

Regulatory Report Automation with R

PhUSE Single Day Event 2019

Shimeng Huang & Yuyao Song

Statistical Programming and Analysis, Roche Canada



Background - Faster Filing

- Optimize the Dossier- Content Standards and Reusability
- Report Automation - Take advantage of R and other Advanced Analytical Tools

Why RMarkdown and RShiny?



R Markdown provides an authoring framework for data scientists. You can use a single R Markdown file to both

- save and execute code
- generate high quality reports

R Markdown documents are fully reproducible and support dozens of static and dynamic output formats.

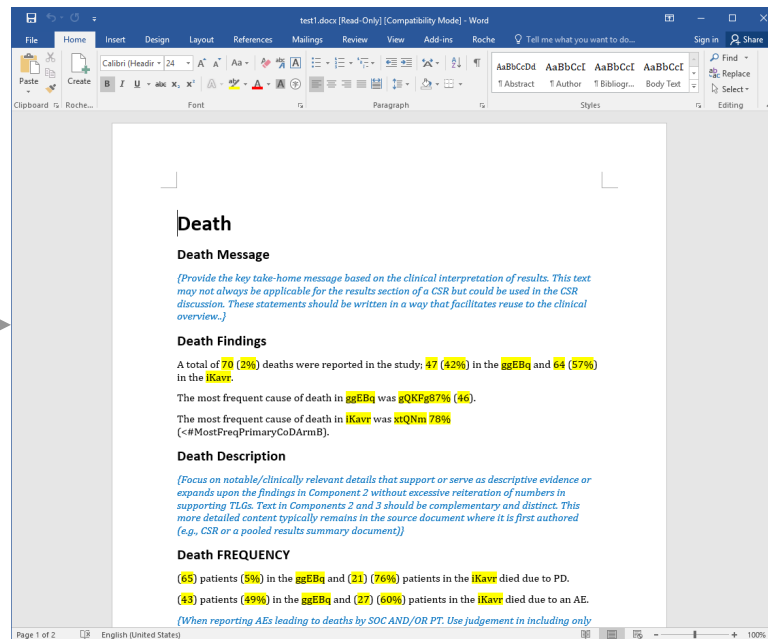
R Shiny provides an elegant and powerful web framework for building web applications using R. Shiny helps you turn your analyses into interactive web applications.

