



Programming techniques to enhance the presentation of forest plot with SAS 9.4

Zhen (Laura) Li

Senior Statistical Programmer II, Early Biometrics & Statistical Innovation,
Data Science & Artificial Intelligence, R&D, AstraZeneca, Gaithersburg, USA

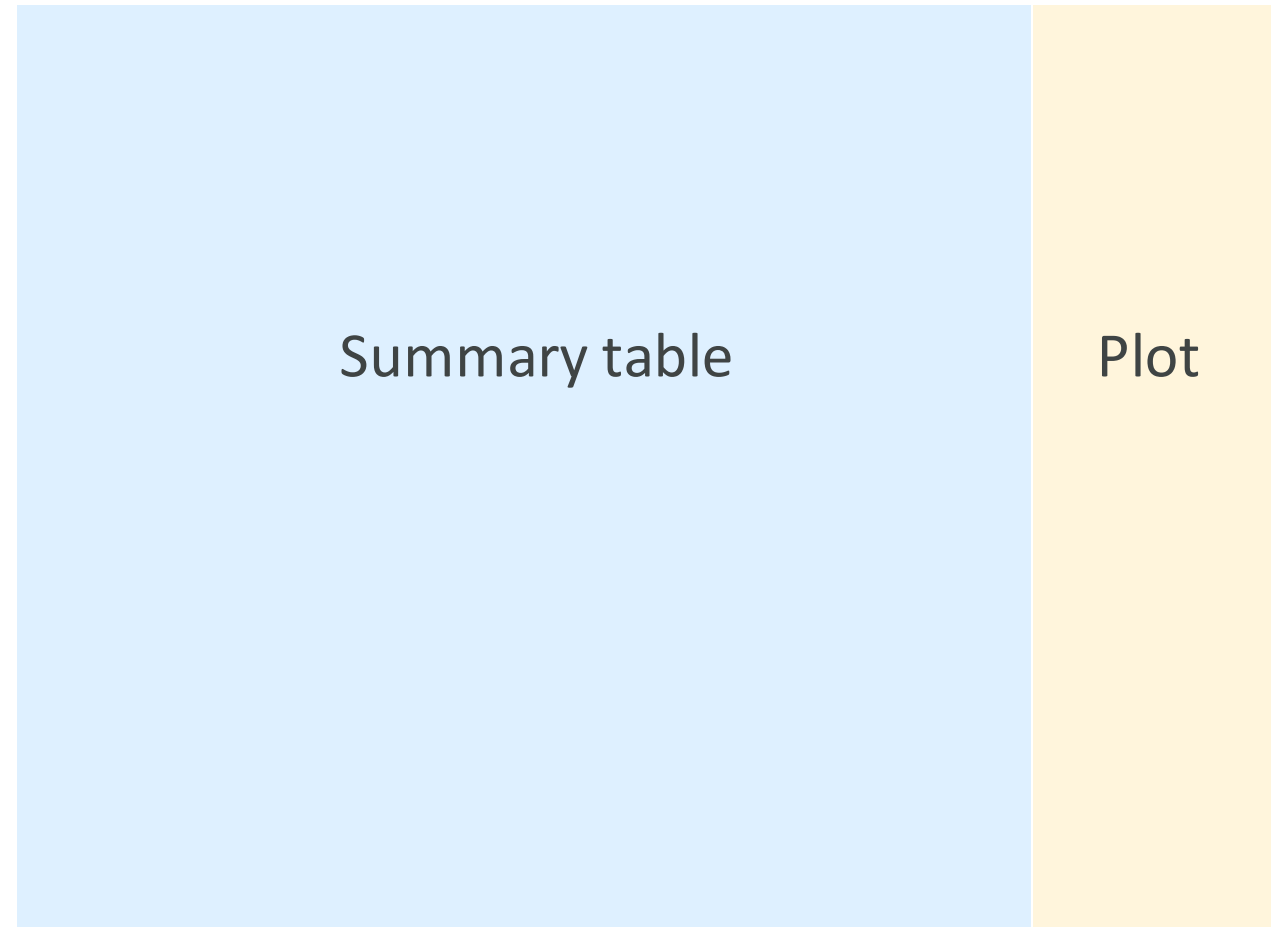
November 15th 2021



Agenda

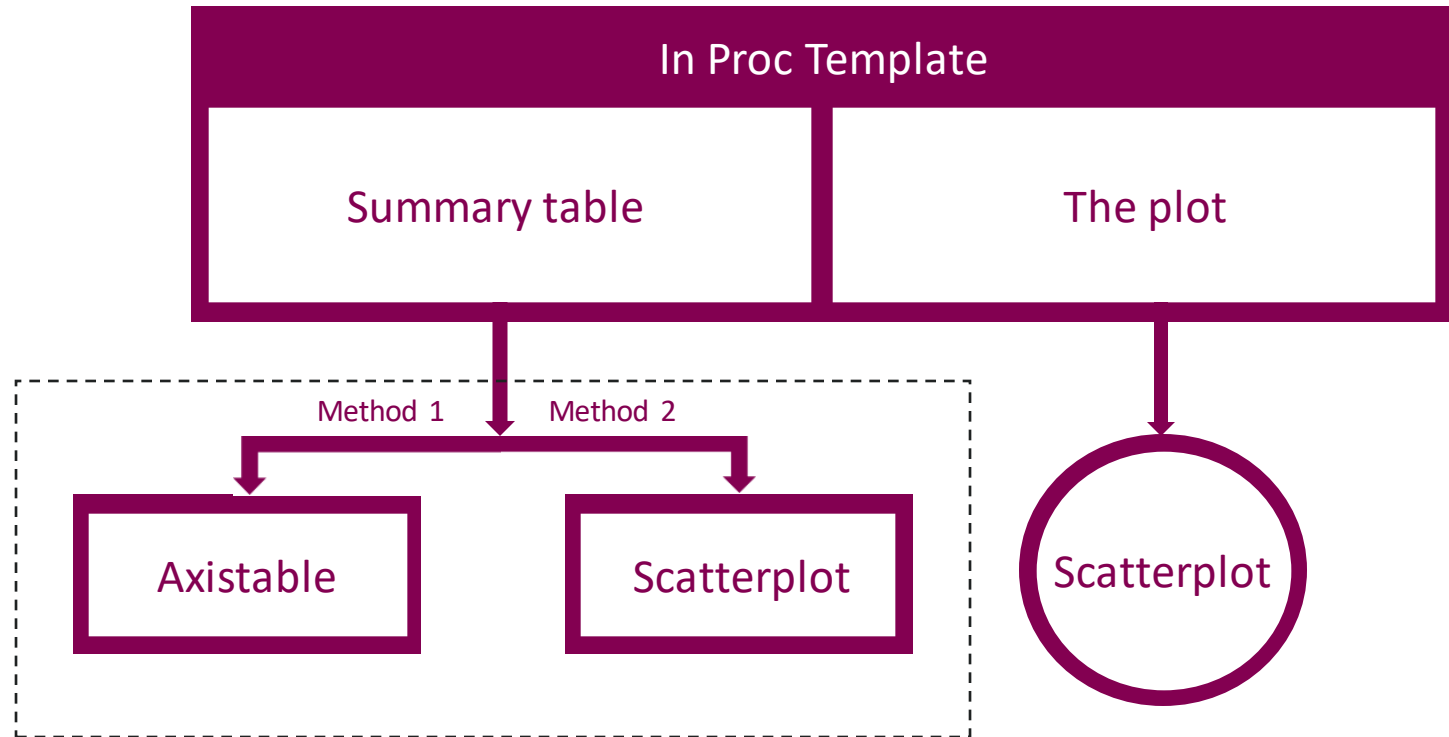
How to generate a forest plot?

