# Automating CDISC Conformance Issue Management

Eric Crockett and Mario Widel, Covance Inc., King of Prussia, PA

## Introduction

Evaluating whether mapped data conforms to the applicable CDISC (SEND, SDTM or ADaM) standard is best done multiple times over the course of a study. This poster describes a VBA macro that facilitates the management of CDISC issues over the life-cycle of a study with support from a curated list of standardized explanations for the issue report.

## Pictorial Process Overview

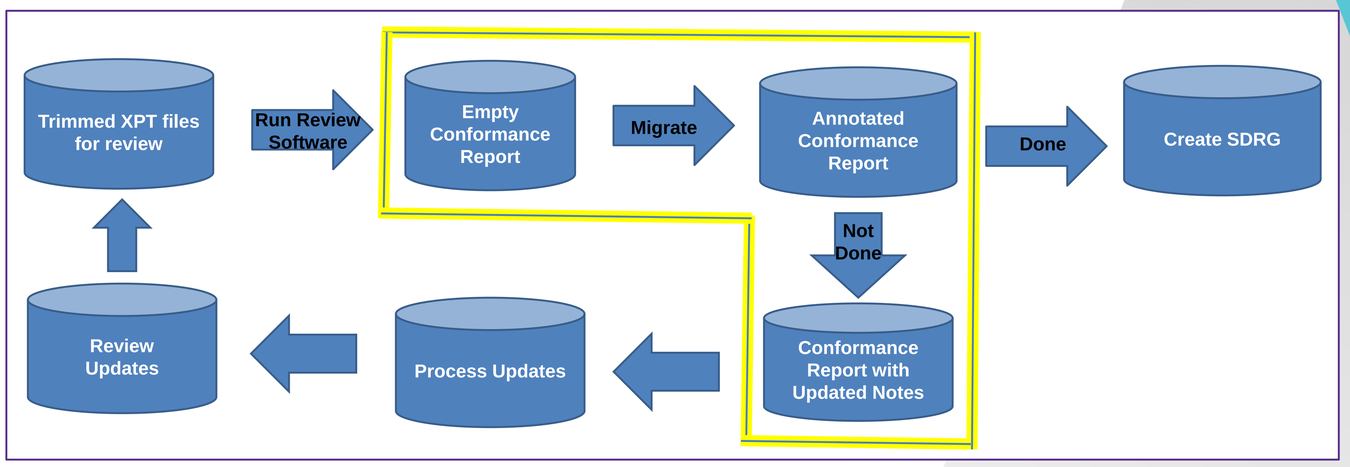


**Figure 1. Diagram of CDISC Conformance Review Lifecycle**

This process starts with the trimmed xpt files and running the CDISC review software. Once the report is generated the VBA macro can be used. The role of the VBA macro is contained within the yellow boxed portion in Figure 1. The macro facilitates migration of comments from an existing report to the newly generated report. Once the comments are migrated, the remaining undocumented issues are documented with a curated list of standard explanations. The tool also flags changes in the number of issues found. At this point, the report can be used to facilitate resolution of resolvable issues by the study team. After updates are completed and reviewed, the xpt files are recreated and the process starts again.

| Common Finding | Domain(s) | Standardized SDRG Comment | Notes |
|---|---|---|---|
| Duplicate records | DM, SUPPDM, TA, TE, TI, TS, TV | N/A | Duplication should not exist in DM, SUPPDM, or the trial design domains |
| Duplicate records | ALL - DM, SUPPDM, TA, TE, TI, TS, TV | These are not duplicates. <variable> varies. | Confirm not a spec or programming error. Verify another variable exists to distinguish records P21 thinks are duplicates. You will have to determine what that variable is. |
| Variable length is too long for actual data | TA, TE, DM, SE | ARMCD, ARM, ETCD, and ELEMENT are set to standard length across domains. | Verify data are trimmed. For example, if STUDYID has length 200, there is likely an issue. Confirm indicated variables are actually the variables set to consistent length across domains. |
| Invalid ISO 8601 value for variable | ALL | N/A | All dates must be in ISO 8601 format. This must be updated. |

