Improving the Reporting of Data Conformance Issues in the Nonclinical Study Data Reviewers' Guide (nsdrg)

The nSDRG team has an objective to improve Section 5 (the "validation section"), for the next update of the PhUSE nsdrg Package (target CSS 2019). We feel it needs to be redesigned to:

- be validation-tool-vendor neutral (i.e. remove specificity to Pinnacle21 as seen in current approach)
- more clearly support the needs and current rules approach of FDA

Goals:

- Increase confidence that reported validation issues are useful and well understood.
- Remove the concept of "Errors" vs "Warnings" as FDA has done.
- Clarify the inclusion, or not, of define file validation results.
- Reduce the necessity to run multiple validators on the same dataset, then parse the results between Section 5 and other sections. (note that some companies feel this is necessary.)
- Understand criteria for what needs to be reported in nSDRG and what, if anything, is unnecessary, not useful or even counter productive.

Plan:

The team agreed we don't have a full picture of what are "the needs of FDA" regarding this section. The team would like to develop a best practice approach for recording/communicating data conformance issues in the next iteration of the nsdrg. To help this development, the team wants to engage with FDA Reviewers to find out what is valuable to FDA consumers of the nsdrg, regarding Section 5. A list of questions has been developed to seek advice.

nsdrg Questions for FDA

Re: Section 5 - Data Standards Validation Rules, Versions, and Conformance Issues

1. *What kinds of information do Reviewers find useful in this section?*
   The information in this section has the most value to the Office of Computational Science's (OCS) Data Loading and KickStart teams. Also, Business Rule errors are very important to the statistical reviewers in the Office of Biostatistics (OB) for the submission of carcinogenicity studies. Nonclinical reviewers will also find value in Business Rule errors related to carcinogenicity studies (*e.g.,* BR FDAB072).

2. *Do you expect that explanations of issues in section 5, when impacting data content are referenced or further explained in the dataset explanations (section 4.2)?*
   Yes, any issues referenced in Section 4.2 should be referenced in greater detail in Section 5. Please also state that there is additional information in Section 5 when listing the issue in Section 4.2.

3. *Are there some examples of what has been seen in nsdrg and considered "too technical"? (in messages or explanations of validation issues.)*
   There is no information that is considered too technical for use by OCS. OCS staff will use the information provided to help with some aspect of the review. It is best to provide as much specific information as possible rather than boilerplate language.

4. *Should all validation issues be described in this section, including those determined to be "false"? (false = errors or warnings that result from bugs in Pinnacle 21 Community or other validators that may be used. Some examples of common rule breaks that are false in nature: Errors: DD0059, DD0028, DD0024, DD0064. Warnings: FDAN037, FDAN169, FDAN212, FDAN218, FDAN341, DD0029, PCO497.).*
   All issues identified by the tool used, regardless of whether it is "false", should be identified by the sponsor in Section 5. We suggest placing any known "false" errors identified by the validator tool in a separate table in the same section.

5. *Are there particular domains which need special consideration regarding the handling of conformance issues (such as MI and Tumor data – where there are a lot (14) of business rules)?*
   Yes, there are times when cross domain checks cannot be automated, or specific Business Rules do not have automated checks. The Business Rules should not be ignored for these non-automated checks. For carcinogenicity studies, these non-automated checks against 14 Business Rules should be performed manually with any errors noted in Section 5. For other types of studies, if a sponsor is confident that the process used to create the SEND datasets aligns with these non-automated rules by establishing that multiple datasets have no errors, manual checking of each dataset is not required.

6. *Do people who do "data loading" use the nsdrg?*
   Yes, if a Data Management and/or Data Loading teams have trouble loading the data, the nSDRG is one of the first places the teams will look for additional information on conformance issues.

7. *Is the FDA tracking validation rule violations by quantity and/or rule ID?*
   OCS staff is tracking validator rule violations by quantity but are not tracking the specific rule ID associated with them.
   Are trends emerging to suggest certain rules could be either A) impractical for companies to adhere to, or B) too vague in meaning. For example, high variety of explanations could indicate wide misunderstanding of the rule or its application.
   The FDA noted several common error trends in datasets submitted by sponsors. The FDA clarifies how to correct the errors in industry meetings and other public forums as well as updates the Technical Conformance Guide (TCG) to ensure the FDA's needs are met.