Improving Data Quality for Risk-based Monitoring and Data Review

With continued efforts to improve clinical processes, risk-based monitoring (RBM) approaches are becoming increasingly important for sponsors and CROs. RBM strategies involve identifying study risk factors including data quality, entry and review cycle times, incidence of adverse events and other measures – all to provide cost-effective decision-making for utilising clinical monitoring resources. As part of a successful strategy for reducing data quality issues over the life of a study, it is important to focus efforts on EDC (electronic data capture) design prior to study start to ensure critical data is reviewed and to encourage continuous and early feedback to sites.

Data Collection

For improved data quality and collection, it's critical to have well-designed case report forms (CRFs), patient visit flow and edit checks. CRF design review should involve a cross-functional team from programming, data management, clinical monitoring, statistics and medical departments, with attention given to areas that could have future data quality issues.

Simplicity: Can data collection design be further simplified?

Overly complicated designs can negatively impact data quality when not readily understood by the site and clinical monitoring, increase training needs, and require potential database revisions to correct unforeseen issues during the study. A high number of database revisions can pose a risk to data consistency, especially if downstream programming departments are not fully aware of the sequence of changes to CRFs and visit structure. While database revisions from protocol amendments are to be expected, careful review of the CRF casebook design prior to the start of the study should strive to prevent and reduce non-protocol revisions. During the design phase, fields should be evaluated for risk of future changes based on protocol and complexity of data collected. For example, if a field has high likelihood of revision in an amendment, a free entry field constrained with edit checks may be more flexible in the long run. As a measure of quality, the number of protocol and non-protocol revisions post go-live can be tracked against data quality issues for impacted fields or for the study in general. This provides a health check of the revisions in comparison to similar studies.

Data Driven: Is the data collection flow self-explanatory?

To guide the site through the entry process, the design should consider the best usage of disabling and enabling fields based on prior selections to reduce entry errors in subsequent fields, as well as the usage of visits and CRFs that trigger based on data responses. For example, if a safety follow-up visit is only required when a lab endpoint has occurred, the visit only becomes available after the endpoint exists. EDC casebook design should be reviewed from the viewpoint of the site for ease of use and to promote data quality upfront.

Critical Data

To target clinical monitoring and data review to the most important data, the cross-functional design review team should identify critical data points for primary and secondary statistical analysis, as well as safety analysis. The list of critical database fields should be prioritised within each category in terms of importance to the analysis. After the list has been agreed upon, the data management plan and monitoring plan should focus on cleaning and reviewing these fields. Classifying the data importance at study start-up provides clarity to the study team, sets the stage for data cleaning and monitoring, and serves as a training tool for new team members to focus their efforts.

With the prioritised list of critical data points in hand, clinical monitoring departments must evaluate whether to source verify the full critical list, a subset, or additional fields based on the clinical monitoring budget. It is recommended that the most important fields receive the greatest source-verification effort, with a portion of the budget assigned for evaluating other data. Setting the EDC system to require source verification and data review for selected fields visually highlights them and enables automatic status tracking from not started, to completed and requiring re-review due to data revision. In addition, drillable search lists of outstanding required work within a subject, site, or across sites allows for better time utilisation.

While there is no single, one-size-fits-all approach for reduced source-verification strategies, some approaches for verifying, such as every fourth patient or every fourth visit, may place a large review cost on descriptive, historical or secondary data not impacting the analysis. It is important to separate the need for evaluating compliance with protocol, and CRF completion instructions from the task of verifying or auditing the data against source documents. To assess completion compliance, the data can be reviewed remotely by data management or in-house monitoring with any data error trends communicated to the monitor or site prior to the monitoring visit. The review, evaluation and re-enforced training steps should be repeated at subsequent visits until site quality is deemed acceptable. A subjective score for the initial and periodic review could be tracked over time.

When the clinical monitor is at the site, the primary focus should be to review required fields for verification; however, additional planned and unplanned reviews may also occur. Planned increased source data verification or review can result from pre-visit planning by utilising data-driven monitoring tools that allow evaluation of site trends. When thresholds for data quality or query rates are exceeded, increased review may be warranted for specific CRFs, fields or in general to further inspect for quality issues. For example, a high rate of SDV (source data verification) queries, a large number of data revisions for a field, or data anomalies, could prompt additional source verification or re-verification of fields. While additional planned work is foreseen, there is a need to allow for unplanned review when the monitor is at the site. During source verification, if a trend is discovered or an issue is suspected, the clinical monitor should have leeway to investigate further and not solely be constrained to required fields. When physically at a site monitoring visit, the monitor is the best person to assess the source documents/data quality and adherence to procedures. This is not to say additional verification should be unlimited, but rather purposeful, as time permits. When exploratory review or additional planned review is deemed necessary, the work performed should be tracked in the EDC system to give further confidence to the data. Each data point verified serves as an audit of the database and should be documented for reporting and evaluation purposes. The goal is to give immediate feedback to the sites to correct the data, prevent future issues moving forward, and update guides/training where appropriate. In comparison to fixed strategies for review of entire patient records or visits, exploratory review or targeted planned additional field review has merits, and empowers the monitor to make decisions based on the available inputs.

Data Quality

Critical data in relation to data management requires special attention. Edit check programming and listing development efforts should support cleaning of the most critical data. When developing edit checks, query text can be prefixed, check code identifiers labelled, and manual queries tagged by type so that priority checks are readily identified by data management and extra care is taken when closing important queries. Similarly, monitoring manual queries can be tagged to identify when the query is attributed to a mismatch between CRF data and source document, due to a completion error or unreported data. Query categorisation allows for reporting and metrics when evaluating site trends.

Similar to clinical monitoring, fields important for data management review and cleaning can be highlighted as data management required on the CRF. While clinical monitors review each data point, data management review is typically based on exception via automated edit checks, listings and specified manual review. Therefore, the ability to mark multiple forms and visits as reviewed increases efficiency.

It is essential for clinical monitoring and data management departments to collaborate and communicate to help reduce data issues and evaluate site quality. To do so, teams should keep the following steps in mind:

- Review query trends across sites and patients: Query information displayed in data dashboards and reports allows teams to react in real time to changes in volume and data quality. The volume of open or answered queries attributes to workload for resource planning and monitoring visit scheduling. Aging metrics provide a measure for site timeliness. Query check frequency and rates of safety, critical data and re-queries highlight areas of data quality risk.
- Review unreported events: The rate of unreported safety events as a result of data review or source verification can also be used as a measure of data quality.
- Review missing data: The usage of field placeholders to describe why data is not available should be captured during site entry to reduce queries and questioning of missing data multiple times by different departments. Ideally, missing critical data should be evaluated for clustering, e.g., higher rates at certain sites, certain CRFs or certain fields.
- Review significant protocol violations: Because significant protocol violations can reduce available per-protocol patients in the analysis, the amount of violations, type, and frequency per patient and site warrants special attention by the study team. Depending on its significance or severity, a violation could trigger an earlier monitoring visit and may require additional data review across patients to rule out similar violations.

In summary, as monitoring and data review strategies evolve, attention should be placed on preventing issues through study design where possible, focusing efforts on critical data and evaluating data quality trends. As technology allows faster review and summarisation of outstanding work and data trends, the keys to improving data quality ultimately lie in quicker feedback to the site to prevent future issues and the ability of clinical monitoring and data management to adapt to the metrics presented.

Petra Cramer, Product Owner, eClinical Data Services, PAREXEL Informatics, has been with the company for over 15 years. Her background is in data management and EDC (electronic data capture). She currently works within product management for the DataLabs® EDC solution.
Email: petra.cramer@parexel.com