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.

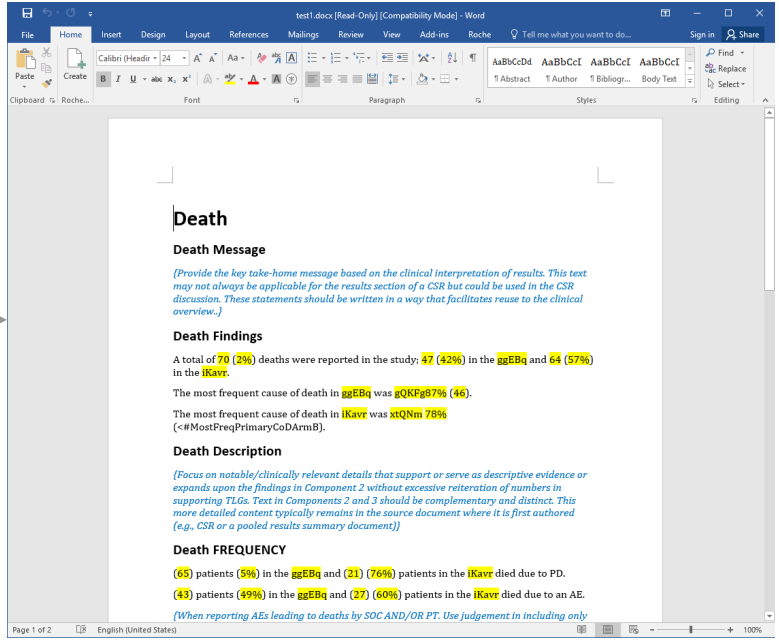
R Shiny provides an elegant and powerful web framework for building web applications using R. Shiny helps you turn your analyses into interactive web applications.