How to enhance its presentation?



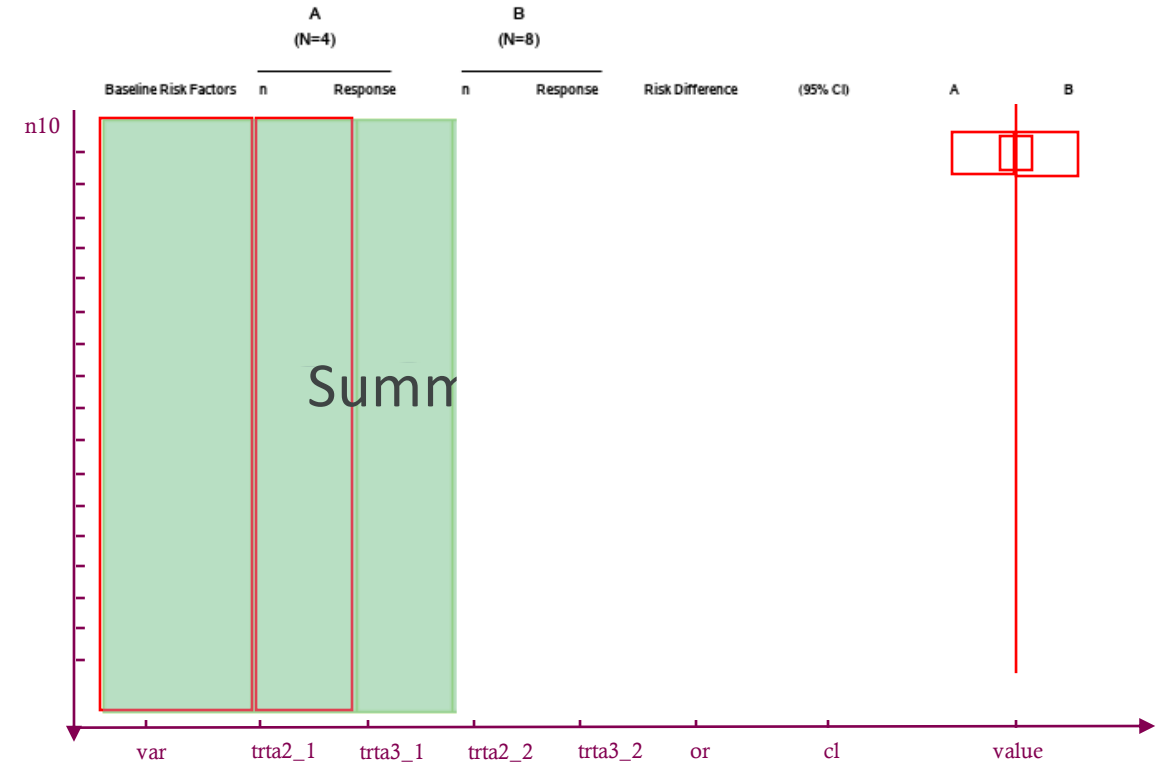
Note: this presentation uses dummy data.

How to generate a forest plot?

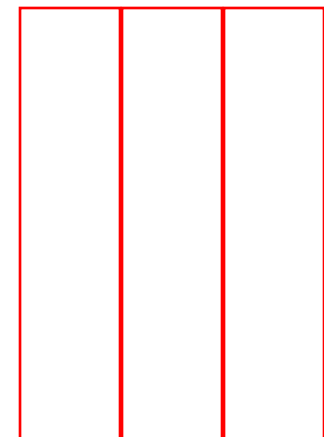


Use Axistable for summary table

```
proc template;
define statgraph forestplot;
begingraph;
  layout lattice / columns=8 columnweights=(0.15 0.1 0.1 0.1 0.1 0.1 0.15 0.2);
  -- code of column headers is omitted --
  /*-- generate Baseline Risk Factors column --*/
  layout overlay / walldisplay=none xaxisopts=( display= none ) yaxisopts=(reverse=true display=none
  tickvalueattrs=(weight=bold));
  referenceline y=0/ lineattrs=(thickness=1px color=black);
  axistable y=n10 value=var /label=" " indentweight=indent textgroup=type;
  endlayout;
  /*-- generate group A' s n column --*/
  layout overlay / xaxisopts=( display=none ) yaxisopts=(reverse=true display=none) walldisplay=none;
  referenceline y=0/lineattrs=(thickness=1px color=black);
  axistable y=n10 value=trta2_1/label=" " ;
  endlayout;
  /*-- generate group A' s Response column --*/
  layout overlay / xaxisopts=(display=none) yaxisopts=(reverse=true display=none) walldisplay=none;
  referenceline y=0 / lineattrs=(thickness=1px color=black);
  axistable y=n10 value=trta3_1 /label=" ";
  endlayout;
  -- code of other columns is similar and omitted --
  /*-- generate risk difference plot --*/
  layout overlay /xaxisopts=(label=" " linearopts=(tickvaluepriority=true tickvaluelist=(-30 0 30 )))
  yaxisopts=(reverse=true display=none) walldisplay=none;
  scatterplot y=n10 x=value / xerrorlower= lowercl xerrorupper= uppercl
  markerattrs=(symbol=squarefilled) yaxis=y sizemin=4 sizemax=12;
  referenceline x=0;
  referenceline y=0 / lineattrs=(thickness=1px color=black);
  endlayout;
endlayout;
endgraph;
end;
run;
```



n10	var	trta2_1	trta3
1	All Patients	4	3
2	Age Group		
3	<65	1	1
4	>=65	3	2
5	Race		
6	White	2	1
7	Asian	1	1
8	Multiple	1	1
9	Baseline ECOG		
10	0	2	2
11	1	2	1
12	Stage		
13	II	4	3
14	III	0	0
15	PD-L1 Status		
16	IC0	4	3
17	IC1/2/3	0	0



