



缔脉生物医药科技(上海)有限公司
dMed Biopharmaceutical Co., Ltd.

SDRG Automation Using R Markdown

Qiang Wang

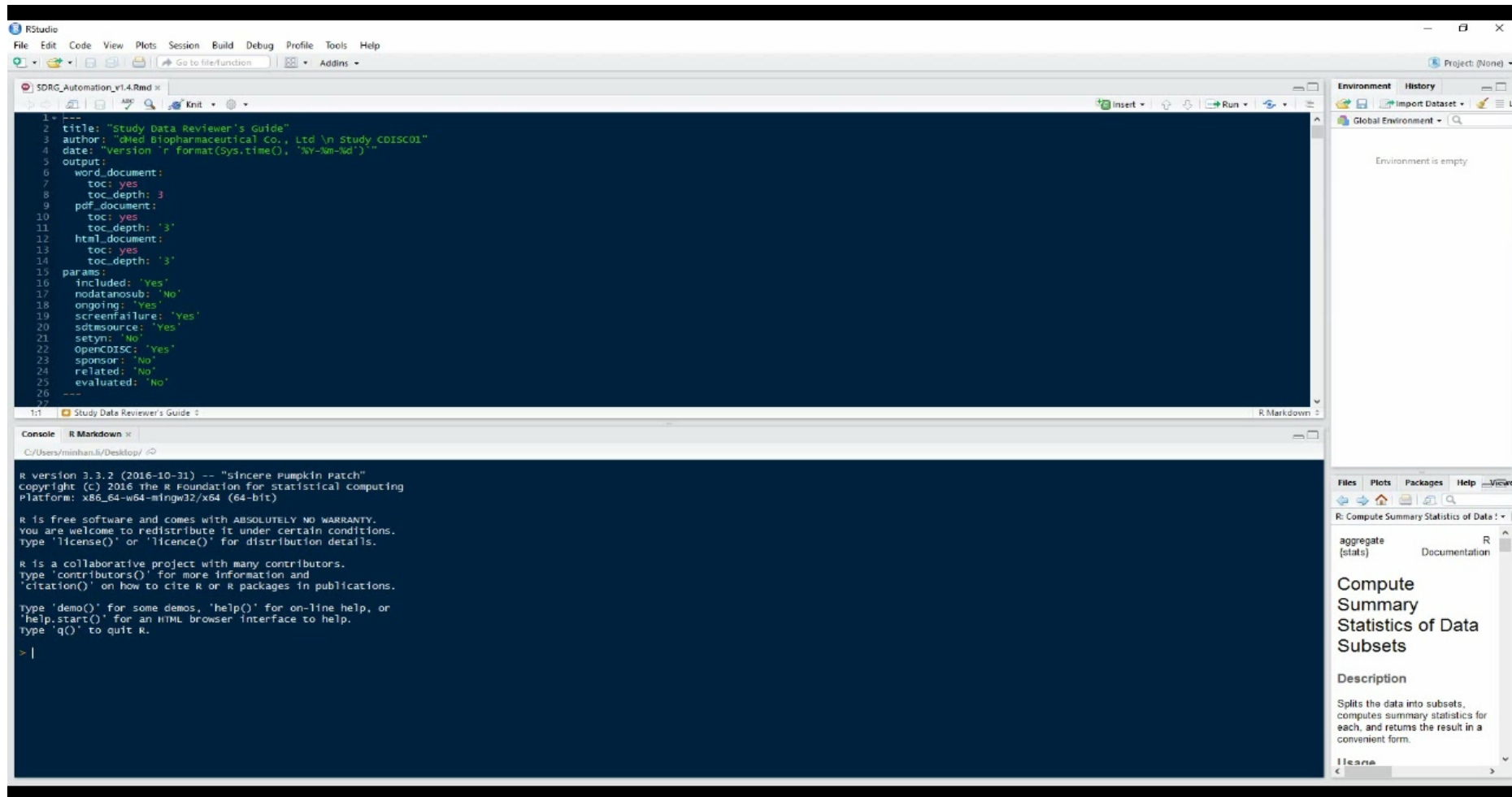
缔结创新之脉，助力成功之道
Increase the probability of your products' success
www.dmedglobal.com

