

Paper DH10

White Paper: Janus Clinical and Nonclinical Loading: The most common issues identified from sponsor submissions

Kendra Worthy, PharmD, MS, Office of Translational Sciences, Office of Computational Science, CDER, US Food and Drug Administration

Acknowledgements:

Mohamed Taraouat, Deloitte, Silver Spring, MD, USA
Yonathan Tadese, Deloitte, Silver Spring, MD, USA
Kunal Modi, Deloitte, Silver Spring, MD, USA
Sai Krishana Borra, Deloitte, Silver Spring, MD, USA
Lilliam Rosario, PhD, Director, Office of Computational Science, CDER, FDA
Qais Hatim, PhD, Office of Computational Science, CDER, FDA

****This white paper reflects the views of the author and should not be construed to represent FDA's views or policies.****

ABSTRACT

The Office of Computational Science (OCS) drives the modernization of the Center for Drug Evaluation and Research (CDER's) scientific review process through the implementation of tools, services, and training to enable reviewers to apply their expertise to information. Janus Clinical is a data repository for subject-level clinical trial data submitted to FDA as part of a marketing application. Janus Clinical accepts Clinical Data Interchange Standards Consortium (CDISC) Study Data Tabulation Model (SDTM) compliant submissions packages. Janus Nonclinical is a data repository for validated CDISC Standard for Exchange of Nonclinical Data (SEND) compliant study data submitted by sponsors in support of Investigational New Drugs (INDs), Biologic License Applications (BLAs), and New Drug Applications (NDAs). We will examine critical and noncritical data loading issues from sponsor-generated SDTM & SEND datasets and the planned improvements to enhance and better categorize data loading into Janus Clinical and Nonclinical.

INTRODUCTION

Janus supports a transparent, consistent, and efficient scientific review process by performing automated extraction, transformation, loading, management, and integration of data to facilitate regulatory review. To provide CDER reviewers innovative and reliable solutions that improve and strengthen the scientific review process by integrating data, tools, and training. OCS works with reviewers to provide comprehensive tools, services, and training to support and improve the review process. OCS helps reviewers understand, analyze, and use their data so they spend more time thinking about the results of the data analyses rather than setting up those analyses. Our understanding of both regulatory review challenges and a range of technical solutions enables us to create solutions that impact the review process.

Implementation of Janus Clinical and Nonclinical affords CDER the opportunity to begin to leverage standard data to support better exploration of safety signals as well as new analysis approaches and methods, creation of standard analyses and graphs, and analyses across studies, drug classes, and

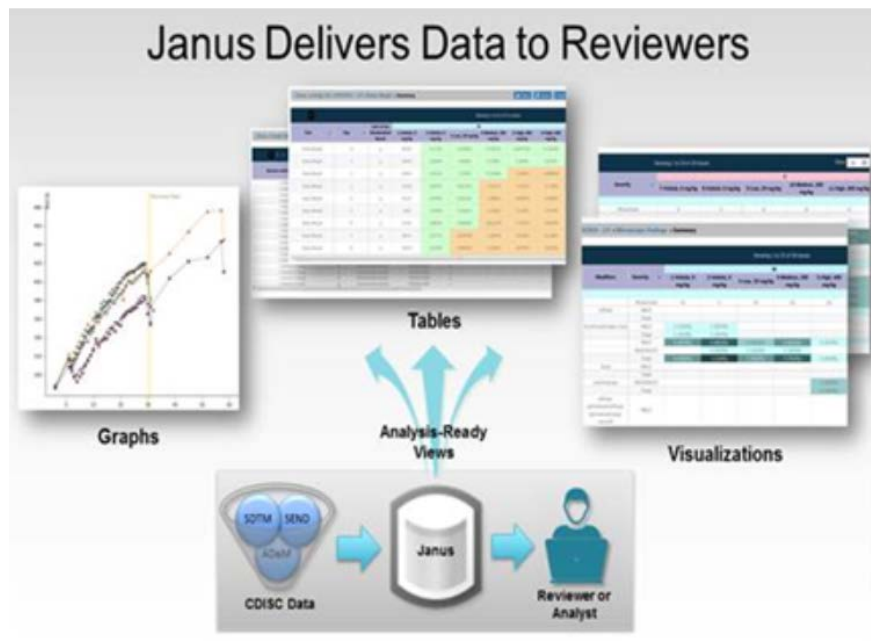
therapeutic products. Capture of standardized study data also provides an opportunity for more proactive regulatory monitoring and oversight of trials activity. These improvements ultimately will result in improved productivity—less time will be needed for routine data management activities to prepare for a review, and reviewers will have more time to perform their analyses and interpret results.

