PhUSE 2018

Paper CD11 **Final 1.0**

How Do You Determine # SDTM's, ADAMs and TLFs and its impact on Outsourcing?

Roy Ward, PPD, Alsace, France

ABSTRACT

We are all familiar with CDISC processes, domain and TLF production to support study reporting and submissions. However how accurate are the SDTM, ADaM and TLF estimates made at the start of the study? Are we aware of the consequences of estimating incorrectly? After all, these estimations strongly determine the costings of a study, whether it be the amount of internal resource required or the dollar cost of outsourcing. As one of the leading sources of "Out of Scope" discussions between CRO's and sponsors are mismatches between predicted and actual numbers of outputs, why are expectations so different and are enough discussions taking place at the bidding stage?
This presentation focuses on the need for good initial estimates and the methods that drive them leading to a discussion regarding variations between companies. The presenter has experience from the pharmaceutical and the CRO world and will share personal opinions.

INTRODUCTION

With spiraling pharmaceutical costs, companies are looking to many ways to reduce costs in order to contribute to high value activities. A contributor to costs are number of SDTM's, ADAM and TLFs produced which appear to be rising in number.

However how accurate are the SDTM, ADaM and TLF estimates made at the start of the study? Are we aware of the consequences of estimating incorrectly? One of the leading sources of "Out of Scope" discussions between CRO's and sponsors are mismatches between predicted and actual numbers of outputs, why are expectations so different and are enough discussions taking place at the bidding stage?

This presentation focuses on the need for good initial estimates and the methods that drive them leading to a discussion regarding variations between companies. The conversation is then around why is this happening? Should it be happening. Do we understand the cost and can it be managed?

The presenter has experience from the pharmaceutical and the CRO world and will share personal opinions.

WHERE ARE WE NOW IN THE INDUSTRY?

Most methods of estimations within the industry appear to be based on internal baselines. For example, the last submission in the CV indication the company has produced X SDTMs, Y ADAMs and Z TLFs per study. This becomes the baseline with a little information "outside in" maybe from new recruits who bring the outside information inside. Let's look at the methods of estimation and control a little more closely.

A basic method of resource estimation is a blanket estimation of resources based on the complexity of the study hence a) low b) moderate c) complex d) very complex e) mega trial. This is usually a very quick statistic to derive and ignores # of TLF's etc. However how broad is the definition of the complexity of a study and who is defining it. For example, one would imagine a phase II hyper-tension study would be of low complexity however a phase III Oncology study would be of high complexity. However, what is the case if the phase III identical Oncology study is the 5th in a series of studies hence programs are set up and studies are standard. Compare this to the phase 2 study which may be a study which follows few standards and in-licensed from a company which no longer exists. This is no longer a low complexity study. Is the statistician/programmer or clinician defining complexity.

An advancement in the level of complexity is to split the deliverable into three categories

a) unique

- creating for the first time.

b) repeat

- taking an already created TLF but sub-setting e.g. demography by gender.

b) re-run.

- Changing no code but re-running the output with new data.

This can be further subset by considering efficacy versus safety tables, safety tables of course being considered more complex.

TLFS

- estimation.

a) Internal benchmark: what was previously done in a similar study in a similar phase.
b) A range is produced expect between n1 and nZ TLFs.

- Driver.

a) Statistician.
b) Programmer
c) Clinician.
d) PM

- Control

a) Statistician policies but against which baseline?
b) Programmer raises when resource needs exceed what has been predicted.
c) PM policies but against which baseline?

ADAM

- Estimation

a) Flat rate
b) Benchmark rate
c) Function of #SDTMs.

- Driver.

a) Statistician.
b) Programmer
c) Standards group
d) Resource lead

- Control

a) Programmer raises when resource needs exceed what has been predicted.

SDTM

- Estimation

a) Flat rate
b) Benchmark rate
c) Function of # unique CRF's

- Driver.

a) Statistician.
b) Programmer
c) Standards group
d) Resource lead
e) DM

- Control

a) Programmer raises when resource needs exceed what has been predicted.

CONSEQUENCIES OF INCORRECT ESTIMATIONS

From an internal perspective the obvious consequences for miscalculating # SDTM's, ADAM and TLF's is not only incorrectly allocating the amount of resource but also the incorrectly allocates the type of resource. As a crude example, internally if 200 TLFS are estimated and 300 are required this will lead to either

a) internal pressure to speed up production
b) work over-time.