Development Workflow

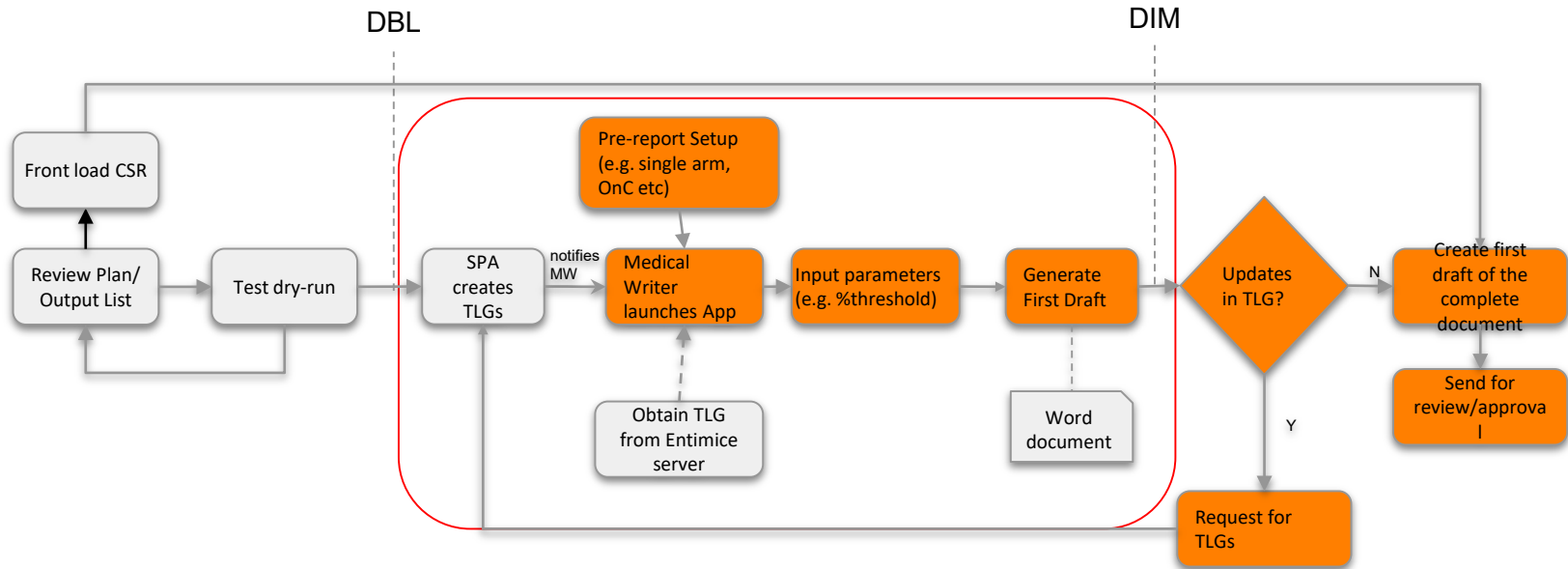
Template from Content Standards

Inputs from Medical Writer

Datasets from Statistical Programmer



Cross-Functional Workflow



DBL- Database Lock
 DIM- Data Interpretation Meeting
 SPA- Statistical Programming Analyst
 TLG- Table Listing Graphic

Report Template with Content Standards

The screenshot shows an Excel spreadsheet with multiple tabs. The active tab displays a table with columns for 'Object ID', 'Object Name', 'Object Type', 'Content Standard', 'Content Standard ID', 'Content Standard Description', and 'Content Standard Usage'. The table contains several rows of data, including 'Death', 'Death Findings', 'Death Description', 'Death FREQUENCY', 'Death RELATEDNESS', 'Death TIMING OF DEATH EVENTS', 'Death EVENTS DUE TO OTHER CAUSES', and 'Death Source'. Each row is associated with a specific content standard and its usage.

Template in Excel sheets

```
> print(tree_l4, "objName", "objtype", "usage")
```

levelName	objName	objtype	Usage
1 Root	Death	Cat - IP	RQD
2 --L5_1	Death Message	Comp	RQD
3 --L6_1		Para - AG	RQD
4 --L5_2	Death Findings	Comp	RQD
5 --L6_1		Para - AG	RQD
6 --L6_2		Para	CND
7 --L6_3		Para	CND
8 --L6_4		Para	CND
9 --L6_5		Para	CND
10 --L7_1		Para - AG	CND
11 --L7_2		Para	CND
12 --L7_3		Para	CND
13 --L7_4		Para	CND
14 --L7_5		Para	CND
15 --L5_3	Death Description	Comp	RQD
16 --L6_1		Para - AG	RQD
17 --L6_2	Death FREQUENCY	Sub-Comp	RQD
18 --L7_1		Para	CND
19 --L7_2		Para	CND
20 --L7_3		Para - AG	RQD
21 --L7_4		Para	RQD
22 --L6_3	Death RELATEDNESS	Sub-Comp	CND
23 --L7_1		Para	CND/OPT
24 --L7_2		Para	CND/OPT
25 --L6_4	Death TIMING OF DEATH EVENTS	Sub-Comp	CND
26 --L7_1		Para	CND
27 --L8_1		Para - AG	CND
28 --L7_2		Para	CND
29 --L8_1		Para - AG	CND
30 --L6_5	Death EVENTS DUE TO OTHER CAUSES	Sub-Comp	CND
31 --L7_1		Para	CND
32 --L5_4	Death Source	Comp	RQD
33 --L6_1		Para - AG	RQD
34 --L6_2		Para	RQD

Template as R objects (data tree)

Report Template with Content Standards (continue)

