PhUSE US Connect 2018

Paper DV09

Supporting A Data Review and Visualization Application With SDTM

Deepika Sunkara, MS, Covance, King of Prussia, Pennsylvania
Jagadish Dhanraj, MS, Medimmune, Gaithersburg, Maryland

ABSTRACT

Clinical research is driven by data. Data from clinical trials supports a wide range of clinical, safety, regulatory, and analytic groups who all share the same basic need: to efficiently access, analyze and review study data. We will provide a practical overview of how Medimmune planned for, launched and maintains a study data review and visualization application.
Topics discussed will include:

- Why leveraging CDISC SDTM was the best choice
- Designing a clinical data stream that supports both review and analysis
- Supporting consistent SDTM mapping with a sponsor implementation guide
- Ensuring that available data content supports all consumer needs
- Building a refresh and review process that makes accurate data available as needed
- Examples of how the application supported efficient use and review of study data

With SDTM standards, proper planning and the right application, sponsor reviewers can efficiently leverage data across studies and therapeutic areas. And integrating SDTM data standards into internal review processes helps insure that the SDTM data delivered to downstream users (including regulatory reviewers) is consistently formed and fit for (efficient) use.

INTRODUCTION

Clinical research is driven by data. Data from clinical trials supports a wide range of clinical, safety, regulatory, and analytic groups who all share the same basic need: to efficiently access, analyze and review study data.
"Standardizing study data makes the data more useful. Data that are standardized are easier to understand, analyze, review, and synthesize in an integrated manner in a single study or multiple studies, thereby enabling more effective . . . decisions. Standardized data also facilitate data exchange and sharing (e.g., between a contract research organization and a sponsor)."*

JREVIEW

Jreview is a tool used to perform data visualization, reporting, and exploration, which is designed explicitly for clinical research by people who are knowledgeable about clinical research data. Using JReview report objects, users can monitor data quality and patient safety, as well as visually explore data to identify trends. At Medimmune it is one of the tools used to review clinical study data received from EDC and external study data sources, i.e., lab, pk, etc. after standardization to SDTM format. JReview reports are generated for internal use only, primarily for data review, and are not used for any official purpose. Study physicians and clinical scientists performing medical monitoring of studies use JReview to facilitate medical safety review, clinical signal detection and to identify data trends and drill down to view supporting clinical data. Clinical Data Managers and project study managers use this tool to perform data cleaning activities. Lab data reviewers use it for aggregate review, to verify the consistency of conversion to standard units, and for grading of events.

DESIGNING A CLINICAL DATA STREAM

The raw data as collected will vary across studies based on the specific collection system and forms used. It is important to ensure that the data from across studies has a consistent, standard format and supports all the consumer requirements.

CDISC SDTM standards are used to harmonize collected data across studies for JReview, analysis, reporting and submission. A simple representation of the process is in Figure 1. The data standardization is only performed once with a single data stream used for JReview, ADaM (analysis) and submission. This approach is cost effective since each consumer group uses the same standardized data.

