PharmaSUG Paper DV-113

A Dynamic Data Visualization Report for Efficacy Endpoints of Early Oncology Study

Lily Zhang, Yixin Ren, Huei-Ling Chen
Merck & Co., Inc., Kenilworth, NJ, USA

ABSTRACT

In early oncology studies with solid tumors, efficacy analysis usually focuses on tumor size change and the best overall response of the drug treatment effect. Clinical clinicians and statisticians analyze the drug effect based on the summarized results and individual patient-level analyses. When they have questions on certain data points, they prefer more detailed individual patient data to understand the data better. The clinical programmers often work with the clinicians and statisticians in the data research effort by providing requested data. This paper presents a macro tool that creates a dynamic visualization specifically for oncology studies with solid tumors to avoid repeated programming efforts to answer requests from clinicians and statisticians. This macro tool connects the patient's solid tumor size, response, and any critical events in the study to produce patient-level listing worksheets in excel file format. This report is well received by clinicians, statisticians, and programmers due to significantly reduced communication and coding time. This paper will discuss the details of this tool and the key syntax.

KEYWORDS

Early Oncology, Solid Tumor, RECIST 1.1, iRECIST, Dynamic Report, ODS TAGSETS EXCELXP, Hyperlink

INTRODUCTION

The efficacy analysis of oncology studies involves various endpoints from multiple domains, such as TU, TR, and RS. These endpoints' results and the date-time variables from multiple domains that we need to carefully evaluate to ensure efficacy data consistency. This paper provides a dynamic visualization specifically for solid tumors in early oncology studies. The report presents each patient's solid tumor evaluation results and any related important events in the study.

RECIST (Response Evaluation Criteria in Solid Tumours) 1.1 and iRECIST are well-established in oncology solid tumor clinical trials, and we already have the standardized data collection and SDTM data domains.

There are three related domains for tumor assessment and disease response: Tumor Identification (TU), Tumor Results (TR), and Disease Response (RS). The detailed introduction of these domains is in section 6.3 Findings of ***CDISC SDTM Implementation Guide Version 3.2***.

TU: Tumor Identification

The TU domain represents the basic data for every individual tumor identified, existing or new. Under the RECIST criteria for solid tumors, the tumor lesions are classified as the target lesions, non-target lesions, and new lesions. Each identified tumor has a tumor ID in variable TULNKID, and the label of this variable is 'Link ID'. This variable is the key that connects the data in the TR domain. Tumor lesions are evaluated repeatedly at scheduled visits and potential non-scheduled visits such as discontinuation visits.

TU – Examples for Tumor Identification Domain Model

TU Example 1

Row	STUDYID	DOMAIN	USUBJID	TUSEQ	TUGRPID	TULNKID	TUTESTCD	TUTEST	TUORRES	TUSTRESC	TULOC	TULAT
1	ABC	TU	44444	1		T01	TUMIDENT	Tumor Identification	TARGET	TARGET	LIVER	
2	ABC	TU	44444	2		T02	TUMIDENT	Tumor Identification	TARGET	TARGET	KIDNEY	RIGHT
3	ABC	TU	44444	3		T03	TUMIDENT	Tumor Identification	TARGET	TARGET	CERVICAL LYMPH NODE	LEFT
4	ABC	TU	44444	4		T04	TUMIDENT	Tumor Identification	TARGET	TARGET	SKIN OF THE TRUNK	
5	ABC	TU	44444	5		NT01	TUMIDENT	Tumor Identification	NON-TARGET	NON-TARGET	THYROID GLAND	RIGHT
6	ABC	TU	44444	6		NT02	TUMIDENT	Tumor Identification	NON-TARGET	NON-TARGET	CEREBELLUM	RIGHT
7	ABC	TU	44444	7	T04	T04.1	TUSPLIT	Tumor Split	TARGET	TARGET	SKIN OF THE TRUNK	
8	ABC	TU	44444	8	T04	T04.2	TUSPLIT	Tumor Split	TARGET	TARGET	SKIN OF THE TRUNK	

