PhUSE 2009

Paper RG06

CRT for eCTD submission

Aurelien Guillouche, Novartis, Basel, Switzerland

ABSTRACT

A CRT is the document we send to the FDA for an electronic submission. Last year, in 2008, new specifications from the FDA have been decided, based on the eCTD format. A delay of one year has been given to be compliant. These specifications, apparently minor, have changed a lot of processes of Novartis for the creation of CRTs. So we will see what we have done to be eCTD compliant and problems we have encountered.

INTRODUCTION

CRTs means Case Report Tabulations. CRTs are required by the FDA. They are part of the electronic submission (currently US submissions only). The CRT specifications are based on the eCTD (electronic Common Technical Document) defined as an interface for industry to agency transfer of regulatory information. You can find guidance on different websites:

- FDA:http://www.fda.gov/Drugs/DevelopmentApprovalProcess/FormsSubmissionRequirements/
- ICH: http://estri.ich.org/eCTD/eCTD_Specification_v3_2.pdf

The Food & Drug Administration have decided that on the 1st of January 2008 electronic submissions will be based on the eCTD format. This format imply two important rules. The first one is the removing of underscore for which Novartis obtained a delay of one year to be compliant (1st of January 2009). The second one is that we have to follow a specific data structure. Files that are relevant to raw data should be in a folder called 'Listings' and files relevant to derived data have to be in a folder named 'Analysis'. In the Novartis structure, we have kept the old structure and became compliant with these new structure when we put data under a Drug Regulatory Affair specific area. In the following paper you will see that these terms are interchangeable:

Raw ⇔Listings
Derived⇔Analysis

CRTs reflect the Statistical Reporting work, that is why all the SR group have to be involved in the creation of the package or at least in the reviewing.

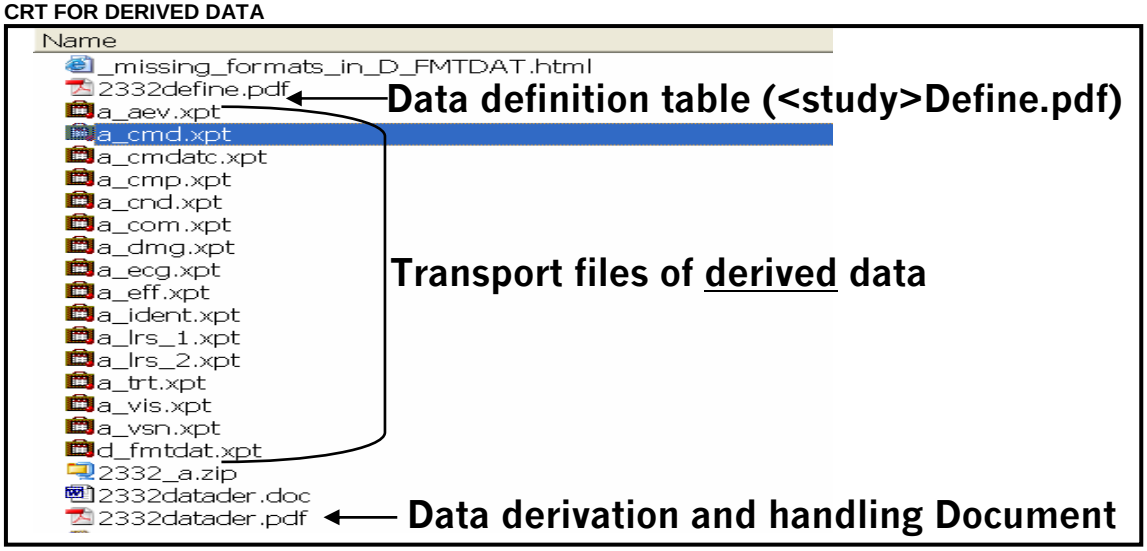
WHAT IS A CRT?

A CRT is the package you send to the FDA when you have to submit a study electronically. We will list what are the documents required in this package and what are the new rules applied to be compliant for an electronic submission.