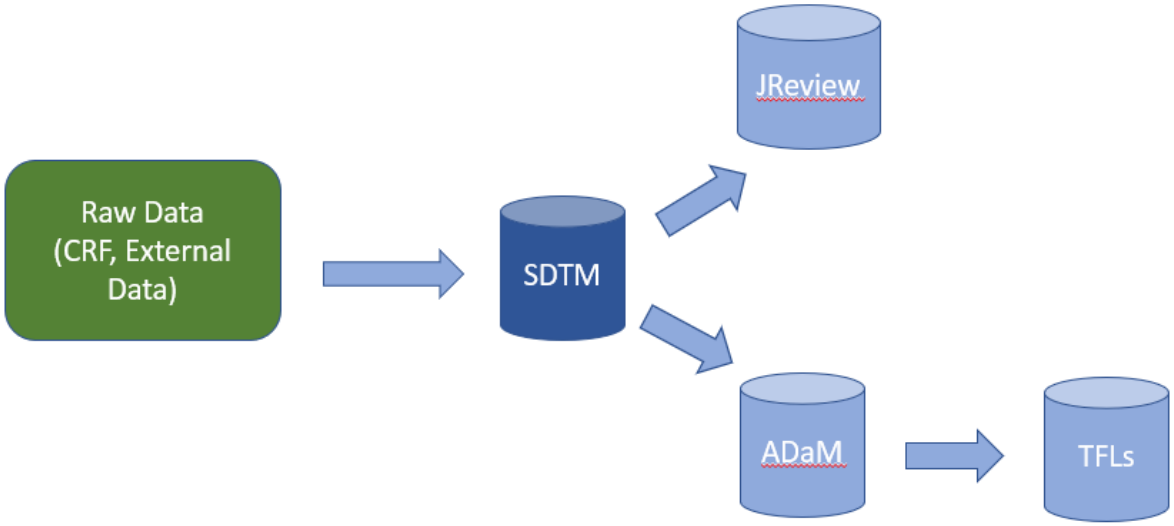

Figure 1: Standardization based on SDTM

LEVERAGING CDISC SDTM IS BEST CHOICE

The data required for JReview purposes and for analysis and submission will often differ. Leveraging CDISC SDTM standards is the best choice since standardized data is needed for submission and can also support JReview. However, CDISC SDTM as submitted is not an operational model and often does not fully support consumer needs. Therefore, it is important to understand the requirements of all the consumer groups before developing sponsor CDISC SDTM implementation standards.

Clinical reviewers prefer to review real time (or near real time) data. If raw data is used for data visualization, then the reports and visualization schema would have to be rebuilt and revalidated for each study. This would be time consuming and inefficient. For example, lab data as collected typically contains multiple units and reference ranges for a specific test across studies. In SDTM, all the lab results, units and reference ranges will be harmonized to standard units. Visualization will be efficient with harmonized lab data.

Clinical teams find it beneficial to review standardized data as it is more efficient. This is the same data that the statistical teams will be reviewing as well as for analysis and submission. This will reduce the effort from having multiple groups reviewing the data in multiple formats.

SDTM data reflects the complete conduct of the study. For JReview purposes, we include additional EDC system variables in SDTM (unique record id and when each record was first created and last updated) which are needed to establish a link to the data as collected, compare data across domains and access the quality of the collected data. The system variables are included in supplemental domains and are removed before the SDTM is used for analysis and reporting.

Since the JReview team needs more information than what will be included for submission, we needed to build the additional supplemental content (EDC system variables) into the standardization process. The additional content can be added either after the SDTM data is generated for submission or can be generated as initial step and later removed. We have designed our SDTM process to incorporate JReview requirements (SDTM + EDC system variables) as part of the initial step. This approach streamlines the path to near real time review of data in JReview and ensures SDTM data is available across all studies even if submission is not planned. This approach is shown in Figure 2 below.

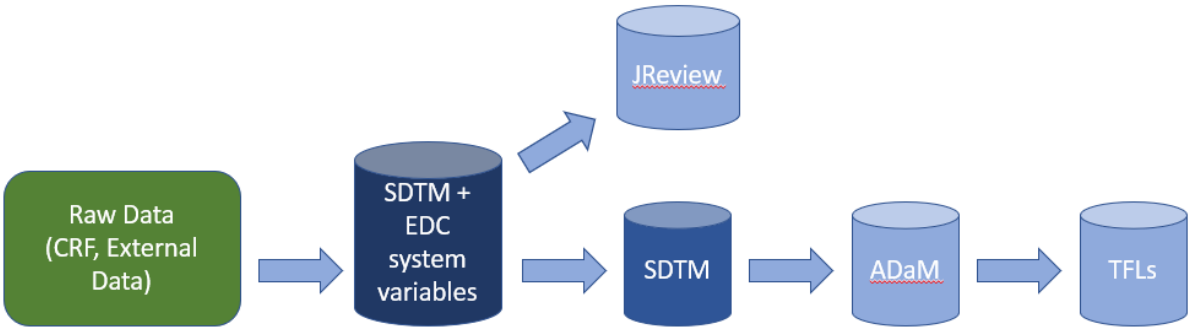

Figure 2: SDTM Standardization + EDC System Variables for JReview created before SDTM

With this approach JReview team needs can be most easily met as the data inputs are one step away from the data as collected.