The speeding up of production will result in on-boarding talented and expensive resource who are probably not familiar with the study which will then result in an immediate dip in efficiency as study based knowledgeable resource is switched away from producing from producing TLFs to educating and leading newly recruited personnel. This leads to inefficiencies.

If expensive contracting resource is brought on board this may result in internal resentment as the permanent and probably more key members of the team are not immediately getting recognized financial compensation whilst the contractors are clearly getting remunerated. Hence contractors getting paid by the hour and permanent folk not.

Working overtime will over a sustained certain period of time will introduce fatigue which will uncharacteristically introduce "school boy errors". Overtime for a short period of time is fine, over a prolonged period it becomes a chronic issue/company expectation.

From a CRO perspective incorrect TLF estimations, lead to incorrect resource estimations will to lead "out of Scope" discussions. It is in the interest of both parties to avoid these discussions. Scope discussions therefore should take place before the work is performed by the CRO thus avoiding any awkward discussions post-delivery. Hence number of SDTMS, number of ADAMs and number of TLFs and complexity should all be considered and agreed between interested parties. CRO's share the same goal as pharmaceutical companies in that they need to get the correct information to the correct people as fast as possible and in high quality. However extra work leads to extra time which results in extra cost which is far more visible and tracked in a CRO. Hence the incorrect estimate will surface earlier in a CRO.

CRO's generally are more resource flexible than pharma e.g. a resource is not assigned to a TA but can flexibly move from one TA to another and across phase. Therefore, whilst resource can be moved faster within CRO's to wherever the need is, CRO's share the same issues as pharma companies as last minute unplanned requests for resources is likely to be a challenge and will cause burn with the leads in attempting to resolve a resourcing challenge. Hence for an acceleration or unplanned work, "costs" needs to be considered. A topic non-Bios folk sometimes struggle to grasp.

Complexity is often overlooked. For example, a simple cardio-vascular efficacy dataset is easier to produce compared to a complex oncology RECIST dataset yet in resource estimations how are they being differentiated? How are unique' s, repeats and re-runs measured all in the same statistic? All we all defining unique, repeats the same and do we need to go to that level of granularity.

CONTROL

Most of the solutions above appear to have as the control a programmer who raises the issues when resources supplied are too minimal for what is truly required. Surely this is too late in the process and fuels inadequate resourcing or Out of Scope requests for CRO work? So, what possible methods are in place to control # deliverables?

Considering TLFs clearly the worst-case scenario is giving a medical colleague a free range of what to request for a deliverable ("Isn't this all push button anyway?"). In almost all situations a statistical colleague is advising the medical colleague of whether the deliverable makes statistical sense and adds value. However, in some cases the statistician has little authority to challenge the medical colleague resulting in this inflated figure merely being passed through to programming to execute. Some companies do apply "an expected range" or even better, a "must be discussed with management if upper range exceeded" process.

Most companies in the business of outsourcing operate their own UNIT GRID of costs e.g. cost for Data Listings, SDTM, ADAMS. TLFS, Project management. However most of these company UNIT GRIDs are unique to that company meaning a mapping exercise must occur from pharm company UNIT GRID to CRO UNIT Grid. How many misalignments occur here during the mapping phase, hence is this truly comparing like for like?

A good number of operational people use the "Cost per TLF" statistic which is a good quick and easy guide to costing, taking the overall cost as a numerator. The question then becomes is the nominator "all TLFs in including re-runs" or just "unique and repeats".

