

Paper SM07

PhUSE EU Connect 2019

Automate Validation of Statistical Reports

Neel Patel

Deepak Kumar

Anik Chatterjee

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

Agenda

- Ideation
- How it works?
- Benefits & Challenges
- Process (demo)
- Conclusion
- Potential
- Questions

Ideation – Problem

- Independent programmers validate reports.
- Visual comparison is performed with original output.
- Some safety reports are 20+ pages long.
- Chances of oversight is high during manual comparison.

Ideation – Problem

- Independent programmers validate reports.
- Visual comparison is performed with original output.
- Some safety reports are 20+ pages long.
- Chances of oversight is high during manual comparison.

Multiply this problem with 10x because we have many studies.

Ideation – Solution

What can be done???

Ideation – Solution

Is it possible to convert RTF reports to SAS dataset?

The answer is YES!

We followed two-step process.

1. Read the RTF reports and convert them to SAS dataset
2. Compare that SAS dataset with independently generated dataset.

Ideation – Solution

xxx -(Week 24 analysis)

Table 14.1-2.8 (Page 1 of 43)

Relevant medical history/current medical conditions, by primary system organ class and preferred term by subgroup
Safety Set

Subgroup: Overall

Primary system organ class/ Preferred term	XXXXXXXXXXXX N=102 n (%)	Placebo N=102 n (%)
Blood and lymphatic system disorders		
Total	2 (2.0)	3 (2.9)
Anaemia	1 (1.0)	0
Antiphospholipid syndrome	1 (1.0)	0
Iron deficiency anaemia	0	3 (2.9)
Leukopenia	0	1 (1.0)
Cardiac disorders		
Total	3 (2.9)	1 (1.0)
Atrioventricular block second degree	1 (1.0)	0
Extrasystoles	1 (1.0)	0
Right ventricular dilatation	1 (1.0)	0
Arrhythmia	0	1 (1.0)

How it works?

Segregating rows into columns.

<code>\par {\f0\fs18 1.0) 0 }</code>	Anaemia	1 (1.0)	0
<code>\par {\f0\fs18 1 (1.0) 0 }</code>	Antiphospholipid syndrome	1 (1.0)	0
<code>\par {\f0\fs18 0 3 (2.9) }</code>	Iron deficiency anaemia	0	3 (2.9)
<code>\par {\f0\fs18 1 (1.0) 0 }</code>	Leukopenia	0	1 (1.0)
<code>\par {\f0\fs18 }</code>			
<code>\par {\f0\fs18 Cardiac disorders }</code>	Cardiac disorders		

How it works?

Removal of controlled characters and making it human readable.

Anaemia	Anaemia 1 (1.0) 0	16	1 (1.0)	0
Antiphospholipid syndrome	Antiphospholipid syndrome 1 (1.0) 0	17	1 (1.0)	0
Iron deficiency anaemia	Iron deficiency anaemia 0 3 (2.9)	18	0	3 (2.9)
Leukopenia	Leukopenia 0 1 (1.0)	19	0	1 (1.0)
Cardiac disorders	Cardiac disorders	21		

Benefits

- Time efficient (human eye vs. machine)
- Additional layer of check
- Saves time in repetitive checking
- Can be reused in multiple studies

Challenges

- Initial time investment for writing up QC code
- Title & footnote can not be checked

What is next?

Imagine if there is a SAS functions which can perform all discussed steps and even perform comparison...

Why FCMP?

- Functions are independent from the main program unlike macros.
- Macro variables can get mixed with global vs. local variables.
- PROC FCMP written in (essentially) in SAS syntax.
- Ease of use in SQL / DATA step etc.

Show me more!

xxx - (Week 24 analysis)

Table 14.1-2.8 (Page 1 of 42)
Relevant medical history/current medical conditions, by primary system organ class and preferred term by subgroup
Safety Set

Subgroup: Overall

Primary system organ class/ Preferred term	xxxxxxx N=102 n (%)	Placebo N=102 n (%)
Blood and lymphatic system disorders		
Total	2 (2.0)	3 (2.9)
Anaemia	1 (1.0)	0
Autophospholipid syndrome	1 (1.0)	0
Iron deficiency anaemia	0	3 (2.9)
Leukopenia	0	1 (1.0)
Cardiac disorders		
Total	3 (2.9)	1 (1.0)
Atrioventricular block second degree	1 (1.0)	0
Extrasystoles	1 (1.0)	0
Right ventricular dilatation	1 (1.0)	0
Arrhythmia	0	1 (1.0)
Congenital, familial and genetic disorders		
Total	7 (6.9)	3 (2.9)
Dolichocolon	1 (1.0)	0
Factor XII deficiency	1 (1.0)	0
Gilbert's syndrome	1 (1.0)	0
Larsen syndrome	1 (1.0)	0
Syringomyelia	1 (1.0)	0
Thalassemia beta	1 (1.0)	0
Thalassemia minor	1 (1.0)	0
Birth mark	0	2 (2.0)
Factor II mutation	0	1 (1.0)

- The above will be summarized by SOC in alphabetic order and PT within each SOC by the descending order of frequency.
- Relevant medical history/current medical conditions will be coded using MedDRA VERSION <22.0>.