Development Workflow

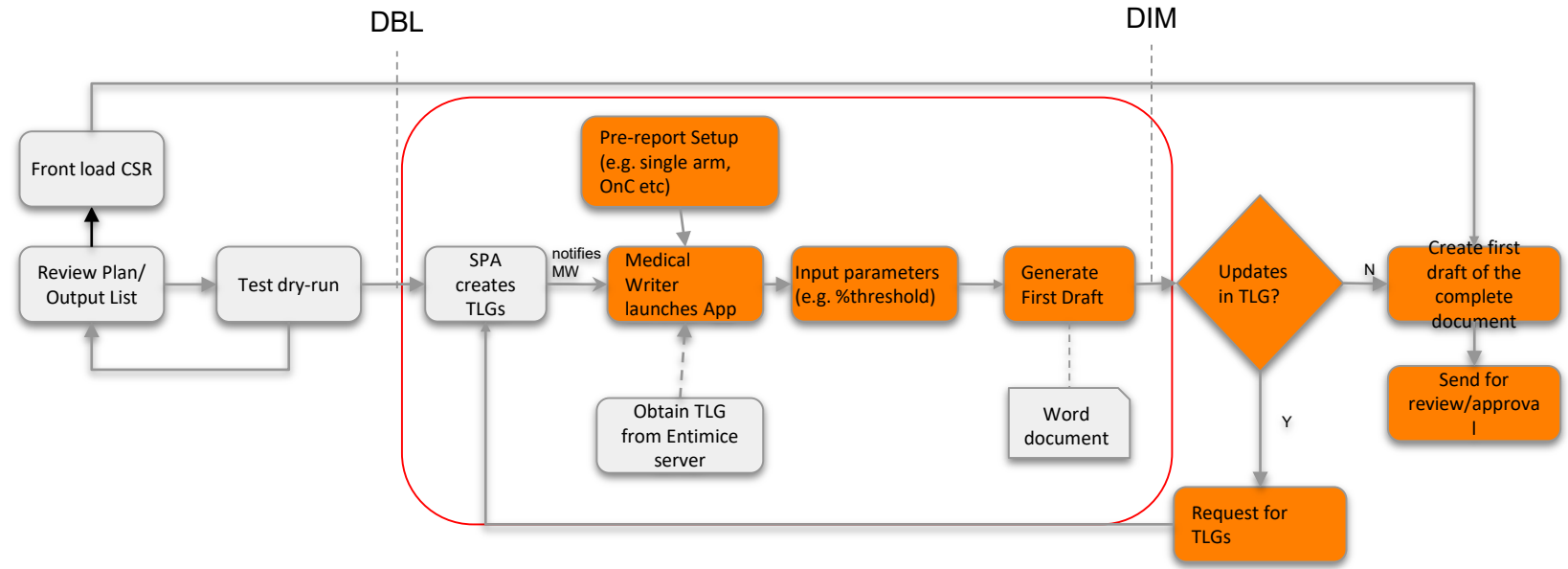
Template from Content Standards

Datasets from Statistical Programmer

Inputs from Medical Writer

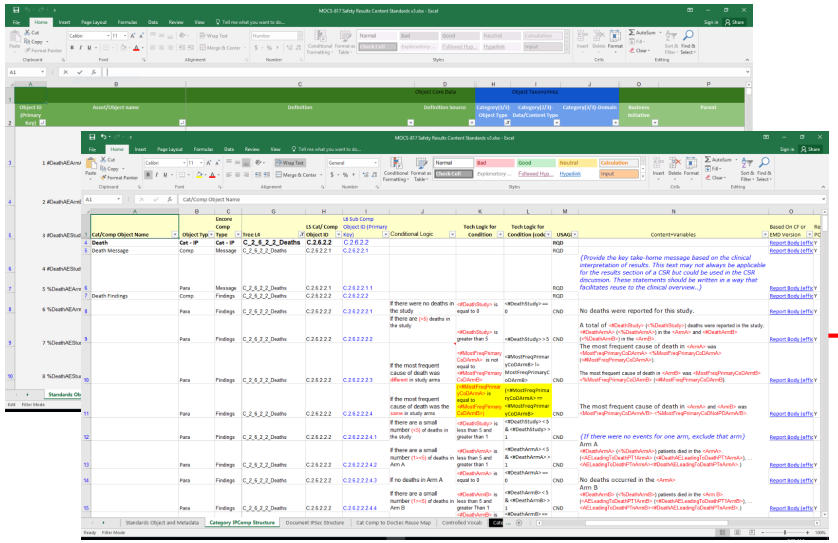


Cross-Functional Workflow



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

Report Template with Content Standards



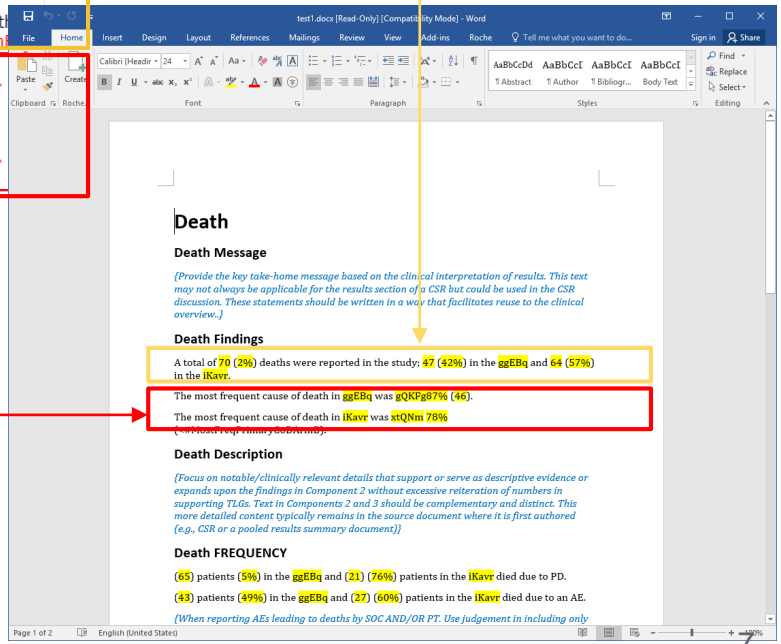
```
> print(tree_l4, "objName", "objType", "Usage")
                levelName                                  objName   objType   Usage
1  Root                                                      Death  Cat - IP     RQD
2   |--L5_1                                          Death Message      Comp     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 in Excel sheets

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 th <ArmA> and <#DeathArmB> (<%DeathArmB>) in the <Armi
C.2.6.2.2.2.3	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> (<#MostFreqPrimaryCoDArmA>). The most frequent cause of death in <ArmB> was <MostFreqPrimaryCoDArmB> <%MostFreqPrimaryCoDArmB> (<#MostFreqPrimaryCoDArmB>).



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 APRI 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	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

SAS Universal Viewer - [c:\users\huangs69\documents\work\auto_report-dev_death\input\test\two_arm\tdtht01absresults.sas7bdat]

File Tools Window Help

Address

Library TDTHT01ABSERESULTS

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

Table View

	__ANALYSIS	__ANALYSIS_FOR	__ANALYSIS_ORD	ARM	ARM_FORMAT	ARM_ORDER	__LEVEL1	__LEVEL1_FORM	__LEVEL1_ORDEI	__LEVEL2	__LEVEL2_FORM	__LEVEL2_ORDEI	__KEYVARS	__STATSVAR	__CONTENT	__DESTINATION	__REPEATNO	__VALUE	__NUM