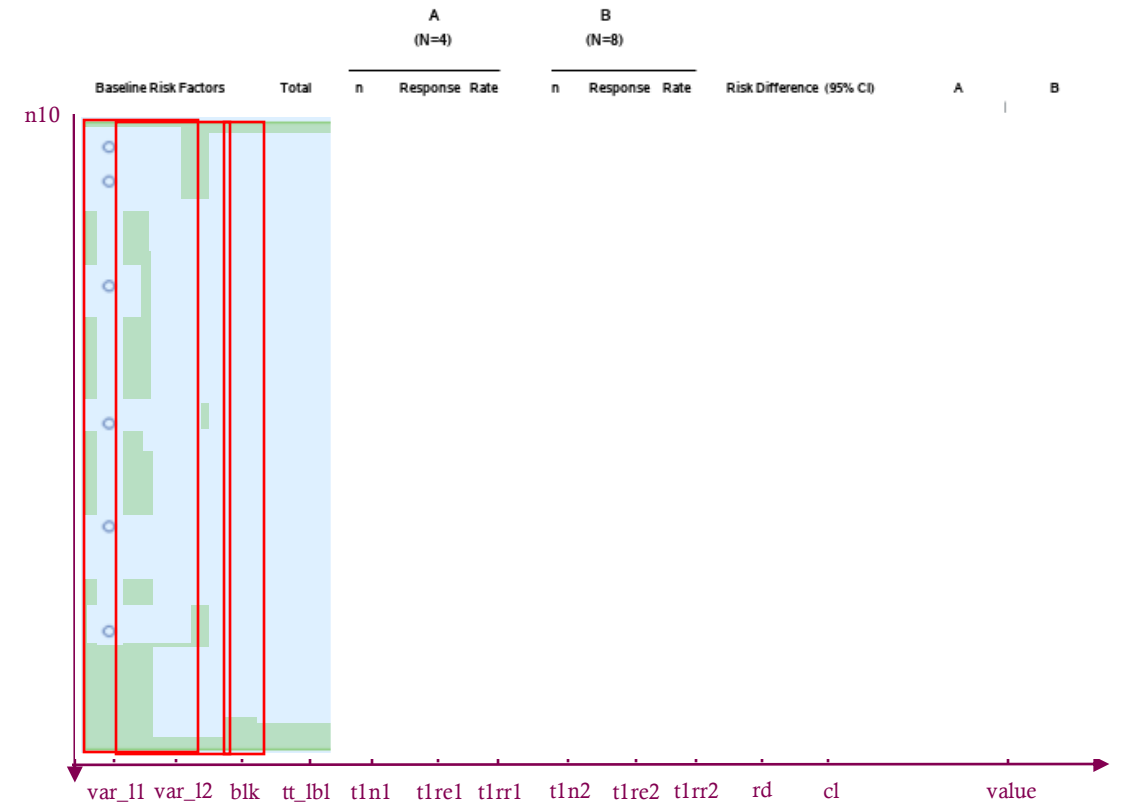


Can we treat the
summary table
part as a whole?

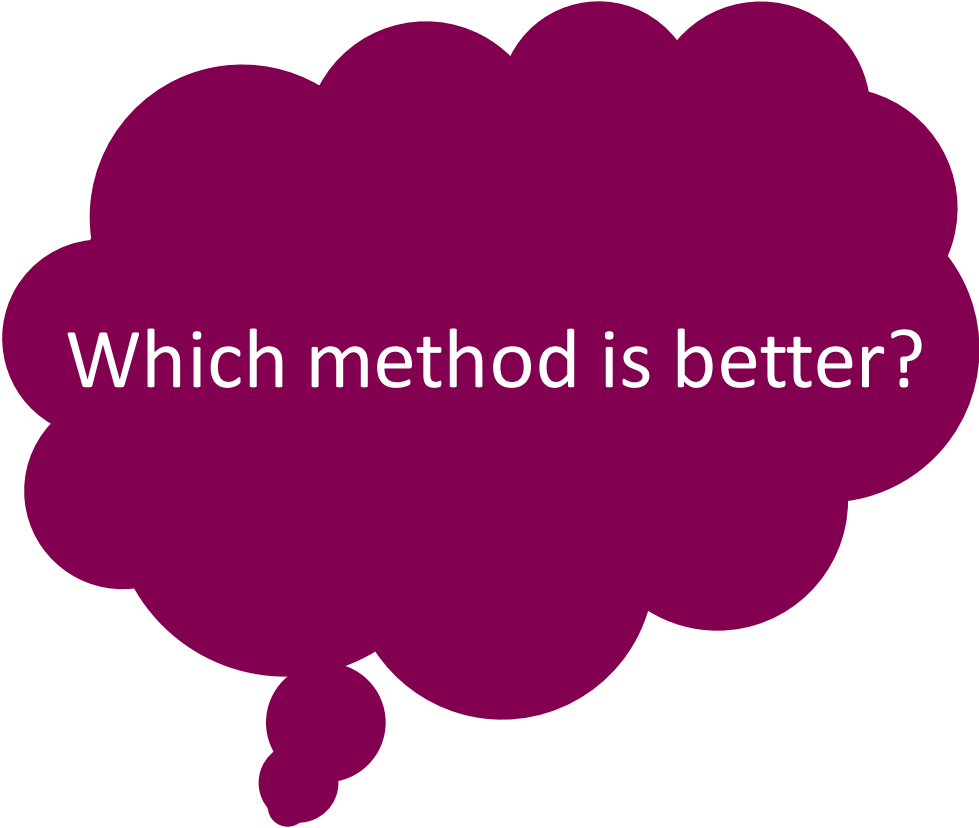
No and Yes

Use Scatterplot for summary table

```
proc template;
define statgraph forplt;
begingraph;
  layout lattice / columns=2 columnweights=(0.89 0.21) shrinkfonts=true;
  layout overlay /walldisplay=none yaxisopts=(offsetmin=0 offsetmax=0.05 display= none reverse=true
  linearopts=(viewmin=0 viewmax=20)) xaxisopts=(offsetmin=0 offsetmax=0.08 display=none);
  -- code of column headers is omitted --
  /*-- generate Baseline Risk Factors -- group column --*/
  scatterplot y=n10 x=var_l1 _____;
  ;
  /*-- generate Baseline Risk Factors -- subgroup column --*/
  scatterplot y=n10 x=var_l2/ markercharacter= var2 markercharacterattrs=(size=16.5spx)
  markercharacterposition=right;
  /*-- add a blank column --*/
  scatterplot y=n10 x=blk/ markercharacter= blank markercharacterattrs=(size=2spx);
  /*-- generate Total column --*/
  scatterplot y=n10 x=tt_lbl / markercharacter=trta1_1 markercharacterattrs=(size=16.5spx);
  -- code of other columns is similar and omitted --
endlayout;
/*-- generate risk difference plot --*/
layout overlay /xaxisopts=(display=(ticks tickvalues line) linearopts=(tickvaluepriority=true
tickvaluelist=( -30 0 30 ))) yaxisopts=( reverse=true display=none) walldisplay=none;
  scatterplot y=n10 x=value / xerrorlower= lowercl xerrorupper= uppercl markerattrs=(symbol=squarefilled)
  yaxis=y sizemin=4 sizemax=12;
  referenceline x=0;
  referenceline y=0 / lineattrs=(thickness=1px color=black);
endlayout;
endgraph;
end;
```



	n10	var	var_l1	var1	var_l2	var2	blk	blank	tt_lbl	trta1_1
1	All Patients	var1	All Patients				balnk		Total	12
2	Age Group	var1	Age Group				balnk		Total	.
3	<65				var2	<65	balnk		Total	8
4	>=65				var2	>=65	balnk		Total	4
5	Race	var1	Race				balnk		Total	.
6	White				var2	White	balnk		Total	7
7	Asian				var2	Asian	balnk		Total	2
8	Multiple				var2	Multiple	balnk		Total	3
9	Baseline ECOG	var1	Baseline ECOG				balnk		Total	.
10	0				var2	0	balnk		Total	9
11	1				var2	1	balnk		Total	3
12	Stage	var1	Stage				balnk		Total	.
13	II				var2	II	balnk		Total	8
14	III				var2	III	balnk		Total	4
15	PD-L1 Status	var1	PD-L1 Status				balnk		Total	.
16	IC0				var2	IC0	balnk		Total	8
17	IC1/2/3				var2	IC1/2/3	balnk		Total	4



Which method is better?

