PhUSE 2014

Paper PD07

Learnings and Evolutions of Risk Assessment Quality Control within Statistical Programming group

Tim Barnett, Roche Products Ltd, Welwyn Garden City, U.K.

ABSTRACT

At the PhUSE conference of 2007, a Roche colleague presented on the "Validation of Programs developed using SAS", focusing on the implementation of a new Risk-assessment approach and highlighting the effectiveness and efficiencies an approach can have. This paper will look to review how the Risk-assessment approach evolved over the last seven years, with changes in industry and programming landscape as well as our understanding of risk, to its current format being used in Roche today. Such areas to review will be the incorporation of smart risk within the business, the requests for transparency and closer inspection from the outside world, learnings we took from an idealized risk approach to the physical implementation within the programming group, and where risk assessment is heading for future iterations.

INTRODUCTION

For the PhUSE Conference of 2007, a Roche Products Ltd. colleague (Nikos Tsokanas) presented on the "Validation of Programs developed using SAS", and specifically on a new approach to validation for the Statistical Programming team in employing a Risk Management Strategy in their validation approach. This approach enabled the programmers to assign risk levels to programs so that the level of quality control could be determined and appropriate resources (in relation to that risk) could be assigned.

Over the last seven years this strategy has been used effectively within the group, with many learnings and adaptations taking place as we learn more on how Risk Management and Statistical Programming can be combined. Over time there has been many factors leading us to review our strategy, including external influences, business changes, as well as the implementation of such a unique process, and we also need to consider what the future holds for such an approach.

THE APPLICATION OF RISK ASSESSMENT FOR QUALITY CONTROL

RECAP

In 2007, the Statistical Programming group within Roche Products Ltd. agreed to implement a Risk Management Strategy to their validation procedures, incorporating the elements of Risk Analysis and Risk Evaluation to help the overall Risk Control. This approach came out of a desire to focus energies of the group on the elements considered the most at risk and therefore help the team allocate resources and time on specific activities wisely. For instance, if a program was considered to have a high risk level, then the validation efforts would need to be reflective of this, and the full validation steps available should be taken. If, however, a program was assessed as a low risk level, then the validation steps could be reduced and significantly less time would be spent on this program.

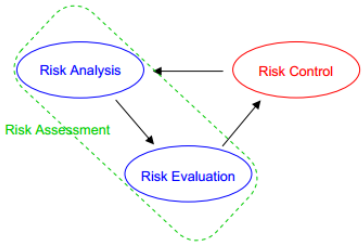

Figure 1 – Risk Management

Key elements of the risk approach taken at the time included:

- Risk evaluation at three levels – Reporting event, report object and the individual programs used. This would then be evaluated for an overall assigned risk level to help determine validation effort of the specific program
- Full documentation of risk level chosen – Assessors should be clear in their level selected and clearly document what has been chosen
- Three levels of risk available – High, Medium and Low, and these levels will decide the level of validation performed always by an independent programmer
- Risk evaluation needed to consider many factors including intended use, audience, criticality of data, formality of analyses, complexity of code and the likelihood of identifying an error.

Practical implementation of the approach above was well received, but learnings over the years have adapted this process and our approach to risk assessment now looks very different to this original conception.

IMPLEMENTATION

The implementation of this process was always seen as a learning curve for the department, and iterations and adjustments were expected.

One of the first adjustments made was to the number of risk levels. At conception the idea of a High, Medium and Low risk felt a good, even coverage, but on application a number of concerns were raised. Some teams assigned medium throughout, seeking the middle ground. Such an approach is common when individuals are maybe confused by the assessment and so choose a safe average, or maybe feel pressed for time and consider a middle score to hopefully reflect what is needed. Other teams never used medium risk preferring a black and white approach, and instead of choosing a steady middle ground preferring to be bold in choice of high and low. This led to inconsistencies in its application and therefore a revision was necessary. On review, it felt that the assignment of medium risk was vague and fell into the hands of those preferring to be safe or not fully consider their options, and therefore the group decided a more decisive approach was needed, and only low and high risk was taken forward. This has led teams to ensure they take the necessary time to consider the risk to apply and be bold in their selection.

A second adjustment was a revision of the concept of low risk and the validation efforts included. A concern around the use of an independent programmer has always been that there becomes a reliance on this “safety-net”, and that programmers will be more likely to perform less checks on a 1st line program if an independent person will be there to review and catch mistakes. In order to help the group place more emphasis on 1st line programmer checks, and place overall accountability of quality for the program on the assigned programmer, the decision was made to make low risk validation checks entirely with the 1st line programmer and not assign an independent review. What we have also noticed is that the introduction of more robust standards, in data collection, tabulation and analysis, has meant more robust, reliable up-front programs. This is either as part of standard macros or reusable from other study teams, so that the risk levels automatically allow a single developer to provide a high enough level of quality foregoing the need for an independent programmer to be involved.