JREVIEW STUDY SETUP PROCESS

The JReview team identifies studies that need to be mapped and provides the details needed to create a project plan to Covance. Once the study is started, the JReview team obtains the protocol (original and all amendments) and tracks all the timelines and study milestones till database Go Live. All the milestone dates and study requirements such as SDTM and ADaM are captured in project plan. Next the CRF (case report form) and external data transfer specifications are developed, and the study build is completed. After the study build process is completed, the rave database is activated and the ALS (architect load specifications) file is generated.

All the study documents (protocol, CRF, ALS file, visit matrix, external data specifications) and raw data (both CRF and external data) are provided to Covance for creation of SDTM. Once all the required study documents and raw data are available, Covance starts authoring SDTM specifications. Senior review is performed after the specifications are authored. After Covance senior review is completed, specifications are sent to Medimmune for approval. Programming is initiated once the specifications are approved by Medimmune. Pinnacle 21 is run on the study once all SDTM domains are programmed and validated. Specifications and programming (primary/validation) are updated as needed based on the Pinnacle 21 conformance report. A clean report is required before delivering SDTM datasets and specifications to Medimmune JReview team.

The JReview team performs post-processing of the SDTM data (merges parent domains with supplemental domains). The lead programmer then registers the study in JReview and loads all post-processed SDTM data into the application. After the initial JReview load, the JReview lead programmer regularly regenerates SDTM (using programs developed by Covance) based on the latest raw data and uploads updated SDTM datasets to JReview. Any issues or errors based on new raw data are documented by the JReview programmer in a communication log and the Covance study lead is notified. Once the SDTMs are updated to resolve the issue, the study lead documents the resolution in a communication log and notifies the JReview team. Then the JReview team reruns the programs to ensure the issue is resolved and closes the issue in the communication log. The entire process is shown in Figure 3 below.

Standard reports and visualizations within JReview are based on the SDTM data structure. Reports are not rebuilt for each study, thereby saving time for programming teams and reviewers. Clinical reviewers will always have access to the latest near real time data for review using this process. Currently dozens of studies are loaded in JReview and utilized by clinical review teams.