Row	TUMETHOD	TUEVAL	VISITNUM	VISIT	TUDTC	TUDY
1 (cont)	CT SCAN	INVESTIGATOR	10	SCREEN	2010-01-01	-3
2 (cont)	CT SCAN	INVESTIGATOR	10	SCREEN	2010-01-01	-3
3 (cont)	MRI	INVESTIGATOR	10	SCREEN	2010-01-02	-2
4 (cont)	PHOTOGRAPHY	INVESTIGATOR	10	SCREEN	2010-01-03	-1
5 (cont)	CT SCAN	INVESTIGATOR	10	SCREEN	2010-01-01	-3
6 (cont)	MRI	INVESTIGATOR	10	SCREEN	2010-01-02	-2
7 (cont)	PHOTOGRAPHY	INVESTIGATOR	40	WEEK 6	2010-02-20	48
8 (cont)	PHOTOGRAPHY	INVESTIGATOR	40	WEEK 6	2010-02-20	48

SUPPTU for Example 1:

Row	STUDYID	RDOMAIN	USUBJID	IDVAL	IDVARVAL	QNAM	QLABEL	QVAL	QORIG
1	ABC	TU	44444	TUSEQ	1	PREVIR	Previously Irradiated	N	CRF
2	ABC	TU	44444	TUSEQ	2	PREVIR	Previously Irradiated	N	CRF
3	ABC	TU	44444	TUSEQ	3	PREVIR	Previously Irradiated	Y	CRF
4	ABC	TU	44444	TUSEQ	3	PREVIRP	Irradiated then Subsequent Progression	Y	CRF

Display 1: TU Example 1 from "*CDISC SDTM Implementation Guide Version 3.2*"

TR: Tumor Results

The TR domain represents the evaluation result of each tumor lesion identified in the TU domain. The result is either in quantitative measurements or qualitative assessments of the tumors. The target lesion size must be measurable and thus would be represented quantitatively.

TR Example 1

Row	STUDYID	DOMAIN	USUBJID	TRSEQ	TRGRPID	TRLNKGRP	TRLNKID	TRTESTCD	TRTEST	TRORRES	TRORRESU
1	ABC	TR	44444	1	TARGET	A1	T01	DIAMETER	Diameter	17	mm
2	ABC	TR	44444	2	TARGET	A1	T02	DIAMETER	Diameter	16	mm
3	ABC	TR	44444	3	TARGET	A1	T03	DIAMETER	Diameter	15	mm
4	ABC	TR	44444	4	TARGET	A1	T04	DIAMETER	Diameter	14	mm
5	ABC	TR	44444	5	TARGET	A1		SUMDIAM	Sum of Diameter	62	mm
6	ABC	TR	44444	6	TARGET	A1		SUMNLNLD	Sum Diameters of Non Lymph Node Tumors	47	mm
7	ABC	TR	44444	7	NON-TARGET	A1	NT01	TUMSTATE	Tumor State	PRESENT	
8	ABC	TR	44444	8	NON-TARGET	A1	NT02	TUMSTATE	Tumor State	PRESENT	
9	ABC	TR	44444	9	TARGET	A2	T01	DIAMETER	Diameter	0	mm
10	ABC	TR	44444	10	TARGET	A2	T02	DIAMETER	Diameter	TOO SMALL TO MEASURE	mm
11	ABC	TR	44444	11	TARGET	A2	T03	DIAMETER	Diameter	12	mm
12	ABC	TR	44444	12	TARGET	A2	T04	DIAMETER	Diameter		
13	ABC	TR	44444	13	TARGET	A2	T04.1	DIAMETER	Diameter	6	mm
14	ABC	TR	44444	14	TARGET	A2	T04.2	DIAMETER	Diameter	7	mm

