Position Paper: Designing Risk Detection Indicators for Risk-Based Monitoring

Authors: Tammy Finnigan, Duncan Hall, Chad Finch, Mike Emanuel, Neil Watkins

Published on 22-Oct-2013
1. Background

Triumph Research Intelligence (TRI) was founded as a subsidiary of Triumph Consultancy Services in 2013, following 12 years of consulting to the clinical trial industry. Over the past 12 months TRI has been evaluating the specific challenges facing the industry when implementing a risk-based monitoring strategy and the various approaches and products being utilized by organizations as they move into this new arena. This paper aims to summarize our findings and provide guidance as to how the main challenges can be overcome.

2. Introduction

In August 2013, the FDA published the final version of the Guidance for Industry paper, Oversight of Clinical Investigations – A Risk-Based Approach to Monitoring. In this paper the FDA were clear that the recommendation to sponsors is to consider a change in approach to monitoring. The FDA goes on to state that they believe a risk-based approach to monitoring could improve sponsor oversight of clinical investigations. The paper specifically refers to the use of centralized monitoring using statistical assessments to prioritize and guide on-site monitoring visits, and that they are anticipating a decreased use of on-site monitoring as monitoring methods evolve and technical capabilities are enhanced.

In the paper the FDA also acknowledge that there is limited empirical data to support the utilization of the various methods employed to monitor clinical investigations, including on-site monitoring. As a result, current risk-based monitoring approaches are largely based on past experience rather than validated evidence.

While it has been shown in some studies, that centralized statistical monitoring could detect more than 90% of the findings identified during on-site monitoring visits, and that there are improvements in data quality as a result of implementing this practice, this is not the same as early identification and prevention of inadequate site behaviors and processes. Many centralized monitoring approaches focus heavily on reducing the amount of SDV required, and this is translated into a reduction in on-site visits by targeting on-site visits to sites where central monitoring is demonstrating there may be a problem. The monitors are then focused on correcting mistakes that have already happened but monitoring should be focused on preventing quality issues from occurring and addressing site processes and behaviors that are putting data quality and subjects at risk. Experience from audit and inspections continue to demonstrate that there are weaknesses in

---

monitoring practices and that many significant and critical findings are as a result of monitoring practices failing to detect deficiencies in site processes.

Our research has shown that the most significant challenges facing the industry in the implementation of risk-based monitoring strategies are:

- Identifying risk indicators which are prospective (identify sites with potential quality issues and inadequate behaviors before they impact data quality/integrity and potentially subject safety) rather than retrospective (highlight areas of issue once issue has occurred)
- Identifying risk indicators that can be used early in the study life cycle
- Identifying risk indicators that can be used with low data volumes
- Separating the signals from the noise
- Operationalizing the risk indicators

While the regulatory agencies are looking to the industry to adapt monitoring strategies, ultimately the industry needs to be able to demonstrate that the efficacy and safety data submitted to the agencies is an accurate reflection of the data from the subjects who took part. As a result monitoring practices are under more scrutiny than ever before.

In this article we aim to examine an alternative approach to defining risk indicators and evaluate how to operationalize the use of risk indicators to enhance monitoring practices and improve site performance.

3. **Defining Site Quality Risks**

Many approaches to risk-based monitoring have focused on traditional site operational indicators to provide insight into site performance. Indicators such as delays in reporting data, query turn-around timelines, high screen failure rates, and high frequency of eligibility violations, are currently being used as a measure of risk to data quality and integrity. It’s our opinion that operational indicators do not specifically map to risk to data quality and are more appropriate for general study oversight and performance management.

When TRI first began to look at risk based monitoring methods, we also started with the approach of operational indicators combined with statistical methods such as standard deviations. However, over the past 12 months, the analyses we have performed on clinical trial data has allowed us to evolve our thinking and modify our approach to designing risk indicators for the purpose of risk-based monitoring. Rather than using static thresholds and identification of outliers, we have isolated a small collection of highly specific and sensitive risk indicators, which have consistently shown to be good indicators of site behaviors that present a risk to the quality and integrity of the data that the site is producing.