It depends

Compare Axistable with Scatterplot

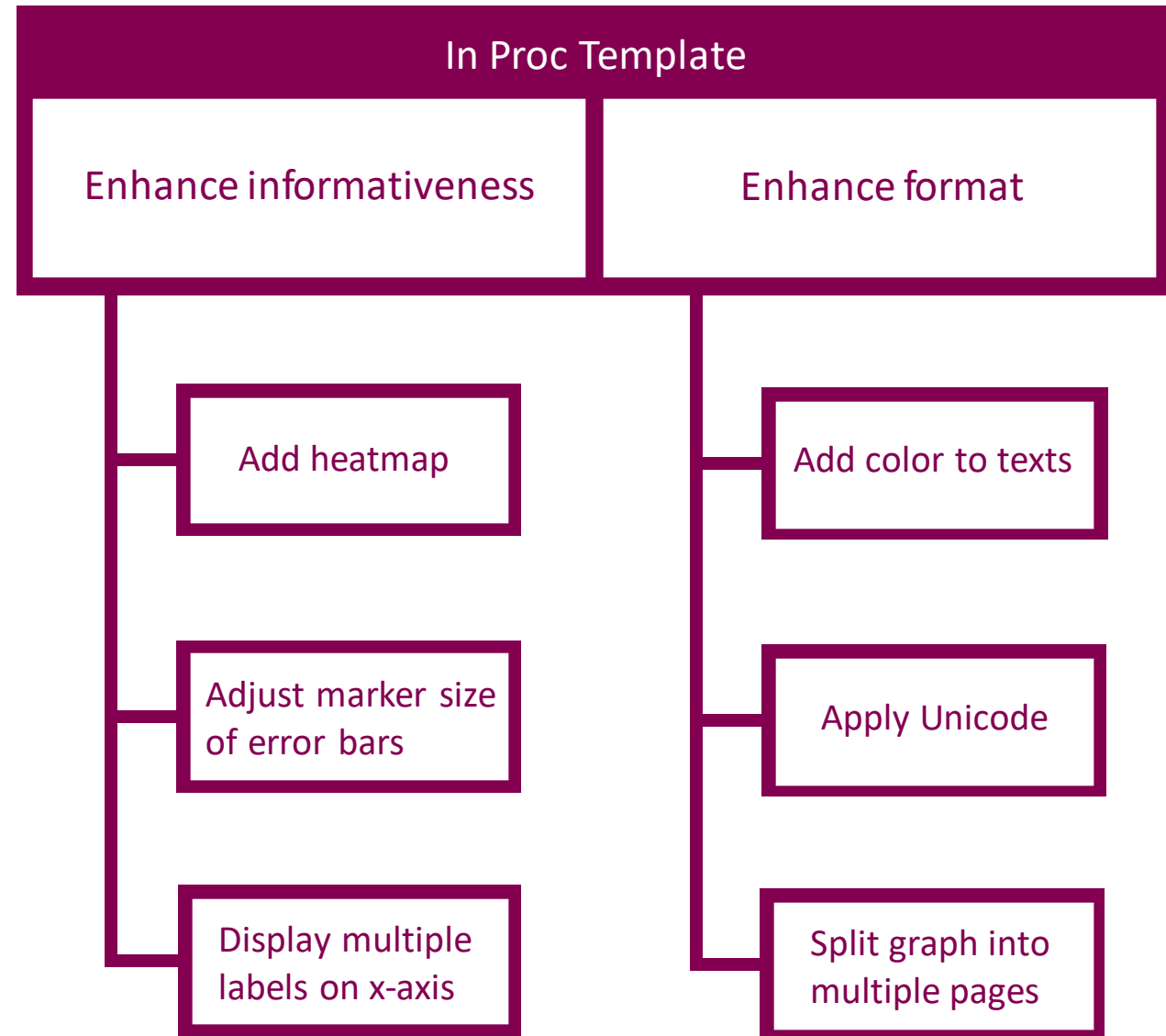
Axistable:

- The code is more straightforward to use and understand.
- It's easier to adjust the layout for each column differently.

Scatterplot:

- The code is more efficient.
- It's easier to nicely fit all columns into one graph.

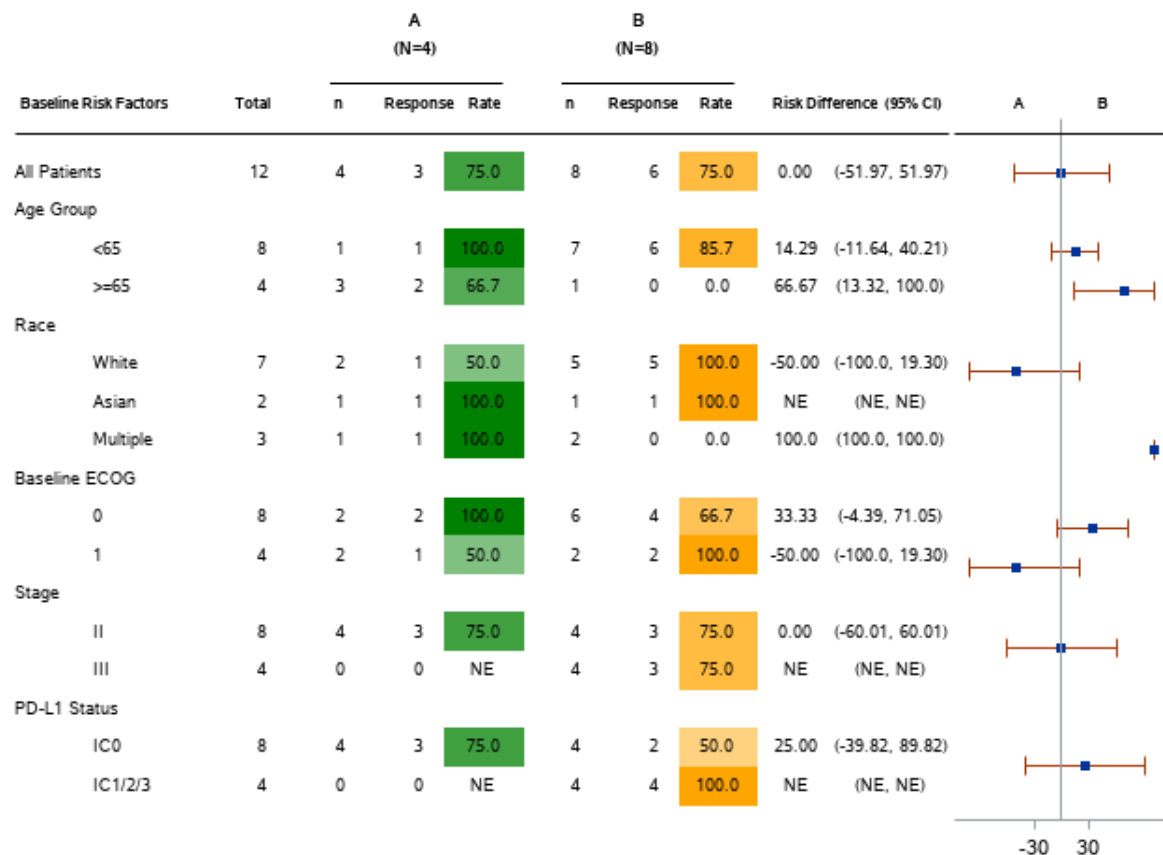
How to enhance its presentation?





