

CDISC Advisory Board Validation Project

Best Practices for Validation
Rules

Lauren Shinaberry
Senior Project Manager
CRO Services
Business & Decision Life Sciences

Tel +32 2 774 11 00
Fax +32 2 774 11 99
Mobile +32 491 86 89 04
Lauren.shinaberry@businessdecision.com

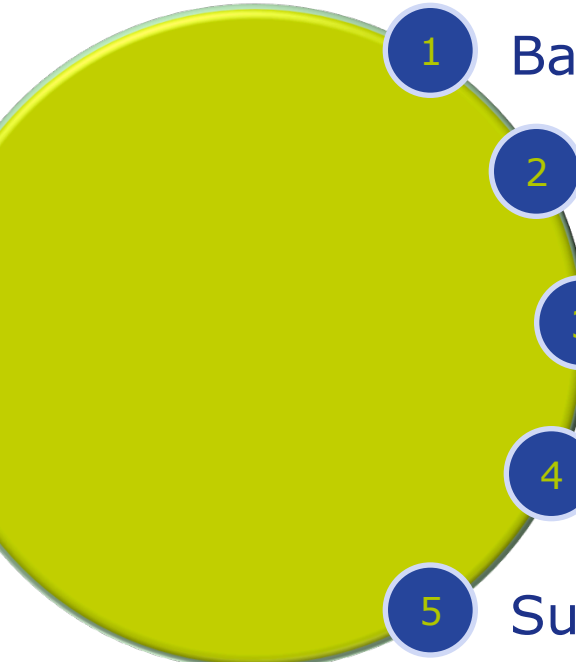
Sint-Lambertusstraat 141 Rue Saint-
Lambert
1200 Brussels

www.businessdecision-lifesciences.com

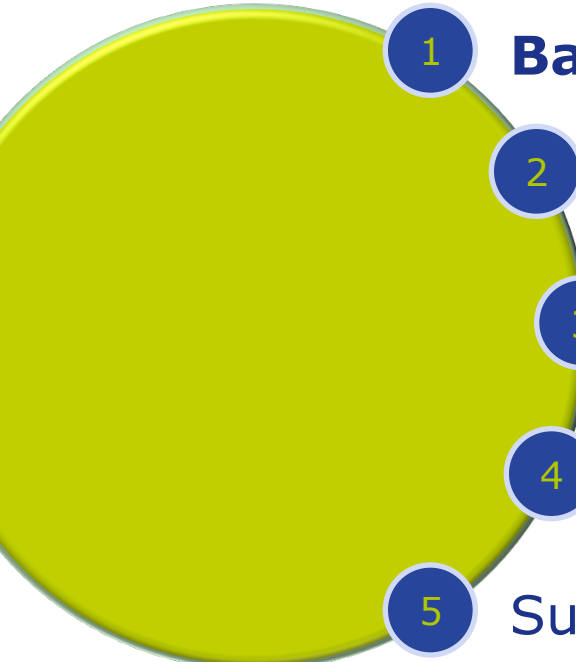


DRIVE YOUR PERFORMANCE

Agenda

- 
- 1 Background
 - 2 Methodology
 - 3 Findings
 - 4 Best Practices
 - 5 Summary & Conclusion

Agenda

- 
- 1 **Background**
 - 2 Methodology
 - 3 Findings
 - 4 Best Practices
 - 5 Summary & Conclusion

Background

- CDISC Advisory Board
 - Representatives from each paying member organization
 - Provide advice to the BoD on the direction CDISC members think should be taken
 - Volunteers from across the industry
 - Vendors, CROs, pharma, biotech, academy, etc.

Goal of project

- Contribute to a better assessment of the quality of clinical data
- In particular:
 - “Provide an agreeable mechanism to assess data quality based on how closely a clinical database follows the CDISC SDTM standard and corresponding Implementation Guide”*

* Courtesy of Trisha Simpson, UCB

How does this improve data quality?