I	J	K	L	M
L6 Sub Comp Object ID (Primary Key)	Conditional Logic	Tech Logic for Condition	USAGE	Content+Variables
C.2.6.2.2.2			RQD	
C.2.6.2.2.2.1	If there were no deaths in the study	<#DeathStudy> is equal to 0	CND	No deaths were reported for this study.
C.2.6.2.2.2.2	If there are (>5) deaths in the study	<#DeathStudy> is greater than 5	CND	A total of <#DeathStudy> (<%DeathStudy>) deaths were reported in the study; <#DeathArmA> (<%DeathArmA>) in the <ArmA> and <#DeathArmB> (<%DeathArmB>) in the <ArmB>.
C.2.6.2.2.3	If the most frequent cause of death was different in study arms	<#MostFreqPrimaryCoDArmA> is not equal to <#MostFreqPrimaryCoDArmB>	CND	The most frequent cause of death in <ArmA> was <MostFreqPrimaryCoDArmA> (<%MostFreqPrimaryCoDArmA>). The most frequent cause of death in <ArmB> was <MostFreqPrimaryCoDArmB> (<%MostFreqPrimaryCoDArmB>).

Death

Death Message

(Provide the key take-home message based on the clinical interpretation of results. This text may not always be applicable for the results section of a CSR but could be used in the CSR discussion. These statements should be written in a way that facilitates reuse to the clinical overview.)

Death Findings

A total of 70 (29%) deaths were reported in the study; 47 (42%) in the ggEBq and 64 (57%) in the iKavv.

The most frequent cause of death in ggEBq was gQKp87% (46).

The most frequent cause of death in iKavv was stQNm 78%.

Death Description

(Focus on notable/clinically relevant details that support or serve as descriptive evidence or expands upon the findings in Component 2 without excessive reiteration of numbers in supporting TIGs. Text in Components 2 and 3 should be complementary and distinct. This more detailed content typically remains in the source document where it is first authored (e.g., CSR or a pooled results summary document))

Death FREQUENCY

(65) patients (5%) in the ggEBq and (21) (76%) patients in the iKavv died due to PD.

(43) patients (49%) in the ggEBq and (27) (60%) patients in the iKavv died due to an AE.

(When reporting AEs leading to deaths by SOC AND/OR PT. Use judgement in including only

Analysis Output Dataset

t_dtht01_AB_SE.out - Notepad

File Edit Format View Help

Deaths, Safety Patients

Protocol: xxxxx

	ARM A (N=292)	ARM B (N=822)	All (N=1114)
Total Number of Deaths	14 (4.8%)	50 (6.1%)	64 (5.7%)
Primary Cause of Death			
n	14	50	64
Adverse Event	6 (42.9%)	17 (34.0%)	23 (35.9%)
Progressive Disease	6 (42.9%)	28 (56.0%)	34 (53.1%)
Other	2 (14.3%)	5 (10.0%)	7 (10.9%)
COR PULMONARE	0	1 (2.0%)	1 (1.6%)
MASSIVE STROKE	0	1 (2.0%)	1 (1.6%)
PATIENT WAS HOSPITALIZED ON 25 APRIL 2018 DUE TO FEBRILE NEUTROPENIA AND DIED ON 28 APRIL DUE TO SEPTIC SHOCK	0	1 (2.0%)	1 (1.6%)
RESPIRATORY FAILURE DUE TO INTERSTITIAL PNEUMONIA (INDUCED BY DOCETAXEL CHEMOTHERAPY)	0	1 (2.0%)	1 (1.6%)
UNKNOWN	1 (7.1%)	1 (2.0%)	2 (3.1%)
Unknown	1 (7.1%)	0	1 (1.6%)
Days From Last Study Drug Administration			
n	14	50	64
<= 30 days	5 (35.7%)	21 (42.0%)	26 (40.6%)
> 30 days	9 (64.3%)	29 (58.0%)	38 (59.4%)
Primary Cause by Days from Last Study Drug Administration			
<= 30 days			
n	5	21	26
Adverse Event	0 (100%)	16 (76.2%)	21 (80.8%)
Progressive Disease	0	5 (23.8%)	5 (19.2%)
> 30 days			
n	9	29	38
Adverse Event	1 (11.1%)	1 (3.4%)	2 (5.3%)
Progressive Disease	6 (66.7%)	23 (79.3%)	29 (76.3%)
Other	2 (22.2%)	5 (17.2%)	7 (18.4%)