Enhance
informativeness

Add heatmap



```
proc template;
define statgraph forplt;
begingraph;
```

trta4_1_ is numeric version
of trta4_1

trta4_2_ is numeric version
of trta4_2

```
layout lattice / columns=2 columnweights=(0.89 0.21) shrinkfonts=true;
layout overlay / walldisplay=none yaxisopts=(offsetmin=0 offsetmax=0.05 display= none reverse=true
linearopts=(viewmin=0 viewmax=20)) xaxisopts=(offsetmin=0 offsetmax=0.08 display=none);
-- code of column headers is omitted --
-- code of Baseline Risk Factors, Total, group A's n and Response columns is omitted --
/*-- generate group A's Rate column --*/
```

```
scatterplot y=n10 x=t1rr1 / markercharacter=trta4_1 markercharacterattrs=(size=16.5spx) yaxis=y;
-- code of group B's n and Response columns is omitted --
/*-- generate group B's Rate column --*/
```

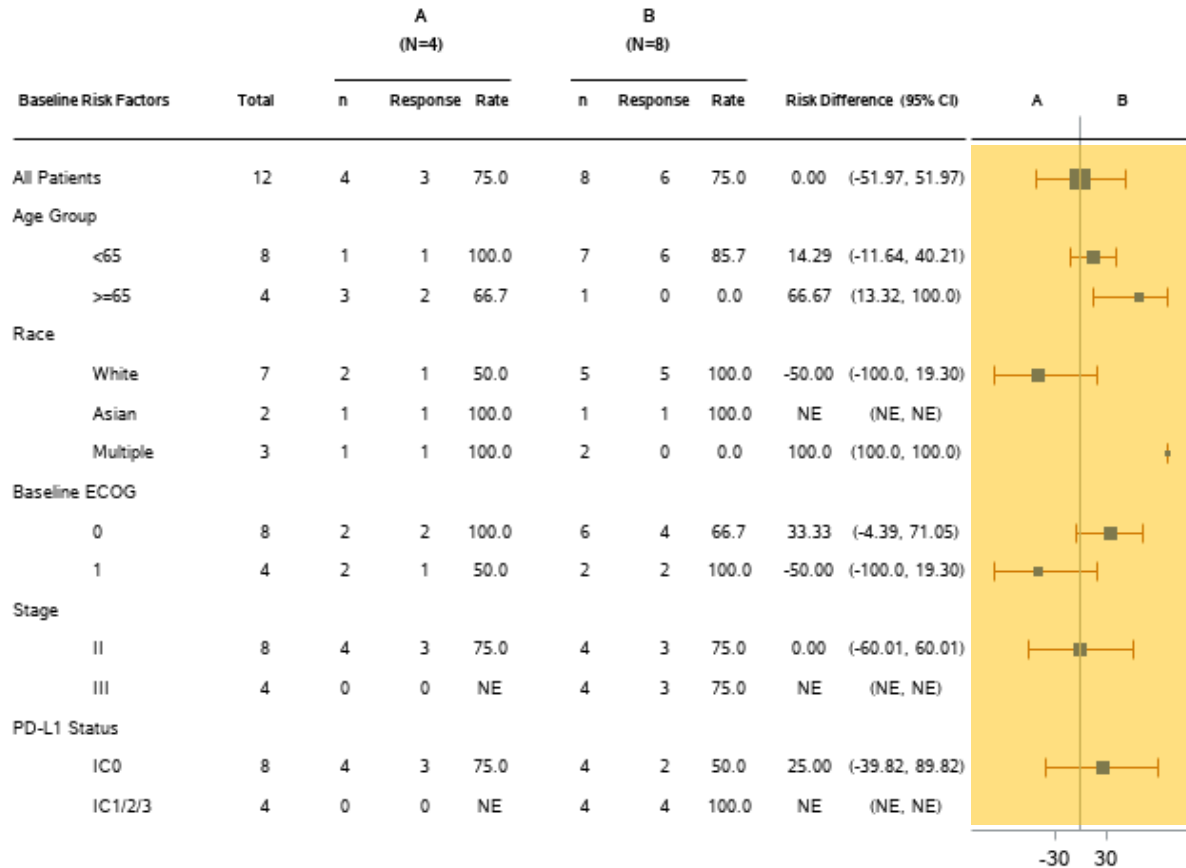
```
scatterplot y=n10 x=t1rr2 / markercharacter=trta4_2 markercharacterattrs=(size=16.5spx) yaxis=y;
-- code of other columns is omitted --
```

```
endlayout;
```

```
/*-- generate risk difference plot --*/
```

```
layout overlay / xaxisopts=(display=(ticks tickvalues line) linearopts=(tickvaluepriority=true
tickvaluelist=( -30 0 30 ))) yaxisopts=( reverse=true display=none) walldisplay=none;
scatterplot y=n10 x=value / xerrorlower= lowercl xerrorupper= uppercl markerattrs=(symbol=squarefilled)
yaxis=y sizemin=4 sizemax=12;
referenceline x=0;
referenceline y=0 / lineattrs=(thickness=1px color=black);
endlayout;
endgraph;
end;
```