Janus Clinical

The Janus Clinical data service model, shown in Figure 1, comprises a series of components that:

- Validates, transforms, and loads study data into a warehouse repository
- Extracts data from the repository to create an analysis-ready database (of enhanced SDTM, Integrated Summary of Safety (ISS), Analysis Data Model (ADaM), and other data standard views) that can be accessed by reviewers using a variety of analysis tools
- Enables reviewers to access summary information about specific studies (study design, trial summary, explanation of codes used, location of specific variables, etc.); select, pool, filter, or subset study data, and/or perform complex queries with logical operators; and create and download customized analysis-ready datasets to support a variety of analytical activities and create and download customized reports to support and document review decisions.

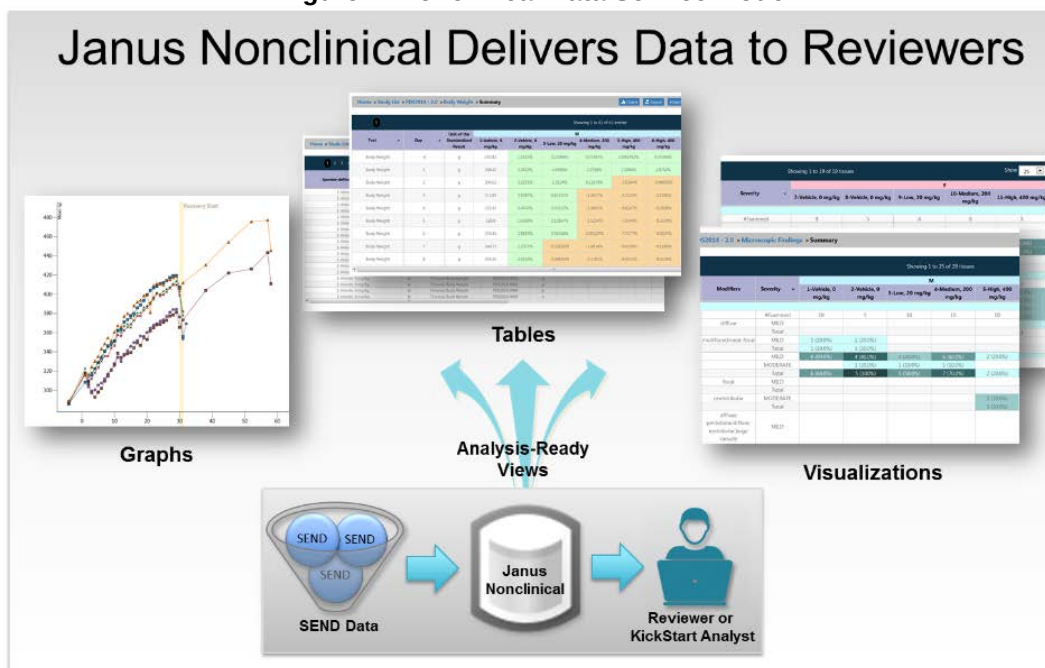
Figure 1. Janus Clinical Data Service Model



The service model, shown below in Figure 1, comprises a series of components that:

