PhUSE US Connect 2019

Paper DH12

Analysis Results that Save Trees - #KillTFLs

Chris Decker, d-wise, Morrisville, NC, USA

ABSTRACT

In today's clinical programming environment, we spend large amounts of time (and money) producing Tables, Figures, and Listings – single use, throw-away artifacts that I would challenge focus more on the cosmetics then the content. Artifacts that, by their own admission, are rarely looked at by regulatory reviewers. We build these reports for one narrowly focused outcome instead of making that result, and many more, available for many different uses. What if we could capture and store analysis results that gave the results context and allowed a clinician or reviewer to dynamically search the results within and across studies without sifting through 300 pages of paper? We can do this! This presentation will provide an overview of a project where a graph-based analysis results data store was designed and populated to capture and make analysis results available to many different stakeholders reducing their workload by 50%. We can leverage new technology and/or methodologies to optimize our archaic processes and #killTFLs.

INTRODUCTION

During a clinical trial we generate a lot of analyses with the goal of proving the safety and efficacy of a product. In most cases we create very prescriptive outputs that are mapped to the key endpoints we have predefined within our protocol and statistical analysis plans. We become very narrowly focused on delivering the infamous section 14 of the clinical study report. We build these reports with one focused outcome and don't think about how the numbers being generated could be potentially be used for other needs.

This presentation will provide you an overview of how we ended up at where we are today, challenge the need to continue down this path, and introduce you to a new way of capturing, storing, and using analysis results. We can move the industry in a new direction if we just push a bit harder.

WE'VE ALWAYS DONE IT THAT WAY

One of the most dangerous phrases in our language is "We've always done it that way". That one phrase shuts down innovation and has proven to be very expensive in business. The only constant in our world is change which is juxtaposed to the natural resistance to change. The resistance to change in the TFL case study can lead to longer timelines, a lack of focus on what is important, and piles of extra work all impacting the cost and profit of the work being done.

IN THE BEGINNING

Twenty-five years ago, when I first started in this industry, one of my first tasks was to adopt a set of existing SAS code that generated a set of tables and incorporate feedback from the clinical review group. As I sifted through the spaghetti code, I learned a lot about data _null_ and placing things on a piece of paper and learned very little about the clinical science or statistics I was working with. I began to wonder why I worked on getting multiple degrees in statistics only to focus most of my time making sure the columns were aligned, I properly counted page numbers and had superscripted my footnotes.

Another fascinating learning in those early days was that when we received comments back from the clinical stakeholders, 80% of those comments again were focused on how things looked and not the value of the results. I was naïve to the industry but still scratched my head trying to understand why.

DOES ANYONE PAY ATTENTION?

Fast forward to 2019 and I haven't generated a TFL in over 15 years. However, in 2018 we engaged in a project with a customer to build them a SAS based TFL generator. And guess what? The solution was developed to target a set of single use outputs and 60% of the code written was to make the single those outputs 'pretty'. So, twenty-five years later we continue to follow the same process, deliver the same outputs, and I continue to scratch my head.

Demographic Summary

Page 1 of 2
9:15 11JAN2018
DDDM

Intent-to-Treat Population

Demographic Parameter		drug c (N=69)	drug b (N=71)
Sex	n*a	69	71
	F	31 (44.9)	39 (54.9)
	M	38 (55.1)	32 (45.1)
Age (yrs)	n*a	69	71
	Mean	56.49	58.24
	Std. Dev.	7.96	9.90
	Median	57.00	59.00
	Min	38.00	30.00
	Max	71.00	77.00
Race	n	69	71
	AMERICAN INDIAN OR ALASKA NATIVE	3 (4.3)	8 (11.3)
	ASIAN	2 (2.9)	0 (0.0)
	BLACK OR AFRICAN AMERICAN	6 (8.7)	7 (9.9)
	WHITE	55 (79.7)	55 (77.5)
	MULTIPLE	3 (4.3)	1 (1.4)
Ethnicity	n*a	69	71
	HISPANIC OR LATINO	20 (29.0)	23 (32.4)
	NOT APPLICABLE	5 (7.2)	4 (5.6)
	NOT HISPANIC OR LATINO	44 (63.8)	44 (62.0)
Weight (kg)	n*a	69	71
	Mean	93.27	92.49
	Std. Dev.	18.32	18.20
	Median	92.80	93.10
	Min	52.20	57.30
	Max	139.00	130.40