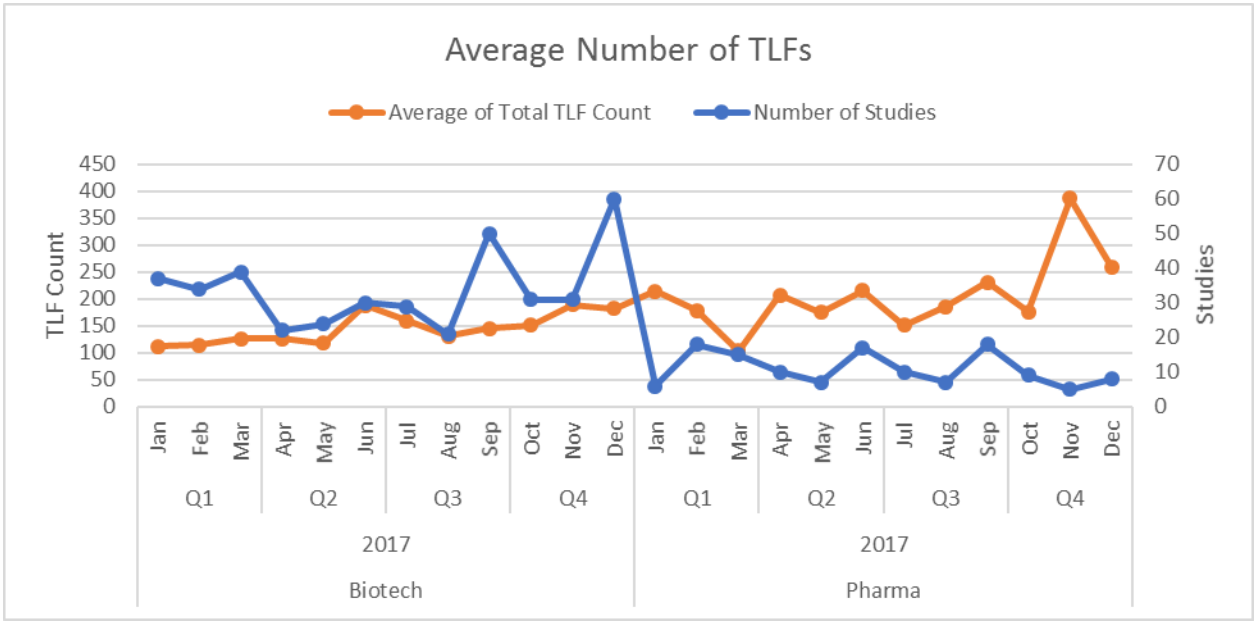
THE VARIANCE EXAMPLES

The observations made below are not claims but are designed to stimulate discussion as opposed to be a statement

The data presented below is taken from the PPD database of bid studies for year 2017. The number of studies is sufficiently large to make general conclusion however when sub-setting one needs to take care on diminishing values of N. The graph below shows a) the average Total TLF count per month b) # of studies bid in that month. For interest the graph is split firstly by Biotech companies and secondly by Pharma companies for ALL TA's. Brief observations are made under each graph/bar chart.

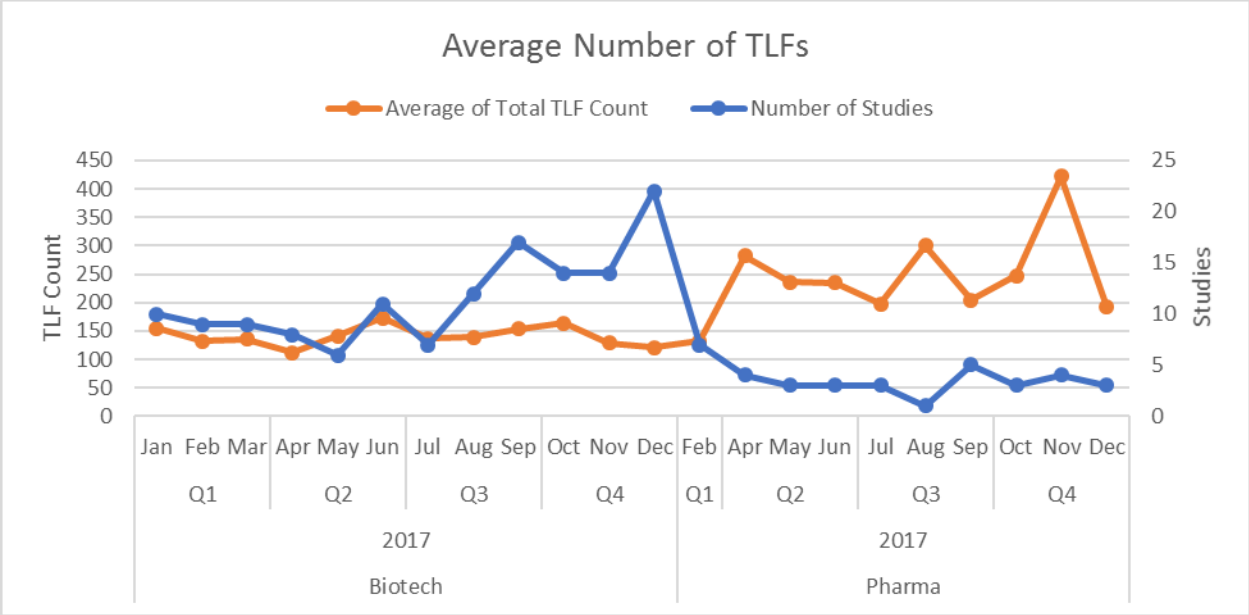
.

