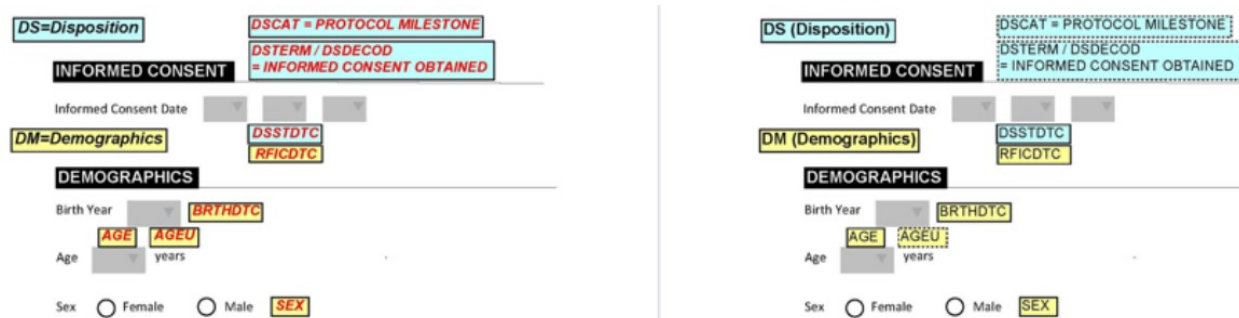


Figure 1. Difference between V1.0 and V2.0

Furthermore, two types of borders are introduced (see figure 1). The solid border, which represent the items that have the origin in the Define-XML as 'collected' and the dashed border, which represents items that are annotated on the CRF to provide the programmer and/or reviewer with some extra information to understand the conversion to the SDTM domains more easily. Even though this seems like a straightforward rule, at OCS we keep having discussions on when to use what type of border.

Example: A pre-printed value on a CRF page.

When only one unit is possible for a question on the CRF, it is logical that this value is pre-printed. This saves time for the investigator when completing the CRF. However, if a value is pre-printed on a CRF, does this SDTM variable (--ORRESU) then have the origin "CRF" or the origin "ASSIGNED"? Even though this does not matter for the final result of the SDTM domain, it does make a difference when describing the metadata for that variable in the Define-XML. Of course, the answer depends on the preferences of the sponsor, but it is good to think about the other metadata when creating the annotated CRF (aCRF).

Another excellent update is the inclusion of recommended colours for the annotation boxes. Four colours are selected to be preferred (i.e. specific shades of blue, yellow, green and red). These colours are light and fresh and above all, they should be distinguishable even to people that suffer from colour blindness. However, from our experience, some of the boxes are still hard to differentiate for people with some form of colour blindness, especially the use of the yellow and green. We would therefore recommend to not use those two colours on one page if it can be avoided.

Regarding the colouring of the boxes, the MSG team advises to use one specific colour per domain. This is not a new concept however in their examples they also set the "NOT SUBMITTED" boxes in the same colour as the domain that is being annotated on the page. In our experience, it is easier to distinguish the information from the CRF to go into the SDTMs from the information that is not submitted in the SDTMs by giving the "NOT SUBMITTED" boxes a different colour. As deviating from the guidelines is generally allowed, we suggest the use of the colour grey for all "NOT SUBMITTED" boxes throughout the aCRF.

LINKING OF NON-UNIQUE CRF PAGES

As the MSG V2.0 mostly deals with how to annotate a CRF consisting of unique page, we would like to offer some tips for dealing with a CRF that for instance consists all forms asked during all visits, e.g. multiple repeated forms. The new guidelines do state what is considered a unique page: a CRF page is unique when that page has a data collection point that is added, removed or otherwise differs from other instances of the same page. Also, minor rearrangements or instructional information of a CRF page that does not affect data collection would not affect the uniqueness.

However, are two pages considered identical when for Visit X you want to store the information of a CRF page in an SDTM, and for Visit Y the same information can be set to NOT SUBMITTED? The pages are the same, the data collection is the same, however the annotations are not. Our advice would be that they are the same page and that we can add a dashed border box to instruct the programmer/reviewer that an annotation is only applicable for Visit X and not for Visit Y.