1	dcau2	Total Number of ...	1	ARM A									arm __repeatno		count	qc	1	14	decimal
2	dcau2	Total Number of ...	1	ARM A									arm __repeatno		percent	qc	1	4.8	decimal
3	dcau2	Total Number of ...	1	ARM B									arm __repeatno		count	qc	1	50	decimal
4	dcau2	Total Number of ...	1	ARM B									arm __repeatno		percent	qc	1	6.1	decimal
5	dcau2	Total Number of ...	1	All									arm __repeatno		count	qc	1	64	decimal
6	dcau2	Total Number of ...	1	All									arm __repeatno		percent	qc	1	5.7	decimal
7	dcau2	Total Number of ...	1	ARM A	ARM A@(N=292)	1							arm __repeatno		countpercent	table	1	14 (4.8%)	countperc
8	dcau2	Total Number of ...	1	ARM B	ARM B@(N=822)	2							arm __repeatno		countpercent	table	1	50 (6.1%)	countperc
9	dcau2	Total Number of ...	1	All	All@(N=1114)	4							arm __repeatno		countpercent	table	1	64 (5.7%)	countperc
10	dcau	Primary Cause of ...	2	ARM A			Adverse Event	Adverse Event	1				arm __repeatno		count	qc	1	6	decimal
11	dcau	Primary Cause of ...	2	ARM A			Adverse Event	Adverse Event	1				arm __repeatno		percent	qc	1	42.9	decimal
12	dcau	Primary Cause of ...	2	ARM B			Adverse Event	Adverse Event	1				arm __repeatno		count	qc	1	17	decimal
13	dcau	Primary Cause of ...	2	ARM B			Adverse Event	Adverse Event	1				arm __repeatno		percent	qc	1	34.0	decimal
14	dcau	Primary Cause of ...	2	All			Adverse Event	Adverse Event	1				arm __repeatno		count	qc	1	23	decimal
15	dcau	Primary Cause of ...	2	All			Adverse Event	Adverse Event	1				arm __repeatno		percent	qc	1	35.9	decimal
16	dcau	Primary Cause of ...	2	ARM A			Progressive Dise...	Progressive Dise...	2				arm __repeatno		count	qc	1	6	decimal