Abbreviations: N = number of subjects in population; n = number of subjects;
NC = not calculable;
*a - number of subjects with non-missing data, used as denominator
Program Execution Location:
Output Execution Location:
Data Set Execution Location:

Figure 1: The old-fashioned Demographics table

The figure above is a symbol of the same 'TFL' we have been generating for the last three decades. My fundamental question is whether anyone pays attention to these hundreds of pages of paper and what value does the targets customer put on the content. During conversations with my FDA informants (who I will not name to protect their identity) they overwhelmingly said they do not review these outputs in any detail and prefer to start with the data and generate their own analysis. In complete transparency, they don't really trust the results sponsors give them because they can't establish rigorous traceability from the results back to the collected data point. The lack of true traceability is a much bigger fish to fry and one for another paper.

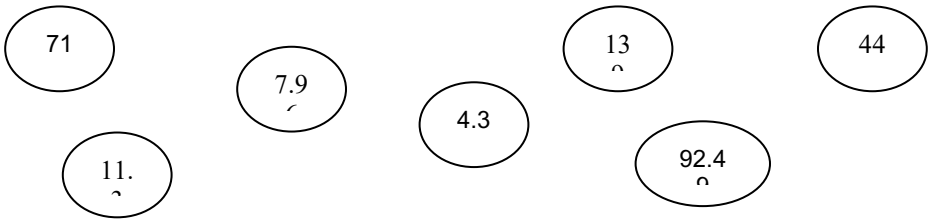
THE REAL USE CASE

So, the real question you are asking now is 'So what'. Why do we care about storing this information? The only goal I have is producing the TFLs for my section 14 of the clinical study report. Again, the figure above is a symbol of the same 'TFL' we have been generating for the last three decades. An artifact that has served one single purpose for just as long with no one pausing to say, "Could this data be valuable and reusable?".

This narrow focus is the fundamental problem that exists in todays way of thinking. If we only focus on a single deliverable, albeit one of the most important, we lose track of how valuable this information could be for other stakeholders.

ANALYSIS FOCUSED AND CONTEXT

As I step back and study this table, I want to share a few key concepts. First, the only way to really understand this table is to have the institutional knowledge of the content. By itself, the information in this table has no meaning. The number 71 in the second column has no context without a human interpreting the table and telling you what it means. Below I selected a random set of numbers from the table and when the numbers are not in the table, a person has no idea what they mean.



However, if we could add context to each of those numbers, and more importantly store that context in a machine, we could reuse those numbers and get significantly more value out of analysis result.

Context/Number	71	7.96	11.3	4.3	139	92.49	44
The number is a count?	✓						✓
The number is for drug C?		✓			✓		✓
The number is in the ITT population?	✓	✓	✓	✓	✓	✓	✓
The number is a statistic?		✓			✓	✓	
The number if a percentage?			✓	✓			
The number is for the Weight parameter?					✓	✓	

If we could easily collect context about each number, we could do so much more with the numbers then the one-time throwaway TFL.

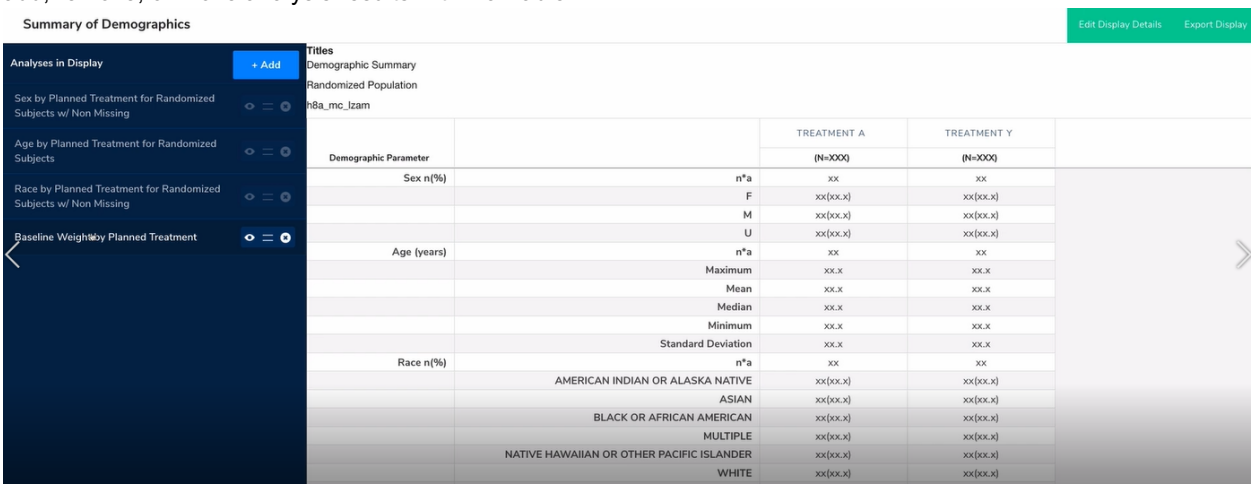
Second, the unit of measure we have always focused on is the table itself. The reality is that this table contains many different analyses, each with its own meaning and context. Summary of Demographics is paradigm that creates a presentation of many different analyses. Through our own doing, we have muddled presentation with the need to define and generate individual analysis results creating a partnership that has no need to exist and limits what we can do.

