

Bootstrap simulations to estimate overall survival based on the distribution of a historical control

Antonio Nieto (PharmaMar, Madrid, Spain) and Javier Gómez (PharmaMar, Madrid, Spain)

ABSTRACT

Following the calculation of the median overall survival (OS) in a clinical trial, it is often desirable to put the estimates into perspective by comparing them with the results of other studies reported in the current bibliography. The main limitation of this comparison is the different distribution of prognostic baseline characteristics between studies. A SAS[®] program to obtain a bootstrap estimation for the median OS, balancing it by the historical distribution, is described herein.

KEYWORDS

Bootstrap; Simulation; Overall Survival;

INTRODUCTION

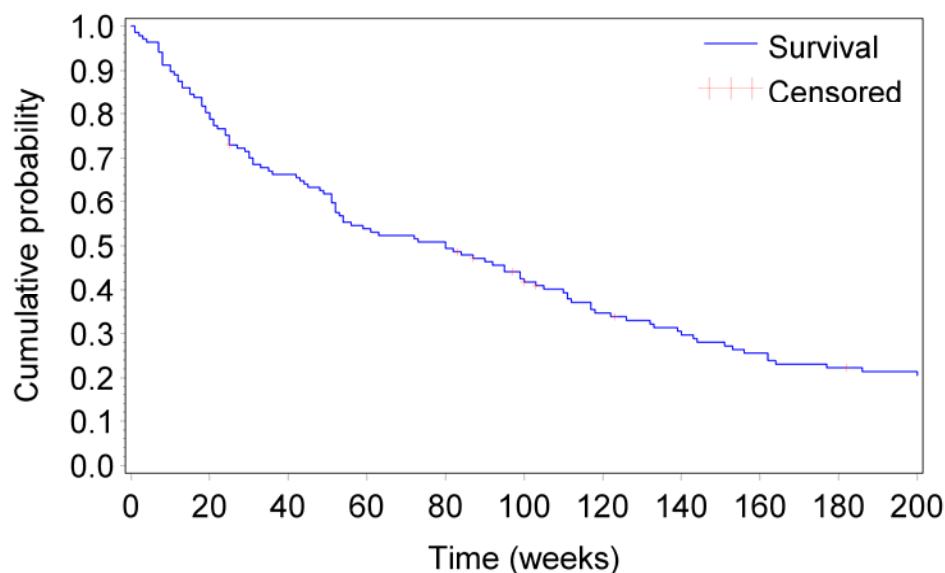
This paper has the aim of helping programmers/physicians to obtain a rough comparative of time-to-event estimation using the historical distribution by means of bootstrap replications.

Most of the background idea is based on the paper “A standardization method to adjust for the effect of patient selection in phase II clinical trials” by Mazumdar M, Fazzari M, Panageas KS (Statistics in Medicine 2001 20:883-892)¹. The code has been simplified and some features have also been added, including summary frequency tables and a plot of the whole bootstrap survival curve to check both the whole data and the particular case of the median estimate.

Standard SAS procedures² were used on a fictitious example based on a slightly modified lung SAS dataset³. This approach might be applied to the analysis of any time-to-event variable.

HYPOTHETICAL SCENARIO

The hypothetical scenario comprises a new compound with promising activity that has been evaluated in a single-arm phase II clinical trial with 137 patients. All data have been collected and are ready to be analyzed and later put into context by comparing it with the available treatment. A conventional Kaplan-Meier plot has been prepared (Figure 1) and the median overall survival (OS) of patients treated with the new compound has been estimated to be 80 weeks. According to the literature, the median OS for the standard treatment is 70 weeks.



Direct head-to-head comparison is the most common way to compare median OS values, but the validity of this method is almost always limited by differences in the distribution of prognostic baseline characteristics between studies. In the hypothetical scenario, the tumor type distribution reported in the literature for the studies that resulted in the median OS estimate for the standard treatment, based on the most relevant covariate “cell type”, was 40% squamous, 10% adenocarcinoma and 50% others. A frequency table by means of a proc tabulate was then created to find distribution imbalances that might have an effect on the median OS estimate (Table 1).

Table 1: Cell type distribution in the trial with the new compound

Cell type	N (%)
Adenocarcinoma	27 (19.7%)
Squamous	35 (25.5%)
Others	75 (54.7%)

Table 1 shows that the tumor type distribution in the phase II trial with the new compound was quite different to the one reported for the standard treatment. This in turn raised the concern about whether our data would improve or not after balancing by the relevant covariate. The following approach was prepared to check this possibility.

BOOTSTRAPPING

Bootstrapping, which was introduced by Professor Bradley Efron, is a technique based on multiple resampling with replacement of a collected sample in order to study uncertainty in the statistical sample estimate.

