

Ensuring Data Quality on Outsourced Studies - The Role of the Data Scientist

Sarah Malbon, Roche , Welwyn Garden City, UK
Vikki Horton, Roche , Welwyn Garden City, UK

ABSTRACT

Ensuring consistently high quality data is delivered to clinical science & statistics for decision making and reporting is the number one priority of a data scientist in Data Management (DM). Maintaining the right level of sponsor oversight to ensure quality, when DM is outsourced to a Contract Research Organisation (CRO), requires a balance of different skills.

In Roche pRED CDM, steps have been taken to ensure quality data is provided. This starts with the data scientist understanding study assessments in the protocol to know what data is required to be collected; knowing the study primary & secondary endpoints; and what makes up composite data etc.

Investing the time to work closely with the CRO to set clear expectations, as early as the CRO selection process, is a crucial first step. Strong relationships will aid/avoid issue resolution later on.

In this presentation we will show some of the processes and tools used to maintain oversight.

INTRODUCTION

The aim of our poster will be to show how we ensure quality data is provided to our stakeholders, and ensuring Data Management (DM) has a correct understanding of the aims of the study. We will also explain the importance of building relationships with both internal and external teams whilst covering the tools which will be used for oversight.

DISCUSSION

How can the Data Scientist (DS) add value to a study team and what are the key requirements of working with a CRO without duplicating or increasing the work of the CRO?

We will cover how to build relationships at all levels, including setting mutual expectations for all functions both internal and external.

We will share the systems and tools used for maintaining oversight (e.g. Excel, Spotfire, guidance documents and metrics), as well as the structured Quality Control (QC) approach taken before delivery of a milestone, whilst aiming not to duplicate the work of the CRO.

1. BUILDING RELATIONSHIPS

Building relationships is a crucial step to the smooth running of the study, along with understanding each other's processes and requirements from the outset, without increasing the burden on either party.

In order to build relationships it is important to take the time to set expectations from both sponsor and CRO (i.e. what is expected, when, and how), and clearly documenting those roles and responsibilities.

It may take time to build good relationships and trust, however this can be lost very quickly on either side, therefore being proactive (e.g. performing and discussing risk assessments), and having strong communication and clear transparency regarding study status, are key in mitigating issues that might arise.

PhUSE 2015

It is important to set mutual expectations for all DM sub functions and to get all counterparts to discuss what is required.

The key ways to building relationships are to:

- Encourage two way discussions
- Create a positive environment
- Build relationships at all levels
- Foster strong relationships, which will help/avoid issue resolution
- Invest into governance for own function for strategic partnerships
- Capture exact scope of work in Work Order

2. UNDERSTANDING THE STUDY AND PROTOCOL

It is important that DM is involved early on in the protocol development and understand the key deliverables for the study. This will help understand what data that is required to ensure it is later collected in an optimal streamlined way, and thus only collecting what is needed for analysis, The Roche DS should review the protocol as well as the CRO to:

- Ensure that Roche specifics around data collection are present in the protocol (e.g. inclusion/exclusion criteria and screen failures are not databased; physical exam abnormalities are captured as Medical History and/or Adverse Events)
- Understand the primary and secondary endpoints
- Understand any composite data, including review of the statistics section for detailed explanation of what is being included in the analysis
- Understand data being collected to ensure collection is optimal and streamlined, i.e. collecting only what's needed for analysis (In line with Roche internal initiative "Optimising Data Value")
- Understand the data collection requirements of different parts of a study, for example, Single Ascending Dose (SAD), Multiple Ascending Dose (MAD), Food Effect components; studies with both healthy volunteer and patient cohorts; studies with both adult and paediatric cohorts etc.

3. CRO SELECTION AND STUDY TEAM INPUT

The CRO Selection is an important part of the process that DM should contribute to, in order to ensure that both parties have the same level of understanding of the study, what is required and what is expected from each side.

This allows the chance for Roche to raise and discuss requirements that may be specific to them as a sponsor or any study specific requirements that may be outside the normal expectations.

3.1 REQUEST FOR PROPOSAL (RFP) REVIEW

For the Roche DS, key points to consider are provided within the guidance documentation that is available within Roche. The RFP should preferably be reviewed by the Roche DS, and relevant questions provided to the CRO ahead of any bid defense meeting.