- Conformance to standards allows faster review of data content
 - Reuse of checking programs
 - Increased familiarity with structure
 - Focus on the unique aspects of a study
 - Enables efficient integration of data
- Non-conformance slows review and inhibits ability to integrate data

Problem

- SDTM Model Document and Implementation Guide have been primary source for assumptions and rules for conformance
- Requires manual review of documents and is subject to interpretation
- There are no concise rule statements available from CDISC for SDTM compliance (although there are for ADaM)

Question

- How can implementers be assured that the tools they are using are reliable and consistent with the CDISC SDTM team's definition of the model?
 - Project started in 2010
 - Assessed the current state of the rules being used to validate SDTM
 - Reviewed the top 5 vendor tools being used
 - Included FDA representation on project

Tools that were included in project

- CheckPoint from Octagon
- Data Model Compliance Checker from Business & Decision Life Sciences
- Open CDISC Validator
- SAS Clinical Standards Toolkit
- WebSDM from PhaseForward

Agenda



Top 5 tools reviewed

- Vendors of top 5 SDTM checking tools provided a list of the checks performed by their tool
- Checks were assessed to identify
 - Differences in meaning/interpretation
 - Categorization
 - Severity
 - Descriptions of rules

Compared to SDTM IG

- Rules were consolidated
 - Remove duplicates
 - Match up rules that were described differently across tools but were essentially checking the same thing
- Consolidated list was compared to the SDTM and IG to identify any gaps
 - This was not an exhaustive comparison
 - Looked for key rules and assumptions

Results of assessment

What is the input for the checks?

Regardless of which is considered the primary input to the checks, the metadata between all these sources must be consistent

Agenda



Overview of findings

- Majority of checks across all 5 tools were identifying the same types of issues
- Each tool included unique checks not found in the others
- Some rules found in the IG were not present in any of the tools

Check Types

- Four types were identified
 - SDTM conformance
 - Data quality
 - System-specific
 - FDA Implementation
- Important to know the type in order to know what action to take if an issue is identified

SDTM Conformance Checks

- Can be traced back to a specific reference in a CDISC document
 - SDTM Model
 - SDTM Implementation Guide
 - Any other CDISC documentation

Data quality checks

- Checks that are Good Clinical Data Management practices
 - Regardless of the structure or standards used
 - E.g. a start date should not be after an end date

System-specific Checks

- Seen primarily in WebSDM
 - Identify requirements necessary for specific features of the system to work properly
 - Related to the fact that this tool is also used to view the data
 - Outside the scope of this project

FDA Implementation Checks

- Can be traced back to a specific reference in the CDER Common Data Issues document AND
- Are not already referenced in a CDISC document
- May be some discrepancies between “pure” SDTM compliance and FDA needs
 - EPOCH in SDTM v1.2 was Permissible but FDA requests it is included

Why are the types important?

- Understanding which type of check is failing is critical to understanding what action to take
- Currently, none of the tools identify checks this way
- Confuses implementers who may think all of the issues are SDTM conformance problems

Severity

- Low, Medium or High vs. Error/Warning
- No consistent assignment of severities was found
- No clear definition on what the severity means
- Project team made recommendations for assigning severity based on Check Type

Severity for SDTM Conformance Types

- **Error:**
 - Must be addressed to conform to SDTM
 - Clearly violates a rule stated in the SDTM or SDTM IG
- **Warning:**
 - May need to be investigated to determine if conformance criteria have not been met

Severity for FDA Conformance Types

- **Error:**
 - Must be addressed or the data cannot be reviewed
- **Warning:**
 - May cause questions from the reviewer if not properly explained in the Reviewer's Guide or other documentation provided with the submission

