PharmaSUG 2016 - Paper SS03

Strategic Considerations for CDISC Implementation

Amber Randall, Axio Research, Seattle, WA
William Coar, Axio Research, Seattle, WA

ABSTRACT

The Prescription Drug User Fee Act (PDUFA) V Guidance mandates Electronic Common Technical Document (eCTD) format for regulatory submissions by May 2017. The implementation of CDISC data standards is not a one-size-fits-all process and can present both a substantial technical challenge and potential high cost to study teams. There are many factors that should be considered in strategizing when and how to implement which include timeline, study team expertise, and final goals. Different approaches may be more efficient for brand new studies as compared to existing or completed studies. Should CDISC standards be implemented right from the beginning or does it make sense to convert data once it is known that the study product will indeed be submitted for approval? Does a study team already have the technical expertise to implement data standards? If not, is it more cost effective to invest in training in-house or to hire contractors? How does a company identify reliable and knowledgeable contractors? Are contractors skilled in SAS programming sufficient or will they also need in-depth CDISC expertise? How can the work of contractors be validated?

Our experience as a data management and statistical CRO has allowed us to observe and participate in many approaches to this challenging process. What has become clear is that a good, informed strategy planned from the beginning can greatly increase efficiency and cost effectiveness and reduce stress and unanticipated surprises.

INTRODUCTION

The PDUFA V Guidance mandates submissions for NDAs, ANDAs, BLAs, INDs and Master Files in eCTD format by May 2017. New data standards and terminology apply prospectively and are currently only required for studies that start within 12 months after the final guidance is issued, but will eventually be required for all clinical research. Standardization is a critical factor in improving the effectiveness and efficiency of the regulatory process by providing better data traceability, facilitating the combination of data and reuse of code, and lowering the learning curve for reviewers beginning to review a new study. The Clinical Data Interchange Standards Consortium (CDISC) has provided guidance on standards that should be used for clinical data. The implementation of CDISC data standards is not a one-size-fits-all process and can present both a substantial technical challenge and pose a potentially high cost to study teams. Errors in the process of creating submission-ready data using CDISC standards can delay submission acceptance, and ultimately delay getting effective drugs to patients. Many factors should be considered in the development of a strategic plan for each study such as: format of source (i.e., raw) data, data flow through production of tables, listings, and figures, final goals for submission, size and experience of a study team, available resources, and many other factors unique to each study team's circumstances. This essentially boils down to a careful consideration of when, who, how, and what is needed for any given team to accomplish submission goals.

The major components of CDISC which should be considered are:

- CDASH (Clinical Data Acquisition Standards Harmonization) – describes standards in CRF design
- SDTM (Study Data Tabulation Model) – describes standards in tabulating raw study data and is useful in data management activities
- ADaM (Analysis Data Model) – describes standards in analysis dataset design
- Define.xml – standardized model for presenting the meta data contained in SDTM and ADaM datasets

The ideal data flow as envisioned by the CDISC committee proceeds from Raw data > SDTM > ADaM > TLFs. However, SDTM and ADaM are submission models and are not always the most efficient data structures for daily operations. So, what is the best approach? That depends on the goals. Does the study team program use the most efficient programming approach to obtain critical answers as soon as possible not following submission standards and only consider SDTM or AdaM once submission is a real possibility? Or, do companies modify all of their existing tools to operate under submission standards? Both have merits but the latter seems far more likely goal than the former.

WHEN SHOULD CDISC IMPLEMENTATION BE CONSIDERED?

Choosing an approach to implement CDISC standards should be driven by considerations of the current stage of the

project, the final goal or purpose of study data, available resources, and even the stage of the developmental program. The process can be both expensive and time consuming and cost-conscious study teams are understandably considering all factors prior to beginning. Different approaches may be more efficient for brand new studies as compared to existing or completed studies. If a late stage trial is being performed and the possibility of submission is high or already planned, then it may make more sense to consider CDISC from the beginning of the study design. In the programming of interim reports for use in activities such as Data Monitoring Committee review, following CDISC standards from the beginning may prove to be very costly as it is quite challenging to program SDTM on early/incomplete data and any changes will need to be reflected in downstream deliverables. As data capture systems evolve, the output of these systems is moving closer to a structure similar to SDTM which should make earlier conversion to fully compliant SDTM easier. Some study teams may choose to implement the ideal CDISC data flow from the beginning while others may delay SDTM implementation until there is some hint of potential benefit that could lead to submission for regulatory review.