Another point we would like to make is that the MSG team advises to include a table of content at the start of the aCRF created from the bookmarks in the CRF. This is fine if your aCRF only includes unique pages. However, when you have linked two CRF pages based on the physical PDF page (which you do when referring to a previous page for the annotations), the physical PDF page you referred to does no longer match with the physical PDF page when you have included a table of content, due to the additional physical pages at the start of the PDF. A solution could be to use NamedDestinations for your linking. However, as for instance the FDA does not require a table of content at the start of the aCRF, we could also not include a table of content, as we feel that the bookmarking alone will suffice for the reviewer to quickly navigate through the CRF.

SMART CRF ANNOTATION

As most sponsors or CRO's have a standard in place to develop their CRF's we should also be able to implement this standardization in the aCRF in such a way that you could either have a library of annotated CRF pages, or a way to re-use the complete CRF annotation.

If you want to create a library of annotated CRF pages, this can be achieved by saving each page separately and when creating the aCRF just copy and paste the separate pages into one PDF.

Another way of re-using the annotations created is by storing the annotated PDF as an XML Forms Data Format (XFDF) file. An XFDF file stores the descriptions of the page form elements and their values, such as the names and values for text fields or comments. And as most annotations are in fact comments in a PDF file, we can store these annotations in the XFDF file. Once the annotations are stored in the XFDF file, they can be imported in another blank CRF form, which is then turned into a annotated CRF form with a few clicks. By the way, this XFDF file can also be used to retrieve the CRF pages of SDTM variables for the Define-XML. Solving two repetitive tasks with one newly created file!

```
>0 G 0 0 0 rg 0 To 0 Tw 100 Tz 0 TL 0 Ts 0 Tr /TimesNewRoman 12 Tf</defaultappearance>
><defaultstyle
>font: 'Times New Roman', serif 12.0pt; text-align:left; color:#000000 </defaultstyle
></freetext
><freetext color="#BFFFFFF" creationdate="D:20181122173818+01'00'" flags="print" date="D:20210624212420+02'00'" name="026c86a8-ab65-45d2-a822-973fca464ab1" page="7" rec
><contents-richtext
><body xmlns="http://www.w3.org/1999/xhtml" xmlns:xfa="http://www.xfa.org/schema/xfa-data/1.0/" xfa:APIVersion="Acrobat:10.1.16" xfa:spec="2.0.2" style="font-size:12.(
><p dir="ltr"
><span style="font-family:'Times New Roman'"
>[NOT SUBMITTED]</span
></p
></body
></contents-richtext
><defaultappearance
```

Figure 2 Annotations in XFDF format

Furthermore, having the annotation in XFDF format can also be useful when you want to update your old MSG V1.0 annotations to match the MSG V2.0 annotations, as the file also stores colours and font sizes and types. You could edit the XFDF file in a text editor and use find and replace to change the attributes of the boxes. This does require you to be able to interpret the XML file.

CRF BOOKMARKING

BOOKMARKING V1.0 VS V2.0: IMPROVEMENTS AND POINTS OF ATTENTION

A great update in MSG V2.0 is the wording of what you need to bookmark. It is still advised to have dual bookmarking, "By Chronology" and "By Topic". However, the MSG team now specifies that the term domain in the MSG V1.0 should not be interpreted as the SDTM dataset, but as the domain/form name. Therefore, in MSG V2.0 the dual bookmarking is based on "Chronology" and on "CRF Topic or Form", see figure 3. As this was a big discussion point within our organisation, we are happy that the new guidelines are more straight forward. This means that a page will be bookmarked twice per visit it is used in as specified by the protocol in the Study Schedule of Activities (SoA).

