Validation Framework for Datasets from QT Studies: Leveraging Data Standards to Enable Automation

Saikrishna Bhavaraju, MSc[1], Lars Johannesen, PhD[1], Jeffry Florian, PhD[2], Mike Li, MSc[1], Jose Vicente, PhD[1]
[1]DCN/OCHEN/OND/FDA, [2]DARS/OCP/OTS/FDA

Abstract

Since 2005, all new drugs are required to undergo a clinical study to assess their potential to prolong the QT interval in the electrocardiogram and cause a potentially fatal arrhythmia. ICH E14 guideline provides recommendations concerning the design, conduct, analysis, and interpretation of this QT assessment. To reduce the data management time required to produce analysis ready datasets from data collected in these clinical studies, the Interdisciplinary Review Team for Cardiac Safety Studies (IRT) recently published a Technical Specifications Document for QT Studies (QT-TSD), which is based on existing data standards. This research investigated whether there are automation opportunities that could reduce the time needed to validate datasets submitted following the QT-TSD. A systematic review of the validation process of QT-TSD datasets documented the steps that could be automated in human- and machine-friendly spreadsheets and python scripts (plugins). The spreadsheets list QT-TSD data elements and their associated validation rules. Plugins implement the algorithm of each validation rule. Using these spreadsheets and plugins, an automatic validation framework in python reproduces the manual validation of any QT-TSD dataset as follows: the validator takes the QT-TSD dataset as input, reads and iterates through the spreadsheet's rows executing the plugin rules, and produces a validation report. To facilitate fixing dataset errors, the report includes description for errors found in the dataset. Lastly, if the validation is passed without errors, the framework also derives an analysis ready dataset that can be used out-of-the-box by IRT's analysis tools. Validation of 6 QT-TSD datasets showed the potential of the framework to speed up the review time of QT-TSD datasets, although further testing is ongoing. In addition, the validation framework can be applied to other therapeutic areas (e.g., ambulatory blood pressure monitoring) by replacing the QT-TSD spreadsheets and plugin rules with those of the therapeutic area of interest. Technical Specifications Documents like the QT-TSD facilitate efficient interchange of clinical datasets, usually with a focus on a specific area. This framework illustrates that TSDs also present an opportunity to reduce the time needed to validate, analyze and review clinical datasets when analysis needs, and processes are coupled and well-defined.

Introduction

- CDER's Interdisciplinary Review Team (IRT) for Cardiac Safety Studies regularly reviews clinical datasets from thorough QT (TQT) or substitute of TQT studies assessing the potential of new drugs to prolong the QT interval in the electrocardiogram and cause Torsade de Pointes, a potentially fatal arrhythmia.
- There are three types of analysis conducted during review of QT studies: by time (i.e., intersection union test or central tendency analysis), concentration-QT, and categorial (i.e., outliers).
- The Technical Specifications Document for QT studies (QT-TSD, https://www.fda.gov/media/128187/download) is based on data standards and specifies how to organize the contents of datasets from QT studies to reduce data management time and to enable generation of analysis-ready datasets.
- A validation framework to assess the adherence of datasets to the QT-TSD along with an application to derive analysis-ready datasets could streamline the review process further.

Materials and Methods

- Systematic review of the manual validation process for the ADEG and ADPC datasets following QT-TSD specifications.
- Summary of identified validation steps, associated data elements and validation rules.
- Plugin-based design to allow for reusability and extensibility of validation rules so that the framework can be customized and run for other therapeutic areas with minimal effort.

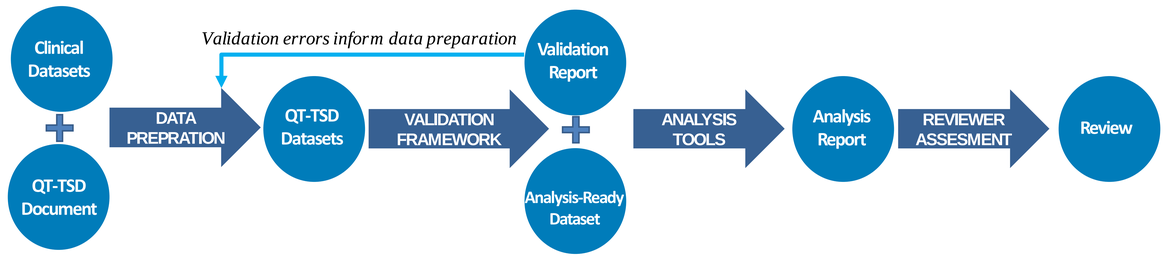

Figure 1. Semi-automatic validation and analysis workflow. The validation framework automatically generates a report that facilitates updating datasets to meet the QT-TSD. If validation passes, the framework derives an analysis-ready dataset that can be used "as-is" by IRT's analysis and reporting tools to inform the review.

