

Are we really doing Independent Validation of Advanced Statistical Procedures in SAS?

Sameer Bamnote

Cytel Statistical Software & Services Pvt. Ltd., Pune, India

15 October, 2013

Disclaimer

Any views or opinions presented in this presentation are solely those of the author and do not necessarily represent those of the company.

Agenda

- 1 Introduction
- 2 Alternative Approach For Validation
- 3 Why R for validation?
- 4 Example
- 5 Summary
- 6 Conclusion

Introduction

- Statistics plays a crucial role in clinical trial reports

Introduction

- Statistics plays a crucial role in clinical trial reports
- SAS is widely used to perform the statistical analysis

Introduction

- Statistics plays a crucial role in clinical trial reports
- SAS is widely used to perform the statistical analysis
- Validation - ensure if the results produced are accurate, consistent and reliable

Introduction

- Statistics plays a crucial role in clinical trial reports
- SAS is widely used to perform the statistical analysis
- Validation - ensure if the results produced are accurate, consistent and reliable
- Present scenario:

Introduction

- Statistics plays a crucial role in clinical trial reports
- SAS is widely used to perform the statistical analysis
- Validation - ensure if the results produced are accurate, consistent and reliable
- Present scenario:
 - Source programmer: performs statistical analysis to produce the required results in SAS

Introduction

- Statistics plays a crucial role in clinical trial reports
- SAS is widely used to perform the statistical analysis
- Validation - ensure if the results produced are accurate, consistent and reliable
- Present scenario:
 - Source programmer: performs statistical analysis to produce the required results in SAS
 - Validator (QC Programmer): independent code in SAS

Introduction

- Statistics plays a crucial role in clinical trial reports
- SAS is widely used to perform the statistical analysis
- Validation - ensure if the results produced are accurate, consistent and reliable
- Present scenario:
 - Source programmer: performs statistical analysis to produce the required results in SAS
 - Validator (QC Programmer): independent code in SAS
 - However, statistical PROC used by both remains same

Introduction Contd...

	Source	Validator

Introduction Contd...

	Source	Validator
Data Processing	SAS Macro/Code	Independent SAS Code

Introduction Contd...

	Source	Validator
Data Processing	SAS Macro/Code	Independent SAS Code
Statistical Analysis	SAS PROC	SAS PROC

Introduction Contd...

- Source programmer and validator uses same PROC

Introduction Contd...

- Source programmer and validator uses same PROC
- Results will be same, as they are using the same PROC (especially while performing advanced statistical analysis)

Introduction Contd...

- Source programmer and validator uses same PROC
- Results will be same, as they are using the same PROC (especially while performing advanced statistical analysis)
- Do we call this independent validation?

Alternative Approach For Validation

- Statistical analysis plays a vital role in decision making

Alternative Approach For Validation

- Statistical analysis plays a vital role in decision making
- Can we strengthen our validation process?

Alternative Approach For Validation

- Statistical analysis plays a vital role in decision making
- Can we strengthen our validation process?
- Use of a different software for validation

Alternative Approach For Validation

- Statistical analysis plays a vital role in decision making
- Can we strengthen our validation process?
- Use of a different software for validation
- More reliable as both softwares will have different processing

Alternative Approach For Validation

- Statistical analysis plays a vital role in decision making
- Can we strengthen our validation process?
- Use of a different software for validation
- More reliable as both softwares will have different processing
- Use R (an open source software) for validation of advanced statistical analysis

Why R for validation?

- An open source software for statistical computing and graphics

Why R for validation?

- An open source software for statistical computing and graphics
- Statistical programming language with a different approach to process large and complex datasets

Why R for validation?

- An open source software for statistical computing and graphics
- Statistical programming language with a different approach to process large and complex datasets
- Highly extensible

Why R for validation?

- An open source software for statistical computing and graphics
- Statistical programming language with a different approach to process large and complex datasets
- Highly extensible
- Latest statistical techniques are available earlier

Why R for validation?

- An open source software for statistical computing and graphics
- Statistical programming language with a different approach to process large and complex datasets
- Highly extensible
- Latest statistical techniques are available earlier
- Major pharmaceutical companies and regulators started using R

Survival Analysis(Hazard Ratio)

- Background: Oncology Trial: A double blind, randomized, phase 3, international, multi-center, examining the efficacy of Treatment A in combination with Treatment B.

Survival Analysis(Hazard Ratio)

- Background: Oncology Trial: A double blind, randomized, phase 3, international, multi-center, examining the efficacy of Treatment A in combination with Treatment B.
- Objective: To demonstrate superiority of Treatment B plus Treatment A compared to Treatment A plus placebo as measured by progression-free survival (PFS), including events of radiographic disease progression per RECIST 1.1, clinical progression or death from any cause in subjects with recurrent ovarian or fallopian tube cancers.