A third adjustment was helping the mindset of a group which was focused on maximum validation efforts. A culture fostered around quality control and maximum efforts to ensure quality made the implementation of risk levels difficult to implement, and teams needed coercion to embrace lower risk levels, and therefore lesser validation, on their study work. One way the group targeted this was to set a cap of a maximum 40% of report objects (TLGs) to be assigned as high risk. The reasoning for this was that datasets will carry the more complex derivations and that TLG programming therefore should be considered less necessary for independent QC, and certainly in instances where iterations of a single template (eg. Subgrouping or differing population analysis) can mean a consistent high risk should not be necessary.

A final adjustment was the review of how we determine our risk assessment, and then document what we have done. We originally conceived an algorithm for determining risk, based on reporting event, the report object and the program being used. We would then look at various factors including an overall reporting event assessment that looked into criticality, intended use and source of programs (if reused), and then look lower at a program level to review complexity, underlying data and likelihood of error capture, to then determine an overall risk.

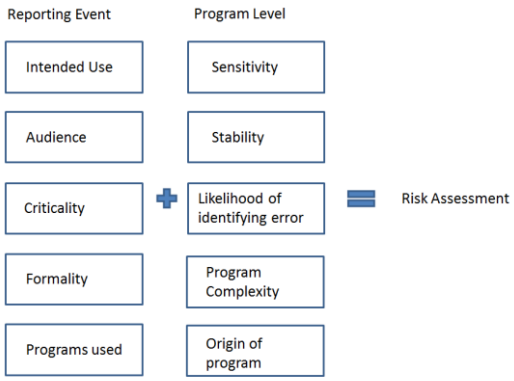

Figure 2 – Original Risk Assessment model

What we found by adding so many layers to risk assessment was that it made the assessment procedure confusing, over-complicated and absorbing too much time and effort. What in the end was decided was to simplify the risk assessment ideology, remove the numerous factors of influence and instead focus on an overall assessment that then helps assign the risk. The model we now work off is a simpler approach on a report object level (the output being generated, as opposed to the program being used) that enables individuals to assess a risk in a quicker and simpler fashion. The focus is broadly what can go wrong - thinking about complexity and underlying factors in data, to then the impact of it being wrong - thinking about the intention of the report object and the chance of spotting an error early. It is an approach that favors experience and therefore is now recommended as a practice by lead programmers to complete or oversee.

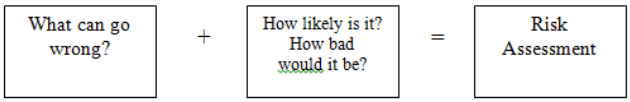

Figure 3 – Latest Risk Assessment model

For documentation, we stress the importance of capturing the risk level assigned. However, we realized that the rationale for this risk was not always clear externally to the team, so the peer review of this risk should be formalized to ensure consistency. We now have in place more guidance that provides rationale around risk for certain types of report objects and scenarios, and allow teams to capture rationale if special circumstances dictate a high or low risk. We also ensure manager approval of risk assignments to give consistency and formality to the process, as well as potential objectivity and a wider context.

BUSINESS CHANGES

The approach to risk assessment within the group was brought in at a time when Risk Management was being viewed with interest within the industry. Since then it has been a great working example of how understanding risk can bring efficiencies, and this is being seen elsewhere in the business.

Changes within the industry, and within the company, has meant broad changes in many aspects of how we work within Statistical Programming. As companies evaluate how to bring drugs to market for less cost, areas of efficiency and time value are promoted and innovative solutions embraced. This has led to more focus in companies with Smart Risk, the concept of taking calculated, controlled risks that bring high gains if conducted effectively. The assessment of risk within statistical validation is seen as part of this Smart Risk mindset – of employing a process that focuses on where resource needs to be applied, and thus getting more value from the activity. Other areas of the business are also adopting this Smart Risk approach, including risk-based monitoring – identifying the study aspects of most importance and applying more time and resource to those at the higher end of the risk scale - and more recently data quality focusing on ‘reviewed’ rather than ‘clean’ data on those data points most critical to the trial outcome.