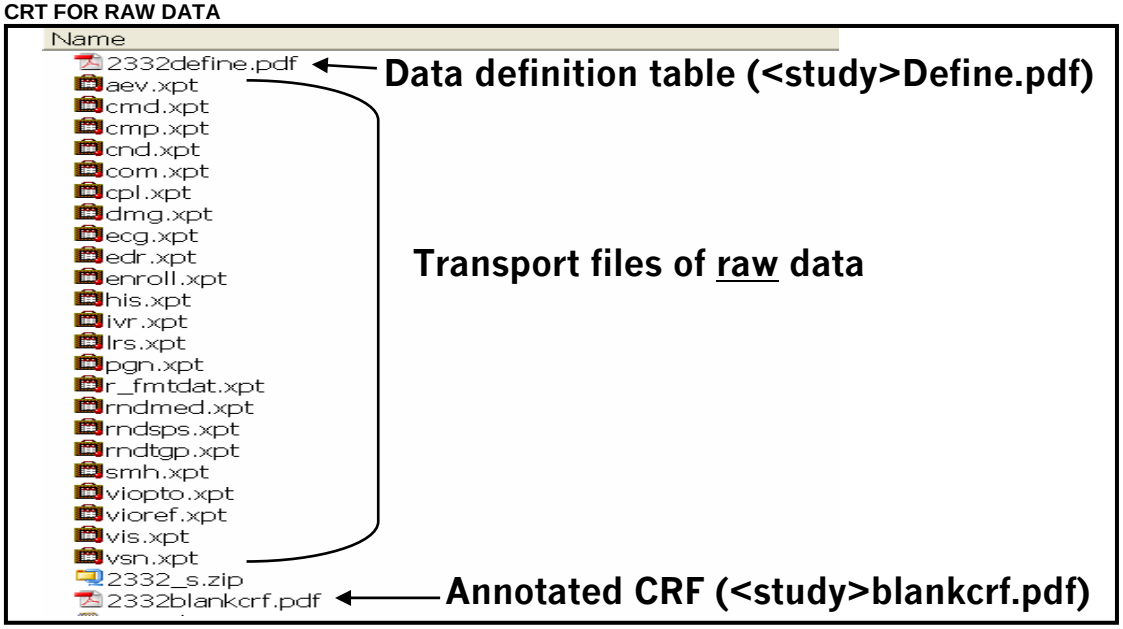
The term “CRTs” refers to the entire package, not just one component and includes:

-SAS® Version 5 transport files (raw and derived datasets)
-Data Derivation and Handling Document
-Annotated CRFs/eCRFs
-Data definition tables (one for raw data and another for derived data).

CRT FOR DERIVED DATA



CRT FOR RAW DATA



SAS Version 5 TRANSPORT FILES

The aim of the CRT is to provide and present datasets on which your study is based upon. Sas datasets have to be transformed into xpt transport files. For a first submission, raw datasets and derived datasets are expected. For an update, FDA should have raw datasets already, that is why only derived dataset are needed. Each transport file has to be provided with meaningful label.

DATA DERIVATION & HANDLING DOCUMENT

The name of this document should be <study>datader.pdf. This document is optional if we follow the specifications of the FDA, but Novartis has decided to consider the datader document as mandatory.

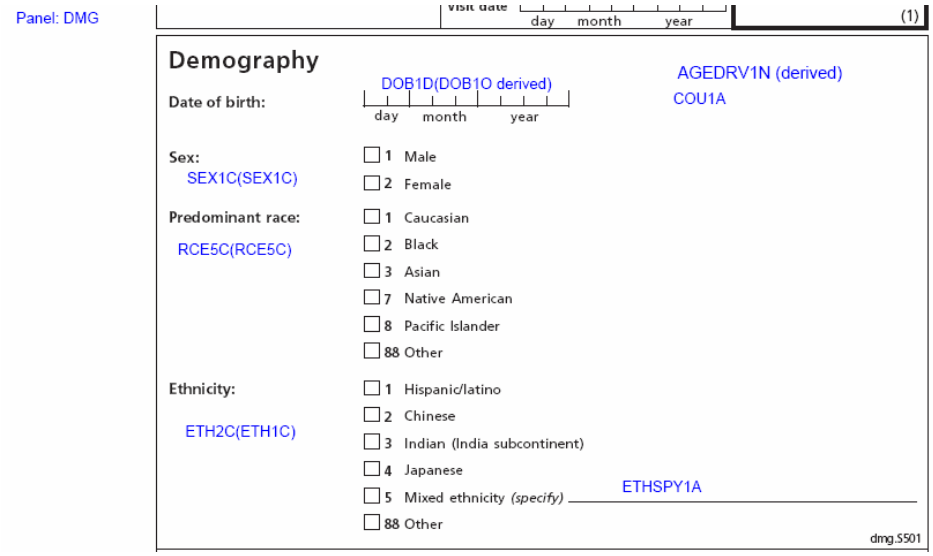
The data derivation and handling document provides further detailed explanation to the reviewer on special data handling issues and other data imputation rules or conventions that were used in the study.
It should contain details which are not self-explanatory by the variable label within the data definition table.
It is provided in order to explain to the FDA how derived datasets are created. Please not that we avoid to put SAS code inside because reviewers of the FDA are not programmers.

