SouthEast SAS® Users Group (SESUG) 2023

An Innovative Approach to Generate a CONSORT Diagram in Clinical Trials

Lydia Li, Shunbing Zhao, Merck & Co., Inc.

ABSTRACT

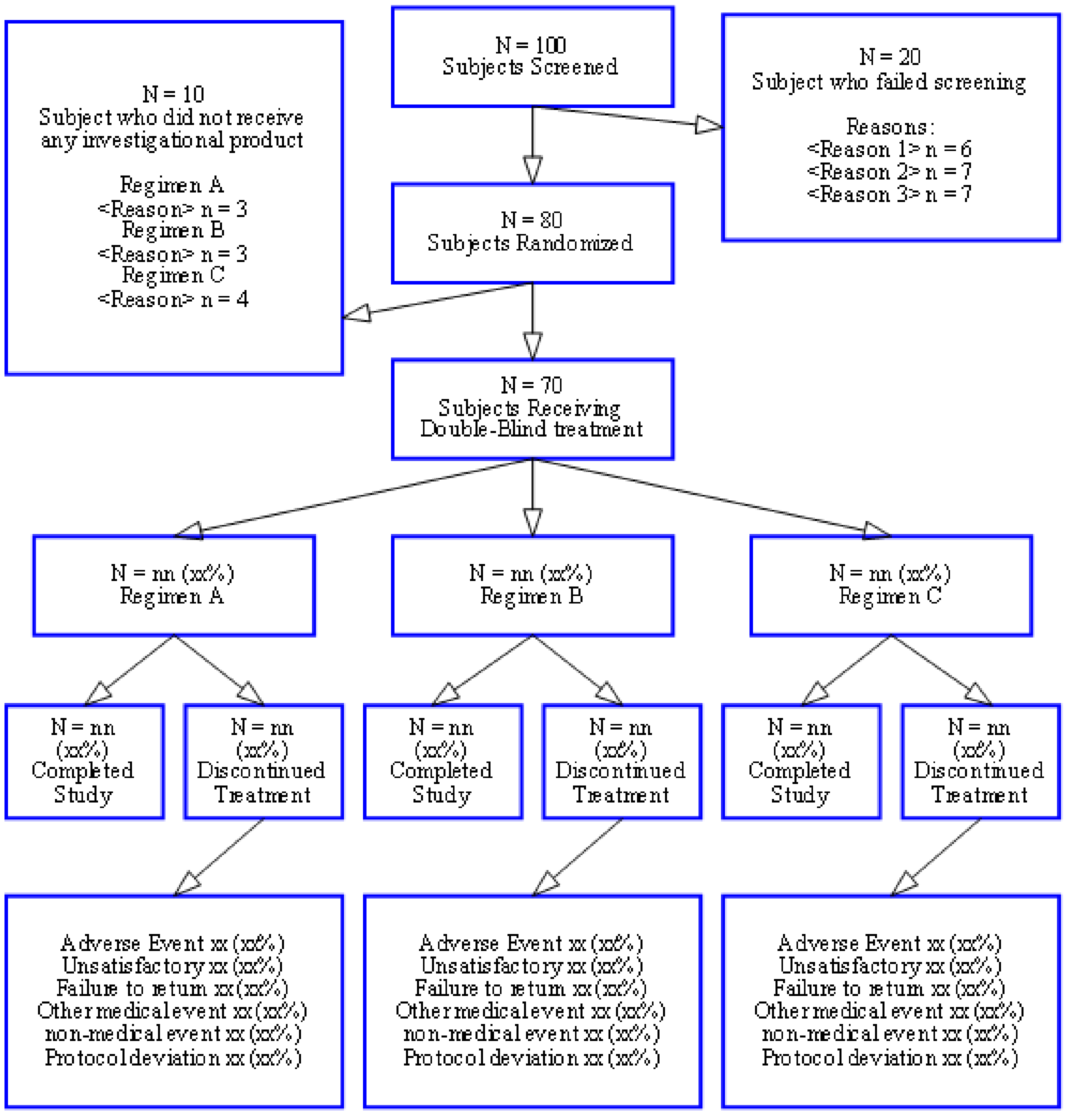
In the reporting of late stage clinical trials data, the consort diagram is a powerful graphic representation for reporting the disposition of each participant, including number of participants screened, number of participants failed screening, number of participants randomized and treated, etc. It gives a vivid overview of each participant's status in the trial and becomes one of the deliverables for Clinical Study Report (CSR) package once the database is locked. This paper illustrates how to use SAS proc template and proc sgrender procedures to create the consort diagram and analyzes the advantages and disadvantages of using this method.