Program: root/clinical_studies/ro_testing/cdp_testing/spa_testing/data_analysis/FF/work/work_all/program/t_dtht01.sas
Output: root/clinical_studies/ro_testing/cdp_testing/spa_testing/data_analysis/FF/work/work_all/output/t_dtht01_AB_SE.out
21FEB2019 16:46

Page 1 of 1

SAS Universal Viewer - [c:\users\huangp\documents\work\auto_report_der_death\input\test\test_arm\adtht01\abresults.csv\Tabd]

File Tools Window Help

Address

Library TCHT01TABDERESULTS

Freeze Hide Show... Format Filter... Font... Find

	ANALYSIS_FOR	ANALYSIS_ORG	ARM	ARM_FORMAT	ARM_ORDER	_LEVEL1	_LEVEL1_FORM	_LEVEL1_ORDER	_LEVEL2	_LEVEL2_FORM	_LEVEL2_ORDER	_KEYWORDS	_STATUS	_CONTENT	_DESTINATION	_APPEND	_VALUE	_NUM_A
1	Total Number of	1	ARM A									am_sepatro	count	0	0		1 14	decimal
2	Total Number of	1	ARM B									am_sepatro	percent	0	0		1 4.0	decimal
3	Total Number of	1	ARM B									am_sepatro	count	0	0		1 50	decimal
4	Total Number of	1	ARM B									am_sepatro	percent	0	0		1 6.1	decimal
5	Total Number of	1	AB									am_sepatro	count	0	0		1 64	decimal
6	Total Number of	1	AB									am_sepatro	percent	0	0		1 5.7	decimal
7	Total Number of	1	ARM A	ARM AB(N=292)	1							am_sepatro	countpercent	table			1 14 (4.8%)	countpercent
8	Total Number of	1	ARM B	ARM B(N=822)	2							am_sepatro	countpercent	table			1 50 (6.1%)	countpercent
9	Total Number of	1	AB	AB(N=1114)	4							am_sepatro	countpercent	table			1 64 (5.7%)	countpercent
10	Primary Cause of	2	ARM A			Adverse Event	Adverse Event	1				am_sepatro	count	0	0		1 6	decimal
11	Primary Cause of	2	ARM A			Adverse Event	Adverse Event	1				am_sepatro	percent	0	0		1 42.9	decimal
12	Primary Cause of	2	ARM B			Adverse Event	Adverse Event	1				am_sepatro	count	0	0		1 17	decimal
13	Primary Cause of	2	ARM B			Adverse Event	Adverse Event	1				am_sepatro	percent	0	0		1 34.0	decimal
14	Primary Cause of	2	AB			Adverse Event	Adverse Event	1				am_sepatro	count	0	0		1 23	decimal
15	Primary Cause of	2	AB			Adverse Event	Adverse Event	1				am_sepatro	percent	0	0		1 35.9	decimal
16	Primary Cause of	2	ARM A			Progressive Disease	Progressive Disease	2				am_sepatro	count	0	0		1 6	decimal
17	Primary Cause of	2	ARM A			Progressive Disease	Progressive Disease	2				am_sepatro	percent	0	0		1 42.9	decimal
18	Primary Cause of	2	ARM B			Progressive Disease	Progressive Disease	2				am_sepatro	count	0	0		1 28	decimal
19	Primary Cause of	2	ARM B			Progressive Disease	Progressive Disease	2				am_sepatro	percent	0	0		1 34.0	decimal
20	Primary Cause of	2	AB			Progressive Disease	Progressive Disease	2				am_sepatro	count	0	0		1 34	decimal
21	Primary Cause of	2	AB			Progressive Disease	Progressive Disease	2				am_sepatro	percent	0	0		1 53.1	decimal
22	Primary Cause of	2	ARM A			Other	Other	3				am_sepatro	count	0	0		1 2	decimal
23	Primary Cause of	2	ARM A			Other	Other	3				am_sepatro	percent	0	0		1 3.4	decimal
24	Primary Cause of	2	ARM B			Other	Other	3				am_sepatro	count	0	0		1 5	decimal
25	Primary Cause of	2	ARM B			Other	Other	3				am_sepatro	percent	0	0		1 6.1	decimal
26	Primary Cause of	2	AB			Other	Other	3				am_sepatro	count	0	0		1 7	decimal
27	Primary Cause of	2	AB			Other	Other	3				am_sepatro	percent	0	0		1 10.9	decimal
28	Primary Cause of	2	ARM A			UNKNOWN	UNKNOWN	4				am_sepatro	count	0	0		1 1	decimal
29	Primary Cause of	2	ARM A			UNKNOWN	UNKNOWN	4				am_sepatro	percent	0	0		1 1.6	decimal
30	Primary Cause of	2	ARM A			Unknown	Unknown	4				am_sepatro	count	0	0		1 1	decimal
31	Primary Cause of	2	ARM A			Unknown	Unknown	4				am_sepatro	percent	0	0		1 1.6	decimal
32	Primary Cause of	2	ARM B			COR PULMONA	COR PULMONA	4				am_sepatro	count	0	0		1 1	decimal
33	Primary Cause of	2	ARM B			COR PULMONA	COR PULMONA	4				am_sepatro	percent	0	0		1 2.0	decimal
34	Primary Cause of	2	ARM B			MASSIVE STRO	MASSIVE STRO	4				am_sepatro	count	0	0		1 1	decimal
35	Primary Cause of	2	ARM B			MASSIVE STRO	MASSIVE STRO	4				am_sepatro	percent	0	0		1 2.0	decimal
36	Primary Cause of	2	ARM B			PATIENT WAS	PATIENT WAS	4				am_sepatro	count	0	0		1 1	decimal
37	Primary Cause of	2	ARM B			PATIENT WAS	PATIENT WAS	4				am_sepatro	percent	0	0		1 2.0	decimal
38	Primary Cause of	2	ARM B			RESPIRATORY	RESPIRATORY	4				am_sepatro	count	0	0		1 1	decimal