17	dcau	Primary Cause of ...	2	ARM A			Progressive Dise...	Progressive Dise...	2				arm __repeatno		percent	qc	1	42.9	decimal
18	dcau	Primary Cause of ...	2	ARM B			Progressive Dise...	Progressive Dise...	2				arm __repeatno		count	qc	1	28	decimal
19	dcau	Primary Cause of ...	2	ARM B			Progressive Dise...	Progressive Dise...	2				arm __repeatno		percent	qc	1	56.0	decimal
20	dcau	Primary Cause of ...	2	All			Progressive Dise...	Progressive Dise...	2				arm __repeatno		count	qc	1	34	decimal
21	dcau	Primary Cause of ...	2	All			Progressive Dise...	Progressive Dise...	2				arm __repeatno		percent	qc	1	53.1	decimal
22	dcau	Primary Cause of ...	2	ARM A			Other	Other	3				arm __repeatno		count	qc	1	2	decimal
23	dcau	Primary Cause of ...	2	ARM A			Other	Other	3				arm __repeatno		percent	qc	1	14.3	decimal
24	dcau	Primary Cause of ...	2	ARM B			Other	Other	3				arm __repeatno		count	qc	1	5	decimal
25	dcau	Primary Cause of ...	2	ARM B			Other	Other	3				arm __repeatno		percent	qc	1	10.0	decimal
26	dcau	Primary Cause of ...	2	All			Other	Other	3				arm __repeatno		count	qc	1	7	decimal
27	dcau	Primary Cause of ...	2	All			Other	Other	3				arm __repeatno		percent	qc	1	10.9	decimal
28	dcau	Primary Cause of ...	2	ARM A			UNKNOWN	UNKNOWN	4				arm __repeatno		count	qc	1	1	decimal
29	dcau	Primary Cause of ...	2	ARM A			UNKNOWN	UNKNOWN	4				arm __repeatno		percent	qc	1	7.1	decimal
30	dcau	Primary Cause of ...	2	ARM A			Unknown	Unknown	4				arm __repeatno		count	qc	1	1	decimal
31	dcau	Primary Cause of ...	2	ARM A			Unknown	Unknown	4				arm __repeatno		percent	qc	1	7.1	decimal