November 03rd, 2017

Table of Contents

- 1. About SDRG
- 2. R markdown and its advantages
- 3. How to do SDRG Automation
- 4. Conclusion

Demo



The screenshot displays the RStudio environment. The main editor window shows a markdown file named 'SDRG_Automation_v1.4.Rmd'. The file content is as follows:

```
1 ---
2 title: "Study Data Reviewer's Guide"
3 author: "dMed Biopharmaceutical co., Ltd \n Study_CDISC01"
4 date: "Version " format(Sys.time(), "%Y-%m-%d")"
5 output:
6   word_document:
7     toc: yes
8     toc_depth: 3
9   pdf_document:
10    toc: yes
11    toc_depth: '3'
12   html_document:
13     toc: yes
14     toc_depth: '3'
15   params:
16     included: 'Yes'
17     nodatanosub: 'No'
18     ongoing: 'Yes'
19     screenfailure: 'Yes'
20     sdmsource: 'Yes'
21     setyn: 'No'
22     OpenCDISC: 'Yes'
23     sponsor: 'No'
24     related: 'No'
25     evaluated: 'No'
26 ---
```

The console window at the bottom shows the R startup message and the prompt '> |'.

```
R version 3.3.2 (2016-10-31) -- "Sincere pumpkin patch"
copyright (c) 2016 The R Foundation for statistical computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
type 'q()' to quit R.

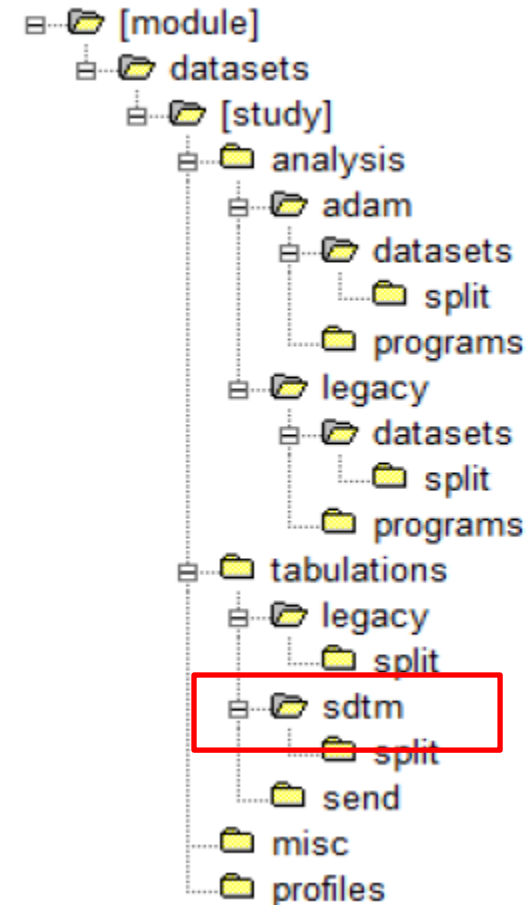
> |
```

The right-hand pane shows the 'Environment' tab, which is currently empty. The 'Files' tab is also visible, showing the file structure of the project.

About SDRG

- What is SDRG? Its purpose
 - Means Study Data Reviewer's in Chinese.
 - As part of a regulatory submission with a single point of orientation
 - Describe SDTM data submitted in the Module 5 clinical section of the submission
 - Should be PDF format and submitted with datasets, aCRF and etc..
 - Replaced by **cSDRG** and nSDRG

Figure 1: Folder Structure for Study Datasets



- Component of SDRG – 4 main sections
 - Introduction – Overview and inventory of standards used on the study.
 - Protocol Description – Brief orientation to the study, additional context about the Trial Design Datasets.
 - Subject Data Descriptions – Additional context for subject-level SDTM domains that are not adequately documented in define.xml or the SDTM Implementation Guide and its supplements. Sponsor-specific annotated CRF conventions.
 - Data Conformance Summary – Documents the validation inputs used to evaluate SDTM conformance and summarizes conformance findings.