INTRODUCTION

The Consolidated Standard of Reporting Trials (CONSORT) diagram is often used to report the flow of participants through each stage of the clinical trial. It is essential to assess the validity of study results and provides a bird's eye view of completeness of information in the clinical trials. The Consort diagram is more than a replica of disposition of participants summary table. It provides information for each participant's status at different stages in the trial, including screen failures.

Below is an example of a consort diagram from a double blinded study which has 3 treatment arms.

Consort Diagram Disposition of Participants


N = 100
Subjects Screened
N = 20
Subject who failed screening
Reasons:
<Reason 1> n = 6
<Reason 2> n = 7
<Reason 3> n = 7
N = 80
Subjects Randomized
N = 10
Subject who did not receive any investigational product
Regimen A
<Reason> n = 3
Regimen B
<Reason> n = 3
Regimen C
<Reason> n = 4
N = 70
Subjects Receiving Double-Blind treatment
N = nn (xx%)
Regimen A
N = nn (xx%)
Regimen B
N = nn (xx%)
Regimen C
N = nn (xx%) Completed Study
N = nn (xx%) Discontinued Treatment
N = nn (xx%) Completed Study
N = nn (xx%) Discontinued Treatment
N = nn (xx%) Completed Study
N = nn (xx%) Discontinued Treatment
Adverse Event xx (xx%)
Unsatisfactory xx (xx%)
Failure to return xx (xx%)
Other medical event xx (xx%)
non-medical event xx (xx%)
Protocol deviation xx (xx%)
Adverse Event xx (xx%)
Unsatisfactory xx (xx%)
Failure to return xx (xx%)
Other medical event xx (xx%)
non-medical event xx (xx%)
Protocol deviation xx (xx%)
Adverse Event xx (xx%)
Unsatisfactory xx (xx%)
Failure to return xx (xx%)
Other medical event xx (xx%)
non-medical event xx (xx%)
Protocol deviation xx (xx%)

There are 4 components in the SAS program for the Consort Diagram. SAS version 9.4 is the tool to develop the program.

1. Calculate number of participants and percentage for each box
2. Assign appropriate position for each box, lines connecting each box and text inside each box
3. Set up style and template to display
4. Apply style and template to prepared data

1. CALCULATE NUMBER OF PARTICIPANTS AND PERCENTAGE FOR EACH BOX

There are various ways to calculate participant count and percentage. For example, use proc sql and macro variables to hold the value of counts:

```
 proc sql noprint;
    select sum(randfl='Y'),
           sum(randfl='N'),
           count(*)
      into :totrand,
           :totnotrand,
    from adsl;
 quit;

%let totrand=&totrand;
%let totnotrand=&totnotrand;
```

Or use proc freq to calculate the counts:

```
proc freq data=adsl(where=(randfl='N')) order=freq noprint;
   tables dcsreas/out=scrnfail;
run;

proc freq data=adsl(where=( trtfl='Y')) noprint;
   tables trt01an*trt01a/out=trted;
run;
```

ADaM dataset ADSL is the input data and population variables in ADSL are the input variables. ADaM stands for Analysis Data Model. ADaM Implementation Guide provides fundamental principles of deriving analysis datasets and variables. ADaM standards are designed to meet the requirements of regulatory submission, such as United States Food and Drug Administration (FDA). ADSL is subject level analysis dataset in the ADaM standard. The variables in ADSL include demographic variables, population flag variables, subgroup analysis variables, etc. These subject-level variables can be copied to other ADaM datasets for analysis.