Row	TRSTRESC	TRSTRESN	TRSTRESU	TRSTAT	TRREASND	TRMETHOD	TREVAL	VISITNUM	VISIT	TRDTC	TRDY
1 (cont)	17	17	mm			CT SCAN	INVESTIGATOR	10	SCREEN	2010-01-01	-3
2 (cont)	16	16	mm			CT SCAN	INVESTIGATOR	10	SCREEN	2010-01-01	-3
3 (cont)	15	15	mm			MRI	INVESTIGATOR	10	SCREEN	2010-01-02	-2
4 (cont)	14	14	mm			PHOTOGRAPHY	INVESTIGATOR	10	SCREEN	2010-01-03	-1
5 (cont)	62	62	mm				INVESTIGATOR	10	SCREEN		
6 (cont)	47	47	mm				INVESTIGATOR	10	SCREEN		
7 (cont)	PRESENT					CT SCAN	INVESTIGATOR	10	SCREEN	2010-01-01	-3

Display 2: TR Example 1 from "*CDISC SDTM Implementation Guide Version 3.2*"

RS: Disease Response

The RS domain represents the response results reported based on the data in the TR domain. At each visit, the sum of all target lesion sizes is calculated and compared with the baseline and nadir (the smallest sum before this visit). The response evaluation in each visit is summarized in the RS domain.

RS – Examples for Disease Response Domain Model

RS Example 1

Row	STUDYID	DOMAIN	USUBJID	RSSEQ	RSLNKGRP	RSTESTCD	RSTEST	RSCAT	RSORRES	RSSTRESC	RSSTAT
1	ABC	RS	44444	1		TRGRESP	Target Response	RECIST 1.1	PR	PR	
2	ABC	RS	44444	2		NTRGRESP	Non-target Response	RECIST 1.1	SD	SD	
3	ABC	RS	44444	3	A2	OVRLRESP	Overall Response	RECIST 1.1	PR	PR	
4	ABC	RS	44444	4		TRGRESP	Target Response	RECIST 1.1			NOT DONE
5	ABC	RS	44444	5		NTRGRESP	Non-target Response	RECIST 1.1			NOT DONE
6	ABC	RS	44444	6		NRADPROG	Non-Radiological Progression	CLINICAL ASSESSMENT	Pleural Effusion	PD	
7	ABC	RS	44444	7	A3	OVRLRESP	Overall Response	CLINICAL ASSESSMENT	PD	PD	

Row	RSREASND	RSEVAL	VISITNUM	VISIT	RSDTC	RSDY
1 (cont)		INVESTIGATOR	40	WEEK 6	2010-02-18	46
2 (cont)		INVESTIGATOR	40	WEEK 6	2010-02-18	46
3 (cont)		INVESTIGATOR	40	WEEK 6	2010-02-18	46
4 (cont)	All targets not assessed	INVESTIGATOR	60	WEEK 12	2010-04-02	88
5 (cont)	Non-targets not assessed	INVESTIGATOR	60	WEEK 12	2010-04-02	88
6 (cont)		INVESTIGATOR	60	WEEK 12	2010-04-02	88
7 (cont)		INVESTIGATOR	60	WEEK 12	2010-04-02	88

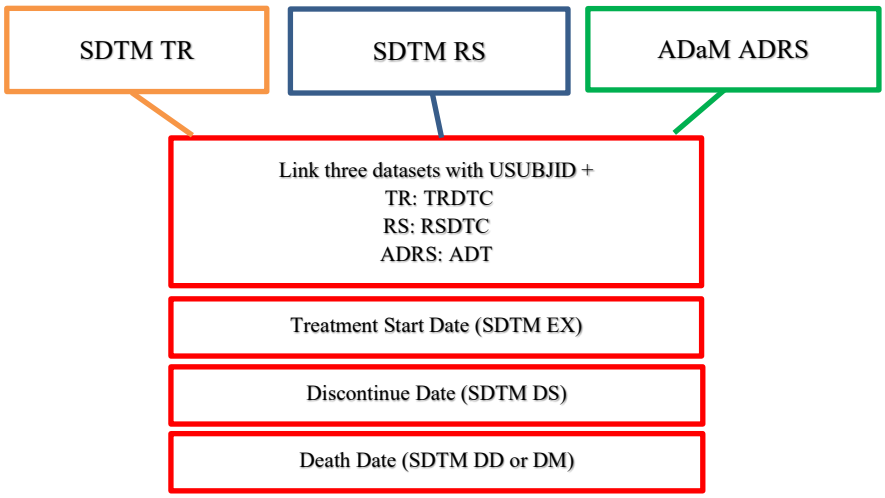
Display 3: RS Example 1 from ""*CDISC SDTM Implementation Guide Version 3.2*""

SOME INTERESTED SCENARIOS

It is important to cross-check the SDTM domains to ensure the following issues won't happen.