LEGACY DATA

The data collected for the majority of completed studies to date probably have not been structured to CDISC standards and analysis datasets are probably not specified based on ADaM principles. Bringing the data to CDISC standards would require retrofitting the database structure to SDTM. This mapping task may prove to be complex since neither the databases or CRF terms were designed with industry-wide standardization in mind. Furthermore, CDISC standards and guidelines evolve over time which necessitates the discussion of which version to choose and whether to update as standards refine. However, for these legacy studies, the study team does have the luxury of knowing the outcome of the trial as well as whether the data will be useful for future analyses. For studies that will not be submitted to the FDA or further explored in meta analyses, the investment to convert the data should be worth the value added in obtaining a standardized database.

ONGOING STUDY

Some study teams may face the challenge of deciding whether or not to convert data for an ongoing study that began prior to the requirement for CDISC standards, but may be submitted after the time when the standards requirements are in place. If the chance of submission is low, then making the effort to adhere to CIDSC may be expensive and time consuming. The study team may instead want to focus their efforts on only the data being analyzed and building analysis datasets in the most efficient way possible to achieve the next steps in the development program.

It is possible to retrofit SDTM into an existing project, but the decision must be made on how to structure the data flow. A key goal of CDISC implementation is traceability from data collection through analysis. CRF data is collected following CDASH guidelines, tabulated using SDTM, and analyzed with ADaM-compliant analysis datasets mapped from the SDTM datasets. If non-ADaM compliant analysis datasets and/or report elements such as tables, listings and figures have already been created, should they be reprogrammed to draw from the new SDTM datasets or should the SDTM be built completely independently of the existing analyses?

While retrofitting existing analysis datasets to pull data from the new SDTM allows traceability and is the most rigorous approach, it is also the most difficult. Put simply, it is extremely challenging to reproduce a set of tables, listings, and figures if the raw data need to be converted to SDTM and analysis datasets converted to ADaM. All existing downstream programming must be considered. Any existing reports which pull from existing analysis datasets must also be considered. The attempt to retrofit SDTM during the course of a study can be a never ending task as new data become available and standard guidelines may be updated. It can be difficult to plan for potential future data anomalies when programming early deliverables. This approach requires constant vigilance to track and update as well as potentially revalidate all existing programming. If analysis datasets for interim analyses have already been created from non-CDISC compliant raw data, some may argue that it may make more sense to continue this approach through the end of the study and program SDTM independently though care needs to be taken in documentation to assure traceability in the absence of standards. If no analysis datasets have been created, then it may be reasonable to plan for the proper data flow prior to their creation. If a final clinical statistical report is being created, it may be worthwhile to formalize the process and create final analysis datasets from compliant SDTM.

PROSPECTIVE DESIGN

The most optimal way to implement CDISC is to plan for it from the beginning. Design CDASH-compliant CRF pages that easily map to SDTM datasets. Disease area-specific variables that may not yet have standard mapping can be identified early and standards can be proposed. Creating ADaM-compliant datasets from clean SDTM should be relatively easy. If multiple similar studies are being conducted, standardization from the beginning can facilitate the reuse of analysis code and report elements and lead to an increase in overall efficiency.

WHERE, HOW, AND BY WHOM SHOULD CDISC STANDARDS BE IMPLEMENTED?

Before taking on the task of creating CDISC-compliant study data, a study team needs to consider whether they already have the in-house expertise necessary to do it correctly and efficiently. This requires a thorough knowledge of both the standards and the process as well as the requisite documentation and disease-area specifics. They should also acknowledge that creating CDISC-compliant study data can take a substantial amount of time and resources. Do they have the time to commit to this process?

CDISC classes can be expensive and may not always be offered in convenient places or at convenient times. In our experience attending a single class is generally not sufficient to feel prepared to lead a team in the production of CDISC-compliant deliverables. There are many factors to consider beyond just knowing the standards including process, validation, and required documentation. The classes do provide a good overview, but real world experience is key to becoming truly well-versed and fluent. The documentation and guidance describing CDISC standards is freely available, but learning on your own can also be time consuming. Programming CDISC standards is learned daily, not in a day. We have attempted this self-learning approach and found that it only works for the right type of individual, others tend to require additional guidance from a mentor. In addition to a substantial amount of time, it also takes a fair amount of maturity, judgment, natural curiosity, and respect for the importance of attention to detail. For a motivated senior-level programmer, becoming fluent in CDISC standards independently is possible, but success is more likely when the programmer has a substantial amount of time and availability to focus solely on that task. A smaller company may not have the resources to devote a full-time senior-level programmer to this. Further, CDISC standards continue to develop and refine. Completing the exercise once is never sufficient. Your CDISC expert will need to continue to stay abreast of new topics, new standards, and new best practices.