This also fits well with the reflection paper from the EMA in 2011 (“Reflection paper on risk based quality management in clinical trials“) that states:

“Since absolute perfection in every aspect of an activity is rarely achievable or could only be achieved by disproportionate allocation of resource, it is necessary to establish clear priorities, to mitigate the significant and

serious risks to those priorities, and to establish tolerance limits within which different processes can operate. By establishing the priorities, mitigating the most significant risks and operating within sensible tolerance limits, the required quality standard can be described, and its achievement (or failure to achieve it) can be more readily measured, reported and recognised."

The search for employing resources effectively is not intended to reduce overall resource, but more to ensure resource is assigned on tasks that bring the most value. As statistical programmers have increased demands to analyze their data more, to re-think analyses and bring in new tools to support our stakeholders, the time a programmer spends needs to be allocated wisely. The evolution of a programmer to what is often described as an analyst means focusing validation efforts, and freeing more time for other activities of value, fits in with the vision of where statistical programming roles are heading.

LANDSCAPE CHANGES

Whilst we see more emphasis on focused resourcing, and efforts to improve assignment of time and ensuring value in the work we apply ourselves to, we also have external pressure to ensure more emphasis on ensuring quality.

The evolution of Data Transparency across the industry, culminating in the release of datasets and analyses to the external world, has meant our identified list of reviewers has increased dramatically. We also now have submissions to Health Authorities where the request for our programs can be made as a deliverable in their own right. The scrutiny on our work is now increasing as the number of customers to our work increases beyond the recognized stakeholders, and this brings with it fresh challenges on how we should perceive quality and the elements of risk we apply. Elements of risk assessment we applied before suddenly need more thought as the boundary potentially pushes beyond what we previously knew. One of the major principles of risk assessment includes what is the impact if this is wrong? Previously we may have thought of our individual study team evaluation, or the delivery to Health Authorities based on the analysis we defined. Now this may be picked up by multiple external groups with different agendas or focus. A seemingly innocuous concomitant medication table may be seen as lower priority and value at individual study assessment level, but when released to a wider community could this be used as a focal point of analysis and therefore have larger repercussions if incorrect. And any potential inaccuracies can always be leapt upon, and potentially the overall perception of quality harmed.

And perception is also an area that risk assessment can compete with. The perception (both internal and external) of risk assessment and application of risk levels, if not properly explained, can lead to misunderstandings of applying lesser validation in some scenarios and compromising quality. This is an all-too-easily reached conclusion, which is why clarity and communication are always needed to embed that low risk does not imply we are any less certain of its quality than high risk, it is just requiring less validation steps to achieve the same quality level.

THE FUTURE

The future for risk assessment within the group will still continue to evolve as the landscape continues to change and the business demands alter.

In the short term, documentation and clarity of our process will probably be reviewed, as the increase in transparency and visibility of our work are requested. This will include stronger guidance documents for individuals to make more consistent risk assessments, and the capture of reasons behind selections of risk made more apparent. It most likely will also incorporate risk assessment capture into standard tools used for programming, and the consistent capture and accessibility of that information allow for stronger control and review of how we assign risk across the group, hopefully allowing more understanding on how we can achieve consistency. It may also be broadly recognized that as more companies come to grips and understand the impact of Data Transparency that an increase in High risk assignment is seen in the short term, as the evaluation of the impact of an error increases.

In the long term, the increase of standards across data capture, tabulation and analysis should eventually lend itself to a more standard risk assessment, allowing teams to apply a default assessment based on study design and an expectation of more low risk programs as standard programs become more useable "off-the-shelf". It may be that in the long term the high risk programs will be those controlling the meta-data around a standard suite of programs, and thus having quality control incredibly stream-lined.

CONCLUSION

The risk assessment assigned to the programming group for assigning quality control has been seen as an asset to the group in assigning resource to the areas of most value, whilst still maintaining the same levels of quality in its deliverables. Over the years adaptions to this approach have been made, to simplify the assessment, to place more emphasis on the 1st line development, and in helping adjust mindset and set documentation levels for the approach. For the future of risk assessment, I can foresee more efforts to align risk levels across teams, and the increase of standards to allow more low risk assignments and therefore more time able to be spent elsewhere. The increase of external scrutiny of our deliverables may cause a swing to more validation checks and higher risk in the short term, but an increase in dataset standards and standard reporting tools should help balance this out over time. As long as risk assessment has clarity and transparency in its use, it is a process to stay for the future.

REFERENCES

EMA Reflection Paper “Reflection paper on risk based quality management in clinical trials “ 2011

CONTACT INFORMATION

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

Tim Barnett
Roche Products Ltd.
6 Falcon Way
Welwyn Garden City
Hertfordhsire
AL7 1TW
Email: timothy.barnett@roche.com