---

Three classifications of site quality risk have been identified against which indicators are mapped:

1. Reporting diligence
2. Data quality
3. Protocol Compliance

**Reporting diligence**
This defines the reporting propensity of a site. For example, if a site identified an adverse event (AE) or an abnormal physical examination, what is the likelihood that the event will be reported? If the site is unlikely to report the event, then there is a higher level of risk that the site is not complying with other regulatory requirements and/or may fail to report more significant data such as serious adverse events (SAEs) or efficacy endpoints.

**Data quality**
This is not purely a measure of data cleaning queries; instead it is a measure of the variability of the data which would indicate that the data is an accurate reflection of the real world. For example, some sites have very high adverse event reporting rates but there is very little variability in the reported events. This may indicate that the events are not being reported accurately or in extreme circumstances are fabricated. The indicators in this category are more aligned with the analyses used in central statistical monitoring, but by focusing on a small, very specific sub-set of the clinical data early in the study, we can get an early and accurate view of the sites risk profile.

**Protocol compliance**
This is a measure of the sites ability to comply with the protocol; some indicators are a direct measure of compliance, for example recruitment of ineligible subjects and in some indications, other indicators, such as high randomization rates, can be surrogate markers for potential non-compliance.

It should be noted that an individual risk indicator can be mapped to more than one classification; the relationship is not necessarily one to one. At TRI we have created risk indicators to support each category and developed proprietary algorithms to present the data in such a way that site demonstrating high risk to data quality and/or integrity can be clearly identified. In the next section we will discuss how these risk indicators have been designed.

### 4. Designing Risk Indicators

There are some key factors to keep in mind when designing risk indicators.

The objective is to ensure that at each phase of the study, e.g. recruitment, maintenance, there is at least one indicator identified in each classification of site quality risks i.e. Reporting Diligence, Data Quality and Protocol Compliance. As an example, randomization rates are only available during the recruitment phase of the study and so may be replaced with AE reporting rates during the
maintenance phase, although this particular indicator can also be used in recruitment phase if the subject and site numbers allow.

The indicators should be based on high frequency data, as the accuracy of the indicator will increase with data volume (and hence the error rate will decrease). Low frequency indicators, e.g. SAEs, while important for the overall safety profiling of the IMP and safety of the subjects, they prevent early utilization and often have limited application in identifying risk to data quality and/or integrity that the site is producing. The higher the event count, the better the indicator and the earlier in the trial it will become valuable as a quality risk indicator.

The source of the indicator data is also an important consideration. A ‘high quality’ data source is data that is part of the clinical trial itself. As we are building a relationship between the risk indicators, and site behavioral patterns and processes, the data should be data generated by the investigator and/or site staff who administer the trial, for example EDC, IxR systems and all related metadata. This typically excludes systems used to manage the trial team e.g. Clinical Trial Management Systems (CTMS), Training Systems. A common misconception is that monitor issue logs, deviation logs, medical queries and manually generated data management (DM) queries are a good source of indicators, but this is not the case. Data sources need to be objective and these indicators are subjective, more a measure of monitor, medical monitor or DM performance, than site performance. This information can have value as supplementary information after the risk evaluation has been completed, but should not be used as the source of risk evaluation. If sites, data and subject profiles are randomly distributed to medics, monitors and data managers, such that the number of queries is not a bias, then this type of data could be used as an indicator, but operationally, this does not happen and therefore the subjective nature of these indicators will introduce bias.

Once the data points and data source have been identified the indicators need to be fully normalized for volume in order that the indicator retains its sensitivity irrespective of data volume (there is a minimum volume below which the error margin is so high it is not worth consideration). This approach generates a much more powerful two dimensional view of the indicator as opposed to the one dimensional view we have seen being adopted by some organizations.