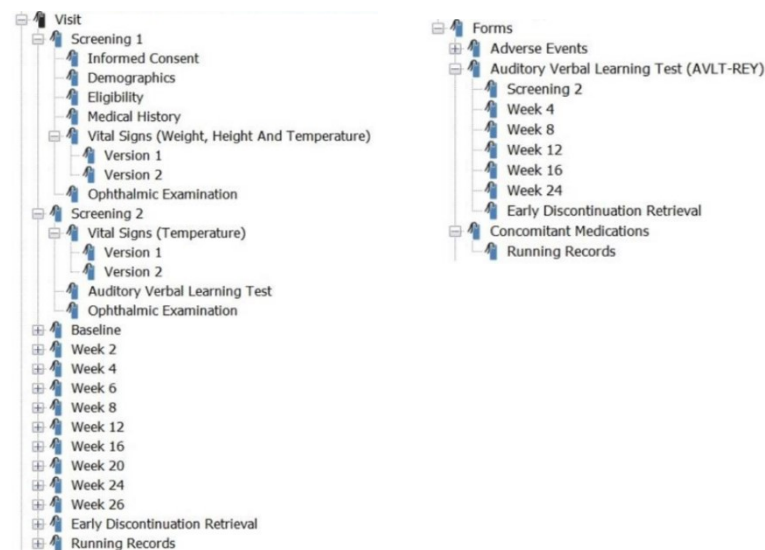


Figure 3 Dual Bookmarking

One thing that we did notice is that when using the CRF form names for the bookmarks, this might hinder the review process as some of the CRF form names are not that descriptive. We would recommend to keep this in mind when the CRF is designed and to use descriptive names for the forms.

Something that does seem a bit redundant, or at least very labour-intensive, is the bookmarking of a unique page for a CRF consisting of only unique pages. The guidelines state that a bookmark needs to be present for every visit in which that page is used. This makes sense in practice, and the bookmarking process is easier than bookmarking a CRF that has all the forms for all the visits. However, it does feel redundant to create multiple bookmarks after each other that will direct the reviewer to the same unique page. For now, we are bookmarking the page for every visit, however a less labour-intensive solution could be to create one bookmark of which the name includes all visits the form is used in.

Lastly, we noticed that it is not clear where the first bookmarks in the hierarchy (the “Visit” or “Forms” bookmark (figure 3)) should link to. For now, we link these first bookmarks to the first visit or topic bookmark for that tree, but we could also have an empty bookmark or bookmark to the first page of the aCRF. We believe all options are acceptable, but for harmonisation between different aCRFs from different sponsors, it would be nice that we all do this in the same manner.

SMART CRF BOOKMARKING

CRF bookmarking in itself is a simple task if you know your way around in your PDF editor. It is however highly repetitive and requires lots of pointing and clicking, which makes it time-consuming. Understanding – and making smart use of – the XML export functionality can help with that. Most PDF editors can export the created bookmarks to an XML file (figure 4). Such an XML file contains information on the page the bookmark points to, the fitetype (zoom level), zoom, and color and style of the bookmark. If you want to bookmark a new aCRF with the same pages and form names, we can import this XML file into the CRF to get a bookmarked CRF.

```
<?xml version="1.0" encoding="UTF-8" ?>
<BOOKMARKS>
  <ITEM NAME="Visit" PAGE="70" FITETYPE="XYZ" ZOOM="0.000000" VIEWRECT="Left=-0.000008;Top=841.980042" COLOR="0,0,0" STYLE="notbold,notitalic" ACTION="" Open="true" INDENT="0">
  <ITEM NAME="Screening" PAGE="70" FITETYPE="XYZ" ZOOM="0.000000" VIEWRECT="Left=-0.000008;Top=841.979980" COLOR="0,0,0" STYLE="notbold,notitalic" ACTION="" Open="true" INDENT="1">
    <ITEM NAME="SignContract registration form" PAGE="70" FITETYPE="Fith" ZOOM="0" VIEWRECT="Top=841.980042" COLOR="0,0,0" STYLE="notbold,notitalic" ACTION="" Open="true" INDENT="2"/>
    <ITEM NAME="Demographic registration form" PAGE="28" FITETYPE="XYZ" ZOOM="0.000000" VIEWRECT="Left=0.000000;Top=841.980103" COLOR="0,0,0" STYLE="notbold,notitalic" ACTION="" Open="true" INDENT="2"/>
  </ITEM>
  <ITEM NAME="Visit 1" PAGE="8" FITETYPE="Fith" ZOOM="0" VIEWRECT="Top=841.799988" COLOR="0,0,0" STYLE="notbold,notitalic" ACTION="" Open="true" INDENT="1">
    <ITEM NAME="BP_HR_sup registration form" PAGE="8" FITETYPE="Fith" ZOOM="0" VIEWRECT="Top=841.799988" COLOR="0,0,0" STYLE="notbold,notitalic" ACTION="" Open="true" INDENT="2"/>
  </ITEM>
  <ITEM NAME="Running Records" PAGE="2" FITETYPE="XYZ" ZOOM="1.178610" VIEWRECT="Left=0.000000;Top=841.980042" COLOR="0,0,0" STYLE="notbold,notitalic" ACTION="" Open="true" INDENT="1">
    <ITEM NAME="AE registration form" PAGE="2" FITETYPE="XYZ" ZOOM="1.178610" VIEWRECT="Left=0.000000;Top=841.980042" COLOR="0,0,0" STYLE="notbold,notitalic" ACTION="" Open="true" INDENT="2"/>
  </ITEM>
</ITEM>
```