Rows 1-133 of 133 Filter Off Sort none

c:\users\huangp\documents\work\auto_report_der_death\input\test\test_arm\adtht01\abresults.csv\Tabd
initializing...

Analysis Output Dataset (continue)

__ANALYSIS_FORMAT	ARM Description of Planned Arm	ARM_FORMAT Description of Planned Arm (format)	__LEVEL1_FORMAT	__CONTENT	__DESTINATION	__VALUE
Total Number of Deaths	ARM A	ARM A@(N=292)		countpercent	table	14 (4.8%)
Total Number of Deaths	ARM B	ARM B@(N=822)		countpercent	table	50 (6.1%)
Total Number of Deaths	All	All@(N=1114)		countpercent	table	64 (5.7%)
Primary Cause of Death	ARM A	ARM A@(N=292)	n	count	table	14
Primary Cause of Death	ARM A	ARM A@(N=292)	Adverse Event	countpercent	table	6 (42.9%)
Primary Cause of Death	ARM A	ARM A@(N=292)	Progressive Disease			
Primary Cause of Death	ARM A	ARM A@(N=292)	Other			
Primary Cause of Death	ARM A	ARM A@(N=292)	COR PULMONARE			
Primary Cause of Death	ARM A	ARM A@(N=292)	MASSIVE STROKE			
Primary Cause of Death	ARM A	ARM A@(N=292)	PATIENT WAS HOSPITALIZED ON 25 APRIL 2018 DUE			
Primary Cause of Death	ARM A	ARM A@(N=292)	RESPIRATORY FAILURE DUE TO INTERSTITIAL PNEU			
Primary Cause of Death	ARM A	ARM A@(N=292)	UNKNOWN			
Primary Cause of Death	ARM A	ARM A@(N=292)	Unknown			
Primary Cause of Death	ARM B	ARM B@(N=822)	n			