For example, a site with 100% randomization rate may seem high risk, but this is not the case if it only screened 1 subject who was successfully randomized. Conversely, a site with 100 screened subjects and a 100% randomization rate is highly improbable. Failure to consider size and volume is a major short coming of many risk-based monitoring approaches.

For each indicator, we need to understand and evaluate the risk of apparently high performing sites as well as low performing sites. For example in AE reporting, the high reporting sites are not necessarily the high performing sites. To know the difference we need to look for variability in the data. If the site is a high reporter with high variability, then it would be considered a high performer, but if it’s high reporting with low variability, it is likely to be a very low performing site. In contrast, when looking at blood pressure measurements, high variability actually makes the data from the site indeterminate of risk, due to the likelihood that they are using digital blood pressure meters. We
also need to consider and understand the impact of the site structure on risk. Based on our experience and the data analyzed to date, our current hypothesis is that those sites with nursing staff are more likely to exhibit high performance in areas usually covered by nurses e.g. blood pressure readings, pre-treatment concomitant medications, and medical history. This is something we will continue to evaluate with further studies. What we have also observed is that nursing performance may be completely disconnected from the doctors performance e.g. AE reporting, physical examination, treatment emergent concomitant medications. If a site shows high performance in some areas, but low performance in other areas, we need to be able to determine why, which will allow us to target the inadequate behaviors that are putting data quality/integrity at risk and take the appropriate actions. This brings us to the point of our next innovation - risk signatures.

**Risk Signatures**

As we continue to analyze clinical data, it is becoming evident that certain performance issues are related. For example, if the doctor is underperforming, there are no nursing or support staff and a lack of robust processes in place at the site, we see issues in all areas, but if there are nursing staff, we tend to see performance issues clustered in AEs, Treatment Emergent Con Meds, and Physical Exams. We speculate that a limited number of these collections or ‘signatures’ exist. For example, we do not often see evidence of poor blood pressure data without poor AE data although we often see poor AE data without poor blood pressure data when a site has a nurse. If we go back to an earlier example of randomization rate, in basic terms, a site with 100% randomization rate and 1 screened subject would have a high (>50%) probability that the rate occurred by chance alone. By comparison, a site with 70% randomization rate and 100 subjects might have a much smaller probability of being real or having occurred by chance. However, when the probability of the randomization rate being accurate is combined with the probability that the AE reporting rate is accurate, and Physical Exam reporting rate is accurate, we can build a composite picture of the site – its risk signature. If the site scores high in a number of risk indicators, this is likely to be indicative of a site representing truly high risk to data quality/integrity, and hence one which should be targeted for monitoring attention.

These signatures are likely to be a key diagnostic criterion moving forward, with the intention being to assign a site a ‘risk signature’ which is indicative of the sites performance and more importantly, its behavior in relation to good clinical practices, which allows us to create a much more effective site monitoring/management plan. We also strongly advocate the use of risk trends over time. This means that as we build up a risk signature for a site and assign a risk score to that site, we can evaluate the efficacy of our interventions with that site by looking for a positive trend in the risk score.

So far we have focused our attention on clinical data, but while clinical data as a risk indicator is powerful, it is also limited due to the fact that it is difficult to obtain a large volume of data early in the study. Furthermore, because the incidence of any clinical event is study specific, it could be misleading to use the frequency data from one study as a baseline from which to assess the performance of another study. As a consequence, the study being monitored requires a relatively
large number of sites so that data from one site can be compared to the performance at other sites in order to identify high risk sites. As a result, we are largely unable to use risk based monitoring in its current form in studies with low site volumes. This speaks to 2 of the main challenges when implementing a risk-based monitoring strategy that were identified in the introduction.

We have 2 hypotheses currently under evaluation to address these.