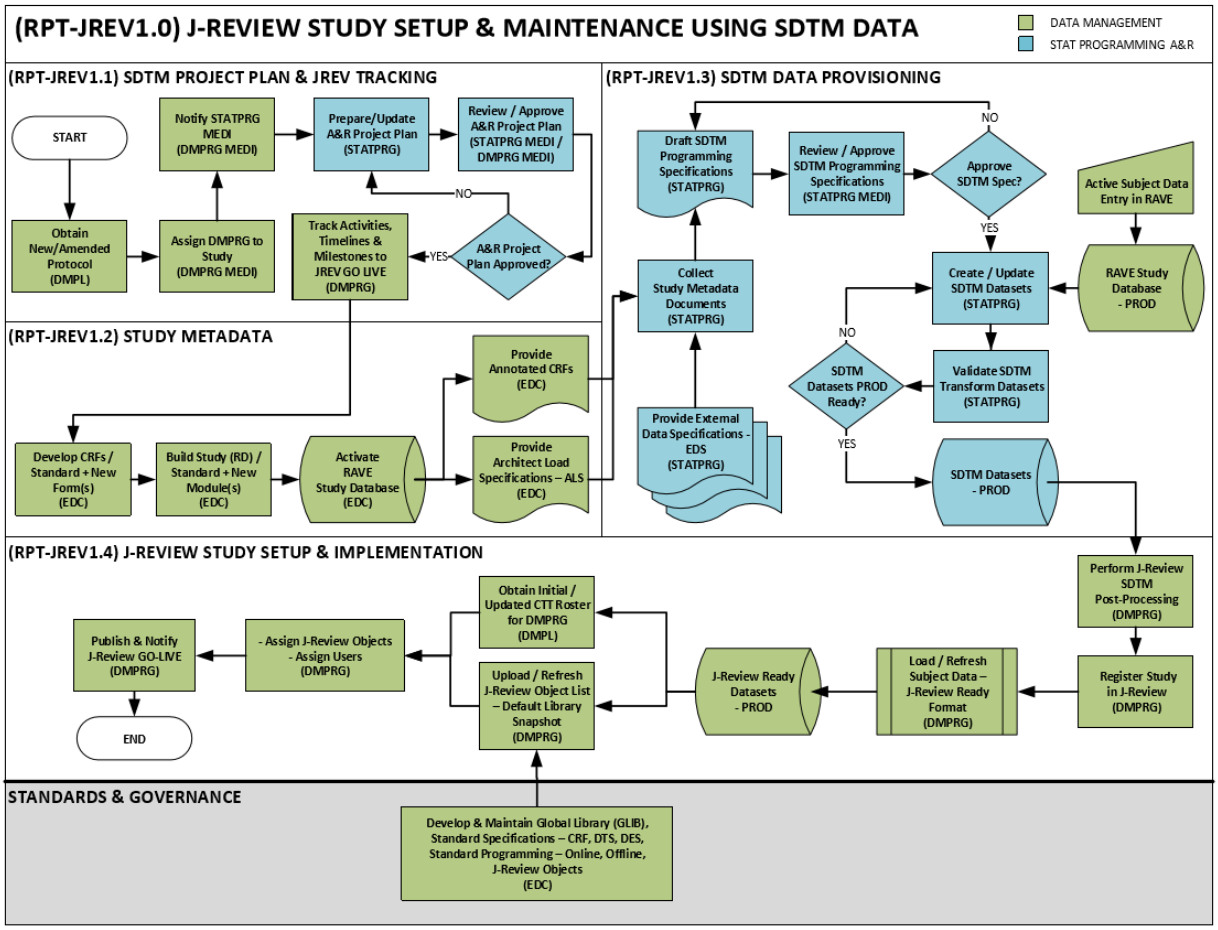

Figure 3: Study Process – JReview/SDTM

CONSISTENTLY IMPLEMENT SDTM

For the same reports and visualizations to be used across all studies the SDTM data must be consistently mapped. To ensure consistent mapping, Medimmune has developed an SDTM Interpretation Guide. The Medi SDTM IG consists of all the mapping standards based on the SDTM IG and Medimmune specific implementation rules. This resource is used to consistently map SDTMs across the studies. The Medi SDTM IG includes standard mapping for all supported domains and includes several operational tabs. The operational tabs target SDTM planning and management and include SDTM Planning and Raw Dataset-Domains Chart.

The SDTM Planning tab is completed as part of the SDTM planning meeting held prior to the start of mapping. It contains key information about the applicable SDTM standard and Pinnacle 21 version. It also contains an overview of the study protocol and basic timing (first subject dose, database lock, etc.) information for SDTM for each end user group (such as JReview and analysis).

The Raw Dataset-Domains Chart is generated prior to the SDTM planning meeting. It provides an overview of how raw datasets and CRF forms relate to SDTM domains. This tab can usefully highlight key mapping decisions that would benefit from discussion. This tab is used by the SDTM team while authoring specifications and programming. It also helps ensure that all the raw datasets and unique CRF forms are mapped to SDTM. Figure 4 shows an example of the raw dataset-domains chart.

CRF Form	CRF Pages	Raw DS	SDTM domain	Notes
Subject	1-2	SUBJECT	Not applicable	
Inform Consent	3-4	IC	DM, DS	
Entry/Randomizaton	5-6	ENTRAND	DM, DS	
Screen Failure	7-8	SF	DM, DS	
Screen Failure Eligibility	9-11	SFIE	IE	
Screen Failure Demographics	12-15	SFDM	DM, VS	
Screen Failure Adverse Event	16-20	SFAE	AE	

Figure 4: Raw Dataset-Domains Chart

The domains tab contains the domain level information required for the define.xml. This tab contains domain name and label (including supplemental domains) in alphabetical order. Key variables used to define a unique record are listed. Domain structure and class information is also available.

Tabs for individual domains consist of all the variable level details required for programming and define generation. All the mapping instructions (based on raw data and study documents) are provided in the programming notes column. Medimmune specific guidelines for mapping are provided for each variable. Figure 5 is an example of variable level content for Adverse Events.