In the hypothetical scenario, the observed trial data sample was replicated using control distribution. Thus, resamplings for the “squamous” subset had size 55 by applying the 40% restriction reported for the standard treatment to the data on the new compound. Likewise, resamplings had size 14 for the “adenocarcinoma” subset and size 68 for the “others” subset. Ten thousand resamplings with replacement were done by means of proc surveyselect⁴.

```
proc surveyselect data=lung method=urs n=SampleSizes out=boots rep=10000
OUTHITS;
    strata Cell;
    id SurvTime censor Cell;
run;
```

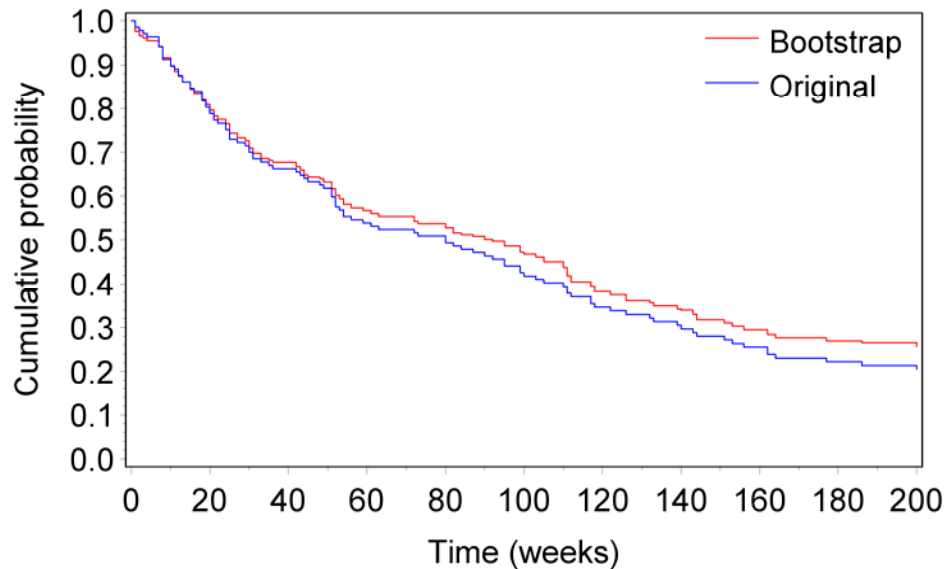
In this case, all the resamplings contained 40% of “squamous”, 10% of “adenocarcinoma” and 50% of “others”. If some uncertainty was added to the category “percentages”, instead of fixing them it would be a good idea to review whether the simulations fell within the pre-specified range.

The next step was to calculate the median OS for all the replicated samples, to store them in a dataset and to calculate the obtained bootstrap median OS estimate as the mean of the median OS in each sample. The 5% and 95% percentile confidence intervals are also presented. Table 2 shows that the median OS after correcting the imbalance in control frequency was 90 weeks.

Table 2: Bootstrap median overall survival estimate in the trial with the new compound

N	Point estimate	p5%	p95%
10000	90 weeks	59 weeks	111 weeks

The survival plots prepared with trial data for the new compound before and after bootstrapping are shown in Figure 2. The bootstrapping steps are calculated using the same approach than OS estimate as the mean of every survival timepoint in each sample.



PROCESS SUMMARY

1. Conventional Kaplan-Meier estimate of time-to-event curve and parameters for the trial.
2. Conventional control treatment estimate for the same parameters (i.e. bibliographic search).
3. Comparison of frequencies of baseline covariates for the trial and the control treatment.
4. Bootstrap replication by means of 'proc surveyselect', applying control frequencies as restrictions.
5. Review of the distribution of the control frequencies in the bootstrap simulations.
6. Proc lifetest of all bootstrap replications.
7. Median point estimate of bootstrap replication, averaging by control frequency.
8. Comparison of the whole survival curve for the original data and the bootstrap estimation.

Appendix 1 shows the main SAS code summary.

As a reminder to the reader, the method would work well when a rationally small number of categories and a reasonably large number of patients in each category are available.

CONCLUSION

This paper describes a SAS® program to obtain a bootstrap estimation for a time-to-event variable by balancing it with a historical distribution. Comparison of results from different studies is often hampered by the different distribution of prognostic baseline characteristics. The code described herein might help to obtain more accurate comparison estimates.

REFERENCES

1. A standardization method to adjust for the effect of patient selection in phase II clinical trials. Mazumdar M, Fazzari M, Panageas KS. *Statistics in Medicine* 2001 20:883-892.
2. <http://support.sas.com/documentation/> accessed on 27June2012.
3. http://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#statug_lifetest_sect018.htm accessed on 28June2012.
4. http://support.sas.com/documentation/cdl/en/statug/63962/HTML/default/viewer.htm#statug_surveyselect_sect026.htm accessed on 06July2012.

Contact Information

Your comments and questions are valued and encouraged.