In addition to the protocol assessments and study timelines, the Roche DS should take into account the following:

- Number and timing of Study Data Tabulation Model view (SDTMv) deliveries
- Meeting requirements with CRO counterpart
- Study specific metrics that may be required
- Roche specific processes (where a niche provider is considered)
- Science data review requirements
- Vendors and electronic data types

3.2 BID DEFENSE MEETINGS, REVIEW OF CURRICULUM VITAE (CVs)

The Roche DS should attend the whole bid defense meeting where multiple functions are being outsourced in order to ensure an understanding of how the DM group works within the study team as a whole. Items that are relevant to DM may well be discussed by other functions. The bid defense meeting can also provide clarification of the requirements of other functions that may be relevant to the study needs, and that could have an impact on the CRO selected.

The Roche DS should use bid defense meetings as an opportunity to ask questions and obtain any information that they feel is relevant and that has not previously been provided. The meeting can also be used to clarify any points provided in the RFP to better understand the CRO, and how they will work on the study, to ensure the CROs understanding of the study requirements and their ability to deliver.

PhUSE 2015

The Roche DS should also use these meetings to ensure that the CRO is aware of any Roche or study specific requirements for the study.

During the CRO selection process, the Roche DS should receive and review the CV of the proposed CRO DM contacts, in order to understand the experience for the proposed team and identify any support that they may require during the study.

3.3 KICK OFF MEETINGS

Every new study team will have a kick off meeting with all functions represented. In addition, a dedicated DM specific kick off meeting is recommended, to allow more time on DM focused topics. The aim of the kick off meeting is for the following:

- Setting expectations across the team (dependant on study type and teams previous relationships)
- Agreeing timelines
- Agreeing metrics that are required, and the appropriate frequency
- Agree on documentation required and responsibilities from both CRO and Roche
- Highlight Roche specific requirements (Sample numbering, Sample and Shipping Reconciliation, Serious Adverse Event (SAE) Reconciliation, and coding)
- Document any actions that are identified as required for the smooth running of the study and who is responsible for the action. This will ensure transparency, and aid any handovers should they be required

4. QUALITY

Data Quality is a measurement of the state of collected data (validity / accuracy / completeness / conformance) for a specific analysis/business need, including acceptance criteria defined by the stakeholders. If data quality is not sufficient, results may be wrong/unfit for the intended purpose.

How do we ensure data quality? In addition to communication and meeting expectations, as mentioned previously, the following should also be considered.

4.1 REGULAR CALLS (SMT AND DM)

It is important that the Roche and CRO DM teams meet regularly. The meeting requirements will have been highlighted at the RFP stage, and should be discussed and documented in the DM kick off meeting.

Regular study team meetings will be held between the CRO and the Roche study team to discuss the study progress, as well as to discuss and mitigate any risks. The CRO would be expected to provide the DM status and provide any pertinent information regarding the status of the data (e.g. explaining inconsistent information due to cleaning) and confirm any follow-up actions on issue resolution (e.g. outstanding data entry and query resolution).

Dependant on the study type, complexity, and duration of the study, regular DM focused meetings may be required to allow the DM functions from both Roche and CRO to meet regularly during the course of the study. The meeting focus and frequency of which may change as needed to ensure the best use of the entire team's time.

4.2 ONGOING COMMUNICATION (EMAILS, PHONE CALLS ETC.)

Communication should not be limited to the scheduled team meetings only, and other methods should be considered as appropriate.

It is important that all members of the study team consider the message to be provided and the most pertinent way to deliver the information required. Communication is key to the CRO and Sponsor working relationship.

Best practices should be considered and a study communication plan implemented and followed to ensure that optimal relationships are maintained.

The team should consider the following:

- How urgent is the information to the intended recipient?
- Is the information sensitive?
- Would a phone call be the best way to deliver the information?
- Decisions and actions agreed on the phone should be documented in some way (a follow up email straight after the call may well suffice)
- When should the information be delivered, can it wait for a study team meeting or is it more urgent?

PhUSE 2015

Email is a good communication tool; however misunderstandings could arise due to assumptions, unclear information, or differences in processes. Everyone should consider if the medium being used is appropriate, and follow up to ensure that the recipient understands the message and any potential implications.