In reality, learning to specify study data to CDISC standards is only the beginning. A stable of programmers is still required to provide programming and validation support as well as produce the required supporting documentation. It is our experience that one CDISC project can tie-up a substantial amount of programming resources for up to several weeks. Many small companies may simply not be able to support this amount of resource allocation.

So, should a company undertake CDISC programming on its own or consider contracting out? This question may be answered differently depending on the type and goals of the company. This is typically a function of the size of the company, type of company, experience, and stage of pharmaceutical development.

STARTUP/SMALL PHARMACEUTICAL

Small pharma and startup companies face a litany of challenges. Many of them may be working to take their first drug to market. Their leadership is often learning to navigate the world of clinical research including IRB and regulatory requirements, process and documentation, setting up clinical sites, designing effective databases and CRFs, and the list goes on. Most are running on a tight budget without many resources to spare. Their employees are working lean and fast. If the company is conducting a Phase I study, then their main goal may be getting quick answers to secure more funding or clarify a direction rather than submission of data to the FDA. In this case adherence to CDISC standards can be delayed for a time. However, if the company is moving into a Phase II, CDISC compliance will need to be part of the discussion.

These companies will need to consider \whether they have the programming resources available for implementing CDISC in-house and whether the investment will be worthwhile in the future. Are there already more products and trials on the horizon? Will this institutional knowledge provide downstream efficiencies and potentially save money and resources? Is the company willing to assume the risk of possibly implementing CDISC incorrectly and potentially experiencing delays in submission as a result? For single product companies hiring specialty contractors or CROs may make more sense.

LARGE ESTABLISHED PHARMACEUTICAL COMPANIES

For large, established pharma companies with multiple trials and products, implementing CDISC in-house generally makes the most sense. These companies often have a large programming force and can afford to hire CDISC experts and dedicate several programmers to the task. They can consider CDISC standards from the beginning, starting with CDASH standards in their CRF design as they tend to build their own library of CRF pages. Programming for report items can be shared between studies because they will be based on known database structures. Once a certain level of standardization has been achieved, it can be applied to all other studies and products within the company, saving substantial, time, money, resources, and effort.

FULL SERVICE CLINICAL CRO

A large full service CRO can provide services including conducting trials, lab services, designing study databases, data management, regulatory advice, and statistical analyses. These companies work with a diverse client base and tend to have a large programming staff and the resources to devote to a CDISC-implementation team.

Standardization is inevitable for a full service CRO. In fact, industry-wide standards are welcomed. Offering CDISC standards from the onset of a study now can be considered value-added. Offering some CDISC compliance within the next two years will be required. These companies should be able to capitalize on the upfront costs of standardization and apply efficiencies across all projects.

SPECIALTY OR BOUTIQUE CRO

The two types of small boutique CROs that are likely to be asked to implement CDISC standards for a client are those that provide data management and those that provide independent statistical consulting. It can be challenging for these business models considering implementation of CDISC is not a simple function of data management or statistics. For those that provide data management as a service, developing processes and expertise to implement CDISC standards, especially CDASH and SDTM, in-house makes sense. Even if many of the studies supported by small data management CROs are early phase and may not be submitted to the FDA, it is likely that more clients will start expecting those standards as the PDUFA V deadline approaches. Implementing standards now can reap efficiencies downstream as it is more likely that standardized code and database design may be reusable between clients. If that CRO also provides statistical services, standardization can make analysis and reporting easier and far more efficient. As EDC systems develop, improve, and evolve to accommodate CDISC standards, it is likely that providing data in SDTM-compliant or SDTM-like standards is going to be easier since the data coming out of the EDC databases will be closer to the content and structure of SDTM.

For those CROs that provide statistical services, the question of whether or not to implement CDISC standards or provide SDTM programming services is a bit more nuanced. Statistical CROs at a minimum need to be familiar with SDTM standards since it is more likely that the client data that they receive will be in a compliant or similar format. It may make more sense for statistical CROs to focus their efforts on developing ADaM standards as standardized analysis datasets will facilitate the programming of report elements. Code for TLFs can be substantially be reused between studies if analysis datasets are in a standard structure. In addition, programmers at statistical CROs are often asked to move between projects to meet deliverable timelines. Standardized dataset structure decreases the time that it takes for a new programmer to become familiar with a project.