Plot 1. Average number of TLFS for all TA's.



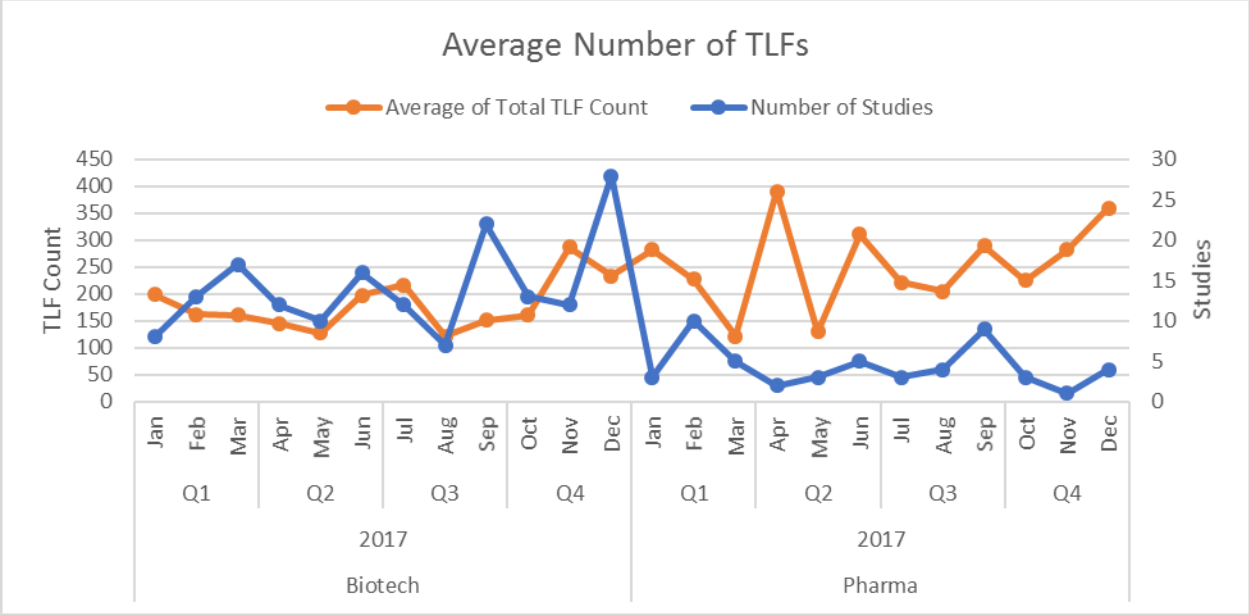
- Little inference can be drawn from this plot other than a certain level of variability is seen month to month.
- Possibly we see Biotech slightly less (~150) than pharma(~200).