Scenarios	Details
Incomplete lesion measurement data at subsequent visits.	There is no missing lesion evaluation at any timepoint in the TR or TU domains. For example, lesion ID 1 in visit 1, visit 3, but missing in visit 2.
Lesion measurement results are collected at different dates for the same visit.	The presence of multiple records of lesion measurement data may not be a data issue. The reason is that the lesion evaluation can be done separately for the same batch of the tumor sample. It is also possible that the tumor samples are taken within a close time window and should be treated as a complete set of samples.
Efficacy data is collected after a patient is discontinued from the study.	This is not necessarily a data issue, yet it is still worth checking further.
Efficacy data is collected after a patient death.	This is a data issue.
Lesion measurement data and overall response results are consistent.	The overall response evaluations from the RS domain should be consistent with the lesion measurement results from the TR domain.

DATA PREPARATION



TR Domain - Transpose dataset from vertical to horizontal

USUBJID	TRTESTCD	TRDTC	TRLNKID	TRSTRESN	VISIT
1234-180001	LPERP	2015-12-24	TN1	29	Screening
1234-180001	LPERP	2016-04-07	TN1	14	Cycle 2 Day 22
1234-180001	LPERP	2016-05-20	TN1	14	Cycle 3 Day 22
1234-180001	LPERP	2016-07-01	TN1	14	Cycle 4 Day 22
1234-180001	LPERP	2016-08-12	TN1	12	Cycle 5 Day 22
1234-180001	LDIAM	2015-12-24	TEN2	21	Screening
1234-180001	LDIAM	2016-05-20	TEN2	2	Cycle 3 Day 22
1234-180001	LDIAM	2016-07-01	TEN2	2	Cycle 4 Day 22
1234-180001	LDIAM	2016-08-12	TEN2	2	Cycle 5 Day 22
1234-180001	LDIAM	2015-12-24	TEN3	18	Screening
1234-180001	LDIAM	2016-04-07	TEN3	7	Cycle 2 Day 22
1234-180001	LDIAM	2016-05-20	TEN3	7	Cycle 3 Day 22
1234-180001	LDIAM	2016-07-01	TEN3	7	Cycle 4 Day 22
1234-180001	LDIAM	2016-08-12	TEN3	7	Cycle 5 Day 22
1234-180001	LDIAM	2015-12-24	TEN4	27	Screening
1234-180001	LDIAM	2016-04-07	TEN4	25	Cycle 2 Day 22
1234-180001	LDIAM	2016-05-20	TEN4	16	Cycle 3 Day 22
1234-180001	LDIAM	2016-07-01	TEN4	16	Cycle 4 Day 22
1234-180001	LDIAM	2016-08-12	TEN4	14	Cycle 5 Day 22

Display 4: TR data in vertical shape

Use PROC TRANSPOSE to transpose the vertical shape data to horizontal data. The data is transformed to one record per USUBJID, TRDTC, and VISIT. The TRLNKID is transferred to the column name, and the lesion size, TRSTRESC, is the data value.

```
proc transpose data=sdtmdata out=tempdata ;
      by usubjid trdtc visit;
      var trstresc;
      id trlnkid;
run;
```

USUBJID	TRDTC	VISIT	TN1	TEN2	TEN3	TEN4
1234-180001	2015-12-24	Screening	29	21	18	27
1234-180001	2016-04-07	Cycle 2 Day 22	14		7	25
1234-180001	2016-05-20	Cycle 3 Day 22	14	2	7	16
1234-180001	2016-07-01	Cycle 4 Day 22	14	2	7	16
1234-180001	2016-08-12	Cycle 5 Day 22	12	2	7	14

Display 5: TR data in horizontal shape

The horizontal data helps eyeball check the incomplete lesion data. From Display 5, The TRLNKID value 'TEN2' has a missing record in the Visit 'Cycle 2 Day 22' batch.