t_dtht01_AB_SE.out - Notepad

File Edit Format View Help

Deaths, Safety Patients

Protocol: xxxxxx

	ARM A (N=292)	ARM B (N=822)	All (N=1114)
Total Number of Deaths	14 (4.8%)	50 (6.1%)	64 (5.7%)
Primary Cause of Death			
n	14	50	64
Adverse Event	6 (42.9%)	17 (34.0%)	23 (35.9%)
Progressive Disease	6 (42.9%)	28 (56.0%)	34 (53.1%)
Other	2 (14.3%)	5 (10.0%)	7 (10.9%)
COR PULMONARE	0	1 (2.0%)	1 (1.6%)
MASSIVE STROKE	0	1 (2.0%)	1 (1.6%)
PATIENT WAS HOSPITALIZED ON 25 APRIL 2018 DUE TO SEPTIC SHOCK	0	1 (2.0%)	1 (1.6%)
RESPIRATORY FAILURE DUE TO INTERSTITIAL PNEUMONIA (INDUCED BY DOCETAXEL CHEMOTHERAPY)	0	1 (2.0%)	1 (1.6%)
UNKNOWN	1 (7.1%)	1 (2.0%)	2 (3.1%)
Unknown	1 (7.1%)	0	1 (1.6%)
Days from Last Study Drug Administration			
n	14	50	64
<= 30 days	5 (35.7%)	21 (42.0%)	26 (40.6%)
> 30 days	9 (64.3%)	29 (58.0%)	38 (59.4%)
Primary Cause by Days from Last Study Drug Administration			
<= 30 days			
n	5	21	26
Adverse Event	5 (100%)	16 (76.2%)	21 (80.8%)
Progressive Disease	0	5 (23.8%)	5 (19.2%)
> 30 days			
n	9	29	38
Adverse Event	1 (11.1%)	1 (3.4%)	2 (5.3%)
Progressive Disease	6 (66.7%)	23 (79.3%)	29 (76.3%)
Other	2 (22.2%)	5 (17.2%)	7 (18.4%)

Program: root/clinical_studies/ro_testing/cdp_testing/spa_testing/data_analysis/FF/work/work_all/program/t_dtht01.sas
Output: root/clinical_studies/ro_testing/cdp_testing/spa_testing/data_analysis/FF/work/work_all/output/t_dtht01_AB_SE.out
21FEB2019 16:46 Page 1 of 1

App Structure

AutoCSR

Home (test)

Home

Content

Safety

Death

Adverse Events

Serious Adverse Events

Adverse Events of Interest

Study Population

Download

Test Data Information

Choose test dataset

Two-arm data

Click to access data:

Access Test Data

Select the name of the investigator arm

ARM A

Select the name of the comparison arm

ARM A

Developers

Shiming Huang: shiming.huang@roche.com

Yuyao Song: yuyao.song@roche.com

User Guide for Testing

- Choose a test case from the drop-down list
- Click Access Test Data
- Go to Content - Death
- Modify control panel selections, these can be changed at any time and report will update itself
- Go to Content - Demographic
- Modify control panel selections, these can be changed at any time and report will update itself
- Go to 'Download' and choose the section(s) and format(s) to download
- The above process should work for the two test cases
- User selections on Home page are saved, user selections on individual sections can be saved but under development

Message

Provide a message for the category

Deaths Interpretive

A total of 64 (8.7%) deaths were reported in this study. The number of deaths was [higher/lower/comparable] in ARM A compared with ARM B (14 (4.8%) and 50 (5.1%), respectively). The most frequent cause of death overall was Adverse Event and this was [higher/lower/comparable] in ARM A compared with ARM B (6 (42.9%) and 17 (34.0%), respectively). [The majority of Fewer] deaths occurred more than 30 days after Last Study Drug Administration (38 (59.4%).