/vob/xxxxx/xxxxxxxxxxxx/csr 2/pgm/saf/c mh 01.sas08/main/10 30SEP19:18:22 Final

PAGE 1 OF 43 9639 WORDS ENGLISH (UNITED STATES)

Conclusion

- Automation helps in faster output review.
- Saves lot of time and efforts (mostly for safety reporting).
- Reduce the chances of oversight during visual review process

Acknowledgment

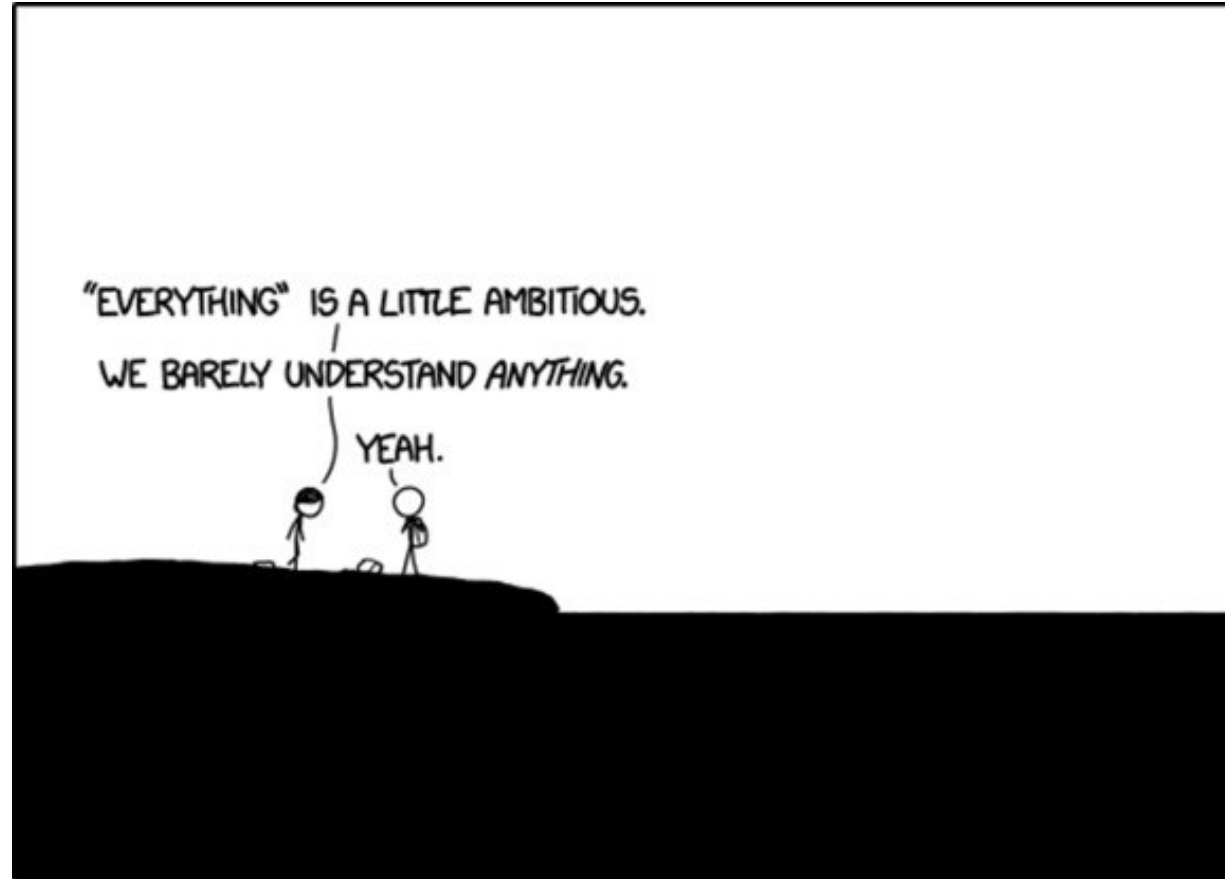
- Anik Chatterjee (Team Lead, Novartis) for original ideation and assistance. His constant guidance helped through ideation to final paper and presentation.
- Sean M. Burke for writing a beautiful RTF Pocket Guide, that small book explains the syntax of RTF with examples in easy language and it helped us a lot.

Potential

- Automation helps in faster output review.
- Saves lot of time and efforts.



Questions



Thanks

Neel Patel

neel.patel@yahoo.com

Deepak Kumar

deepak306@gmail.com

Anik Chatterjee

anik.chatterjee@novartis.com