RS Domain - Take Overall Response result from the dataset

USUBJID	RSTESTCD	RSCAT	RSDTC	RSSTRESC	VISIT
1234-180001	OVRLRESP	RECIST 1.1	2016-04-07	PR	Cycle 2 Day 22
1234-180001	OVRLRESP	RECIST 1.1	2016-05-20	PR	Cycle 3 Day 22
1234-180001	OVRLRESP	RECIST 1.1	2016-07-01	PR	Cycle 4 Day 22
1234-180001	OVRLRESP	RECIST 1.1	2016-08-12	PR	Cycle 5 Day 22

Display 6: RS data

For studies with both RECIST 1.1 and iRECIST criteria to evaluate the overall response, the RS domain has multiple records of the overall response. In this case, the PROC TRANSPOSE procedure should be used to transpose the vertical shape data to horizontal data.

ADRS - Take Best Overall Response result from the dataset

USUBJID	paramcd	adt	avalc
1234-180001	BORCFINV	2016-04-07	PR
1234-180001	BORINV	2016-04-07	PR

Display 7: ADRS data

The ADaM dataset ADRS is our company's tumor response analysis data set. It contains the best overall response, a critical endpoint. Since one study can use different criteria, response evaluation source or methods, there could be one best overall response for each of these criteria. Dataset ADRS can have multiple PARAMCDs which contain best overall response information.

CREATE VISUALIZATION REPORT IN EXCEL FILE FORMAT WITH DYNAMIC LINKS

Step 1: Create Data for the Cover Page

Create a USUBJID listing.

```
proc sort data=dm out=subj(keep=usubjid ) nodupkey;
   by usubjid;
run;

data cover;
   set subj;
   label usubjid='USUBJID List';
run;
```

Step 2: Initial ODS TAGSETS.EXCELXP to create an excel file

```
TITLE;
FOOTNOTE;

ods _all_ close;
ods listing close;
ods tagsets.excelxp path="&outpath.\" file="&outfilename" style=Printer
    options(absolute_column_width='10'
            autofit_height='yes'
            autofilter='all'
            Center_Horizontal = 'Yes'
            Embed_Titles_once = 'yes'
            embedded_titles="yes"
            Embedded_Footnotes = 'yes'
            fittopage='yes'
            frozen_headers='yes'
            gridlines='yes'
            orientation='landscape'
            pagebreaks  = 'yes'
            pages_fitwidth='1'
            Print_Footer = ' '
            Row_Repeat='header');

ods tagsets.excelxp
       options(sheet_name   = "Cover Page"
               absolute_column_width='20');
```

Step 3: Create a cover page with a dynamic link connecting to the ' 'patient's profile

```
proc report data=cover ;
     column usubjid flag ;
     define usubjid / display;
     define flag / display;
     compute usubjid / character length=200;
         urlstring= "#'"||strip(usubjid)||"'!A1";
         call define(_col_, 'URL', urlstring);
     endcomp;
run;
```

Step 4: Assign individual macro parameters to every USUBJID

```
proc sql noprint;
   select count(distinct usubjid) into :subn
   from subj;

   select usubjid into :usub1 - :usub%eval(&subn.)
   from subj;

quit;
```

Step 5: Create patient's individual profile with dynamic link connecting back to cover page

```
%do i=1 %to %eval(&subn.);
   data p&i.;
      set subj (where=(usubjid="&&usub&i."));
   run;
```