Adjust marker size of error bars

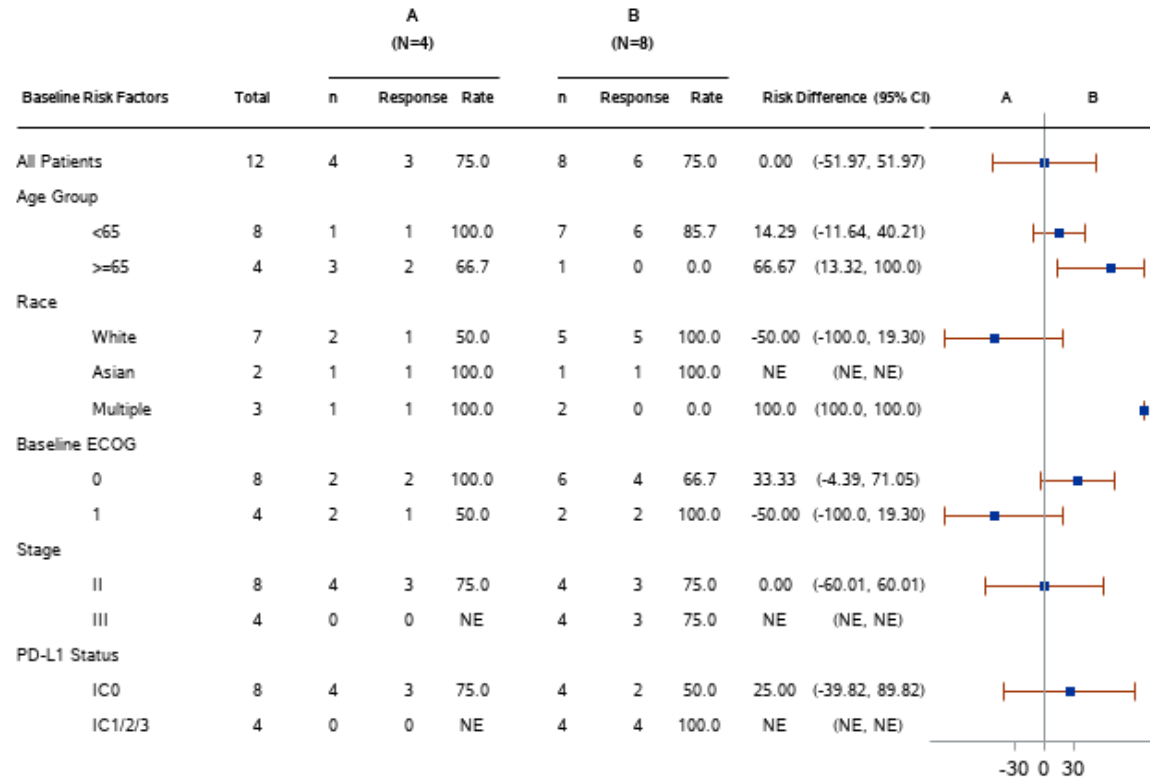


```

proc template;
define statgraph forplt;
begingraph;
layout lattice / columns=2 columnweights=(0.89 0.21) shrinkfonts=true;
layout overlay / walldisplay=none yaxisopts=(offsetmin=0 offsetmax=0.05 display= none reverse=true
linearopts=(viewmin=0 viewmax=20)) xaxisopts=(offsetmin=0 offsetmax=0.08 display=none);
-- code of column headers is omitted --
/*-- generate Baseline Risk Factors -- group column --*/
scatterplot y=n10 x=var_11/ markercharacter= var1 markercharacterattrs=(size=16.5spx)
markercharacterposition=right;
/*-- generate Baseline Risk Factors -- subgroup column --*/
scatterplot y=n10 x=var_12/ markercharacter= var2 markercharacterattrs=(size=16.5spx)
markercharacterposition=right;
/*-- add a blank column --*/
scatterplot y=n10 x=blk/ markercharacter= blank markercharacterattrs=(size=2spx);
/*-- generate Total column --*/
scatterplot y=n10 x=tt_lbl/ markercharacter=trta1_1 markercharacterattrs=(size=16.5spx);
-- code of other columns is similar and omitted --
endlayout;
/*-- generate risk difference plot --*/
layout overlay / xaxisopts=(display=(ticks tickvalues line) linearopts=(tickvaluepriority=true
tickvaluelist=( -30 0 30 ))) yaxisopts=( reverse=true display=none) walldisplay=none;
scatterplot y=n10 x=value / xerrorlower= lowercl xerrorupper= uppercl
markerattrs=(symbol=squarefilled) markersizeresponse=trta1_1
yaxis=y sizemin=4 sizemax=12;
referenceline x=0;
referenceline y=0 / lineattrs=(thickness=1px color=black);
endlayout;
endgraph;
end;

```

Display multiple labels on x-axis

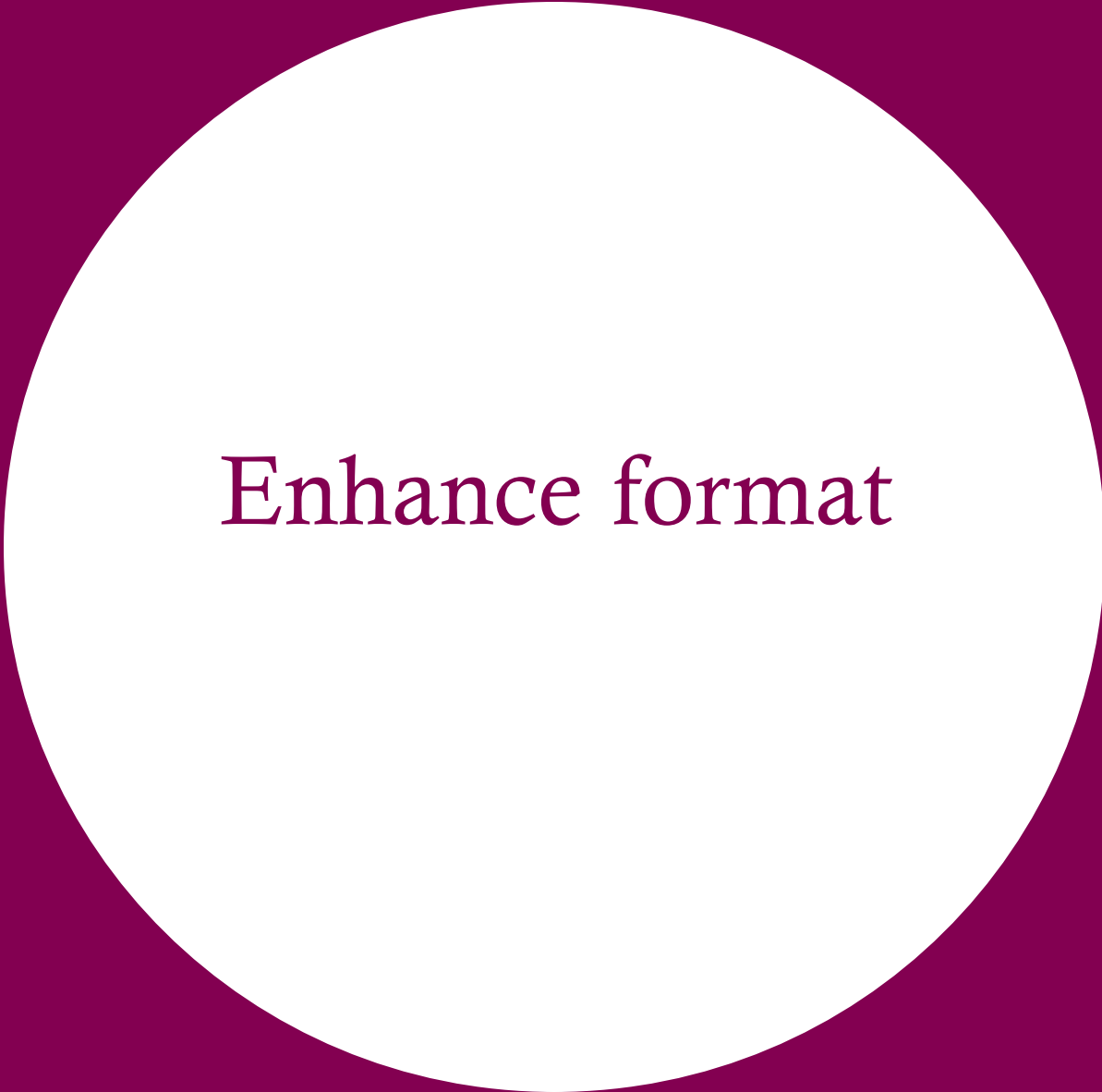


```

proc template;
define statgraph forplt;
begingraph;
layout lattice / columns=2 columnweights=(0.89 0.21) shrinkfonts=true;
layout overlay /walldisplay=none yaxisopts=(offsetmin=0 offsetmax=0.05 display= none reverse=true
linearopts=(viewmin=0 viewmax=20)) xaxisopts=(offsetmin=0 offsetmax=0.08 display=none);
-- code of column headers is omitted --
/*-- generate Baseline Risk Factors -- group column --*/
scatterplot y=n10 x=var_11/ markercharacter= var1 markercharacterattrs=(size=16.5sp)
markercharacterposition=right;
/*-- generate Baseline Risk Factors -- subgroup column --*/
scatterplot y=n10 x=var_12/ markercharacter= var2 markercharacterattrs=(size=16.5sp)
markercharacterposition=right;
/*-- add a blank column --*/
scatterplot y=n10 x=blk/ markercharacter= blank markercharacterattrs=(size=2sp);
/*-- generate Total column --*/
scatterplot y=n10 x=tt_lbl / markercharacter=trta1_1 markercharacterattrs=(size=16.5sp);
-- code of other columns is similar and omitted --
endlayout;
/*-- generate risk difference plot --*/
layout overlay /xaxisopts=(
display=(ticks tickvalues line )
linearopts=(tickvaluepriority=true tickvaluelist=( -30 0 30 ))) yaxisopts=( reverse=true display=none)
walldisplay=none;

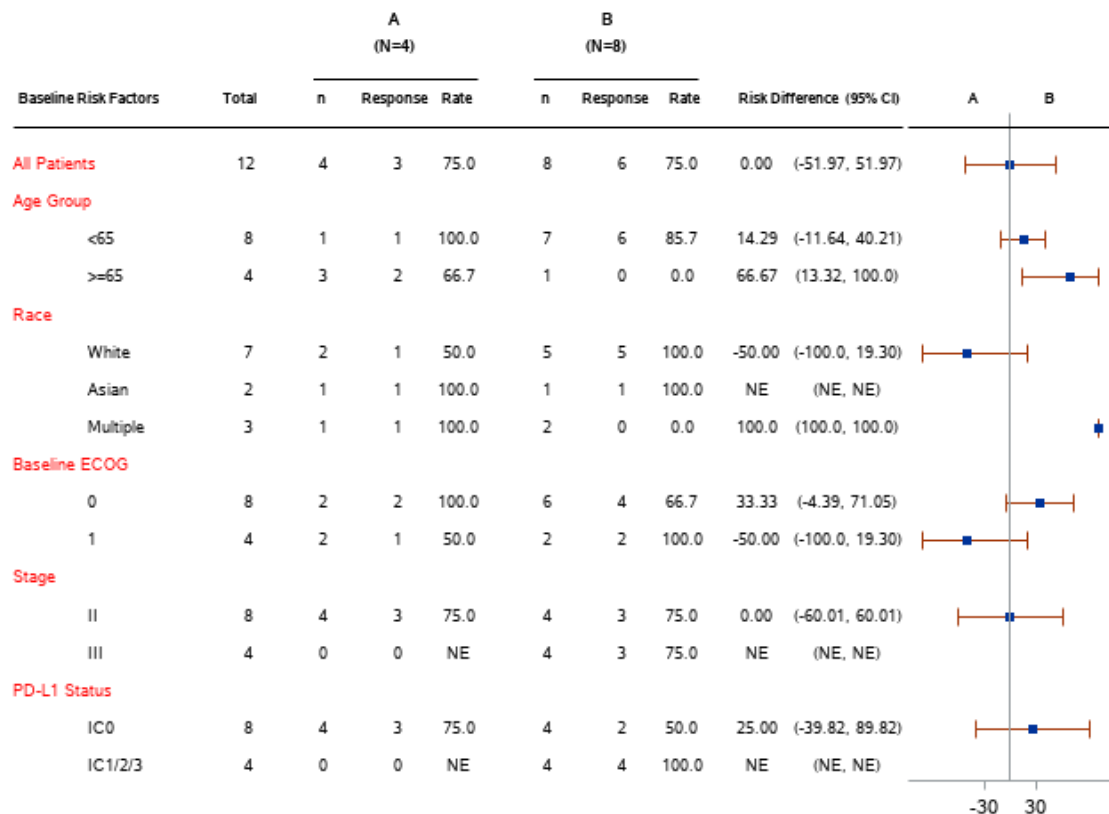
scatterplot y=n10 x=value / xerrorlower= lowercl xerrorupper= uppercl markerattrs=(symbol=squarefilled)
yaxis=y sizemin=4 sizemax=12;
referenceline x=0;
referenceline y=0 / lineattrs=(thickness=1px color=black);
endlayout;
endlayout;
endgraph;
end;