- Component of SDRG – 2 optional appendices
 - Appendix I: Fully describe the inclusion/exclusion criteria or as a hyperlink to the full criteria in aCRF.
 - Appendix II: A detailed record-level description of conformance issues. Not suggested.

About SDRG

- http://www.phusewiki.org/wiki/index.php?title=Study_Data_Reviewer%27s_Guide



Home
Categories
Recent changes
New Pages
Popular Pages

Projects

CSS Working Groups
Annual Conference
Therapeutic Areas
Good Programming Practice
Data Transparency
QSPI

PhUSE Wiki Help
Help

Page Discussion

Navigation trail: Study Data Reviewer's Guide

Study Data Reviewer's Guide

Study Data Reviewer's Guide Final V

Name

 Samples

 Readme

 SDRG_Completion_Guidelines_2015-01-26

 SDRG_Template_2015-01-26

h PHU!

Version	Release Date	Downloadable Work Package	Changes from Previous Version
Clinical Study Data Reviewer's Guide			
v1.2	26-Jan-2015	SDRG Package v1.2 2015-01-26	- Removed Trial Design Dataset navigation table from Section 2.3 - Improved SDRG Template usability - Minor revisions to instructions in SDRG Completion Guidelines - SDRG Examples updated to match revised SDRG Template
v1.1	03-May-2013	SDRG Package v1.1 2013-05-13	Initial Version
Nonclinical Study Data Reviewer's Guide			
v1.0	03-Mar-2016	nSDRG Package v1.0 2016-03-03	Initial Version
v1.1	19-Mar-2017	nSDRG Package v1.1 2017-03-19	- More description added to Guideline from public and FDA reviews - Improved consistency of template and guideline to SENDIG - nSDRG samples not updated yet for this package

Applicable Federal Register Notices:

[FR Notice 2015-18027](#) applied to the Clinical SDRG only. The public review period on this notice has ended.

[FR Notice 2016-04791](#) applied to the Nonclinical SDRG only. The public review period on this notice has ended.

Thinking: What information can be obtained automatically?

- 1. Standards used on the study
 - 1.1 SDTM version – Pinnacle report
 - 1.2 Medical Events Dictionary – Pinnacle report or SDTM.SUPPAE
- 2. Protocol description
 - 2.1 Protocol number and title – SDTM.TS
 - 2.2 Protocol design – From protocol, usually insert as an image file
- 3. Subject data descriptions – SDTM mapping specification
 - 3.1 Information of supplemental datasets – SDTM.SUPPXX
- 4. Data conformance summary – Pinnacle report
-

A large green arrow pointing to the right, with the word 'Summarize' written inside it in red.

Summarize

As the data source for SDRG Automation

Table of Contents

- 1. About SDRG
- 2. R markdown and its advantages
- 3. How to do SDRG Automation
- 4. Conclusion

R markdown and its advantages

- What is R markdown?

R markdown is an extensible markdown language under RStudio which can easily use R code to generate Word format report. It includes the core markdown syntax and displays the running results of the R code that are inserted into the final document.

R markdown and its advantages dMed 缔脉

- 1) Can connect to almost any data sources, including SQL server, SAS datasets, excel files, CSV files.
- 2) Can use xtable or knitr::kable to insert tables.
- 3) Support ggplot2.
- 4) Can use internal and external links.
- 5) Support external image file.
- 6) Inline R code chunk.
- 7) It's plain text file, good to be in version control.
-