The single analysis needs to be our focus and the core of what we define. By creating this more granular unit of measure we can define context as described in the table above. The number 139 is for the Weight parameter, is define for drug C, was within the ITT population, and is a statistic. By defining and storing this context systematically, there is so much more value we can gain from the information.

DYNAMIC AND MULTIPLE NEEDS

If we step back and forget about the need to produce those pesky section 14 TFLs for the clinical study reports that might never get reviewed, we should think about the other stakeholders that might benefit from having these numbers at their fingertips. In the project we engaged with a customer they had three key stakeholders who needed to consume the analysis results downstream.

The first stakeholder was the traditional need for the TFLs within the regulatory group. Within their current process they went through a labor-intensive process to compile and insert the TFLs into their final report. There were many iterations with the programming group making the TFLs 'just right' and most of those changes were cosmetic with little to no impact on the results. When we prototyped the ability to build their own TFL based on the more granular analysis results they immediately saw the efficiency and value it would bring for them to 1) significantly reduce the iterations and 2) build custom TFLs to meet their needs. The figure below shows a prototype where the user could add, remove, or move analysis results within a Table.



The second stakeholder was the medical review team who had put another time intensive process in place to extract numbers from the hundreds of TFLs to populate the CSR, manuscripts, and other external publications. Within their process they would manually transcribe the numbers from the table each time they were generated into a fancy Excel tool which would then populate the various downstream needs. When we showed them that we could give them the individual numbers in more detail with context where they could build rules to automate the transcription, they profusely thanked us for eliminating days of work and reducing significant risk with human intervention.

The final stakeholder we explored was an analytical group that was investigating methods they could use to review meta-analysis across studies. Within their current environment, they had a long-drawn-out process which was resource intensive to generate integrated databases to perform cross study analyses. They wanted an option to look at analysis results within their tables across studies. When we showed them that we could store these analyses within and across studies and provide context it met many of their objectives. They could answer questions easily such as:

- Which studies analyzed blood pressure for females and display those analyses?
- Which of my studies collected a LOCF pain score analysis at week 8 and combine those analyses?
- How many subjects were in my safety population across these 3 pivotal studies?

These three case studies were just a sample of what we tackled within the project and all of them showed significant value to the end users.

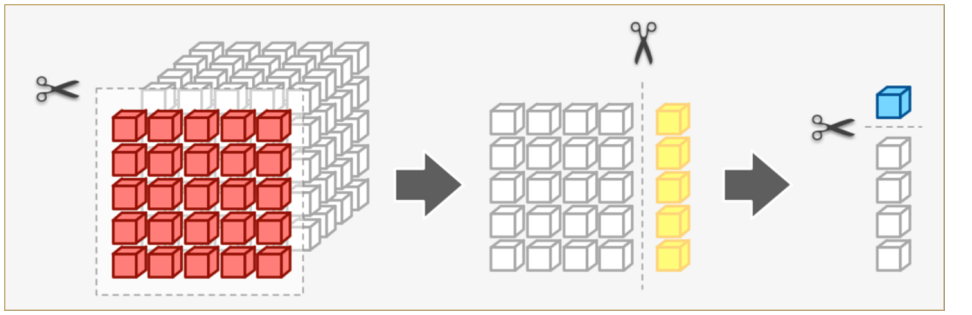
WHERE DO WE BEGIN

So now that you are starting to believe in the value of thinking about analysis results in a different way, the next step is to realize that the technology part is actually easy. Well sort of. If we continue to try to develop our analysis results using the same process, then we will struggle to conform our current data processing outputs into a new model, but we must start somewhere.