```



Enhance format

Add color to texts

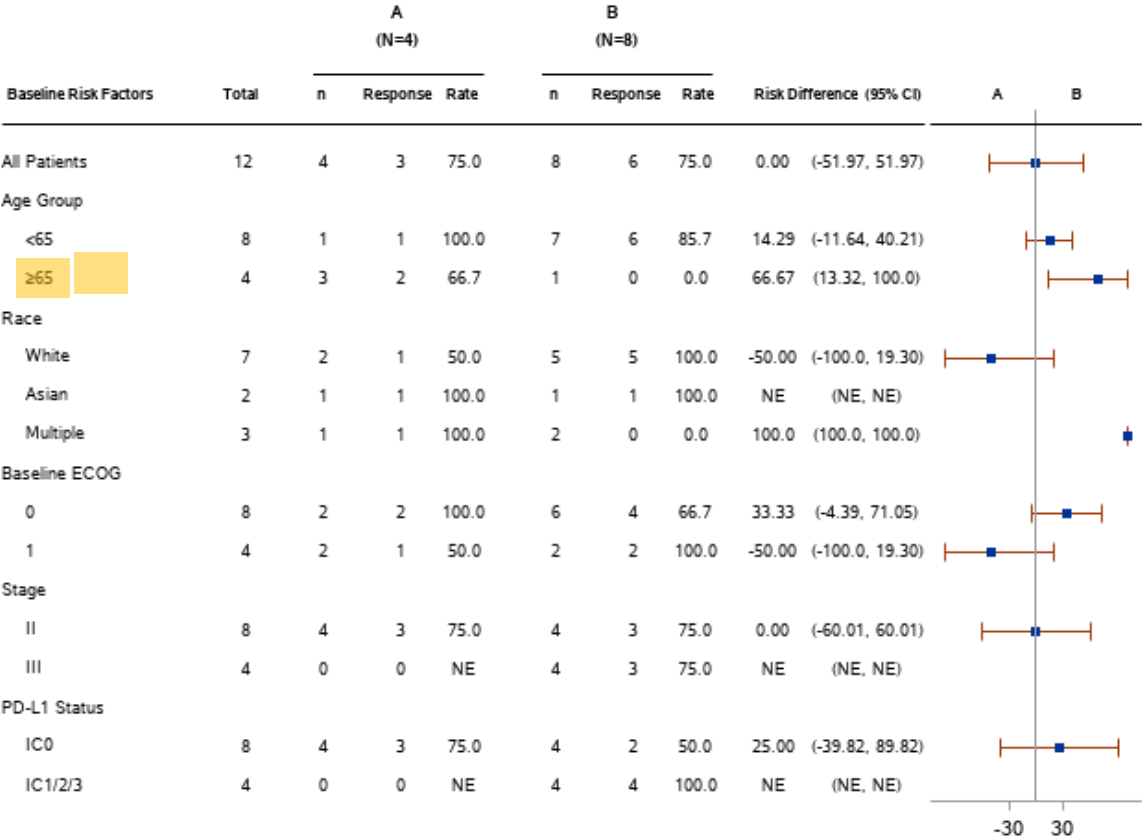


```

proc template;
define statgraph forplt;
begingraph;
layout lattice / columns=2 columnweights=(0.89 0.21) shrinkfonts=true;
layout overlay /walldisplay=none yaxisopts=(offsetmin=0 offsetmax=0.05 display= none reverse=true
linearopts=(viewmin=0 viewmax=20)) xaxisopts=(offsetmin=0 offsetmax=0.08 display=none);
-- code of column headers is omitted --
/*-- generate Baseline Risk Factors – group column --*/
scatterplot y=n10 x=var_l1/ markercharacter= var1 markercharacterattrs=(size=16.5spx color=red)
markercharacterposition=right;
/*-- generate Baseline Risk Factors – subgroup column --*/
scatterplot y=n10 x=var_l2/ markercharacter= var2 markercharacterattrs=(size=16.5spx)
markercharacterposition=right;
/*-- add a blank column --*/
scatterplot y=n10 x=blk/ markercharacter= blank markercharacterattrs=(size=2spx);
/*-- generate Total column --*/
scatterplot y=n10 x=tt_lbl / markercharacter=trta1_1 markercharacterattrs=(size=16.5spx);
-- code of other columns is similar and omitted --
endlayout;
/*-- generate risk difference plot --*/
layout overlay /xaxisopts=(display=(ticks tickvalues line) linearopts=(tickvaluepriority=true
tickvaluelist=( -30 0 30 ))) yaxisopts=( reverse=true display=none) walldisplay=none;
scatterplot y=n10 x=value / xerrorlower= lowercl xerrorupper= uppercl markerattrs=(symbol=squarefilled)
yaxis=y sizemin=4 sizemax=12;
referenceline x=0;
referenceline y=0 / lineattrs=(thickness=1px color=black);
endlayout;
endlayout;
endgraph;
end;