- *Tumor size data - one row per [patient + date]*
- *Create warning flag for missing lesion data*
- *Create warning flag when multiple dates for the same visit - may not be an issue*
- *Combine Best Response data with Tumor Size data by [patient + date]*
- *Create a warning flag for unreasonable death or discontinuation date*
- *Append Treatment Start Date + Response / Tumor Size Data + Discontinue Date + Death Date*

```
%do i=1 %to %eval(&subn.);
   title link="#'Cover Page'!A1" "Back to Cover Page";

   footnote1 j=l f='Times New Roman' h=10pt bc=y "Data Issue 1: …";
   footnote2 j=l f='Times New Roman' h=10pt bc=y "Data Issue 2: …";
   footnote3 j=l f='Times New Roman' h=10pt bc=y "Data Issue 3: …";
   footnote4 j=l f='Times New Roman' h=10pt bc=y "Data Issue 4: …";

   ods tagsets.excelxp
       options(sheet_name   = "&&usub&i."
               print_header = "&&usub&i."
               absolute_column_width="&&column_width__&i"
               );

   proc print data=all_&i noobs label
        style(header)={font_size=8pt}
        style(data)  ={font_size=8pt
                       font_style= Roman
```

```
                                font_face = "marriott light"};

   run;
%end;

ods tagsets.ExcelXP close;
ods listing ;
```

MACRO FEATURES

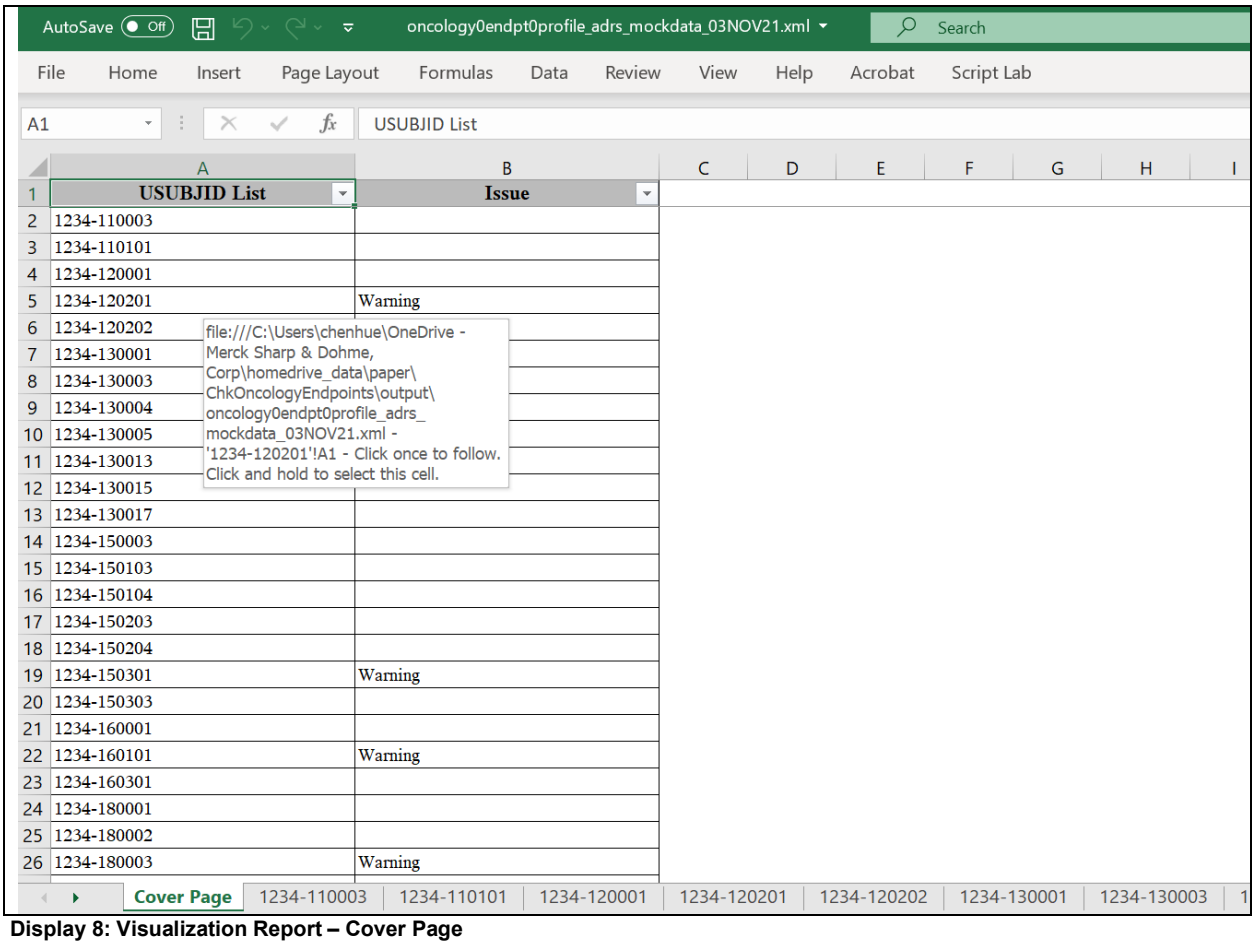
This macro consists of the following parameters.