Analysis dataset	**pcrpopul**
Creating program	pcrpopul.sas
Description	Population according to several criteria, by line
Usage	To select sub-populations for statistical tables
Unique identifier	SID1A, TIMEPNT
Sorted by	SID1A, TIMEPNT
Notes	One observation per subject, line Only ITT patients The IDENT dataset will be merged with PCRPOPUL dataset by sid1a.
Include Common variables	yes

Variable name	**Label**	**Dataset source**	**Variable source derivation**
SID1A	Subject identifier	pcrvis	
TIMEPNT	PCR during 1st or 2nd line	pcrvis	
Populations			
ITT	Intent-to-treat	pcrvis	
ITT_S	ITT sub-study Population	a_sbjpop pcrvis	Patients randomized from Australia/New-Zealand/Germany sub-study : Patients with ITT=1 and cou1a in ("AUS" "DEU" "NZL")=> ITT_S=1 Otherwise ITT_S=0

ANNOTATED CRFs/eCRFs

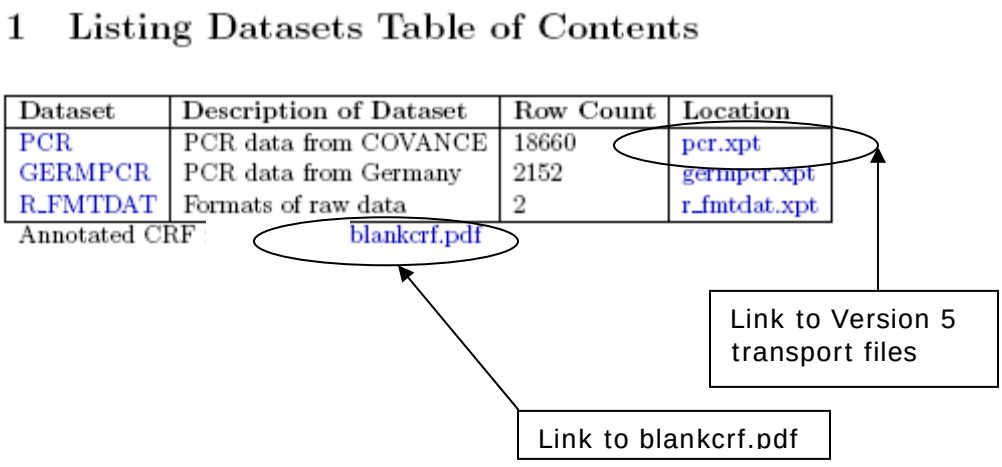
Case report forms (CRFs) need to be sent. In order to show links between CRF questions and the database, the CRF provided to the FDA must contain names of variables. The annotated CRF is named <study>blankcrf.pdf.



DATA DEFINITION TABLE

The name of this document should be <study>define.pdf. The define is a 'pdf' document. It contains a table of contents of the raw and derived datasets. It also contains hyperlinks to the SAS V5 transport files, format files, annotated CRF (for the raw define) and data derivation and handling document (for derived data). The define also provide an alphabetical list of variables.

- A list of contents of Raw datasets (for the raw define)



- A list of contents of Derived datasets (for the derived define)

1 Analysis Datasets Table of Contents

Label of dataset

Dataset	Description of Dataset	Row Count	Location
ASPVIS	Aspirate dates and PH+metaphases	21154	aspvis.xpt
IDENT	Subject population	1106	ident.xpt
PCRVIS	BCR-ABL by PCR sample date	23362	pcrvis.xpt
PCRPOPUL	Patient characteristics by line	1480	pcrpopul.xpt
PCRALL	Merge PCRPOPUL and PCRVIS data	23362	pcrall.xpt
D_FMTDAT	Formats of derived data	840	d_fmtdat.xpt