```

Apply Unicode

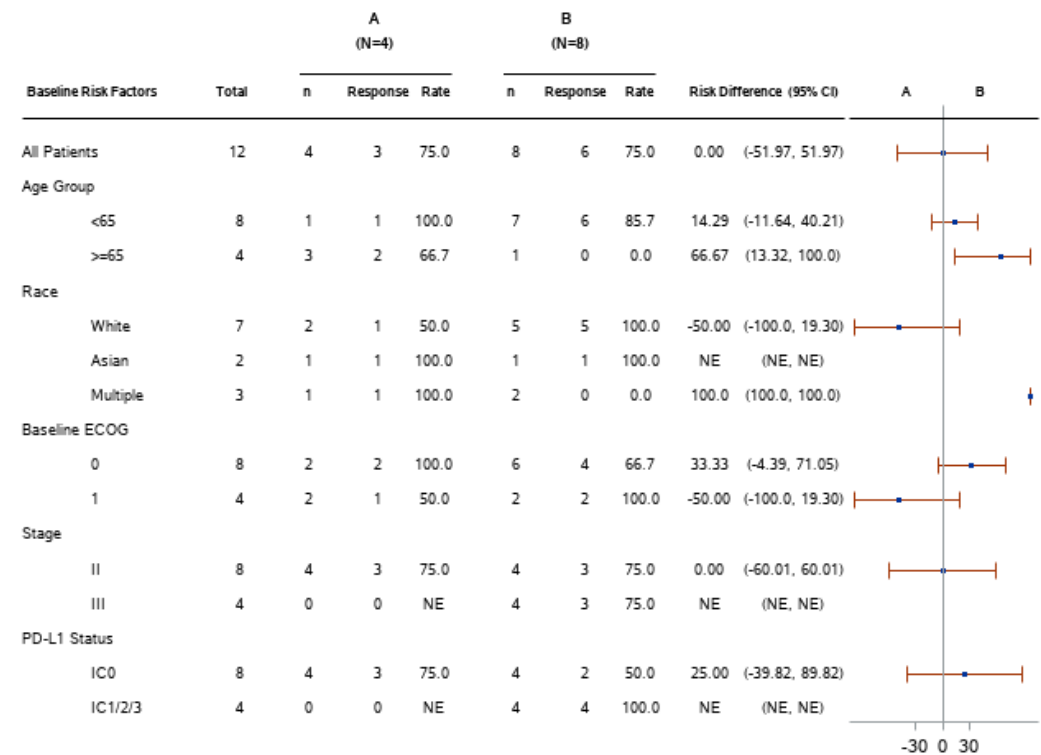


```
proc template;
define statgraph forplt;
begingraph;
  layout lattice / columns=2 columnweights=(0.89 0.21) shrinkfonts=true;
  layout overlay /walldisplay=none yaxisopts=(offsetmin=0 offsetmax=0.05 display= none reverse=true
  linearopts=(viewmin=0 viewmax=20)) xaxisopts=(offsetmin=0 offsetmax=0.08 display=none);
  -- code of column headers is omitted --
  /*-- generate Baseline Risk Factors – group column --*/
  scatterplot y=n10 x=var_l1/ markercharacter= var1 markercharacterattrs=(size=16.5sp)
  markercharacterposition=right;
  /*-- generate Baseline Risk Factors – subgroup column --*/
  scatterplot y=n10 x=var_l2/datatransparency=1;

  -- code of other columns is similar and omitted --
endlayout;
/*-- generate risk difference plot --*/
layout overlay /xaxisopts=(display=(ticks tickvalues line) linearopts=(tickvaluepriority=true
tickvaluelist=( -30 0 30 ))) yaxisopts=(reverse=true display=none) walldisplay=none;
  -- code of risk difference plot is omitted --
endlayout;
endgraph;
end;
```

```
proc sgrender data=all template=forplt ;
run;
```

Split graph into multiple pages



n10	var	var_I1	var1	var_I2	var2
1	All Patients	var1	All Patients		
2	Age Group	var1	Age Group		
3	<65			var2	<65
4	>=65			var2	>=65
5	Race	var1	Race		
6	White			var2	White
7	Asian			var2	Asian
8	Multiple			var2	Multiple
9	Baseline ECOG	var1	Baseline ECOG		
10	0			var2	0
11	1			var2	1
12	Stage	var1	Stage		
13	II			var2	II
14	III			var2	III
15	PD-L1 Status	var1	PD-L1 Status		
16	IC0			var2	IC0
17	IC1/2/3			var2	IC1/2/3
18	Strat 2	var1	Strat 2		
19	A2			var2	A2
20	B2			var2	B2
21	C2			var2	C2
22	Strat 3	var1	Strat 3		
23	A3			var2	A3
24	C3			var2	C3

```
proc sql;
create table all as
  select *, ceil(divide(monotonic(), 17)) as pg from all_abc;
select distinct max(pg) into: maxpg from all;
quit;
```


Contact Information:

Zhen (Laura) Li

E-mail: zhen.li@astrazeneca.com



Thank You

Confidentiality Notice

This file is private and may contain confidential and proprietary information. If you have received this file in error, please notify us and remove it from your system and note that you must not copy, distribute or take any action in reliance on it. Any unauthorized use or disclosure of the contents of this file is not permitted and may be unlawful. AstraZeneca PLC, 1 Francis Crick Avenue, Cambridge Biomedical Campus, Cambridge, CB2 0AA, UK, T: +44(0)203 749 5000, www.astrazeneca.com