Macro parameter	Parameter Value
sdtmlib	%str(c:\datasdtm)
adamlib	%str(c:\dataadam)
population_from	adamlib.adsl
population_where	%str(where trtfl='Y')
adrs_from	adamlib.adrs
adrs_where	%str(where paramcd in ("BORCFINV", "BORINV"))
rs_where	%str(where rstestcd = "OVRLRESP" and rseval="INVESTIGATOR" and rscat in ("RECIST 1.1","iRECIST"))
tr_where	%str(where trtestcd in ("LDIAM", "LPERP") and treval="INVESTIGATOR")
discon_criteria	%str(dscat='DISPOSITION EVENT' and dsscat='DISCONTINUED')
death_domain	sdtmlib.dm
death_criteria	%str(dthfl='Y')
death_var	dthdtc
outpath	%str(c:\output)
outfilename	%str(oncology0endpt0profile _&sysdate..xml))

VISUALIZATION REPORT

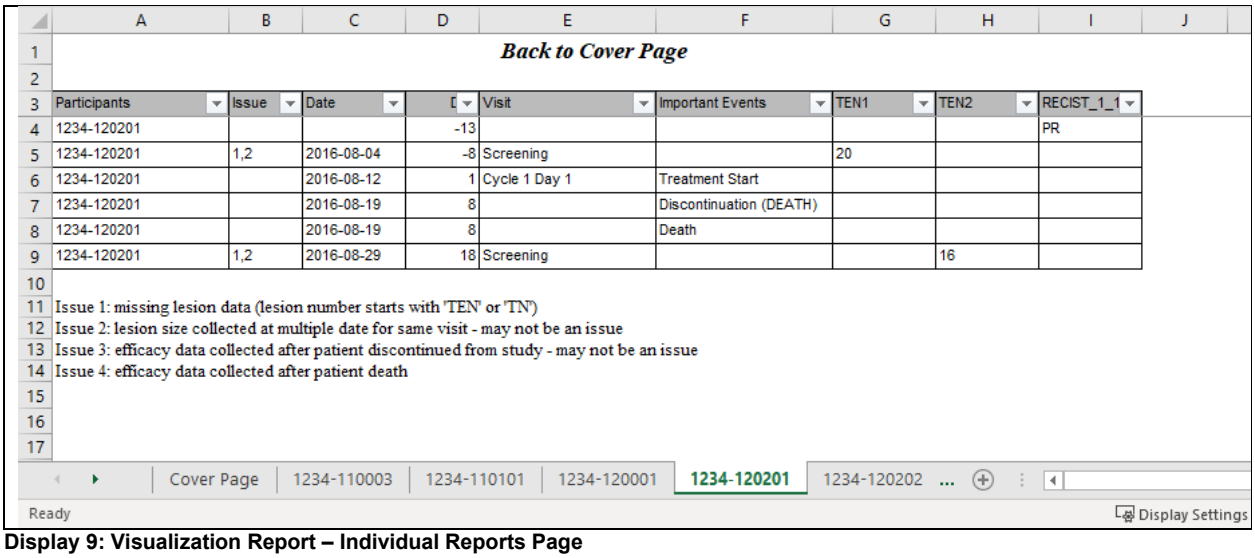
The visualization report contains one **cover page** and **individual reports** for each patient.

The cover page has two columns, ”USUBJID List" and "Issue". The "USUBJID List" column lists each patient ID with a hyperlink. Upon clicking the patient ID, an individual report will open for that patient. The "Issue" column would show a "Warning" when a data issue is detected.