Also, I have always wondered why we are stuck generating the presentation layer of our TFLs with SAS. SAS is a powerful programming and analysis tool when it is used for the right purpose. However, this is another example of using a hammer for everything because it's the only tool within the clinical programmer toolbox. There are significantly better technologies out there for reporting, yet we continue to use PROC Report, data _null_ and other tools within SAS. Part of the reason is that it's all we have ever known and the concept of extracting the generation of the analysis results from the presentation is something people struggle to grasp.

MODELING AND STORING THE CONCEPT

The first and hardest part is defining the model to capture the contextual information tied to an analysis result. While there are conceptual frameworks out there for defining an analysis (RDF Data Cube referenced at the end of the paper), the real work is defining what the concepts are within that framework. You can define an initial model based on a handful of TFLs, but as you add more unique types of analyses you will need to iterate and add/update the model to accommodate and optimize your model. This process takes time but will add value if done right. The idea is to break down each table into the analyses and then define the context for each value in the analysis.



There are many different technologies that can be used to store this contextual information, but the most robust method is to use a technology that support the concept of semantic based triple stores. This can be done with RDF (as referenced below), property graphs, or rudimentary excel spreadsheets (which will limit the power but can be done). In all cases though you need to store the information in a way that captures the contextual information for each analysis and value. Below are examples of how you could model the information but there are many ways to tackle this information.

The table below shows what a set of dimensions might look like for a small subset of TFLs. These are the levels of the context that might address TFLs such as demographics, parameter and/or visit based analyses, or event driven (e.g. adverse events) activities.

Dimension	Label
sponsorqb:dim.population	Population
sponsorqb:dim.treatment	Treatment
sponsorqb:dim.parameter	Analysis Parameter
sponsorqb:dim.visit	Visit
sponsorqb:dim.baseline	Baseline Indicator
sponsorqb:dim.statistic	Statistic
sponsorqb:dim.sex	Sex
sponsorqb:dim.agecat	Age Category
sponsorqb:dim.soc	System Organ Class
sponsorqb:dim.pt	Preferred Term

Once you have defined all the dimensions (and you will need to iterate on the dimensions as you add more analyses) you can use those to define each analysis/value in your TFLs.

qb:Observation	qb:Table	dim.population	dim.treatment	dim.parameter	dim.sex	dim.agecat	dim.statistic	analysisResult
1001	dm.summary	enrolled	Treatment.A	param.subjects	sex.ALL	agecat.ALL	stat.freq	100
1002	dm.summary	enrolled	Treatment.A	param.subjects	sex.F	agecat.ALL	stat.freq	60
1003	dm.summary	enrolled	Treatment.A	param.subjects	sex.F	agecat.ALL	stat.percent	60
1004	dm.summary	enrolled	Treatment.A	param.subjects	sex.M	agecat.ALL	stat.freq	40
1005	dm.summary	enrolled	Treatment.A	param.subjects	sex.M	agecat.ALL	stat.percent	40
1006	dm.summary	enrolled	Treatment.B	param.subjects	sex.ALL	agecat.ALL	stat.freq	50
1007	dm.summary	enrolled	Treatment.B	param.subjects	sex.F	agecat.ALL	stat.freq	30
1008	dm.summary	enrolled	Treatment.B	param.subjects	sex.F	agecat.ALL	stat.percent	60
1009	dm.summary	enrolled	Treatment.B	param.subjects	sex.M	agecat.ALL	stat.freq	20
1010	dm.summary	enrolled	Treatment.B	param.subjects	sex.M	agecat.ALL	stat.percent	40
1011	dm.summary	enrolled	Treatment.ALL	param.subjects	sex.ALL	agecat.ALL	stat.freq	150
1012	dm.summary	enrolled	Treatment.ALL	param.subjects	sex.F	agecat.ALL	stat.freq	90
1013	dm.summary	enrolled	Treatment.ALL	param.subjects	sex.F	agecat.ALL	stat.percent	60
1014	dm.summary	enrolled	Treatment.ALL	param.subjects	sex.M	agecat.ALL	stat.freq	60
1015	dm.summary	enrolled	Treatment.ALL	param.subjects	sex.M	agecat.ALL	stat.percent	40
1016	dm.summary	itt	Treatment.A	param.age	sex.ALL	agecat.ALL	stat.freq	100
1017	dm.summary	itt	Treatment.A	param.age	sex.ALL	agecat.ALL	stat.mean	40.7
1018	dm.summary	itt	Treatment.A	param.age	sex.ALL	agecat.ALL	stat.stdev	10.7
1019	dm.summary	itt	Treatment.A	param.age	sex.ALL	agecat.ALL	stat.median	37.0
1020	dm.summary	itt	Treatment.A	param.age	sex.ALL	agecat.ALL	stat.min	21.0
1021	dm.summary	itt	Treatment.A	param.age	sex.ALL	agecat.ALL	stat.max	66.0
1022	dm.summary	itt	Treatment.B	param.age	sex.ALL	agecat.ALL	stat.freq	50
1023	dm.summary	itt	Treatment.B	param.age	sex.ALL	agecat.ALL	stat.mean	41.2
1024	dm.summary	itt	Treatment.B	param.age	sex.ALL	agecat.ALL	stat.stdev	10.3
1025	dm.summary	itt	Treatment.B	param.age	sex.ALL	agecat.ALL	stat.median	36.0
1026	dm.summary	itt	Treatment.B	param.age	sex.ALL	agecat.ALL	stat.min	23.0
1027	dm.summary	itt	Treatment.B	param.age	sex.ALL	agecat.ALL	stat.max	67.0
1028	dm.summary	itt	Treatment.ALL	param.age	sex.ALL	agecat.ALL	stat.freq	150
1029	dm.summary	itt	Treatment.ALL	param.age	sex.ALL	agecat.ALL	stat.mean	40.9
1030	dm.summary	itt	Treatment.ALL	param.age	sex.ALL	agecat.ALL	stat.stdev	10.4
1031	dm.summary	itt	Treatment.ALL	param.age	sex.ALL	agecat.ALL	stat.median	37.0
1032	dm.summary	itt	Treatment.ALL	param.age	sex.ALL	agecat.ALL	stat.min	21.0
1033	dm.summary	itt	Treatment.ALL	param.age	sex.ALL	agecat.ALL	stat.max	67.0