DOMA	ORDER	KEEP_VAR	VARIABLE	LABEL	MEDI_GUIDANCE	PROG_NOTES	SOURCE	DEFINE_COMMENT	DATATYPE	LENGTH	TERMINOLOGY	CORE
AE	1	Y	STUDYID	Study Identifier	Must match SDTM DM.studyid	Set to STUDYID from SDTM DM		Set to STUDYID from SDTM DM.	Char	200		Req
AE	2	Y	DOMAIN	Domain Abbreviation	Set to AE	Set to AE		Set to AE.	Char	2	(DOMAIN)	Req
AE	3	Y	USUBJID	Unique Subject Identifier	Must match SDTM DM.usubjid	Set to USUBJID from SDTM DM		Set to USUBJID from SDTM DM.	Char	200		Req
AE	4	Y	AESEQ	Sequence Number	Based on "keys" variables in the domains tab; should define unique records within a subject	Sort by key variables, assign sequential integer by USUBJID.		Sort by key variables and assign sequential integer within each USUBJID.	Num	8		Req
AE	5	N	AEGRPID	Group ID	Typically not needed/populated. Can be used to associate related AEs for a subject. This is for categories that vary across subjects.	<drop>			Char	200		Perm
AE	6	Y	AEREFID	Reference ID	Typically used for references that are not on the CRF; use for the variable value that is used as a key reference for linking AE records and associated coding records.	Set to AE.recordid, SFAE.recordid, AECD.recordid, SFAECD.recordid (this is a key merge variable between AE, and AECD, SFAE and SFAECD)	AE.recordid | SFAE.recordid | AECD.recordid | SFAECD.recordid		Char	200		Perm

Figure 5: SDTM Specification

The Medi SDTM IG also contains tabs that associate standard raw variables with SDTM values for applicable domains.

DATA VISUALIZATION

Based on the data review considerations outlined by the clinical teams and available standardized SDTM data, we utilized the interactive dashboard functionality of JReview 12.0.1 to visualize data. The Dashboard is an assimilation of individual report objects developed and presented to meet the needs of reviewers and other data consumers.

The below image (figure 6) is an example of a composite Hy's law dashboard. This visualization schema displays composite Hy's Law plots in the middle section and supporting data listings with details of Liver function tests and adverse events on the left. The right side of the dashboard has a Graphical Patient profile that provides additional information needed and associated with a subject. Each object in the dashboard can be expanded or minimized as needed. The objects within the dashboard are interactive, allowing the reviewers to drill down to the subject level.