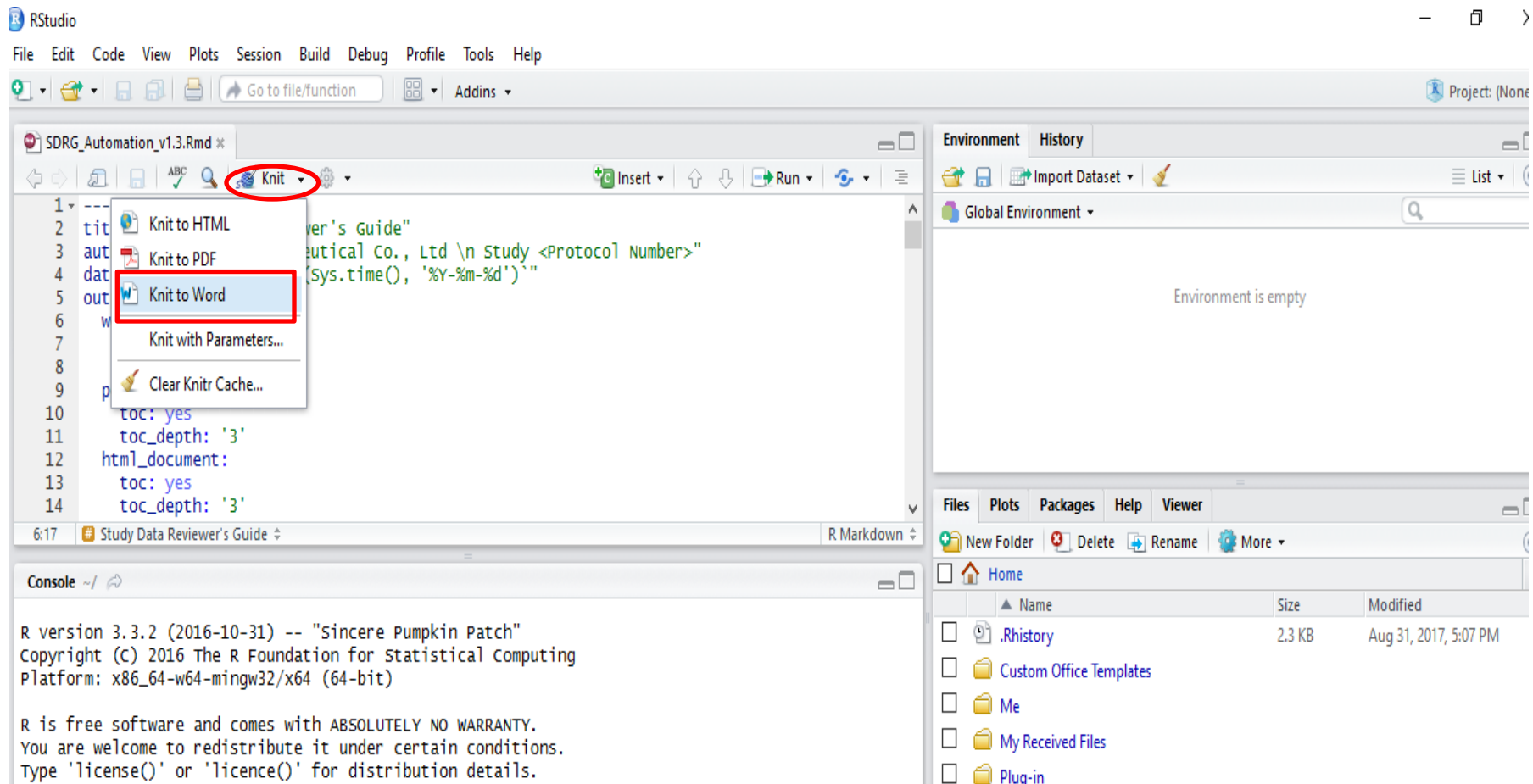
Dynamically generate documents, can be reused.

Table of Contents

- 1. About SDRG
- 2. R markdown and its advantages
- 3. How to do SDRG Automation
- 4. Conclusion

How to do SDRG Automation

- Open the program under RStudio



How to do SDRG Automation

- Example for the output – Basic information of SDTM dataset

3.3 SDTM Subject Domains

Dataset - Dataset Label	Efficacy	Safety	Other	SUPP-	Related Using RELREC	Observation Class
AE-Adverse Events		X		X	CE, CM	Events
CM-Concomitant Medications	X	X		X	AE	Interventions
DA-Drug Accountability						Findings
DM-Demographics						Special Purpose
DS-Disposition						Events
EG-ECG Test Results						Findings
EX-Exposure						Interventions
IE-Inclusion/Exclusion Criteria Not Met						Findings
LB-Laboratory Test Results						Findings
MH-Medical History						Events
PE-Questionnaires						Findings

