For the second year in a row, Genzyme had the winning poster at the FDA-DIA Computational Sciences March Meeting. So, we asked the winner to write us a blog about the poster content. Thank you, Julia Zhang, for this blog!

We would also like to thank Julia Zhang’s leader, Sue Dubman, for her vision, her adoption of CDISC standards end-to-end and her service on the CDISC Board. Sue is now in her fourth year and leader of the Strategy Committee. We believe that this poster won the contest since this is not dissimilar to what FDA is striving to do – improve their process by receiving quality data in standard formats in eSubmissions.

Quality by Design: Automating Validation of Standards-based Data, End-to-End
Julia Zhang and Sue Dubman, Genzyme

Have you experienced data quality issues? What is your strategy for improving your data quality?

Organizations make strategic and operational decisions based on data and patients rely on decisions based upon data. Poor quality data can negatively influence how a company is perceived in the marketplace since the outcome could potentially be a therapy that is unsafe; therefore, it is critical to ensure data quality is given the highest priority. Using standards can increase process efficiency and effectiveness, thus saving resources allocated to the clinical trial data process life cycle; standards can also improve compliance. Genzyme is in the process of implementing CDISC standards end-to-end, including the Protocol Representation Model (PRM), CDASH, LAB, SDTM, ADaM and Controlled Terminology as well as utilizing BRIDG as its underlying information model.

We are building a Metadata Repository (MDR) to govern data collection, data processing and data submission and to leverage the usage of different standards enterprise-wide. In order to increase efficiency and effectiveness, a Data Validation Tool (DVT) is needed for improving data quality and ensuring that data provided by one of Genzyme’s many partners or by internal teams matches all specified requirements; this ensures ‘quality by design’.

Data validation includes the processes and technologies involved in ensuring the conformance of data values to business requirements and acceptance criteria. It uses routines, often called "validation rules" or "check routines", that check for correctness, meaningfulness, and security of data that are input to the system. The ultimate goals resulting from implementing this Data Validation Tool (DVT) are to:

1) be able to check CRF, LAB, SDTM, ADaM data and define.xml files against CDISC standards and Genzyme-specific requirements to ensure that Genzyme receives, produces, and submits quality data;
2) align with the Genzyme Metadata Repository to ensure metadata validation;
3) automate and streamline data validation processes.

The data validation rules will be based on industry standard data validation checks that have already been developed, such as those for SDTM, ADaM, define.xml, and SEND, plus newly developed CDASH data validation checks and Central LAB data validation checks. In addition, we will develop company-specific data validation requirements based on Genzyme’s business needs. The DVT can play a role as a gateway ‘guard’ for outsourced study activities. Before data are delivered to Genzyme by CROs, EDC vendors or other third parties, these data will be recommended to load to the DVT. If the data passes the DVT checks,
it will be delivered to Genzyme; otherwise, the loaded data will be returned to the sender. The senders must ensure that data satisfies the requirements, and correct any quality issues. The DVT can also help us improve communications among different functional groups, CROs, EDC Vendors, and Central Labs. We also plan to add some more features, such as customized GUIs, which will allow users to add study-specific checks.

Using a well-designed DVT will help us reduce potential risks in data processes, improve data quality and smooth communications, thus improving our business process efficiency and effectiveness. The result is an anticipated ROI. This is a customized tool and we plan to build it based on OpenCDISC Validator technology in different phases.

Special thanks to:

Dana Soloff; Judith Goud; Max Kanevsky (OpenCDISC); Mihaela Simion, and Genzyme GetSMART team.