5.1) Risk-based CDM approaches

Over the past decade, regulators have issued several guidance documents such as ICH E6 (R2)[7] which define risk-based principles and advocate for the use of risk-based approaches. Embracing these methods, the industry has already successfully implemented risk-based approaches in the site monitoring and system validation spaces for several years. As a result, our traditionally risk-averse industry has become more comfortable with strategies that match efforts and focus commensurately to the risks.

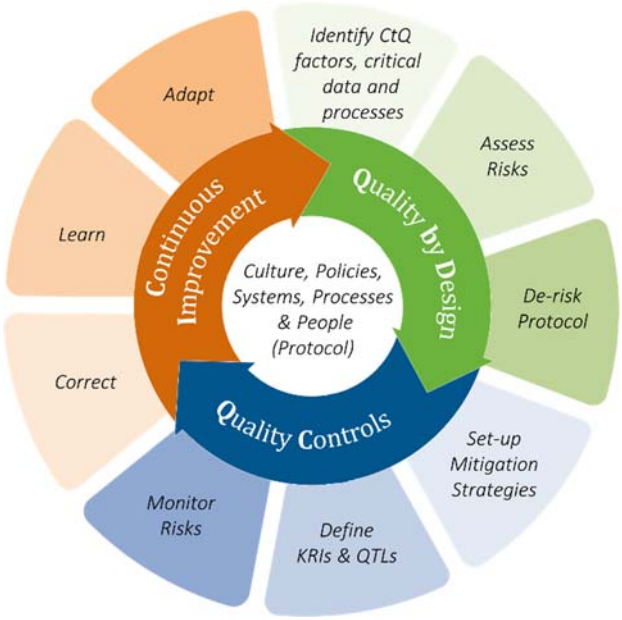

Fig 4. CDS RBQM framework

We now have a meaningful opportunity to expand the use of risk-based quality management (RBQM) principles to encompass all study design and execution aspects within CDM. As articulated in Part 1[1], CDS must redefine its processes and roles to control risks to activities essential to ensuring human subject protection and the reliability of trial results[7]. Learning from the evolution of traditional to risk-based site monitoring, we must realize that this is a fundamental culture and role change. This means moving from a one-size-fits all process based on fixed standards to a new paradigm where quality is infused at the design stage to proactively prevent risks to arise as much as possible (see figure 4).

a) Study Quality by Design

ICH E6 (R2) dedicated a new section on quality management (section 5.0) focusing on risks lifecycle management. It covers the quality controls used from the identification to the reporting of the risks. But rather than controlling the risks by implementing mitigation and monitoring strategies, we should simply use QbD to avoid them in the first place which means we should start with the end in mind.

QbD stands on the assumptions that quality should be planned proactively and not be the act of retrospectively perform Quality Controls (QC). First, QbD must rely on the appropriate foundation including the company culture, policies, systems, processes and people. Next, it relies on optimal protocol design, a critical step with a huge influence on the ultimate success or failure of the trial.

So, consistent with the QbD principles, risks and mitigations should be identified prior to the protocol finalization. If possible, the protocol should be adjusted to prevent the risks. If de-risking the study protocol is not possible, the study team must implement timely mitigation strategies to manage risks and prevent issues from occurring. Of particular importance as part of this proactive risk mitigation process is the need to ensure that "all aspects of the trial are operationally feasible" and "avoid unnecessary complexity, procedures, and data collection" [7]. The Clinical Data Scientist must steer the study team to only perform procedures that are essential to the outcome of the clinical trial. Ultimately, the study team must proactively confirm that the planned protocol is operationally acceptable for all sites and countries involved.

As a key member of the study team, the role of the Clinical Data Scientist is to drive the conversation and proactively manage risks that matter most, for example CtQ factors associated to critical data and processes. The Clinical Trials Transformation Initiative (CTTI)[8] introduced the CtQ factors in 2015 and organized them around the six (6) major categories (see figure 5) with strong emphasis on protocol design. Subsequently, they became a central theme in the ICH E8 draft guidance on the general considerations for clinical trials. The CtQ factors included in Annex 3 of the draft guidance are almost identical to those from the CTTI with one notable difference (i.e., changing the focus from data quantity to data quality during protocol design). In the draft guidance these factors are "*considered to be critical because, if their integrity were to be undermined by errors of design or conduct, the reliability or ethics of decision-making would also be undermined*"[9].