2. ASSIGN APPROPRIATE POSITION FOR EACH BOX, LINES CONNECTING EACH BOX AND TEXTS INSIDE EACH BOX

```
data hTextC1;
xHtc=(35+60)/2; yHtc=195; htextc="N=&totscrned^Subject screened"; output;
xHtc=(65+90)/2; yHtc=180; htextc="N=&totnotrand^Subject who^failed
screening"; output;

xHtc=(35+60)/2; yHtc=170; htextc="N=&totrand^Subject randomized"; output;
xHtc=(35+60)/2; yHtc=143; htextc="N=&tottrted^Subject receiving double-^blind
treatment"; output;
xHtc=(5+30)/2; yHtc=188; htextc="N=&totnottrted^Subject who did not
receive^any Investigational Product"; output;

run;
```

The above SAS code is for the position for the text inside each of the top 5 boxes (XHTC and YHTC values). Since the 5 boxes are in the middle of the page, XHTC is assigned as the middle point of the left border line and right border line. YHTC is assigned from higher value to lower value, representing from top to bottom locations.

Set up empty box data and links data:

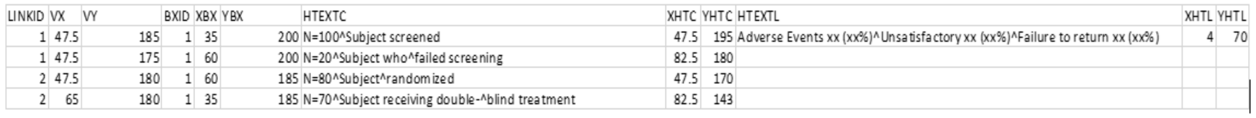
```
/* Empty Box Data and Links */
data emptyBoxes(keep=bxid xbx ybx) linksCoord(keep=linkid vx vy);
  bxid=1;xbx=35;ybx=200;output emptyBoxes;
  bxid=1;xbx=60;ybx=200;output emptyBoxes;
  bxid=1;xbx=60;ybx=185;output emptyBoxes;
  bxid=1;xbx=35;ybx=185;output emptyBoxes;
  bxid=2;xbx=35;ybx=175;output emptyBoxes;
  bxid=2;xbx=60;ybx=175;output emptyBoxes;
  bxid=2;xbx=60;ybx=160;output emptyBoxes;
  bxid=2;xbx=35;ybx=160;output emptyBoxes;
  linkid=1; vx=(35+60)/2; vy=185; output linksCoord;
  linkid=1; vx=(35+60)/2; vy=175; output linksCoord;
  linkid=2; vx=(35+60)/2; vy=(175+185)/2; output linksCoord;
  linkid=2; vx=65; vy=(175+185)/2; output linksCoord;
  ……

run;
```

In this consort diagram, there are a total of 17 boxes. Hence, BXID has a total of 17 distinct values. There are a total of 16 lines connecting the boxes. Hence, LINKID has a total of 16 distinct values.

In the next step, all the input datasets which are used to draw the CONSORT diagram are merged. As illustrated in the table screenshot below, VX and VY are the positions for each line connecting each box. XBX and YBX are the positions for each box's four corners. XHTC and YHTC are the positions for the text inside each box. XHTL and YHTL are the positions for additional text.

```
/* Combine all graph data */
data consort;
merge linksCoord emptyBoxes hTextC hTextl;
run;
```



LINKID	VX	VY	BXID	XBX	YBX	HTEXTC	XHTC	YHTC	HTEXTL	XHTL	YHTL
1	47.5	185	1	35	200	N=100^Subject screened	47.5	195	Adverse Events xx (xx%)^Unsatisfactory xx (xx%)^Failure to return xx (xx%)	4	70
1	47.5	175	1	60	200	N=20^Subject who^failed screening	82.5	180			
2	47.5	180	1	60	185	N=80^Subject^randomized	47.5	170			
2	65	180	1	35	185	N=70^Subject receiving double-^blind treatment	82.5	143			