32	dcau	Primary Cause of ...	2	ARM B			COR PULMONA...	COR PULMONA...	4				arm __repeatno		count	qc	1	1	decimal
33	dcau	Primary Cause of ...	2	ARM B			COR PULMONA...	COR PULMONA...	4				arm __repeatno		percent	qc	1	2.0	decimal
34	dcau	Primary Cause of ...	2	ARM B			MASSIVE STRO...	MASSIVE STRO...	4				arm __repeatno		count	qc	1	1	decimal
35	dcau	Primary Cause of ...	2	ARM B			MASSIVE STRO...	MASSIVE STRO...	4				arm __repeatno		percent	qc	1	2.0	decimal
36	dcau	Primary Cause of ...	2	ARM B			PATIENT WAS ...	PATIENT WAS ...	4				arm __repeatno		count	qc	1	1	decimal
37	dcau	Primary Cause of ...	2	ARM B			PATIENT WAS ...	PATIENT WAS ...	4				arm __repeatno		percent	qc	1	2.0	decimal
38	dcau	Primary Cause of ...	2	ARM B			RESPIRATORY ...	RESPIRATORY ...	4				arm __repeatno		count	qc	1	1	decimal

Rows 1-153 of 153 Filter Off Sort: none 1,1

c:\users\huangs69\documents\work\auto_report-dev_death\input\test\two_arm\tdtht01absresults.sas7bdat

Initializing...

Analysis Output Dataset (continue)

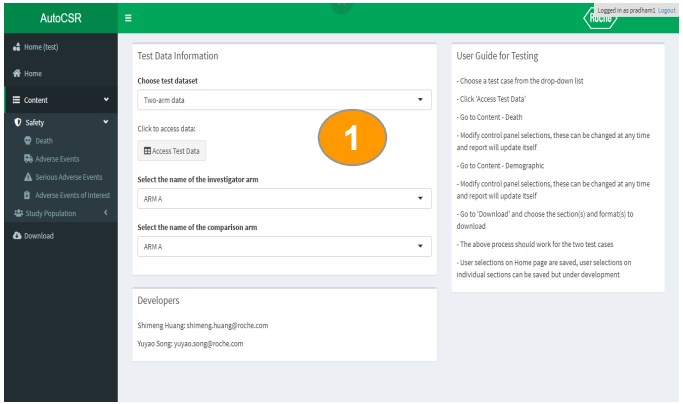
_ANALYSIS_FORMAT	ARM Description of Planned Arm	ARM_FORMAT Description of Planned Arm (format)	_LEVEL1_FORMAT	_CONTENT	_DESTINATION	_VALUE
Total Number of Deaths	ARM A	ARM A@(N=292)		countpercent	table	14 (4.8%)
Total Number of Deaths	ARM B	ARM B@(N=822)		countpercent	table	50 (6.1%)
Total Number of 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 APRI 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			

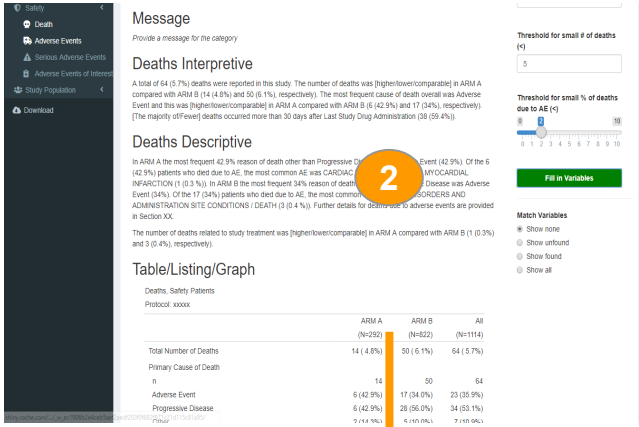
t_dtht01_AB_SE.out - Notepad

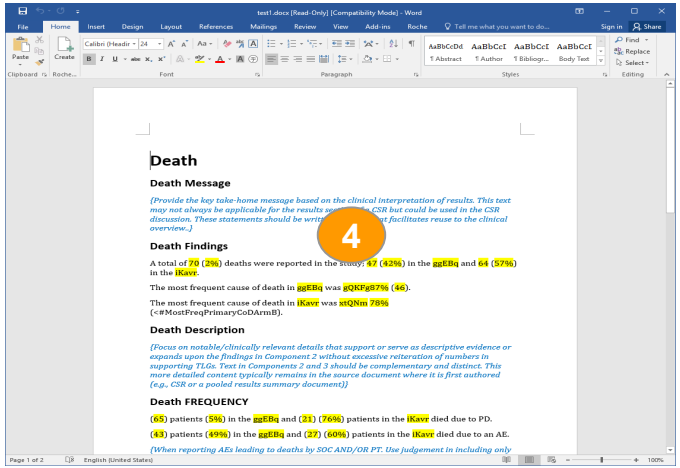
File Edit Format View Help

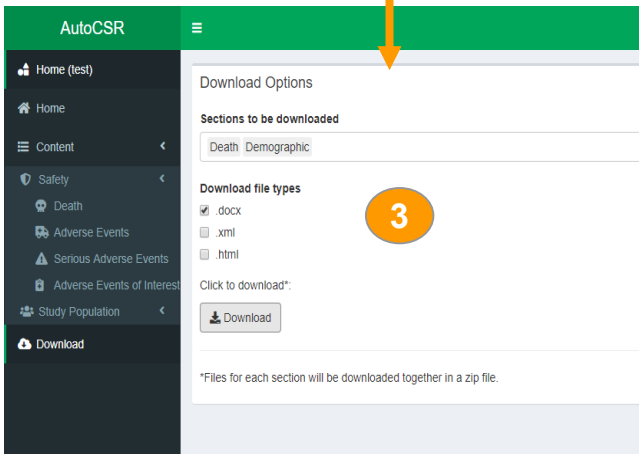
```
Deaths, Safety Patients
Protocol: xxxxx