Deaths Descriptive

In ARM A the most frequent 42.9% reason of death other than Progressive Disease was MYOCARDIAL INFARCTION (1 (3.3 %)). In ARM B the most frequent 34% reason of death was MYOCARDIAL INFARCTION (17 (34.0 %)). Of the 17 (34.0 %) patients who died due to AE, the most common reason of death was MYOCARDIAL INFARCTION (17 (34.0 %)). Further details for deaths due to adverse events are provided in Section XX.

The number of deaths related to study treatment was [higher/lower/comparable] in ARM A compared with ARM B (1 (0.3 %) and 3 (0.4 %), respectively).

Table/Listing/Graph

Deaths, Safety Patents

Protocol: xxxxx

	ARM A (N=259)	ARM B (N=1022)	All (N=1114)
Total Number of Deaths	14 (4.8%)	50 (5.1%)	64 (5.7%)
Primary Cause of Death			
n	14	50	64
Adverse Event	6 (42.9%)	17 (34.0%)	23 (35.9%)
Progressive Disease	6 (42.9%)	28 (56.0%)	34 (53.1%)
	7 (14.3%)	4 (8.0%)	7 (10.9%)

Threshold for small # of deaths (<)

5

Threshold for small % of deaths due to AE (<)

1 2 3 4 5 6 7 8 9 10

Fill in Variables

Match Variables

- Show none
- Show unbound
- Show bound
- Show all

Death

Death Message

(Provide the key take-home message based on the clinical interpretation of results. This text may not always be applicable for the results you are viewing in CSR but could be used in the CSR discussion. These statements should be written in a way that facilitates reuse to the clinical overview.)

Death Findings

A total of 14 (5.7%) deaths were reported in the study: 14 (4.8%) in the [ARM A] and 50 (5.1%) in the [ARM B].

The most frequent cause of death in [ARM A] was MYOCARDIAL INFARCTION (1 (3.3 %)).

The most frequent cause of death in [ARM B] was MYOCARDIAL INFARCTION (17 (34.0 %)).

(<#MostFreqPrimaryCoDArmB)

Death Description

(Focus on notable/clinically relevant details that support or serve as descriptive evidence or expands upon the findings in Component 2 without excessive reiteration of numbers in supporting TLEs. Text in Components 2 and 3 should be complementary and distinct. This more detailed content typically remains in the source document where it is first authored (e.g., CSR or a pooled results summary document))

Death FREQUENCY

(65) patients (5.8%) in the [ARM A] and (32) (7.9%) patients in the [ARM B] died due to PD.

(43) patients (4.9%) in the [ARM A] and (27) (6.9%) patients in the [ARM B] died due to an AE.

(When reporting AEs leading to deaths by SOC AND/OR PT, Use judgement in including only

AutoCSR

Home (test)

Home

Content

Safety

Death

Adverse Events

Serious Adverse Events

Adverse Events of Interest

Study Population

Download

Download Options

Sections to be downloaded

Death Demographic

Download file types

☒ .docx

☐ .xml

☐ .html

Click to download*:

Download

*Files for each section will be downloaded together in a zip file.

Summary of UI Functionalities

- Report Pre-specification
 - E.g., General Study Information, Number of arms, Investigator Arms, etc.
- Interactive customization with medical writer's inputs and judgement
 - E.g., Threshold for small number of deaths, threshold for common AE.
- Dynamic generation of downloadable reports in different formats
 - MS Word, HTML, and xml
 - Select Sections for downloading
- User inputs can be saved for next time

Demo



https://shiny.roche.com/3.5.1/users/huang69/auto_report-dev_death/

Apps programming company studywork tool other

AutoCSR



Logged in as huang69 Logout

Home (test)

Home

Content

Download

Study Information

Molecule #

RO9999999

Theme #

CDPT9999

Protocol #

CO999999

Path to data on entimICE

root/clinical_studies/RO9999999/CDPT9999/CO999999/data_analysis/...