- An alphabetical list of variable

2 Alphabetical List of Variables

2.1 Variables beginning with A

Alphabetical List of Variables		
Variable	Label	Dataset
ABL	ABL transcripts	PCRALL
ABL	ABL transcripts	PCRVIS
ACCEL_N	Progression to accel. phase date (num)	PCRALL
ACCEL_N	Progression to accel. phase date (num)	PCRPOPUL
ACTINT	Actual dose intensity	PCRALL
ACTINT	Actual dose intensity	PCRPOPUL
AGECAT	Age category	ASPVIS
AGECAT	Age category	IDENT
AGECAT	Age category	PCRALL
AGECAT	Age category	PCRPOPUL
AGEDRV1N	Age (years)	ASPVIS
AGEDRV1N	Age (years)	IDENT
AGEDRV1N	Age (years)	PCRALL
AGEDRV1N	Age (years)	PCRPOPUL
AGESEXR	Age/Sex/Race	ASPVIS
AGESEXR	Age/Sex/Race	IDENT
AGESEXR	Age/Sex/Race	PCRALL
AGESEXR	Age/Sex/Race	PCRPOPUL
APRDT_N	Aspirate date	ASPVIS
ASP_USE	Value to use by TW(Aspir. after Trt st.)	ASPVIS

Link to the dataset page, with all variables included in this one.

- A link to the Data derivation and handling document

3 Data derivation and handling

Clarifications on how the data derivation was performed and the handling methods can be found in the following 6datader.pdf

Link to datader.pdf

NOVARTIS PROCESSES

NEW eCTD RULES

The new electronic format of CRT required additional rules to be followed:

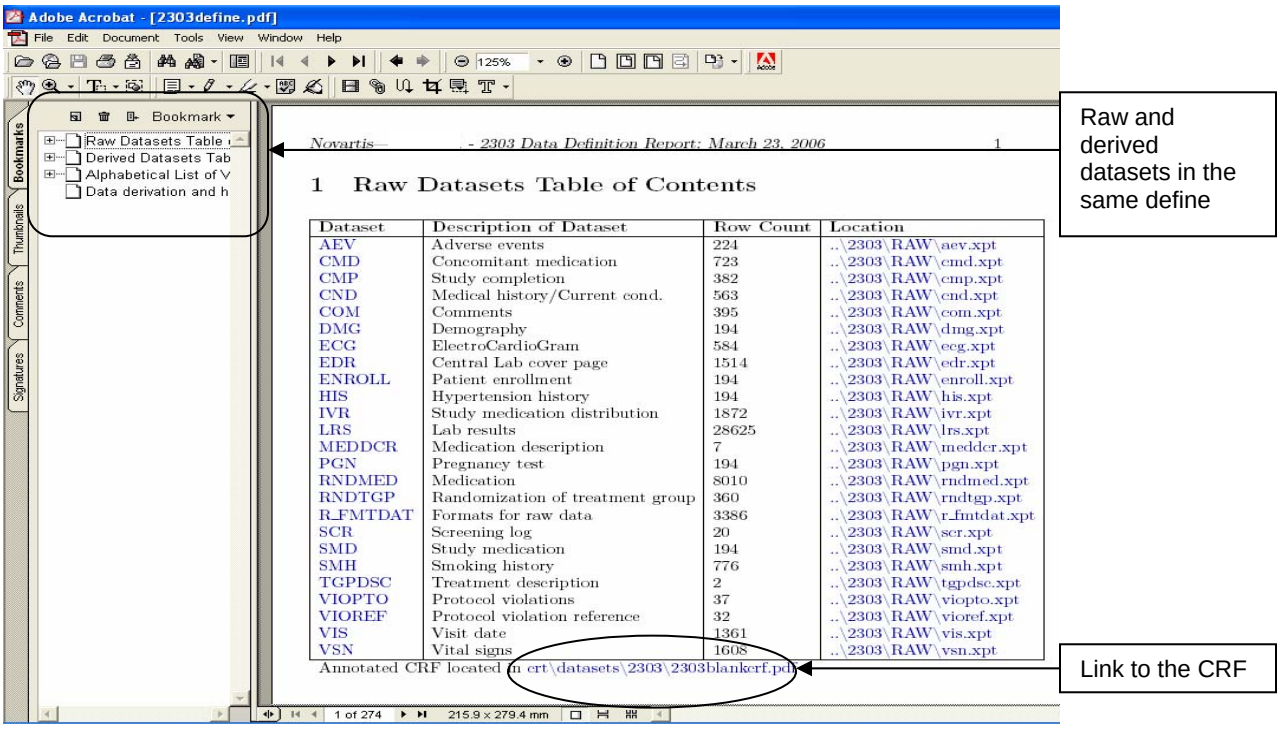
- Naming conventions (no underscores)
- Specific folder structure (listings, analysis)

These two rules required a lot of changes. For the first implementation of the eCTD format in 2008, Novartis adapted the CRT creation process.