                                                                                                ARM A        ARM B        All
                                                                                              (N=292)      (N=822)     (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 APRI 2018 DUE TO FEBRILE NEUTROPENIA AND DIED ON 28 APRIL DUE 0        1 ( 2.0%)    1 ( 1.6%)
  TO SEPTIC SHOCK
  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





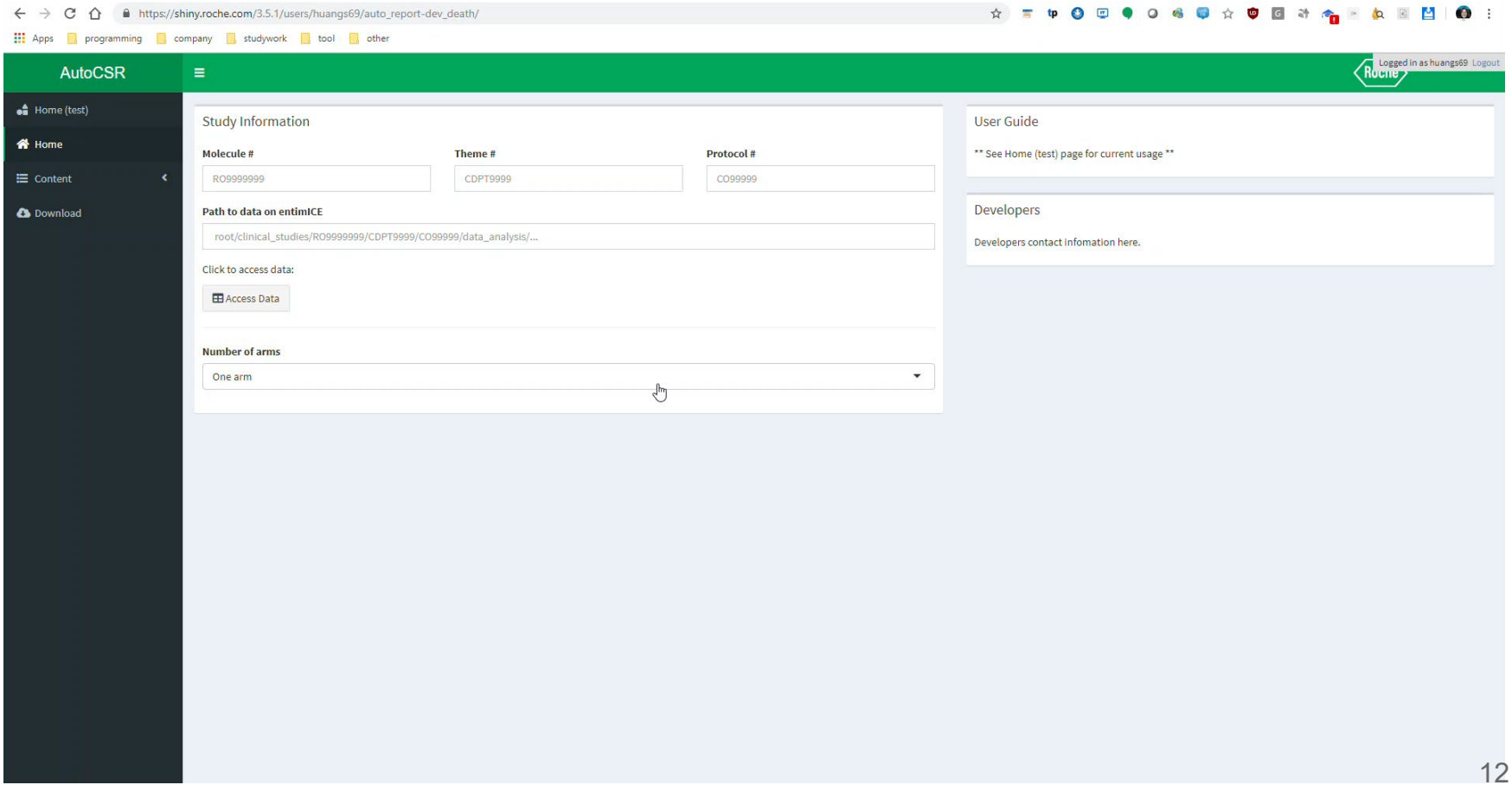