Plot 2. Average number of TLFs for Phase 2 for all TA's



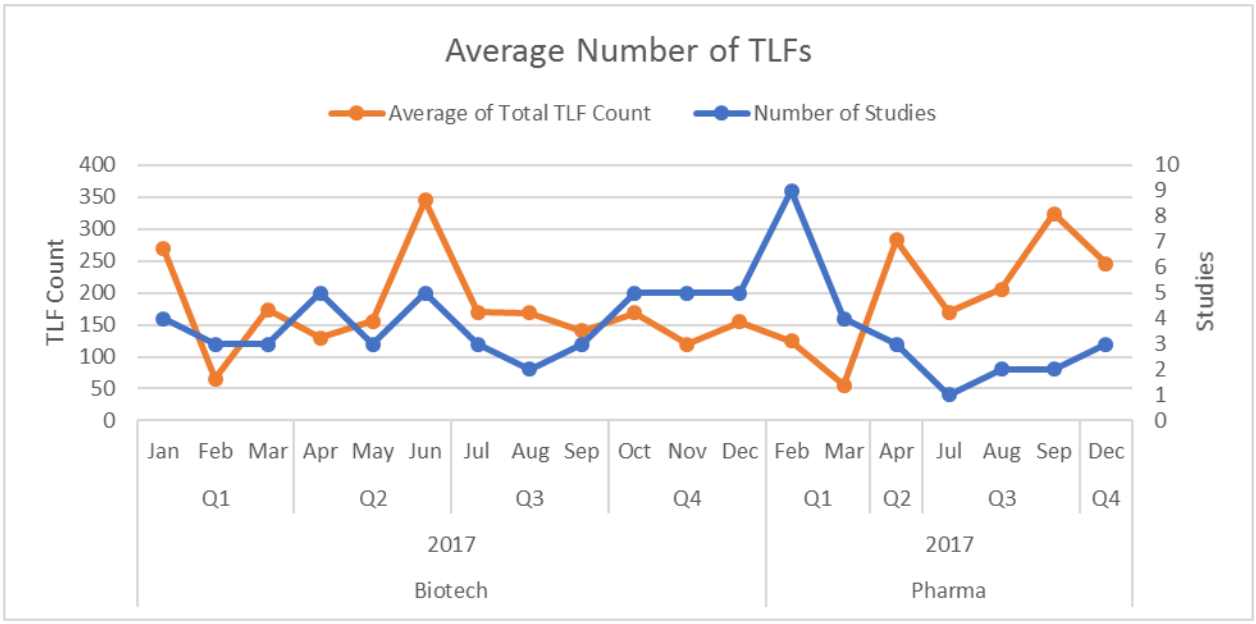
- Begin to see that pharma companies(strategic) have a higher TLF count than Biotech (~250 versus 150).
- Quite some variability in pharma scores but these are "all TA's" so may be bias towards a certain TA heavy weighting for pharma compared to biotech (eg. Maybe pharma is all Oncology and biotech all hypertension).

Plot 3. Average number of TLFs for Phase 3 for all TA's



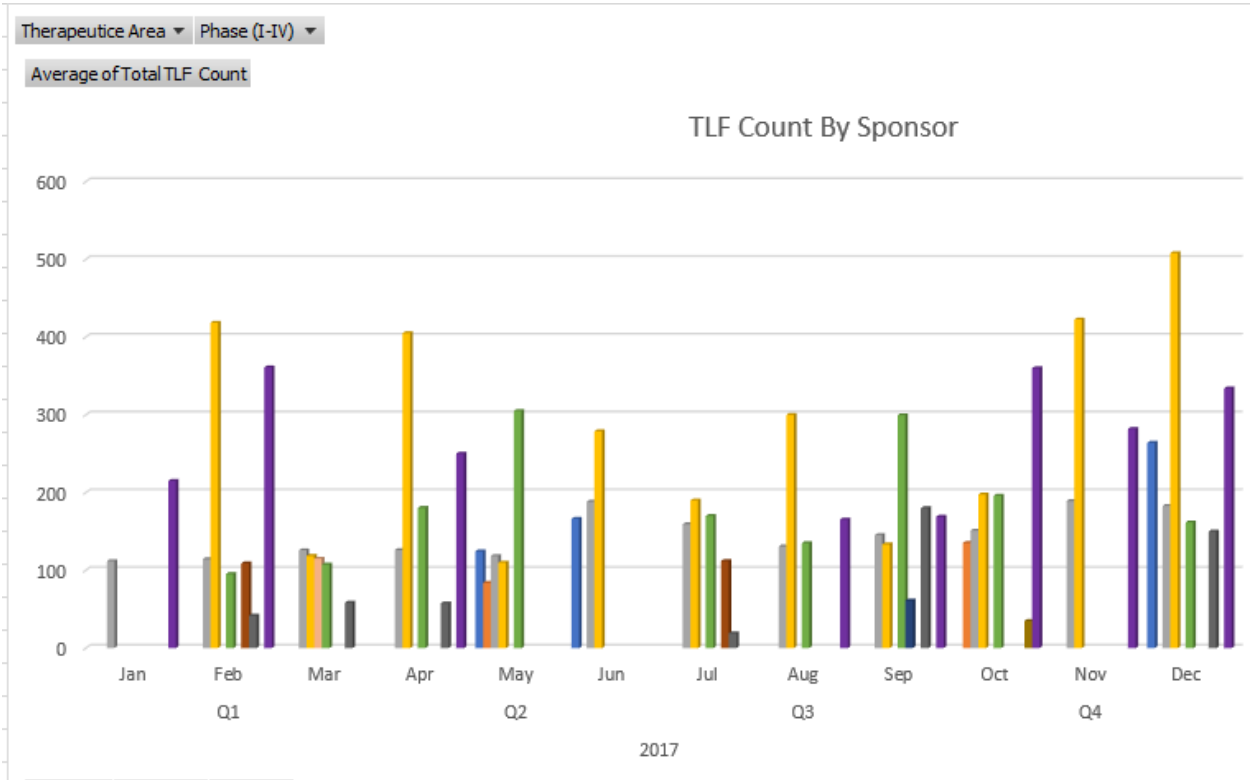
- Trend continues that pharma companies have a higher TLF count compared to Biotech. (250 vs 150).
- Higher variability for pharma scores but recall all TA's are represented so a certain TA may bias the plot.
- Slightly higher counts for phase 3 compared to phase 2 as expected but not as much as expected.