There is a philosophical debate on whether or not CDISC conversion is a data management responsibility or a statistics responsibility. Many statistical CROs are gaining much expertise in CDISC conversion projects, better understand the downstream effects (in production of analysis datasets and TLFs) due to changes to SDTM, and are more likely to be familiar with submission requirements. One can easily argue that it makes sense for statistical programmers to assume responsibility of SDTM. However, it is also arguable that implementation of CDISC should start with protocol design, CRF design, and development of EDC systems. This naturally suggests that data management play the key role of implementation of SDTM.

A less definitive question for a statistical CRO is whether they should consider offering programming services to convert existing study data into compliant SDTM as well as whether they should incorporate the ideal CDISC dataflow into their daily programming process. As noted earlier, the investment for SDTM can be expensive and time consuming. By taking on the task of converting existing study data into a compliant format, the CRO is also taking on the responsibility for doing it correctly or potentially risking submission delays for their clients. One consideration for bringing compliant data flow into the daily programming process is the type of deliverable being programmed. Final clinical study reports that may be used for submission should be programmed using a process as close to the ideal compliant data flow as possible. However, if interim analyses are being programmed, utilizing a compliant data flow may take more effort and cause more work than is necessary to accomplish the goal. Interim reports are often first programmed using very immature data. As early data become more complete, the SDTM will change and mature and the implications to all downstream analysis dataset and report element programming must be considered. This effort may not bring any additional value to the goal of the interim analyses.

A small statistical CRO really needs to consider whether CDISC-compliant programming either for legacy data conversion or in support of other deliverables is a service that they want to offer. It is clear there is a business need for this, but do they have the requisite programming resources to devote a substantial amount to this type of programming project? How will that additional resource requirement affect other projects? As an alternative to providing the service in-house, is it reasonable to consider the use of contractors to cover overflow programming requirements so that a statistical CRO can accommodate all of their client's CDISC-related needs?

CONTRACTORS AS A SOLUTION FOR CDISC IMPLEMENTATION

Contractors can be a desirable option for teams without the time, resources, or desire to take on implementing CDISC on their own. The right programming CRO can provide compliant deliverables on a tight timeline. Individual contractors can guarantee dedicated time to a project that an in-house programming team may not have available. Contractors can vary substantially in cost depending on skill level and experience. Hiring a CRO rather than just individual contractors comes with the benefit of having the option of transferring the management and assembly of

project teams to another organization. Hiring individuals may be more attractive for a project team that has some amount of availability and wants to either maintain more control over the project or gain experience while working with a limited number of CDISC experts.

WHO WILL BE YOUR EXPERT?

If your team decides to proceed with hiring individual programmers, you need to decide who will serve as your content expert. Will you look for an expert level of knowledge in your contractor or will your team keep the expertise and specification in-house and just hire out the actual programming and validation tasks. Both options have merit.

HOW DO YOU IDENTIFY RELIABLE AND KNOWLEDGABLE CONTRACTORS?

If a company decides to proceed with using contractors to specify or program CDISC datasets, the first challenge lies in identification and vetting of those contractors. Where should one begin looking for the right individuals? A search can start with a quick internet search query. Networking sites such as LinkedIn® may also be useful, especially when looking to identify someone for whom you may have a direct reference. Reputable recruiters and staffing agencies, specifically those focused in the pharmaceutical industry may prove to be more successful than independent searches.

Word of mouth networking and personal recommendations have been invaluable in our search for contract programmers – especially when it comes to CDISC expertise. On the surface there appears to be many contractors available for this work, but we have found a high percentage of those CVs to be inflated and actual legitimate experience and knowledge to be lacking. Those contractors with solid known reputations are often overbooked and unavailable. Simply looking for submission or CDISC experience in a resume may not be sufficient for identifying true expertise. It is our observation that implementing CDISC standards extends far beyond simply programming from an existing specifications document. Many potential contractors may have experience in working with a team implementing CDISC, but not in making critical decisions or in really understanding the importance of each detail or in creating the required CDISC documentation or Define files. Often disease-specific knowledge can also be important in proper implementation which adds another layer to the search.

It is important to define a project team model as well as immediate and long term goals prior to hiring contractors. Is your company looking for a contractor team that can complete the entire programming project from specification through programming and validation or will specification be done in-house with only the actual programming and validation assigned to contractors. Will one piece of the programming or validation step remain in-house to verify the work of the contractor? Who will give final assurance that the project has been completed correctly? Deciding up front will give your team a better idea of the level of experience and expertise that you will need to identify in your potential contracting pool.