Click to access data:

Access Data

Number of arms

One arm

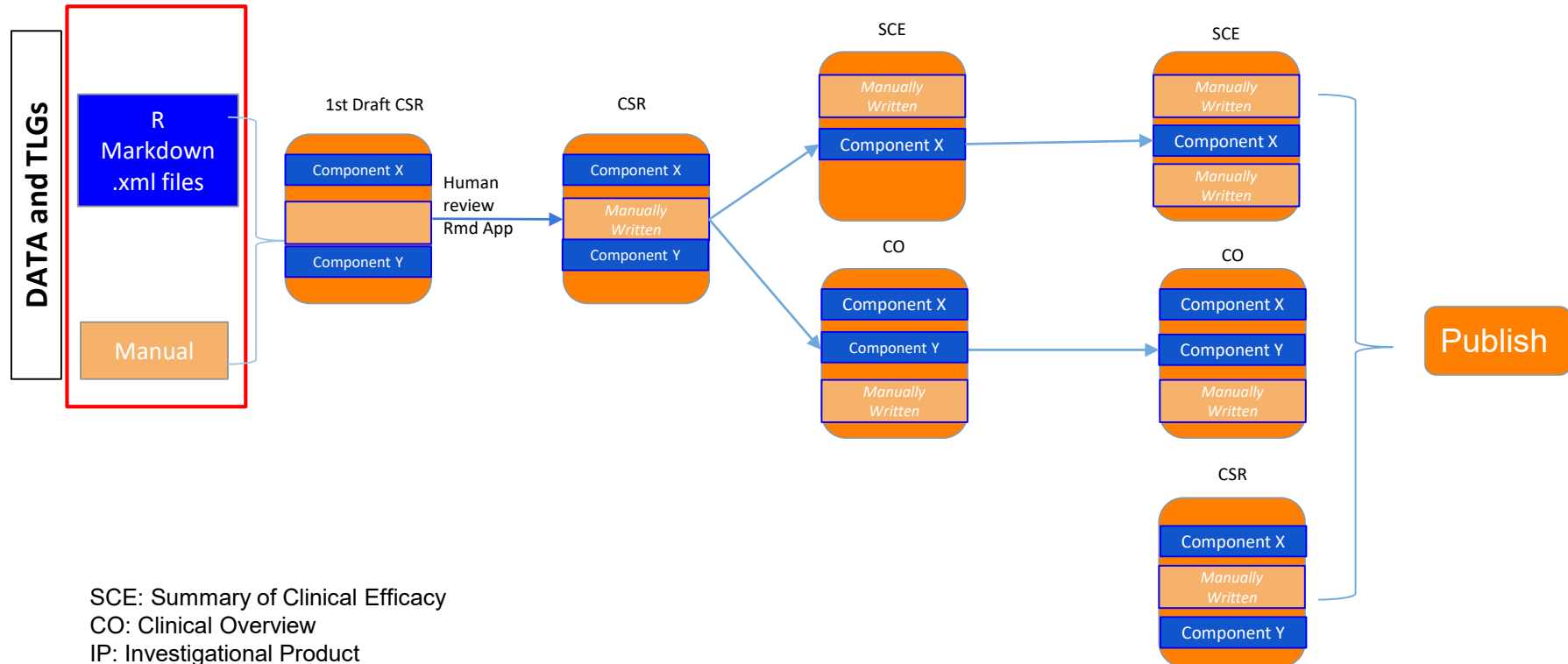
User Guide

** See Home (test) page for current usage **

Developers

Developers contact information [here](#).

Reusability of Content - Beyond CSR



SCE: Summary of Clinical Efficacy
CO: Clinical Overview
IP: Investigational Product

Challenges and Next Steps

- More sections
- Quality of the content standards spreadsheet
- Standardization of SAS output dataset structure. E.g., Consistent column names and orders
- Prevent numbers from being modified by users at adjudication
- MS Word add-in - launch the app inside Microsoft Word

Doing now what patients need next