Where email starts to become a long trail the recipients should consider if a meeting would be more appropriate for the discussions, to avoid frustrations which could be easily avoided. The meeting can then be documented with clear decisions and actions after the discussion has concluded.

4.3 REVIEW OF DM DOCUMENTS

The Roche DS should review the DM study documents during study set-up, including, but not limited to, the eCRF, Data Management Plan, SDTMv mapping specifications etc, to ensure that the cleaning processes match the expectations of the study team. The documentation for each CRO will differ; therefore the Roche DS should ensure that the Roche minimum requirements are covered within the documents (e.g. coding and SAE Reconciliation process).

The Roche DS should take into account that the CRO database and processes may differ to the Roche processes and that this is not an issue as long as the appropriate data is collected and cleaning is planned to be performed as expected. For example, listings may be used by a CRO for data review and cleaning, whereas Roche might program a complex edit check or a custom function. All cleaning steps should be clearly documented to prevent any misunderstanding between Roche and the CRO.

The Roche DS may not be required to sign off all of the study documents that are created, however they should receive a copy of the final document for reference at Roche. The documents should be available on the study shared drive.

The documents that will be created, the level of review and any approvals required per CROs Standard Operating Procedures (SOPs) should be discussed and documented at the study kick off meeting, to ensure that the Roche and CRO requirements for documents are met.

4.4 QC OF SDTMv DOMAINS BY SPONSOR

SDTMv (SDTM view) consists of Clinical Data Interchange Standards Consortium (CDISC) SDTM and all Roche-defined extensions of the model.

A review process has been put in place for Roche DS to check the content of the SDTMv against the raw data at key stages in the study. This allows the Roche DS to check the volume of data expected per protocol, based on information provided via metrics during the study. This is an independent check from any quality checks that the CRO performs. Any findings from the Roche QC are documented on a log and communicated back to the CRO. This log is stored in a central location so that it can be accessed by all parties.

4.5 METRICS

Standard metrics have been agreed with the pRED DM preferred provider CROs, which provide a standardised way to provide metrics to Roche DM and Study Teams. The requirements can be adaptable depending on study type, duration and size.

A key metric for pRED DM is the percentage of missing primary and secondary data-points; whether missing because it has not yet been entered or missing because assessment has not been performed at site and so will never be provided.

The Roche DS should discuss the standard metrics, what is required, frequency and timing, as part of the kick off meeting. This should be reviewed during the course of the study in case the requirements change or more frequent updates are required, for example approaching an interim analysis or database lock.

5. DATA HANDLING AND CLEANING RULES

It is important to understand each CROs data handling rules and how their database flows in order to ensure that appropriate cleaning is in place. This will also help in reducing confusion in documentation review due to differences in processes.

The Roche DSs understanding of this information is key to their interpretation of the metrics and awareness of study status during conduct.

PhUSE 2015

Examples of this include:

- Protection of pre-filled or derived data to prevent the site amending the data
- Use of eCRF triggers (e.g. pregnancy for females)
- Understanding the CROs process for complex cross form edit checks versus the utilisation of data listings
- Understanding the data capture requirements within a form (e.g. questionnaires where answers to questions are dependent on how previous questions are answered)

5.1 ELECTRONIC DATA RECONCILIATION

The Roche DS should ensure that the CRO has an understanding of sample/shipping, as well as data reconciliation requirements of Roche. This should be made clear in the RFP/scope of work and discussed and documented during the kick off meeting, along with the frequency and associated metrics that will be required to demonstrate that this is being performed on an ongoing basis.

Data reconciliation should be performed to the analyte level, and not just the header level, taking into consideration any composite data.

Sample/shipping reconciliation needs to be performed to ensure samples taken per the eCRF have arrived at the relevant lab.

How, and when, to deal with any non-CRF blinded data should also be discussed and documented early on.

6. TOOLS

There are many tools that can be utilised to ensure that quality data is delivered to key stakeholders. This can take many forms from simple word and excel guidance documents to sophisticated computer programs for data review.

6.1 GUIDANCE DOCUMENTS

In order to aid the Roche DS in working on an outsourced study a number of guidance documents have been created. The aim of these is not to answer all questions and provide a rigid framework within which the Roche DS must work, but rather to provide guidance on what to expect and relevant questions to ask along with hints and tips on how the relationship and trust can be fostered and maintained.