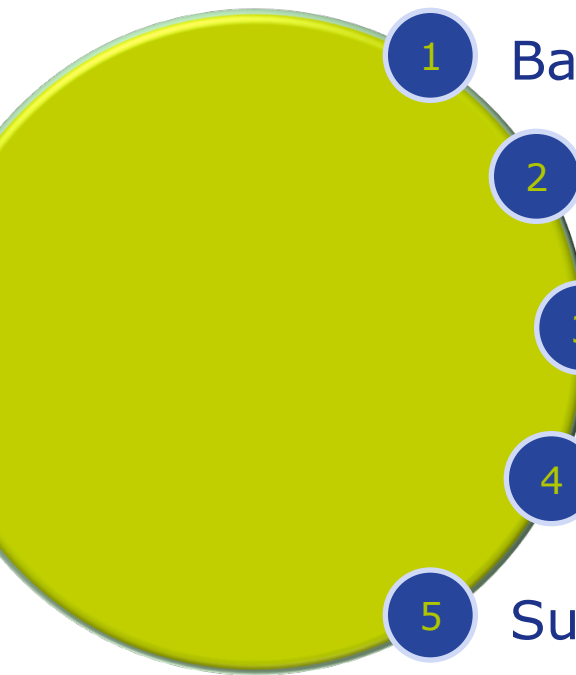
Severity for Data Quality Types

- Not applicable
 - Data quality issues should be addressed through GCDMP and are not specifically SDTM conformance
 - If you wait until you are submitting your data to look for data quality issues it is too late to address them appropriately
 - These should be part of your DM process

Categories for SDTM Conformance

- Structural
 - Compares domain- and variable-level metadata
 - Names, labels, types
- Controlled Terminology
 - Are the correct terminology being used
 - Have non-extensible codelists been adhered to
- Content
 - Usually involved comparison of two or more variables
 - E.g.; check for duplicate –SEQ within a USUBJID, if RELTYPE is present, IDVARVAL should be null

Agenda

- 
- 1 Background
 - 2 Methodology
 - 3 Findings
 - 4 **Best Practices**
 - 5 Summary & Conclusion

References

- Provide a reference for each check
 - Makes it clear why the check is there
 - Directs users to a detailed explanation of the issue in a non-programmatic way
 - Eliminates confusion on exactly what is required to conform to SDTM vs. to FDA needs
 - Can use the text from the reference as a “plain-language” description

Simplify Severity

- A check should either be an Error or a Warning
 - Use the definitions provided in previous slides for different check types
 - Focuses on the action that must be taken
 - Less ambiguous than High, Medium, Low, etc.

Consistent Message Text

- Clearly describe the issue in plain language, not pseudo-code
 - “Variable values should be populated with terms found in 'Race' (C74457) CDISC controlled terminology codelist”
 - “Identifies records where value for [Race] is not found in Codelist [Race], limited to records where [Race is not null]”

Do not duplicate issues

- One issue should not trigger multiple error messages
 - “Race must be populated using the Race codelist”
 - “All variables must use the assigned controlled terminology”
 - Only one of these should be triggered if a non-conformant value is found

Maintenance of Checks

- Have a well-defined process for maintaining checks
- References should include the version of the document used
- Coordinate updating checks with new releases of SDTM or SDTM IG

Agenda



Conclusion & Summary

- CAB VP officially concluded in April 2012
 - SDTM rules that were consolidated during the project provided to the SHARE team for future inclusion with the standards
 - FDA conformance and other checks being discussed in working group

Summary

- There are many options available for checking conformance to SDTM
- Also many organizations are programming their own checks
- Best Practices are intended to make it easier for users to
 - understand the rationale behind each check
 - Understand the implications of not addressing a finding

Acknowledgements

Name	Company	Name	Company
Borowiak , Ken	PPD	Kanevsky , Max	OpenCDISC
Cassells , Sally	Independent Consultant	Koranda , Adrienne	Premier Research
Casteris , Christina	Biogenidec	Labout , Shannon	CDISC
Chau , Albert	Datacision	Lightfoot , Gene	SAS
Chen , Hong	McDougall Scientific	Malla , Amy	FDA, CBER
Chow , Anthony	Allergan	Qian , Wie	Takeda Global
Freimark , Nate	Omnicare	Shinaberry , Lauren	Business & Decision
Fulda , John	Premier Research	Simpson , Trisha	UCB
Gibson , Bill	SAS	Smith , Steve	SAS
Hernandez , Joyce	Merck	Van Manen , Robbert	Phase Forward
Hungria , Marcelina	Independent Consultant	Van Reusel , Peter	Business & Decision
Hurst , Don	Premier Research	Witczak , Bobbie	FDA, CDER
Kalfas , Thomas	Premier Research		