Plot 4. Average number of TLFs for Phase 2&3 for all IRD



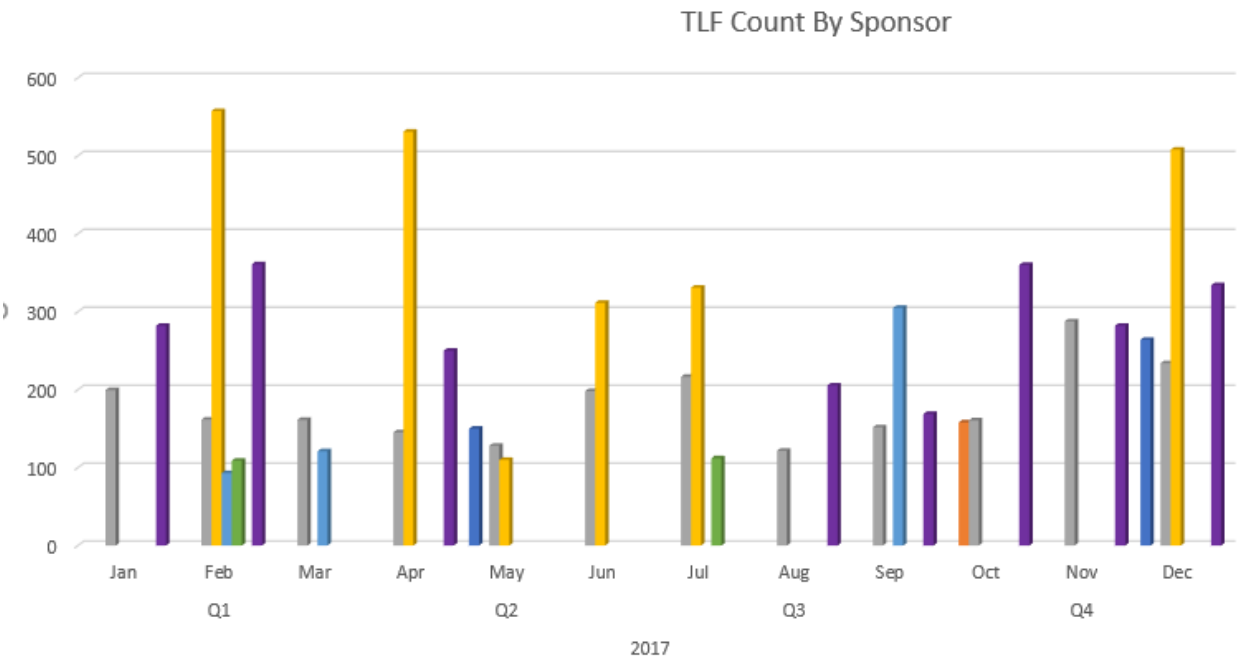
- Pharma higher than biotech (200 vs 150)

Bar Chart 1: TLF count by sponsor all TA , all phases.
(grey is Biotech).