The following guidance documents have been created:

6.1.1 CRO SELECTION (INCLUDING RFP) GUIDANCE

The intention of this guidance document is to detail the internal process of creating an RFP to be sent out to the selected CROs to start the bid defense. It also contains any departmental requirements that should be considered that may not be included in the templates.

This document includes how to review a CRO completed RFP to ensure that the understanding of what is required and what is anticipated is consistent across the board, where possible, from the outset.

This guidance document also covers additional areas such as conduct of bid defense meetings, to give guidance on the relevant questions to be asked, to ensure that any areas that may lead to confusion are covered.

6.1.2 KICK-OFF MEETING GUIDANCE DOCUMENT

Once a CRO has been selected and study teams have been formed, often a Study team level kick off meeting is held. However, depending on study type, there may be the need for a dedicated DM kick off meeting, to allow Roche DS to meet with their counterparts outside the full study team setting at the beginning of the study. A dedicated meeting will allow sufficient time to discuss DM topics in depth.

The kick-off meeting guidance document has been split out from the more general set-up phase guidance document as it has been found to be a key process in establishing the working relationship from the outset.

The dedicated DM meeting should occur as early as possible after the CRO is selected. The purpose of the meeting is to foster or build on existing relationships within the team, along with allowing discussion with counterparts to ensure that there is mutual understanding and expectations for the study. Decisions and actions such as review of certain elements of the scope of work can then be performed at this stage, rather than at a critical point in the study when it may have a greater impact.

PhUSE 2015

The meeting can be documented with the key information that is needed for the CRO and Sponsor for the study, along with any actions that are required during the course of the study.

6.1.3 STUDY SET-UP, CONDUCT AND LOCK GUIDANCE DOCUMENTS

Each phase of the study has been split up into 3 separate guidance documents (Study set-up, Study conduct, and Study lock and post lock activities) which form an overall guidance to Roche DSs.

These documents hold the key information that the Roche DS should bear in mind which going through the study from review and understanding the Scope of work right the way through to the database lock and post lock activities.

The documents provides guidance on the level of oversight that is required from the Roche DS ensuring that the Roche DS allows the CRO to follow their own processes, while ensuring that any Roche specific requirements are covered.

6.1.4 SDTMv QC GUIDANCE

A guidance document has been put in place to guide the Roche DS through an SDTMv QC process.

This process is not intended to duplicate the CRO cleaning or QC processes, but more to allow the Roche DS to maintain the appropriate oversight of the study and ask any pertinent questions that may be required during the course of the study, and to spot check that study data is complete and accurate.

The review is not required across all domains and should be performed on key primary and secondary end point data, prior to any critical data delivery for the study,

The Roche DS is expected to be aware of the study stage when performing this review and only ask pertinent questions relevant at this stage of the study. For example, where the study is in conduct and the level of cleaning for certain domains for a deliverable have been agreed within the study team, any queries that are anticipated and allowed to be open, on ongoing data cleaning should not be raised.

6.1.5 THESAURUS STANDARDS CODING GUIDANCE

It is a Roche requirement that all studies where the DM function is performed by a CRO are still to be coded by the Roche Thesaurus group (There may be exceptions to this rule). The coding should then be provided back to the CRO for inclusion in the relevant SDTMv domain. This document covers the process which should be followed, file formats required and also provides general guidance regarding coded terms that should be considered.

This is in addition to the study set-up, conduct and lock guidance documents that also provide relevant information regarding timelines and requirements for coded data prior to lock.

6.1.6 SAMPLE RECONCILIATION GUIDANCE

A specific guidance document has been created to cover the topic of shipping, as well as data, reconciliation, which should aid the Roche DS to work with the CRO in raising awareness of what is required for their study and associated vendor data types.

The general outsourcing guidance documents contain information regarding timelines and the data transfer process, however the dedicated sample reconciliation guidance document should be used in proving oversight and guidance to the CRO.

6.1.7 CLINICAL DATABASE PROGRAMMER (CDP) SDTMv CREATION AND USER ACCEPTANCE TESTING (UAT) GUIDANCE

This document has been created for the Clinical Database Programmer to use to guide the CRO through the process they need to follow for the creation of the SDTMv mapping specification.