OPERATIONALIZING THE PROCESS

So up to this point we have discussed the concept of breaking down the analysis results into more granular components, talked about use cases where this would add value or increase efficiency in the process, and showed a conceptual model. However, the part of implementing a solution like this is the always the hardest– the people and processes. Determining how and when to operationalize a new approach described above takes time, effort, and change management.

Pick the Right Stakeholder

The first step is picking the right stakeholder to 'experiment' with a new process as described above. In general, we always recommend not trying to boil the ocean in the first pass and instead target a smaller focused group that can benefit from the new process. In this case, trying to implement the analysis results database in the core TFL development workflow that is on the critical path for submission is probably not the best option. We recommend targeting secondary stakeholders such as clinicians, medical writers, or other groups who would gain value from

being able to explore the analyses in a more interactive way but whose needs are not urgent. You can also use these groups to work out the kinks in defining, populating, and making the analysis results available to them.

Start Small – Quick Wins

The next step is to start small and find quick wins that show value. For example, define the analysis results for all your variation of your adverse events and then be able to answer questions about this topic within and across studies showing the value of combining this data. Start with the analyses that are well defined and use most often across studies. This will create the most bang for the buck and provide a good foundation for the model you are building. You can continue to expand from that core set of analyses still focused on the most used analyses across studies. Then when your foundation is solid, you have worked out all the kinks, and you have a more mature operational process, you can begin to expand the capability to therapeutic area analyses building on what you have put in place.

CONCLUSION

In summary, despite the title of my paper and my opening statement, I am not actually trying to kill the concept of presenting analyses in a way that stakeholders can visually digest them. I am trying to encourage us to rethink the path of how we get there. Let's not focus only on the piece of paper, but more importantly think about the analyses we are trying to generate, the myriad of stakeholders that need this information, and how having the ability to reuse this information within and across studies will significantly add value to understanding the safety and efficacy of products we are delivering.

REFERENCES

PhUSE Analysis Results Model Project, 2016. Tim Williams and Marc Andersen.
https://www.phusewiki.org/wiki/index.php?title=Analysis_Results_Model

ACKNOWLEDGMENTS

- Frederik Malfait for introducing semantics to our industry and bringing us into the 21st century.
- Tim Williams and Marc Andersen for introducing our industry to the concept of treating analysis results with a different mindset.
- Scott Bahlavooni for incubating the idea with me and making it real versus an academic exercise.

RECOMMENDED READING

W3C RDF Data Cube Vocabulary. https://www.w3.org/TR/2014/REC-vocab-data-cube-20140116/

CONTACT INFORMATION

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

Chris Decker
d-wise
1500 Perimeter Park Drive
Morrisville, NC USA
Work Phone: 919-995-0957
Email: chris.decker@d-wise.com
Web: www.d-wise.com