3.3.1 AE - Adverse Events	
QNAM	Description
AETRTEM	Treatment Emergent Flag
HLGT	Higher Level Group Term
HLT	Higher Level Term
LLT	Lowest Level Term

3.3.2 CM - Concomitant Medications	
QNAM	Description
ATC4CODE	WHO_DRUG ATC Level 4 Code
ATC4TERM	ATC4 Drug Class Term
CPREFCD	WHO_DRUG Drug Code
PDDREAS	Reason for Discontinuation
PDRESP	Response

How to do SDRG Automation

- An example for Data Conformance Summary report

Pinnacle 21 Validator Report						
Configuration: I:\pinnacle21-community_2016-08\components\config\SDTM 3.2 (CN).xml						
Define.xml: I:\pinnacle21-community_2016-08						
Generated: 2016-08-12T00:05:28						
CDISC CT Version: 2016-06-30						
MedDRA Version: 17.1						
UNII Version: 2016-07-06						
NDE-RT Version: 2016-07-05						
Software Version: 2.1.3						
Issue Summary						
Source	Pinnacle 21 ID	Publisher ID	Message	Severity	Fou	Comments
AE						
	CT2001	FDAC340	AESCAN value not found in 'No Yes Response' non-extensible codelist	Error	1191	This is acceptable.
	CT2001	FDAC340	AESCONG value not found in 'No Yes Response' non-extensible codelist	Error	1191	This is acceptable.
	CT2001	FDAC340	AESDISAB value not found in 'No Yes Response' non-extensible codelist	Error	1191	This is acceptable.
	CT2001	FDAC340	AESDTH value not found in 'No Yes Response' non-extensible codelist	Error	1191	This is acceptable.
	CT2001	FDAC340	AESER value not found in 'No Yes Response' non-extensible codelist	Error	1191	This is acceptable.
	CT2001	FDAC340	AESHOSP value not found in 'No Yes Response' non-extensible codelist	Error	1191	This is acceptable.
	CT2001	FDAC340	AESLIFE value not found in 'No Yes Response' non-extensible codelist	Error	1191	This is acceptable.
	CT2001	FDAC340	AESOD value not found in 'No Yes Response' non-extensible codelist	Error	1191	This is acceptable.
	SD0008	FDAC346	Value for AEDECOD not found in MedDRA dictionary	Error	1191	This is acceptable.
	SD1082	FDAC036	Variable length is too long for actual data	Error	18	This is acceptable.
	SD1090	FDAC131	Incorrect value for AESTDY variable	Error	1	This is acceptable.
Dataset Summary						
Issue Summary						
Details						
Rules						

How to do SDRG Automation

- Example for the output – Data conformance summary

4.Data Conformance Summary

4.1 Conformance Inputs

Was OpenCDISC used to evaluate conformance?

If yes, specify the versions of OpenCDISC and the OpenCDISC validation rules:

Pinnacle 21 Community 2.1.3, SDTM v3.2 (CN)

Were sponsor-defined validation rules used to evaluate conformance?

If yes, describe any significant sponsor-defined validation rules:

(Text here)

Were the SDTM datasets evaluated in relation to define.xml?

Was define.xml evaluated? |

Provide any additional compliance evaluation information:

(Text here)