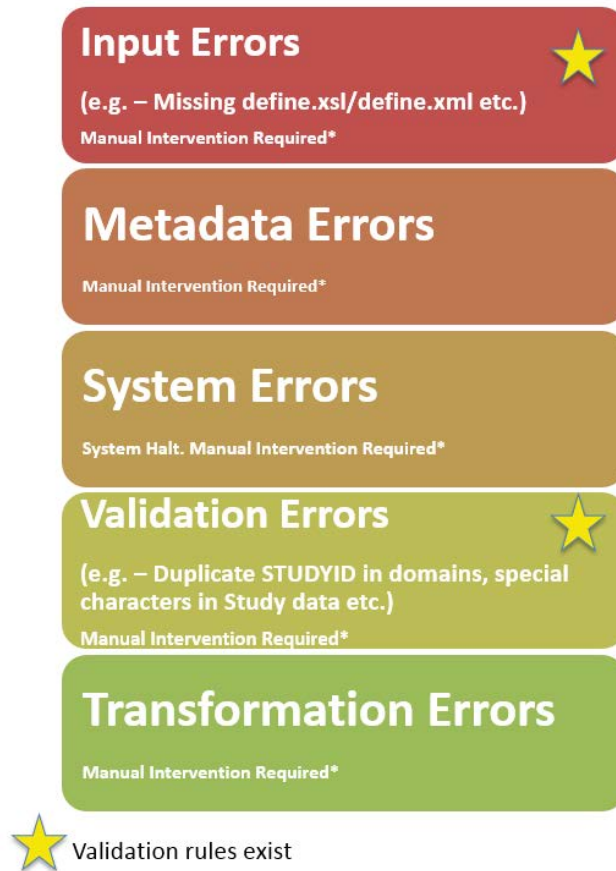
- Validate, transform, and load study data into a database repository
- Extract data from the repository to create a data mart to support nonclinical reports, graphs, charts and other visualizations.
- Enable reviewers and KickStart analysis to access summary information about specific studies (study design, trial summary, explanation of codes used, time points, etc.); create tables, graphs, and visualizations; and filter or subset study data.

Janus Nonclinical Delivers Data to Reviewers



Data loading errors relate to how an applicant implemented the standard or could be technical in nature. Historically, the majority of clinical and nonclinical datasets required at least two attempts, along with curation, to load into Janus. Figure 3 shows the categories of data loading errors encountered since October 2017. Usually, issues can be attributed to missing reference values, the presence of special characters or commas, missing, invalid, or incorrect values in the define.xml or define.xsl, missing references, or missing required data files or domains. Data managers curate minor errors, such as the case with the presence of special characters or commas. However, as a rule, datasets are not altered or changed so as not to affect the integrity of the data. Currently, data loads are in support of Jumpstart and Kickstart services, including Janus development to support future Jumpstart and Kickstart services.

Figure 3: Categorical summary of loading errors



Technical improvements are being developed this year, including:

- Speeding up loading times using a Multi Threaded load approach based on XPT Domains.
- Ability to monitor load progress and get real time ETA via the Admin UI Console.
- Eliminate Manual intervention due to job failures by catching and logging any errors while allowing the good data to proceed.

Along with technical improvements, we will continue to identify errors in implementing CDISC data standards that would preclude loading into Janus. Sharing the types of errors will improve the data conformance to CDISC standards and increase usability and reviewability within the Agency in a more automated, consistent, efficient manner. There is a concurrent effort to enhance validation rules for SDTM and SEND datasets. We will continue to identify a subset of validation rules that supports automated loading efforts. Additionally, we will continue to communicate our findings to policy groups to incorporate into the Technical Conformance Guide (TCG) if clarification is needed.

Conclusion

As part of a larger data management improvement effort, OCS will continue to automate portions of the data loading process to support Janus. In the future, loading trends will be reporting to highlight

differences in number of loading errors, loading times, and types of errors encountered prior to the loading enhancements and afterward.