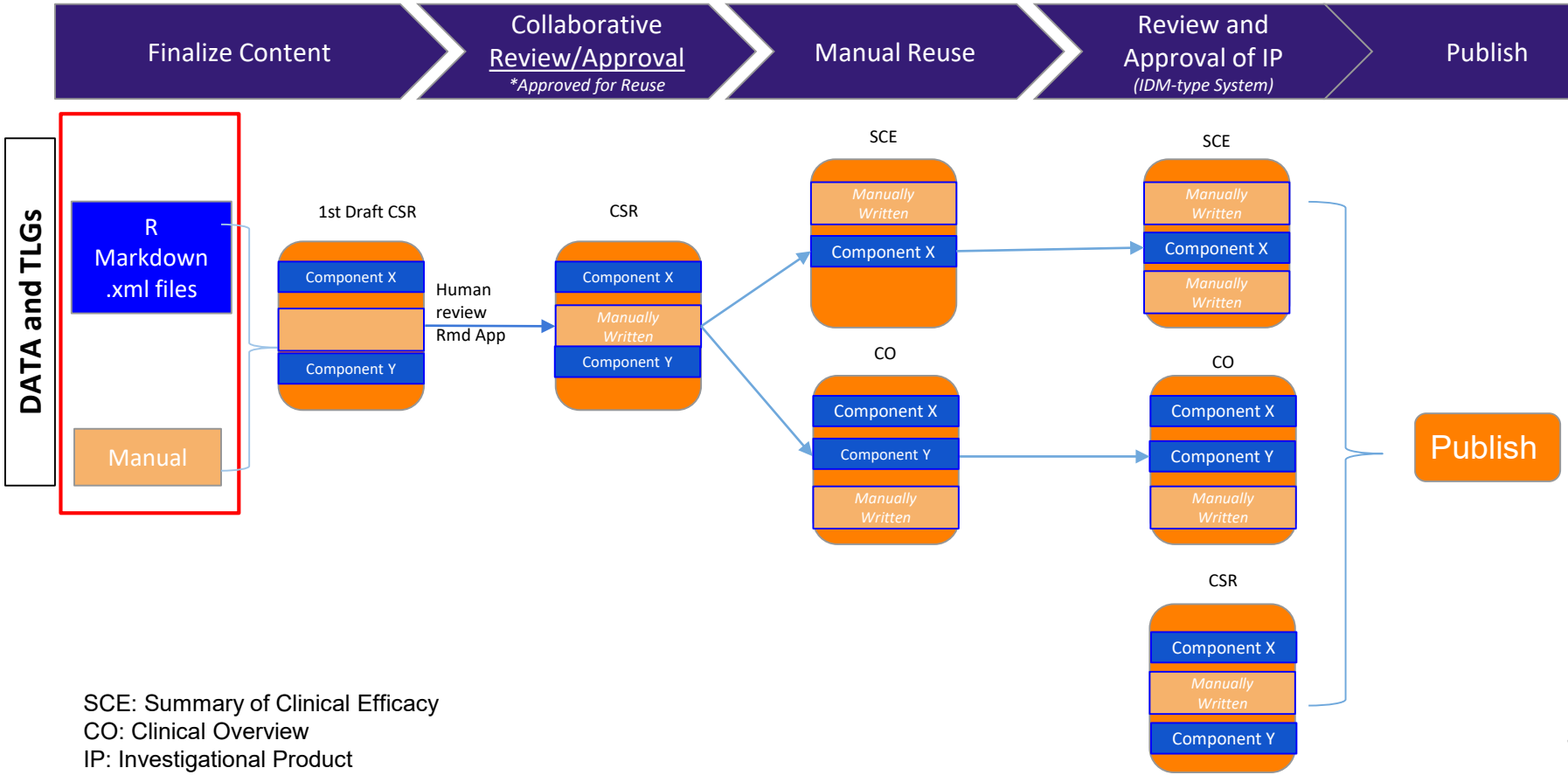
Summary of UI Functionalities

- 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/huangs69/auto_report-dev_death/
Apps
programming
company
studywork
tool
other
AutoCSR
Logged in as huangs69 Logout
Home (test)
Home
Content
Download
Study Information
Molecule #
RO9999999
Theme #
CDPT9999
Protocol #
CO99999
Path to data on entimICE
root/clinical_studies/RO9999999/CDPT9999/CO99999/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 infomation 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