WORKING WITH CONTRACTORS

In our experience very few individual contractors have the ability to simply take an assignment and run with it. Most continue to require some level of supervision from the employing company and/or CRO study team. Those that are able to work completely independently are often very expensive and difficult to procure – though the efficiency and peace of mind of knowing that you have really hired an expert that needs minimal supervision may be worth the investment. Accepting a bit more in-house management responsibility or CDISC expertise may allow your team to hire a less experienced and less costly contractor. You should consider where you want the bulk of the contractors' time to be applied – in specification or in programming.

Once a contractor is chosen a decision must be made on whether to allow that contractor to work remotely or require her to work onsite. Do your IT infrastructure and internal communication systems allow for easy remote work? Will your company allow an external contractor access to internal networks and file systems? Will your company provide equipment to that contractor if the choice is made to allow remote work? How will your data and files be kept secure?

The number of both programmers and contractors from countries outside of North America continues to grow. While this can be a benefit to the industry in providing a greater pool of resources, it also can present the challenges of both communication issues as well as different standards in training and education. These challenges can be successfully addressed, but they should be acknowledged from the start. We have found that assignments, requests, and expectations can be easily misunderstood in interactions with non-native speakers. It is best to communicate using multiple forms of media – generally both verbal and written – and be very clear in expectations up front. If you are looking for specific skill sets, be very direct and thorough in your interview and vetting process so that there are minimal surprises once the work begins.

With any contractor, it is also difficult to really know if the amount of time and/or work that your company is billed for is really the amount of time spent or work being done. And further it is difficult to know if time spent is commensurate with the assignments given. For these reasons, it is imperative to find contractors you know and trust. Efficiency will likely vary based on the skill level and experience of your contractor. In our experience, contractors often took longer

to complete an assignment than was initially expected during the proposal stages. Some of this time was spent learning company-specific procedures and becoming familiar with the project. More than a few times we were told that an assignment had been completed, only to find that it had not been or had not been completed thoroughly or up to CDISC standards. In these times, we needed to maintain an internal expert to provide final review and validation of the work. Some efficiency was gained in hiring the same contractors to do multiple projects once they had learned our company-specific procedures and expectations. At some level experience in working with several different contractors will be necessary for your project management team to get a sense of reasonable output, billing, and workload expectations.

Unless you have the luxury of working with a true industry expert or are a full service contracting organization, it is important to maintain some level of project management within your internal study team. We have found that having an internal point person to assign work and regularly check in, review work for correctness and adherence to internal procedures, and answer questions is very important. Your internal manager should also track timelines and deliverables. It is important to remember that at the end of the day, you are responsible for the work of the contractors that you hire. Consider who will bear the responsibility of the final validation and approval. Tools such as OpenCDISC make validation a bit easier, but should not be fully relied upon. We would suggest that at a minimum one person in your organization maintain at least a reasonable level of knowledge of CDISC standards as a final check before submission.

CONCLUSION

PDUFA V mandates standard electronic submissions by May 2017. Those involved in clinical research will need to consider and adhere to CDISC standards in the very near future. While the initial investment in standardization can be costly and time consuming, there are options for all study teams. Careful planning upfront can facilitate the process. The efforts to implement CDISC early on will potentially lead to additional efficiencies later in the study and in future studies. Standardization will make the review and approval process easier. As a service provider, we understand and appreciate industry wide standards just as much as FDA. We also understand that implementing such standards is challenging. Models for implementation exist, but will surely vary among organizations as a function of internal expertise, long term business goals, and clientele. A non-CRO perspective may differ greatly from our experiences. Regardless of your perspective, we have learned the most important thing is to continue to learn and be involved. Continue to network and continue to have an open mind and consider all available options in pursuit of your final submission goals.

REFERENCES

"CDISC Foundational Standards" Clinical Data Interchange Standards Consortium March 2016. Available at http://www.cdisc.org/foundation-standards.

FDA guidance for industry, Providing Regulatory Submissions in Electronic Format — Submissions Under Section 745A(a) of the Federal Food, Drug, and Cosmetic Act. Available at http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM333969.pdf.

ACKNOWLEDGMENTS

Thank you to Anna Leonen for review, insight, and continuing to push us toward the goal of CDISC compliance and Ruth McBride for guidance, experience, and the willingness to try new approaches to all of our challenges.

CONTACT INFORMATION

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

Name:	Amber Randall
Enterprise:	Axio Research
Address:	2601 4th Ave, Suite 200
City, State ZIP:	Seattle, WA 98121
Work Phone:	(206) 577-0177
Fax:	(206) 547-4671
E-mail:	amberr@axioresearch.com