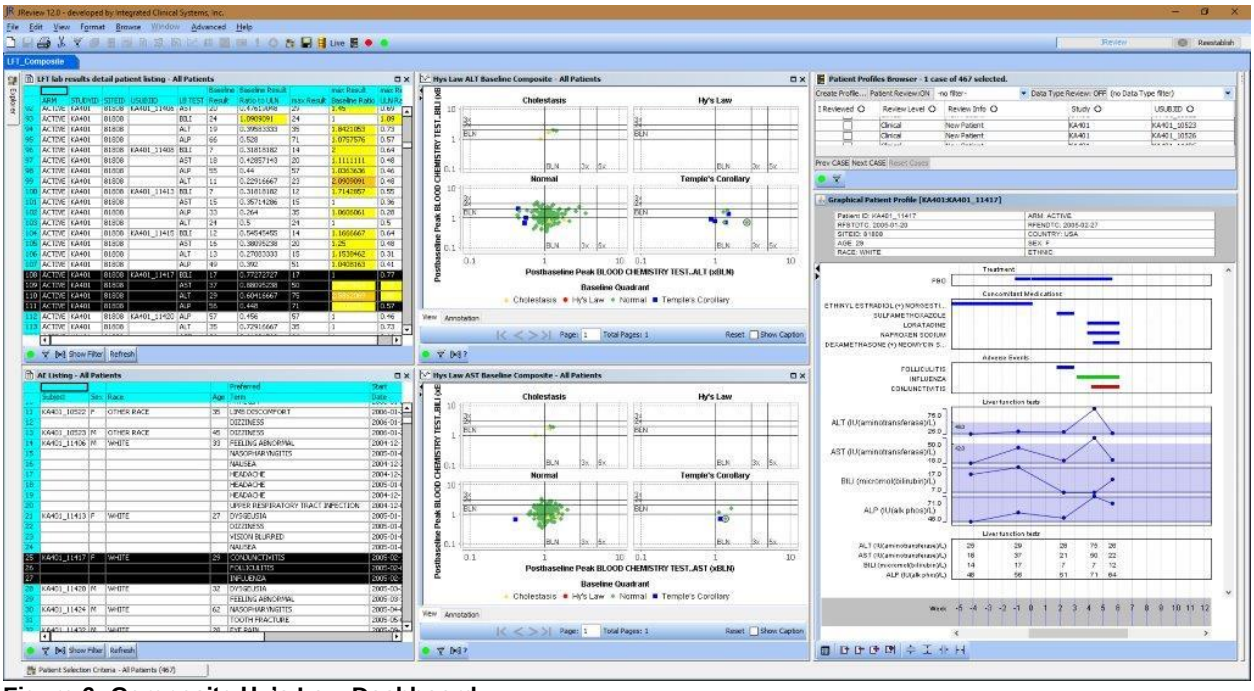

Figure 6: Composite Hy's Law Dashboard

This output is specific to Hy's Law parameters plotting ALT (aspartate aminotransferase) values against BILI (total bilirubin) values. The graph is divided into 4 equal quadrants using the values of 3 times the upper limit of normal for ALT and 2 times the upper limit of normal for BILI. The upper right quadrant will then contain data points for subjects

that are Hy's Law candidates. This graph should be used dynamically by selecting data points in this quadrant which will then interact with other open outputs for further analysis of these subjects. The output is further color-coded.

The patient profile output graphically shows how adverse event, concomitant medication, study treatment, clinical lab and vital signs data line up chronologically over the duration of a study for a single subject at a time. Demographic information is also displayed in the header information.

This is just one example of the graphical support that can be provided to data consumers based on standardized data. Other available outputs target disposition, clinical labs, adverse events, and demographics.

CONCLUSION

Clinical trials that are in the conduct phase require data review in a timely manner for early signal detection, which helps clinical teams to make informed decisions on patient safety. Reviewers often review/analyze data through unstructured reports from discrete sources in a non-standard way. That approach is time-consuming, resource-intensive and can produce inconsistent results. Leveraging SDTM to standardize data across all studies ensures that available data supports all the consumer requirements for JReview, analysis and submission and facilitates efficient use.

Using a linear approach that provides the end user with the ability to easily consolidate data and correlate vast amounts of information from multiple sources supports efficient data review. This optimized process helps to identify the relationship between events and findings and allows users to take advantage of the myriad information at their disposal to make appropriate assessments of patient safety. We successfully implemented JReview (fully integrated client/server data review and visualization tool) on data from studies across different therapeutic areas with a global library of standard reusable reports. Summaries and visualizations are performed via Graphic Patient Profiles, Cross tabulations, Listings and Graphs.

Automating the process by using Application Programming Interface (API) to obtain live data from clinical databases and integrating with JReview, enables individuals/teams to explore the data through visual analytics. This enhancement enables users to review and analyze data in near real time, facilitating a medical monitoring process which is efficient and cost effective since the same standardized data is used by all data consumers.

REFERENCES

Kirby S, Peterson T, Moving from Data Collection to Data Visualization and Analytics: Leveraging CDISC SDTM Standards to Support Data Marts, Presentation, PharmaSUG 2016]. Available at: https://www.pharmasug.org/2016-proceedings.html

*FDA, "Guidance for Industry Providing Regulatory Submissions in Electronic Format — Standardized Study Data [DRAFT]." FDA. FDA.gov. February 2012. Available at: http://www.fda.gov/downloads/BiologicsBloodVaccines/GuidanceComplianceRegulatoryInformation/Guidances/General/UCM292169.pdf

ACKNOWLEDGMENTS

The authors would like to thank Steven Kirby, Covance for supporting CDISC standards implementation and Medimmune Instream data review team for JReview tool (and this paper).

CONTACT INFORMATION

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

Name: Deepika Sunkara, MS
Enterprise: Covance
Work Phone: 610-324-9581
E-mail: deepika.sunkara@Chiltern.com

Name: Jagadish Dhanraj
Enterprise: Medimmune
Work Phone:
E-mail: dhanrajj@MedImmune.com