The removing of underscores was an important problem for Novartis, because the previous naming convention was that all derived datasets had to be called A_Dataset. The structure of all studies in development and begun before 2008 can not be modified. Because of that, it is not possible to change the name of old datasets in all programs, it has been decided that we delete the underscore just before the translating in SAS v5 export files. A delay of one year has been given to us to apply this rule, we had to remove underscores only for January 2009.

Then, having this specific folder structure means separates derived and raw define document. Before this new rule, we only provide one Data Definition Table for the whole study, this one contained documentations for both derived datasets (now called analysis) and raw datasets (listings), as shown thereafter.

Before:



The Data Definition Table is created by a standard macro called createCRT.sas

PROCESS : TOOLS CREATED BY NOVARTIS

When the programmer have checked some points concerning the quality of data to be sent (consistency of variable's labels, formats, etc.), a macro (createCRT.sas), developed by IT, create the Data Definition Table

automatically. This macro creates links to all XPT files and PDF documents that are in the directory. The only additional document we have to create is a TXT file containing names and labels (and, if relevant, pages of CRF to be linked to) of each XPT files. The update of the macro have added a new parameter called “eCTD”. Other parameters are “DVD” or “NODVD” for the Data Derivation and Handling document, and “RAW” or “NORAW” for the Raw Data definition table or the Derived Data Definition Table.

So to create the Raw (Listings) define document, input parameters should be: “NODVD”, “RAW” and “eCTD”. Because we don’t have to put the Data Derivation & Handling Document linked in this define, it pertains to the derivation of data.

Then to create the Derived (Analysis) define document, input parameters should be: “DVD”, “NORAW” and “eCTD”. Raw data have not to be included, but the Data Derivation & Handling Document has to.

Note that there is no parameter for the CRF. It is because the CRF is necessary linked to the raw datasets and not to the derived ones.

DRA PROCESS

The work of the programmer is to deliver the CRT to DRA (Drug Regulatory Affair). DRA is the primary contact with the FDA, it is their responsability to transfer CRTs and check the compliance of the package. That is why the name of folders change: Raw become Listings and Derived become Analysis as explained in the eCTD rules. So we have to send CRTs into a DRA area and reviewed that all works well (hyperlinks, open xpt files…).

REVIEWING PROCESS

As CRTs are the last step of all the Statistical Reporting work, these have to be fully reviewed by both the programmer and the statistician. The final review is done by the project programmer and the global manager.

ISSUES

The way we work do not provide us with the possibility to modify the layout and the contents of the define. All is done automatically. Few months ago, FDA asked us to provide more information in the define. We can’t modify the macro because it is a standard, and it is really a huge work to change a standard due to IT rules of conduct and validation. The short term solution we have is to add new information in the Data Derivation and Handling Document. Technically, it is the only document in which we can openly modify something.

==> Here is what the **Novartis CRT tool** produce (extra

1.34 A_PTR—Prior Therapy:Radiotherapy

A_PTR—Prior Therapy:Radiotherapy				
Variable	*Label*	*Type*	*Length*	*Format*
AGE_1N	Age at Baseline	Num	8	
BESRES1C	Best response	Num	8	BESR11_
COU1A	Country code	Char	3	

==> Here is the **formatting expected by the FDA**:

Variable Name	Label	Type	Formats(Codes)	ORIGIN	ROLE	Comments
ACYCLE	Analysis Cycle	char		Derived		See acyclec.pdf
ACYCLEC	Analysis Cycle Code	num		Derived		See acyclec.pdf
AGE	Age (years)	num		Derived		(date of randomization – date of birth +1)/365.25
AGEGRP	Age Group	char		Derived		Group by age
AGEGRP2	Age Group2	char		Derived		Group by age

CONCLUSION

The CRT is the final step of the programming work. It is the SR document other department and authorities sees. The realization of a CRT is easier if we have anticipated its creation (check labels, formats…). I have noted that unfortunately the working time to create the CRT is often underestimated. We have to make it quickly at the end of the study because time pressure. But that is one of the most important document created by programmers. It is a final outcome of our work, we should ensure that all details are correct. If the CRT is not compliant or not enough clear, it is the whole project and the image of your department, of the drug and of your company that is undermined.
Being compliant with the new eCTD rules have been a huge work because it have changed a lot of processes. As you can imagine with the last part on this paper concerning the remark of the FDA, our way is not optimal, but the next step is already in development with CRTs provided in the XML format as we will provide CDISC compliant submissions at a later stage.

CONTACT INFORMATION

Aurelien Guillouche
Novartis Pharma AG
Basel, Switzerland
Email: aurelien.guillouche@novartis.com