**Identifying risk indicators that can be used early in the study**

Earlier we mentioned that the subset of data can be significant when normalizing indicators and reducing the ‘noise’ of the clinical data. One sub-set of data we are evaluating is pre-treatment data and we believe this could be a good leading indicator of risk to data quality/integrity and will focus on the site processes and behaviors in the very early part of the study. For many therapy areas a typical subject profile will exist, for example in a phase III oncology trial, it would be unlikely that subjects would have a ‘clean’ medical history or no concomitant medications. It is our assumption that setting an informed threshold for baseline clinical data will provide early warnings of site risk. This prior assessment of quality risk and the resulting threshold can then be refined as we collect more data during the trial and the standard error reduces. It should also be noted that as this is baseline data, there is less influence by the protocol, and therefore these could be used to build up an informed view of early risk indicators by therapy area/indication.

**Identifying risk indicators that can be used in studies with low data volume**

As mentioned previously, the cornerstone of a successful indicator is its frequency – high frequency is an essential pre-requisite for a successful indicator. Low frequency indicators are of little or no value in most situations. We are proposing that a solution to this may be to use metadata from the source system, specifically times and speed at which data is entered. Because metadata is by definition higher volume than the core clinical data and unlikely to be study dependent, it could become a very early indicator of risk to data quality, and allow risk-based approaches to be used in studies with lower site and/or lower subject per site volumes.

5. **Signal and Noise**

As an industry, the volume of data we are collecting is increasing exponentially but the objective truth is a relative constant. This means that whilst the proverbial ‘needle’ remains the much the same size, the ‘haystack’ is getting bigger by the day. Risk-based monitoring doesn’t make the haystack any smaller, but it gives us insight as to where in the haystack we should look first. In operational terms this is incredibly important. A poorly implemented risk-based monitoring approach can in fact be more detrimental than to not implement risk-based monitoring at all. An operational team can easily become consumed trying to respond to a multitude of indicators or indicators which do not actually indicate the most important risks. This can result in an increase in workload and a high chance that the true signal will actually be missed in the operational chaos
which can result. Interpreting noise as signals can also lead to inappropriate action being taken. In the case of an intervention or monitoring visit with a site, this may not only be a waste of time and manpower, but may also lead to degradation in the relationship with the site. Companies implementing risk-based monitoring must therefore be capable of assessing each indicator or combination of indicators in terms of their correlation with quality risk and maintain focus on the signal and be prepared to discount the noise.

6. Operationalizing Risk Indicators

In the introduction, we mentioned that one of the biggest challenges to risk-based monitoring approaches was being able to incorporate the risk indicators into the study management process and make them operational. We have focused much attention to this area as we believe it is the key to a successfully realizing the benefits of a risk-based monitoring approach.

To turn risk indicators into an operational tool, we need to be able to understand the results and visualizations and present them in a way that is easily consumable by study teams. Study teams are already overloaded with data and need small packets of information that are easy to understand and translate into actions. At TRI this is referred to as ‘early, frequent and light touch intervention’, which is explained later.

A key part of a risk-based monitoring strategy is to prevent inadequate behaviors and detect deficient processes in place at the site that risk compromising data quality/integrity. The investigative skills required to detect deficient process and less desirable behaviors at the site has already been identified as a failing of most monitoring practices, it is therefore important to have data analysts who can direct monitors to the subjects most likely to provide evidence of the risk, not just detect that there is risk. Not all source document verification is the same, we often hear of risk-based monitoring approaches that use partial/reduced SDV e.g. 20% SDV, where the percentage represents the number of pages verified, or only critical data being verified. In fact, in the recent industry wide survey issued by the MCC\(^4\), the preliminary results showed that more than 50% of the organizations who took part in the survey were operating a reduced SDV approach to risk-based monitoring.