Lastly, the ICH E8 draft guidance is reemphasizing the need to ensure the scientific and operational feasibility of the protocol and fit-for-purpose processes considering the diversity of data sources.

CtQ Categories	CtQ factors
Protocol Design	*Eligibility Criteria* *Randomization* *Masking* *Types of Controls* *Data Quantity (CTTI) – Data Quality (ICH E8)* *Endpoints* *Procedures Supporting Study Endpoints and Data Integrity* *Investigational Product (IP) Handling and Administration*
Feasibility	*Study and Site Feasibility* *Accrual (i.e., Enrollment Strategy)*
Patient Safety	*Informed Consent* *Withdrawal Criteria and Trial Participant Retention* *Signal Detection* *Safety Reporting* *Data Monitoring Committee (DMC) / Stopping Rules (if applicable)*
Study Conduct	*Training* *Data Recording and Reporting* *Data Monitoring and Management* *Statistical Analysis*
Study Reporting	*Dissemination of Study Results*
Third-party Engagement	*Delegation of Sponsor Responsibilities and Collaborations*

Fig 5. CTTI and ICH E8 CtQ categories and factors

Considering this frame of reference and to ensure QbD, the Clinical Data Scientist must strongly contribute to, if not lead, the mitigation of some the risks associated with CtQ factors.

There are many risk areas associated with the CtQ factors, including but are not limited to the:

- Complexity of protocol designs such as umbrella, basket, platform and adaptive
- Vulnerability of the patient population (e.g., elderly, pediatric)
- Complexity of enrollment procedures (e.g., consent, eligibility, stratification and randomization)
- Deviations from standard of care

- Characteristics of the participating countries (e.g., Standard of care, customs, dialects)
- Planned rate and distribution of enrollment
- Number, profile and experience of the study sites and countries
- Nature of the protocol-required procedures, with specific emphasis on the burden they may place on patients and sites
- Organization of the trial (e.g., site-centric vs. decentralized) with telemedicine and home nursing
- Planned technologies used to collect data including when patients bring their own device (BYOD)
- Complexity of the data flow, including variety of the data sources
- Oversight of the capture and modification of the eSource data owned by the sites
- Number and experience of the data and operational vendors
- Any other study execution activities which may lead to data errors that could negatively impact the credibility and reliability of the trial results

While many risks would be evaluated and accounted for in the overall risk mitigation plans from the multidisciplinary study team, some risks such as the complexity of the data, data flows, third party vendors and planned technologies used for data collection would be the primary focus of CDS.

To foster study QbD, Clinical Data Scientists and study teams must understand the advanced concepts introduced in Part 1[1] and 2[2]. This may require changes in both the composition of the protocol review team as well as the process for developing protocols. To ease this evolution, CDS organizations could pre-define guidance for the mitigations of risk associated with standard CtQ factors by leveraging historical information on process and data issues. As the example in figure 6 suggests, the Clinical Data Scientists will need to plan for a robust risk monitoring strategy using analytics tools including key risk indicators (KRIs) and quality tolerance limits (QTLs) to proactively identify trends resulting from known risks such as enrollment speed. Those monitoring strategies could be documented in the sponsor's Integrated Quality Management Plan (IQMP) or in the Data Management Plan (DMP).

	Scenarios for enrollment speed Very fast or very slow enrollment impacting data review strategies	
Potential Data Quality Risks	**Fast Enrollment (Scenario #1)** Unable to match the speed / frequency of data reviews with speed/volume of data collection	**Slow Enrollment (Scenario #2)** Higher risks of sites closure prior to DB Lock impacting data review and reconciliation strategies
Mitigations	• Leverage technologies (e.g., EDC, IRT, eCOA) to identify error as fast and as close as possible from the data source • Focus review on critical data points impacting eligibility • Prioritize the readiness of data review tools that matter most (Incl. KRIs & QTLs)	• Trend Screen Failure metrics • Monitor site closures • Fine-tune data review frequency considering early sites closure risks

Fig 6. Example of CDS risk mitigation guidance

Moreover, those mitigation measures should be pre-identified and implemented to the extent possible prior to study start to allow for the optimum quality controls during study conduct.

b) Risk-based study execution (i.e., The Quality Control stage)

During study conduct, the Clinical Data Scientist should continuously monitor the data related risks by:

- Monitoring the risks identified during the study design phase using the defined KRIs and QTLs
- Performing holistic data reviews including review of the various data audit trails
- Monitoring for the possible emergence of any new risks, including but not limited to:
  - Risk to database availability which could delay study start
  - Risk to study timelines which could negatively impact the availability of study results for safety reviews, the potential submission and product approval
  - Impact of protocol amendments
- Assessing the effectiveness of the implemented risk mitigations
- Adjusting or augmenting risk mitigations as necessary

The identification of outliers and atypical data patterns frequently rely on KRIs and statistical methodology tools. This means the resolution of "issues" is no longer as simple as using SDV and queries from edit checks to verify the accuracy of the data. Instead, the study team must understand the signal generated and apply critical reasoning to analyze the likely root causes. Once a signal is determined to be an issue, the underlying process or data issue needs to be addressed. Lastly, to close the loop, teams must follow-up to make sure the issue has been fully resolved.

Below are some examples of signals that can be found with the potential responses made by teams.

1. All patients at a site in Puerto Rico are Hispanic: An atypical proportion of one ethnicity at the site may be statistically outlying compared to other study sites outside South America but not unexpected in this case. The team does not need to act on the signal but should follow up until the site has finished recruiting to see if the pattern evolves.
2. Many patients at a site have the same respiratory rate: Rather than questioning if the value was correctly entered into the source document, teams should think about how this lack of variability occurred. It is possible, but highly unlikely, that many patients at a site have the same respiratory rate. It is more likely that something was wrong with how the measurements were taken and/or recorded. In this case, the process used to collect and record the rate should be reviewed, and the importance of accurate data collection and recording reiterated to the site personnel. Since the current data is not going to change, any issue with the process in taking measurements should be addressed, fixed, and monitored.
3. Patients on an oncology trial have either no or a very low number of adverse events (AEs): This is statistically unlikely. The study team should ensure the site personnel understand how to collect AEs, and possibly use source data review (SDR) to check for unreported AEs. The site personnel may need retraining, and the study team must follow up to make sure the situation is resolved. Current data might not change, but the process must be fixed and then tracked for ongoing correctness.

To address the examples above, the Clinical Data Scientists and study team must dig deep into the data to understand the root cause of the issues. They need to perform detailed analysis of KRIs and data review findings to resolve them. Occasionally, the team will need to go through multiple iterations of analysis and follow-up to fully understand the root cause. This requires a focus on details and strong communication skills as most findings will not result in queries but rather in addressing systematic process issues and site behaviors.

c) Continuous CDS process improvement

Clinical Data Scientists and study teams must leverage the lessons learned during study execution and adapt the processes to prevent reoccurrence of the issue moving forward. For systematic issues, the mitigation of a specific risk may involve a corrective and preventive action (CAPA) to do so.

Although the CAPA process is usually driven centrally by the quality organization, Clinical Data Scientists should be familiar with it. They, as SMEs in the risk management lifecycle, should also be comfortable contributing to the process through the characterization of the risk and suggesting pragmatic and robust remediations and preventive actions.

d) Impact of Risk-based approaches to CDS role

CDM must evolve substantially if it is to support QbD and risk-based study execution. This will have a dramatic impact on the CDS roles. Overall, the end to end management of the operational and scientific risks must be embedded throughout the entire process. Figure 7 below is an example of a risk-based CDS process where new tasks depicted in green are added to the traditional CDM steps depicted in blue:

Fig 7. Example of a risk-based CDS process flow

To support such process, the scope of traditional CDM vs. risk-based CDS would be as follow:

Traditional CDM Scope	Risk-based CDS Scope
Focused on logical thinking (Output)	Focused on critical thinking (Outcome)
Study set-up upon protocol finalization	Quality by Design
Standard processes across studies	Risk-based processes tailored for each study
Focused on data integrity	Focused on data quality (i.e., data reliability)
Reviews of data after their collection	Risk-based data monitoring

This evolution would require the following CDS roles requirements:

Best Practices	Soft Skills
• Risk-based study execution • KRIs and QTLs life cycle	• Critical thinking and root cause analysis • Adaptability • Pragmatism • Influential leadership
Competencies	**Foundational Knowledge**
• Risk lifecycle management • Advanced analytics • Process management	• New research methodology (adaptive, master protocols) • Decentralized clinical trials approaches and technologies • Risk-based methodologies and regulations • Strong data flow and system literacy to investigate multifaceted issues