3. SET UP STYLE AND TEMPLATE TO DISPLAY

Use proc template procedure to set up style for title, label, footnote, etc. The statements define the basic elements of graph and create graph style for the output appearance.

```
proc template;
  define style mystyle;
   parent=styles.listing;
   class graphwalls / frameborder=off;
   class graphbackground / color=white;

  Style GraphTitleText    / color=black fontsize=8pt fontfamily='Thorndale
AMT';
……
  end;
run;
```

Use proc template define statgraph to set up a graph template, and use seriesplot, polygonplot, and textplot statements to plug in VX and VY, XBX and YBX, XHTC and YHTC, and XHTL and YHYL's values.

```
proc template;
 define statgraph consortdiagram;
 begingraph;
  dynamic _ticklist_ ;
   layout overlay /
   ……
        seriesplot  x=vX y=vY / group=linkid lineattrs=(thickness=3
color=blue pattern=solid);

        polygonplot id=bxid x=xbx y=ybx / display=(outline)
           fillattrs=(color=yellow transparency=0.75)
           label=label labellocation=outsidebbox labelposition=ymax;

        textplot x=xHtc y=yHtc text=hTextC / splitchar='^'
splitpolicy=splitalways;

        textplot x=xHtl y=yHtl text=hTextl / splitchar='^'
splitpolicy=splitalways position=bottomright;
      endlayout;

 endgraph;
 end;
run;
```

4. ASSIGN TEMPLATE TO PREPARED DATA

Sgrender with proc template is a powerful tool to create a customized graph using a user-defined template. Use the proc sgrender procedure to associate the graph template to the consort diagram data and produce the graph. Consort is the dataset containing all the necessary values for boxes, lines, and texts. Consortdiagram is the template defined in the previous proc template statements.

```
proc sgrender data=consort template=consortdiagram;
```

CONCLUSION

This paper discusses how to use the SAS proc template and proc sgrender procedures to create a consort diagram. In order to use these procedures, some preparation work needs to be done: set up box dataset, links dataset, text inside the boxes dataset, style and template dataset.

The advantages of this method are that it is efficient to add or reduce boxes depending on the scenario of the data. As long as setting up positions for boxes, lines and texts are done well, boxes with appropriate size can be placed nicely with the lines and text in the page. Moving boxes is not difficult when updating the positions of the four corners. Once the positions of boxes are set up well, it will not be difficult to provide the positions of lines and texts. This “what-you-see-is-what-you-get” approach is more visually direct.

Another advantage is that this consort diagram was originally manually created by the medical writer who wrote CSR. The manual creation easily introduces errors or miscounts. Having programming create the diagram can potentially reduce these errors or miscounts, which saves time and effort for the medical writer and benefits the CSR writing and publishing process.

At the beginning, program code will be a little more complicated and time consuming to develop. To set up the links between boxes, one needs to manually adjust the values of the positions, so that the lines can link the boxes, and text can be placed inside the boxes and all the components can be aligned together. These manual adjustments take a great deal of time. Once the program is developed, the code can be written as a macro and reusable for other studies.

Overall, using this approach a CONSORT diagram can be generated in a timely and well-controlled way. It provides clear, concise, and precise reporting of the flow of participants in the clinical trials.

REFERENCES

PharmaSUG 2021 - Paper DV-082 Automated CONSORT Flow Diagram Generation by SAS® Programming

https://www.lexjansen.com/pharmasug/2021/DV/PharmaSUG-2021-DV-082.pdf

PhUSE 2016 Poster PP03 CONSORT Diagram: Doing it with SAS

https://www.lexjansen.com/phuse/2016/pp/PP03.pdf

Paper 1090-2017 A Step by Step Solution to Create a Customized Graph for Grouped Data Using SAS® Graph Template Language

https://support.sas.com/resources/papers/proceedings17/1090-2017.pdf

Analysis Data Model Implementation Guide Version 1.1

https://www.cdisc.org/standards/foundational/adam/adamig-v1-1

CONTACT INFORMATION

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

Lydia Li

Merck & Co., Inc.

lydia.li@merck.com

Shunbing Zhao

Merck & Co., Inc.

Shunbing.zhao@merck.com