**Challenges with reduced SDV**

Once a monitor commences SDV on a subject, it is very difficult to focus on only the percentage or data points identified for verification and it virtually becomes 100% SDV. Another common approach is to randomly sample subjects for SDV, where the subjects are identified up front in the monitoring plan. The concept of reducing SDV is not simple, and both of these approaches are flawed as neither ensures that the monitoring activities will identify inadequate processes and behaviors that risk compromising data quality. As mentioned before, risk-based monitoring has an

\(^4\) Metrics Champion Consortium Risk-Based Monitoring Survey
important role in directing the monitor to the subjects to conduct SDV, not only the issues. A good data analyst will identify the subjects most likely to demonstrate that there is a problem with the site processes, for example, if the site is under reporting AEs, then a subject with high duration of exposure and a complex medical history will be the one to focus on when conducting SDV. So the SDV becomes specifically targeted for an individual site, rather than a blanket reduction and we also reduce potential for the variability of CRA performance to have an impact on data quality and site performance.

_**Early, frequent and light touch intervention**_

To address the challenges with reduced SDV, the first part of our ‘**early, frequent and light touch intervention**’ approach is for a data analyst to review the data for the sites in the bottom 10% of the combined risk indicator scores; even if the risk indicators demonstrate that the sites are performing within acceptable standards. This review would typically be performed monthly and the outcome provides the study team with a list of actions to follow up with the sites. The actions will range from phone calls to site to discuss observations to onsite visits with information on which specific subjects to target source data review to highlight issues. This approach is very simple to implement and psychologically, very effective at improving site performance and consequently the overall quality of the study.

The second part of the approach is to turn the 2 dimensional graphs into one dimensional scores which corresponds with the risk estimation as seen on the two dimensional graphs. One major advantage of this is that a standard dashboard can be created with Red, Amber and Green (RAG) flags for each indicator, as described in the position paper⁵ published by TransCelerate BioPharma in Jun 2013. This will make the interpretation of the data easier for study teams, in particular monitors. This simple approach allows the respective RAG flags for each risk indicator to be associated with appropriate actions in the monitoring plan. This means that the study teams do not need to wait on monthly reviews from 3rd parties before taking actions, which will satisfy the regulatory requirement of being able to demonstrate timely and appropriate intervention. This approach can also support the site quality risk trending, discussed earlier in this paper.

**7. Conclusion**

We have seen many different approaches being taken to risk-based monitoring, but believe that there is a core set of capabilities which need to be prevalent in order to truly realize the quality and cost benefits that risk based monitoring can deliver:

- Risk Indicators for RBM need to be designed to be proactive indicators of data quality/integrity risks and specifically highlight sites exhibiting less desirable behaviors and/or inadequate processes

⁵ Position Paper: Risk-Based Monitoring Methodology, 2013 TransCelerate BioPharma Inc.
• Risk indicators must be based on objective data and not subjective data
• Risk indicators should be two dimensional to take account of data volume and size, which will increase the sensitivity
• A small number of risk indicators is desirable to allow separation of the signals of risk from background noise
• Most importantly, the risk indicators must be easy to interpret and be linked to specific action and escalation paths for the study teams to follow

Risk-based monitoring is a significant diversion in approach to running clinical trials from the more traditional 100% SDV supported by a pre-defined regular site monitoring schedule. It is therefore important that companies treat the adoption of RBM as a business and organizational change project and not just the implementation of a new technology.

About Triumph Research Intelligence

TRI's sole purpose is to help industry achieve the quality and cost improvements than can be associated with a well implemented RBM approach. We realize that many companies do not have the resource, experience or desire to develop their own approach and tool set for RBM. In order to allow companies to quickly adopt and benefit from all the elements of RBM discussed in this paper, TRI have developed an entirely new operational platform called Visual OPRA. Visual OPRA is a cloud based RBM solution which is available on a study by study or cross-enterprise basis. Visual OPRA will utilize data already being captured in your EDC/IxR systems and has its own web based user interface, and hence there is no need for implementation of additional data capture or business intelligence tools.

For those companies wishing to develop or improve on their own capabilities, TRI are also able to share the benefit of our extensive experience and knowledge through provision of consultation, implementation, change management and training services.

For more information go to www.triumphconsultancy.co.uk