Table 1. Abbreviations used in this poster.

Abbreviation	Reference
QT-TSD	QT-Technical Specifications Document
TQT	Thorough QT/QTc
ECG	Electrocardiogram
PC	Pharmacokinetic
CDISC	Clinical Data Interchange Standards Consortium
ADEG	Analysis Dataset for Electrocardiogram
ADPC	Pharmacokinetic Concentrations Analysis Dataset
STUDYID	Study Identifier
ANLzz	Analysis zz
ANLzzFL	Analysis Flag zz
ABPM	Ambulatory Blood Pressure
XPT	SAS xport file format

Results and Discussion

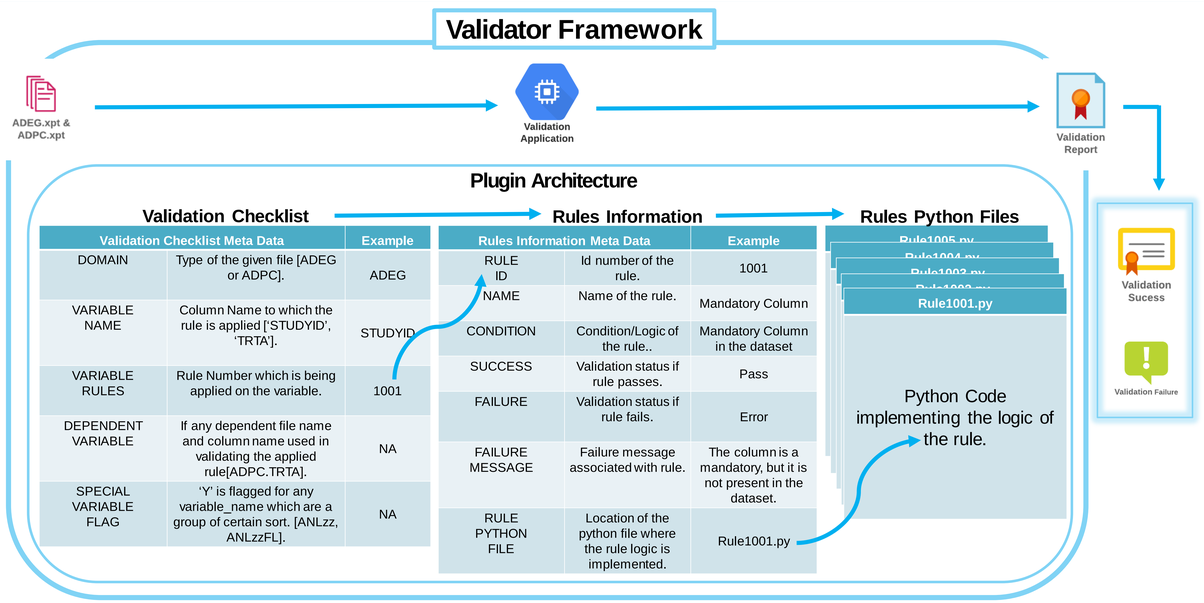

Figure 2. The validator framework is part of a data pipeline which takes in ADEG and ADPC files, applies validation rules, and generates a validation report. In addition, the validation uses a validation checklist, rules information and python rules files (i.e., plugins) as input and then iterates through the list applying each rule to ADEG and ADPC data generating a validation report that can be used to amend identified data issues.

- A semi-automatic pipeline that validates clinical datasets, generates a validation report and, if validation passes, an analysis-ready dataset that can be used "as-is" by IRT's analysis and reporting tools to inform the review.
- Validation rules coded in python scripts (plugins) so that each rule can be automatically applied to QT-TSD (ADEG and ADPC) datasets.
- Test cases for each rule as well as a QT-TSD compliant example dataset.
- Although further testing is ongoing, additional testing was performed with six different study designs showed the potential of the framework to speed up the review time of QT-TSD datasets.
- The validation framework can be customized for other therapeutic areas (e.g., ambulatory blood pressure monitoring) by replacing the validation checklist and providing dataset-specific rules as plugins.

Conclusion

The presented validation framework for datasets from QT studies illustrates that Technical Specifications Documents present an opportunity to reduce the time needed to validate, analyze and review clinical datasets when analysis needs, and processes are coupled and well-defined.

Acknowledgments

This Project was supported in part by an appointment of the Research Fellowship Program at the Division of Cardiology and Nephrology, Office of New Drugs, CDER, U.S Food and Drug Administration, administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the U.S Department of Energy and FDA.