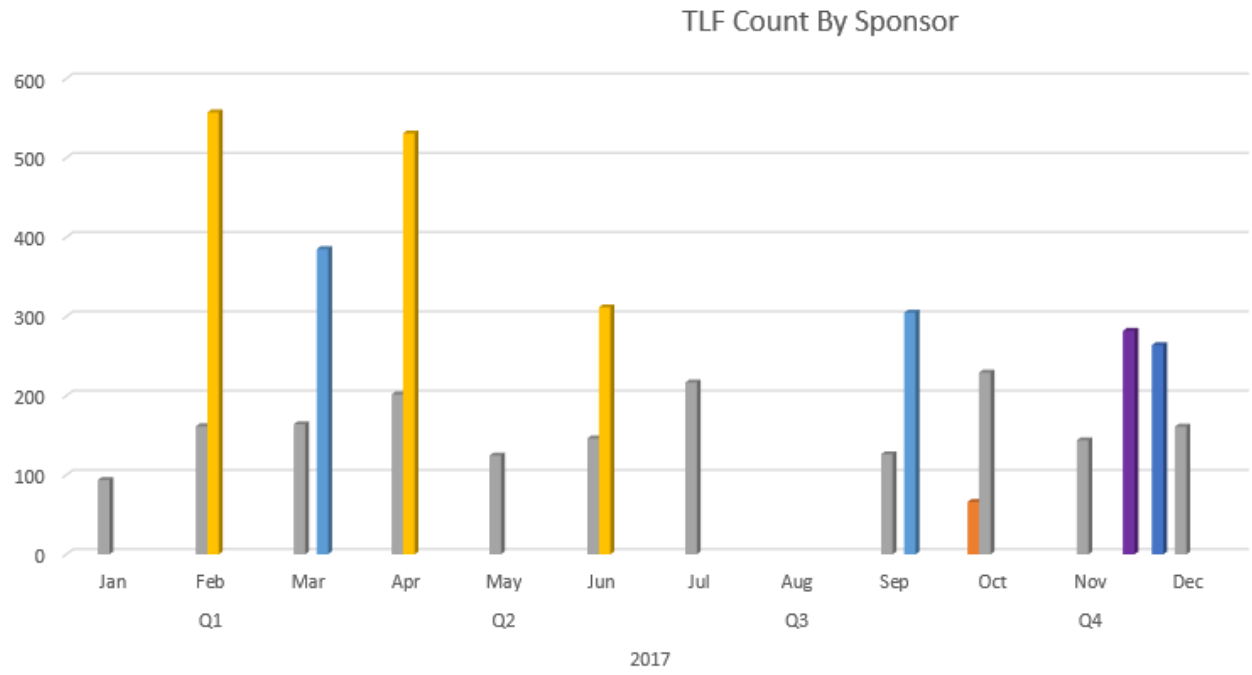
- Biotech has smaller counts compared to pharma.
- Yellow and Purple company consistently hitting high counts, why is that?

Bar Chart 2: TLF count by sponsor all TA, phase 3.



- Yellow and purple companies have high counts.

Bar Chart 3: TLF count by sponsor, Gen Med, phase 3.



- Yellow and light blue company have high counts.

From the analysis above we identify that “generally” biotech companies request less TLF’s than pharma companies, so why are biotech companies more “lean & mean” than pharma? Is it because they outsource more work so are exposed more to costings than pharma? Is it because Biotech companies have tighter control of costs compared to pharma? Some pharmaceutical companies appear to be real outliers with the numbers of TLFs created, if these companies are made aware of this then the process of requesting these TLFs need to be questioned.

CONCLUSION

- A potential cost saving exercise is paying closer attention to the estimations regarding # SDTM’s, ADAM’s and TLFs. Incorrectly estimating these can lead to incorrect resourcing or dreaded “out of scope” discussions if out-sourced.

- Too often the realization of incorrect estimations is only realized when the programming team have inadequate resources to complete a task which will then lead to spiraling costs to operationally achieve the original goal.

- Different methods of estimation are used by companies and with different levels of control/sense check.

- The functional driver of the number TLF’s can be the medic and/or the statistician.

- Biotech companies appear to have lower TLF counts than pharma companies, why is that?

- Some companies have extremely high TLF counts compared to the mean why is that?

Recommendations are

- Scope discussions must take place concerning number of SDTMs, number of ADAMs number of TLFs before study starts to avoid incorrect resource estimations or painful “out of scope” discussions if outsourced, ideally level of complexity is considered at least for TLFs.
- collect historical data regarding number of SDTM’s, Adam’s and TLFs per TA per phase.
- A range is developed for #TLFS per TA/Phase and if these are exceeded they should be discussed with management for verification.
- The lead programmer should be invited to give an opinion on estimations, resource needs and calculations.
- Maybe an independent body such as PHUSE should collect a database of these figures?

CONTACT INFORMATION

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

Author Name Roy Ward
Company PPD
Email: roy.ward@ppdi.com