Figure 4 XML format Bookmarks

Furthermore, the XML file can also help you when bookmarking your CRF. For instance, when a bookmark is created the zoom level is never set to “Inherit Zoom”. This is a requirement of the FDA, so when submitting your study to the FDA you will need to update the zoom level of every bookmark manually, and unfortunately this cannot be done with one click. If you want to do it manually in the PDF, you will need to update the properties of each bookmark separately. Alternatively, you can also update the properties of the bookmarks by updating the information in the XML file. In a decent text editor this should be a fairly straightforward find and replace action. FITETYPE should be set to “XYZ” and zoom to “0.00000” and once this XML is saved you could just import it back into the document you exported the bookmarks from. Do not forget to remove the bookmarks with the wrong zoom first, or you will end up with twice the amount of bookmarks!

Lastly, we can even take the use of the XML file a step further. We could generate the bookmark XML file from scratch from a custom-made spreadsheet in which the bookmarks per visit and CRF form are entered in a matrix, see Figure 5. Using SAS we can create XML code from this matrix and upload the created XML file into the aCRF PDF to create the bookmarks in a more automated manner. This results in a much more convenient way to create the bookmarks than the manual labour in a PDF editor.

Nr	Forms	Page	Not Submitted	Screening	Visit 1	Visit 2	Follow-up	Running Records
1	Adm_IV registration form				x			
2	AE registration form							x
3	Appoint registration form			x		x		
4	Arrival registration form			x	x		x	
5	BegUrCol registration form				x			
6	Bladder registration form				x			
7	BP_HR_sup registration form			x	x	x	x	

Figure 5 Spreadsheet for Automation of Bookmarking

CONCLUSION

We are very happy with the updated MSG V2.0. It gave us clarity in the way annotation and bookmarking should be done and offers great improvement over MSG V1.0. Of course, every sponsor or CRO will have their own preference and not everyone will agree with all recommendations, but that is the reason these are guidelines and not strict rules.

REFERENCES

1. CDISC SDS Metadata Team (2011-12-31), Study Data Tabulation Model Submission Guidelines (SDTM-MSG), Final version 1.0.
2. CDISC SDS MSG Team (2021-03-30), Study Data Tabulation Model Metadata: Submission Guidelines (SDTM-MSG): Human Clinical Trials, Final version 2.0.
https://www.cdisc.org/system/files/members/standard/foundational/SDTM-MSG_v2.0.pdf

CONTACT INFORMATION

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

Author Name	Caro Sluijter
Company	OCS Life Sciences
Address	Ruwekampweg 2G
City / Postcode	's-Hertogenbosch / 5222 AT
Work Phone	0031 (0)73 523 6000
Email	info.nl@ocs-consulting.com
Web	www.ocs-lifesciences.com

Brand and product names are trademarks of their respective companies.