Survival Analysis(Hazard Ratio)

Mock Shell:

	PFS Status by eCRF at Baseline			
	PFS >= 0 to <= 6 months		PFS > 6 to <= 12 months	
	Placebo + Trt B (N=xxx)	Trt A + Trt B (N=xxx)	Placebo + Trt B (N=xxx)	Trt A + Trt B (N=xxx)
Hazard Ratio (Trt A + Trt B /Placebo + Trt B)		x.xx		x.xx
(95% CI)		(x.xx, x.xx)		(x.xx, x.xx)

Survival Analysis(Hazard Ratio)

	Source	Validator
Data Processing	SAS Macro/Code	Independent SAS Code
Hazard Ratio	PROC PHREG	PROC PHREG

SAS Code:

```
proc phreg data = phreg1;  
    model months*pfsc(1) = trt / alpha = .05 ties=exact risklimits;  
    ods output ParameterEstimates=coxph;  
run;
```

Survival Analysis(Hazard Ratio)

R Code:

```
#Calling Required Library (Package)
library(survival)
```

```
#Cox Proportional Hazard Model
```

```
HR<-coxph(Surv(MONTHS, PFSC) ~ TRT_N, data=ph, method='exact')
```

```
#Printing default results
```

```
HR
```

```
#Summary of the Object HR
```

```
summary(HR)
```

Survival Analysis(Hazard Ratio)

SAS Output:

Summary of the Number of Event and Censored Values

Total	Event	Censored	Percent Censored
37	25	12	32.43

R Output:

```
coxph(formula = Surv(MONTHS, PFSC) ~ TRI_N, data = ph, method = "exact")
```

```
n= 37, number of events= 12
```

Survival Analysis(Hazard Ratio)

SAS Output:

Analysis of Maximum Likelihood Estimates								
Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limits	
TRT_N	1	0.21475	0.40498	0.2812	0.5959	1.240	0.560	2.741

R Output:

```

      exp(coef) exp(-coef) lower .95 upper .95
TRT_N      1.928      0.5186      0.5767      6.447

```

What we observe

- # of events in SAS & R outputs do not match

What we observe

- # of events in SAS & R outputs do not match
- Hazard Ratio & 95% CI do not match

What we observe

- # of events in SAS & R outputs do not match
- Hazard Ratio & 95% CI do not match
- # events in R & # censored in SAS match

What we observe

- # of events in SAS & R outputs do not match
- Hazard Ratio & 95% CI do not match
- # events in R & # censored in SAS match
- Why **MISMATCH**?

What we observe

- # of events in SAS & R outputs do not match
- Hazard Ratio & 95% CI do not match
- # events in R & # censored in SAS match
- Why **MISMATCH**?
- Reason: **Difference in the way SAS & R defines the model**

What we observe

- # of events in SAS & R outputs do not match
- Hazard Ratio & 95% CI do not match
- # events in R & # censored in SAS match
- Why **MISMATCH**?
- Reason: **Difference in the way SAS & R defines the model**
- Solution: Modify the model defined in R

Survival Analysis(Hazard Ratio)

- Surv function in R define the survival model as
Surv(time, event, ...)

R Code:

```
#Cox Proportional Hazard Model  
HR<-coxph(Surv(MONTHS, PFSC) ~ TRT_N, data=ph, method='exact')
```

R Code (Model Modified):

```
#Cox Proportional Hazard Model  
HR<-coxph(Surv(MONTHS, 1-PFSC) ~ TRT_N, data=ph, method='exact')
```

Survival Analysis(Hazard Ratio)

SAS Output:

Analysis of Maximum Likelihood Estimates								
Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limits	
TRT_N	1	0.21475	0.40498	0.2812	0.5959	1.240	0.560	2.741

R Output:

```
> summary(coxph(Surv(MONTHS, 1-PFSC) ~ TRT_N, data=p
Call:
coxph(formula = Surv(MONTHS, 1 - PFSC) ~ TRT_N, data
      n= 37, number of events= 25

              coef exp(coef) se(coef)      z Pr(>|z|)
TRT_N 0.2135      1.2380    0.4084 0.523    0.601

      exp(coef) exp(-coef) lower .95 upper .95
TRT_N      1.238      0.8077    0.5561    2.756
```

Survival Analysis(Hazard Ratio)

SAS Output:

Analysis of Maximum Likelihood Estimates							
Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limits
TRT_N	1	0.21475	0.40498	0.2812	0.5959	1.240	0.560 2.741

R Output:

```
> summary(coxph(Surv(MONTHS, 1-PFSC) ~ TRT_N, data=pl
Call:
coxph(formula = Surv(MONTHS, 1 - PFSC) ~ TRT_N, data
      n= 37, number of events= 25

      coef exp(coef) se(coef)      z Pr(>|z|)
TRT_N 0.2135  1.2380  0.4084 0.523    0.601

      exp(coef) exp(-coef) lower .95 upper .95
TRT_N    1.238    0.8077    0.5561    2.756
```

Summary

- R can be used to validate the results of SAS

Summary

- R can be used to validate the results of SAS
- Platform independent validation

Summary

- R can be used to validate the results of SAS
- Platform independent validation
- Strengthen the validation process

Summary

- R can be used to validate the results of SAS
- Platform independent validation
- Strengthen the validation process
- Ensures the validity of the results

Summary

- R can be used to validate the results of SAS
- Platform independent validation
- Strengthen the validation process
- Ensures the validity of the results
- Similarly other softwares can be explored

Conclusion

Use of another statistical software for replicating results of statistical analysis strengthens the validation process.



Questions?

Contact Detail:

Sameer Bamnote
sameer.bamnote@cytel.com