This document also provides guidance on the UAT process the CDP performs on the created SDTMv domains including use of LBPREF and controlled terminology. This UAT is performed programmatically and the output reviewed for anticipated findings (such as for study specific pages) against unexpected findings (such as controlled terminology not used or not consistent with the Roche standards).

The CDP is then responsible for providing comments back to the CRO via a shared questions and issues log which documents the back and forwards conversation regarding the data during the study.

PhUSE 2015

The CDP may also need to request advice regarding data mapping from the internal Roche data standards group in order to maintain consistency across studies and therapy areas.

6.2 OTHER TOOLS

Other tools have also been developed to aid the Roche DS and the CRO to ensure that the required data is delivered at the appropriate time points for the study. These include the following:

6.2.1 DATA DELIVERY EXPECTATIONS (DDE) DOCUMENT OR CRO EQUIVALENT (OPTIONAL)

The Data Delivery Expectations Document is an excel spreadsheet that should be used to promote and document discussion of the data requirements for key decision points in a study (For example, Interim Analysis etc.). It should be developed in conjunction with the DMP or CRO equivalent. A DDE meeting should take place during study set-up, which should include key stakeholders to ensure all data requirements for each milestone are discussed and agreed.

The DDE document should include the data required, the delivery type (for example from the eCRF or electronic transfer) and blinded requirements for the data review.

6.2.2 SDTMv MAPPING SPECIFICATION (INCLUDING LBPREF etc)

Roche need to understand the raw data being collected and the differences between this and the mapped SDTMv data (e.g. laboratory data transformed from vertical to horizontal format, use of controlled terminology). The raw data can vary between CROs and also from study to study dependant on the team and Electronic Data Capture (EDC) system utilised. It is important that there is an understanding not just of the mapped data, but also the raw data, how it is collected and the implications that this may have on the mapping process.

The current SDTMv mapping specification should be provided to the CRO during study set-up and this should be used to specify how the raw data will be mapped to the Roche SDTMv domains.

The LBPREF is a metadata tool to allow for standardisation of Laboratory (LB) test names and ensures the correct use of controlled terminology (CT).

6.2.3 CONFORMANCE CHECKER

The conformance checker is a standardised program that allows the CDP to check the SDTMv domains received against the SDTMv mapping structure. This structural check of the mapping is performed programmatically and highlights any possible issues or discrepancies that need to be reviewed or corrected.

This is purely a structural check and does not check the data content, which is covered in a separate process by the Roche DS.

6.2.4 TIBCO™ SPOTFIRE

Spotfire is a tool which is utilized internally at Roche for review of the data, either by listings or creating data visualizations. The SDTMv for all studies are loaded into a study specific folder (using a restricted access folder for blinded studies) to allow the Roche study team to be able to access the data in one location. This ensures everyone is working from the same, latest version, of the SDTMv domains.

Spotfire can be utilised in a number of ways internally to look at the data during the course of the study, this includes review by the internal science team when making key decisions (e.g. dose decisions), or when performing ongoing science review to ensure oversight from a science perspective.

It is important that the Roche DS and CDP maintain oversight of the SDTMv quality, to ensure that the team understand the data that is provided to them and the status of that data. The Roche DS should also ensure that the CRO is aware of how the data will be utilised internally when provided on a regular basis, and the key stakeholders who will be looking at the data.

CONCLUSION

In conclusion, the key to good quality data delivered to stakeholders is good collaboration and communication, with clear expectations set at study start, which are documented and agreed.

It can take a long time to build a good relationship and trust, which could be lost very quickly. Work from both sponsor and CRO is required. Proactivity, communication and transparency regarding study status are key in mitigating issues that might arise.

PhUSE 2015

Simple solutions include, but are not limited to, capturing the exact requirements in the work order; investing the time in setting expectations; and checking and measuring delivery according to expectations. In addition, building relationships at all levels, asking questions, and encouraging two-way discussion can be extremely beneficial.

CONTACT INFORMATION

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

Sarah Malbon / Vikki Horton
Roche Products Ltd
Falcon Way
Welwyn Garden City, AL7 1TW
Work Phone: 017070 366582 / 01707 361070 :
Email: sarah.malbon@roche.com / vikki.horton@roche.com

Brand and product names are trademarks of their respective companies.