Antonio Nieto Archilla
Clinical Development. PharmaMar S.A.
Avda. de los Reyes, 1
Polígono Industrial La Mina
28770 Colmenar Viejo. Madrid (SPAIN)
anieto@pharmamar.com

Javier Gómez García
Clinical Development. PharmaMar S.A.
Avda. de los Reyes, 1
Polígono Industrial de la Mina
28770 Colmenar Viejo. Madrid (SPAIN)
jgomez@pharmamar.com

Appendix 1

```

/*Executed with SAS v9.2*/

/*Dataset
http://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#statug\_lifetest\_sect018.htm
*/

*Minor edits;

Data lung;
length Therapy Cell $15.;
set VALung;
Therapy='Experimental';
If Cell='adeno' then Cell='Adenocarcinoma';
else If Cell='squamous' then Cell='Squamous';
else Cell='Other';
run;

*Conventional Kaplan-Meier estimate;
Proc lifetest data=lung OUTSURV = lung2;
time SurvTime*censor(1);
title 'Kaplan-Meier (fictitious example)';
run;title;

/*GPlot for figure 1 (omitted)*/

/*TABULATE for cell freqs (omitted)*/

/*Historical distribution 40% squamous, 10% adeno and 50% other*/
Data SampleSizes;
INPUT @1 Cell $15. @16 SampleSize @19 Proportion;
DATALINES;
Adenocarcinoma 14 0.10
Other 68 0.50
Squamous 55 0.40
;
run;

/*http://support.sas.com/documentation/cdl/en/statug/63962/HTML/default/viewer.htm#statug\_surveys\_elect\_sect026.htm*/
Proc sort data=lung;
by Cell;
run;

proc surveyselect data=lung method=urs n=SampleSizes out=boots rep=10000 OUTHITS;
strata Cell;
id SurvTime censor Cell;
run;

*Check control percentage;

Proc sort data=boots;
by Replicate;
run;

Proc freq data=boots;
table cell;
by replicate;
ods output OneWayFreqs=OneWayFreqs_cell;
run;
PROC TABULATE DATA=OneWayFreqs_cell format=comma5.0 missing;
class F Cell;
var Percent;
table F_Cell,Percent*(n median min p5 p95 max)/box='' condense indent=1;
KEYLABEL min = 'Minimum'
max = 'Maximum' all='Total';
title 'Check % in Bootstrap replications';
RUN;

*Bootstrap estimate;

ods exclude all;
PROC LIFETEST DATA=boots OUTSURV = boot2;
ods output Quartiles=Quartiles;

```

```

time SurvTime*censor(1);
by Replicate;
run;
ods select all;

PROC TABULATE DATA=Quartiles format=comma5.0 missing;
  var Estimate;
  table Estimate*(n mean p5 p95)/box='' condense indent=1;
title 'Bootstrap median OS estimation';
where Percent=50;
RUN;

*Joint of both data (Original KM and bootstrap);

Proc sort data=boot2; by SurvTime;run;
Proc means data=boot2 mean;
var survival;
by SurvTime;
output out=boot3 mean=mean;
run;

Data boot3b;
retain mean_b 1;
set boot3;
If mean<=mean_b then mean_b=mean; /*To avoid increasing steps*/
run;

Proc format;
value group
0='Bootstrap'
1='Original';
run;

Data boot4;
set boot3b;
group=0;
format group group.;
keep group SurvTime mean_b;
rename mean_b=survival;
run;

Data lung3;
set lung2;
group=1;
format group group.;
keep group SurvTime survival;
run;

Data plotdata;
set boot4 lung3;
run;

*Plot of both data;

AXIS1 COLOR=BLACK WIDTH=3
LABEL=(FONT='Arial' HEIGHT=20pt JUSTIFY=CENTER
      'Time (weeks)' )
ORDER= 0 TO 200 BY 20
MAJOR=(COLOR=BLACK NUMBER=21)
MINOR=(NUMBER=1);

AXIS2 COLOR=BLACK WIDTH=3
LABEL=(FONT='Arial' HEIGHT=20pt JUSTIFY=CENTER
      ANGLE=90 'Cumulative probability')
ORDER= 0 TO 1 BY 0.1;

symbol1 interpol=join i=stepj ci=red cv=blue l=1 width=3;
symbol2 interpol=join i=stepj ci=blue cv=blue l=1 width=3;

legend1 label=none
value=(h=2.1 j=1)
shape=symbol(5,3)
across=1
position=(TOP RIGHT INSIDE)
mode=share
offset=(-1 pct,-1 pct)
;
```

PhUSE 2012

```
proc gplot data=plotdata;  
  plot survival*SurvTime=group / vaxis=axis2 haxis=axis1 legend=legend1;  
run;  
  
QUIT;  
TITLE;  
footnote;  
RUN;  
goptions RESET=ALL;
```