4.2 Issues Summary

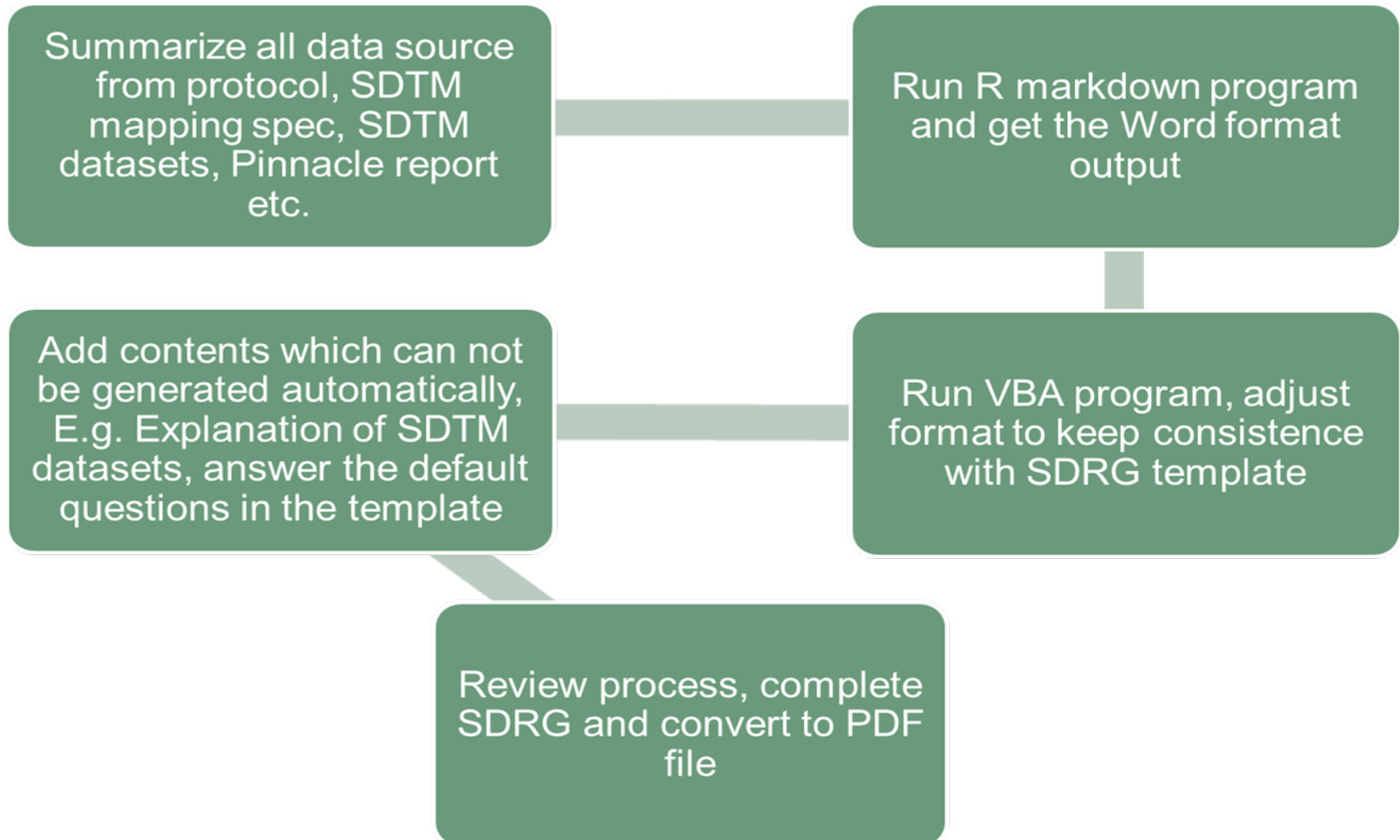
Dataset	Diagnostic Message	Severity	Count	Explanation
AE	AESCAN value not found in 'No Yes Response' non-extensible <u>codelist</u>	Error	1191	This is acceptable.
CM	Missing value for CMDOSU, when CMDOSE, CMDOSTXT or CMDOSTOT is provided	Error	80	This is acceptable.

How to do SDRG Automation

- R markdown limitations, for example:
 1. Header and footer.
 2. No page break, no section break.
 3. No tables' border.

.....
- Solution: Run VBA program and adjust format to keep the consistence with SDRG template.

How to do SDRG Automation



- R markdown can embed R code to handle the data source from protocol, SDTM mapping specification, SDTM datasets, pinnacle report etc. and output to Word document automatically which can save time for manual work, improve efficiency and reduce human error.
- SDRG Automation's extension — ADRG Automation.

Contact Information



Qiang Wang

dMed Biopharmaceutical Co., Ltd.

Email: qiang.wang@dmedglobal.com

Thank you



QA