Display 8: Visualization Report – Cover Page

The individual report includes three parts. The first part is a top row “Back to Cover Page” with a hyperlink. Upon click, it will close the individual report and open to cover page. This hyperlink helps to check other patient information conveniently. The second is the patient profile information, including patient ID, data issue, date, day relative to the treatment start date, important events (treatment start, death, discontinuation), lesion size from TR domain, RECIST1.1 and/or iRECIST result from RS domain. In addition, there is more information from ADRS, such as confirmed best overall response and unconfirmed best overall response. The third includes comments for data issues, which will explain the four types of issues.



Back to Cover Page

Participants	Issue	Date	D	Visit	Important Events	TEN1	TEN2	RECIST_1_1
1234-120201			-13					PR
1234-120201	1,2	2016-08-04	-8	Screening		20		
1234-120201		2016-08-12	1	Cycle 1 Day 1	Treatment Start			
1234-120201		2016-08-19	8		Discontinuation (DEATH)			
1234-120201		2016-08-19	8		Death			
1234-120201	1,2	2016-08-29	18	Screening			16	

Issue 1: missing lesion data (lesion number starts with 'TEN' or 'TN')
Issue 2: lesion size collected at multiple date for same visit - may not be an issue
Issue 3: efficacy data collected after patient discontinued from study - may not be an issue
Issue 4: efficacy data collected after patient death

Display 9: Visualization Report – Individual Reports Page

APPLICATIONS OF THE VISUALIZATION REPORT

The visualization report can help statistical programmers to check the key oncology efficacy analysis results, such as overall response and tumor size results during their development work. In our company, the ADRS analysis dataset has overall response results, and the ADTL analysis dataset has tumor size results. Discrepancies in any dataset between the developing programmer and the validation programmer may consume both programmers' time and effort to investigate the causes of the differences. When a patient's sum value of target lesions in longest diameter in ADTL from the development dataset is different from the validation dataset, a direct and efficient debug way is to look at all lesion size evaluation results across different visits of that patient. Simply pulling up this visualization report can quickly identify the issue's root.

CONCLUSION

The visualization report is a powerful tool for clinical programmers to understand the oncology study efficacy data when working on oncology efficacy datasets. Also, it helps clinicians and statisticians track patients tumor assessments history and the related disease responses more quickly. This paper introduces the related oncology efficacy data, the possible data issues in-and-between domains, and presents the macro developed to create efficacy visualization report in excel format.

REFERENCES

CDISC SDTM and ADaM Standards

- SDTM - CDISC SDTM Implementation Guide 3.2

Oncology Criteria

- RECIST 1.1 and iRECIST **-** RECIST (eortc.org) - The official site of the RECIST Working Group

Proceedings

- Chen, J., 2018. "SAS Macros to check and solve common problems in Oncology studies" Proceedings of PharmaSUG China 2018. (https://www.lexjansen.com/pharmasug-cn/2018/AD/Pharmasug-China-2018-AD08.pdf)
- Wang, H., Xiao, K., 2017. "Where Is the Link Broken – Another Look at SDTM Oncology Tumor Packages" Proceedings of PharmaSUG 2017. (https://www.lexjansen.com/pharmasug/2017/DS/PharmaSUG-2017-DS07.pdf)
- Wu, W., Dai, X., Gao, L., 2014. "Graphical Representation of Patient Profile for Efficacy Analyses in Oncology" Proceedings of PharmaSUG 2014. (https://www.lexjansen.com/pharmasug/2014/DG/PharmaSUG-2014-DG02.pdf)

ACKNOWLEDGMENTS

The authors would like to thank the management team of Merck & Co., Inc., Kenilworth, NJ, USA, for their advice on this paper/presentation.

CONTACT INFORMATION

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

Lily Zhang
Merck & Co., Inc., Kenilworth, NJ, USA
e-mail: lily_zhang2@merck.com

Yixin Ren
Merck & Co., Inc., North Wales, PA, USA
e-mail: yixin.ren@merck.com

Huei-Ling Chen
Merck & Co., Inc., Kenilworth, NJ, USA
e-mail: huei-ling_chen@merck.com

TRADEMARK