**Figure 2. Subset of standardized issues. Yellow – Avoidable. Blue – Sometimes Avoidable. Green – Unavoidable.**

## Curated List of Standard Explanations

A subset of the standardized explanations is on the left. The "Standardized SDRG Comment" column contains default explanatory text and accompanying notes for issues that must be resolved or cannot be resolved. The notes column provides context and helps guide review and updates. This content is used to populate the initial report; and then used on subsequent runs to populate default explanations for new issues. The "Domain(s)" column and the "Common Finding" column are used to associate the appropriate SDRG comment and notes onto a newly generated report.

## VBA Macro Inputs and Outputs

| Issue Summary | | | | | | | |
|---|---|---|---|---|---|---|---|
| Source | Pinnacle 21 ID | Publisher ID | Message | Severity | Found | Standardized SDRG Comment | Notes |
| AE | | | | | | | |
| | CT2001 | FDAC340 | AE start date is after the latest Disposition date | Error | 10 | Consistent with source data. | Confirm not a programming or spec error. Specific details should be added to the SDRG comment if informative. (e.g. Data collection is still ongoing). The issue count is expected to decrease over time. |
| DM | | | | | | | |
| | SD0035 | FDAC183 | Invalid ISO 8601 value for variable | Error | 12 | N/A | All dates must be in ISO 8601 format. This must be updated. |
| | SD0070 | FDAC053 | No Exposure record found for subject | Warning | 2 | Consistent with raw data. Data collection is still ongoing. | |

**Figure 3. Example of a report that has already been documented.**

| Issue Summary | | | | | |
|---|---|---|---|---|---|
| Source | Pinnacle 21 ID | Publisher ID | Message | Severity | Found |
| AE | | | | | |
| | CT2001 | FDAC340 | AE start date is after the latest Disposition date | Error | 3 |
| | SD0002 | FDAC018 | Duplicate records | Warning | 2 |
| DM | | | | | |
| | SD0002 | FDAC018 | Duplicate records | Warning | 4 |
| | SD0035 | FDAC183 | Invalid ISO 8601 value for variable | Error | 16 |
| | SD0070 | FDAC053 | No Exposure record found for subject | Warning | 2 |

**Figure 4. Example of a newly generated report.**

Figure 3 represents a report that has already been documented. Figure 4 is a newly generated empty report. The VBA macro pulls the comments from the documented report into the newly generated report. The curated list of standard explanations in Figure 2 is used to support any missing documentation from the prior report. If there is no prior report, only the curated list is used by the macro to document the empty report.

## The Final VBA Output

| Source | Pinnacle 21 ID | Publisher ID | Message | Severity | Found | Standardized SDRG Comment | Notes |
|---|---|---|---|---|---|---|---|
| AE | | | | | | | |
| AE | CT2001 | FDAC340 | AE start date is after the latest Disposition date | Error | 3 | Consistent with source data. | Confirm not a programming or spec error. Specific details should be added to the SDRG comment if informative. (e.g. Data collection is still ongoing). The issue count is expected to decrease over time. |
| AE | SD0002 | FDAC018 | Duplicate records | Warning | 2 | These are not duplicates. <variable> varies. | Confirm not a spec or programming error. Verify another variable exists to distinguish records P21 thinks are duplicates. You will have to determine what that variable is. |
| DM | | | | | | | |
| DM | SD0002 | FDAC018 | Duplicate records | Warning | 4 | N/A | Duplication should not exist in DM, SUPPDM, or the trial design domains |
| DM | SD0035 | FDAC183 | Invalid ISO 8601 value for variable | Error | 16 | Previous Report Found: 12 N/A | All dates must be in ISO 8601 format. This must be updated. |
| DM | SD0070 | FDAC053 | No Exposure record found for subject | Warning | 2 | Consistent with raw data. Data collection is still ongoing. | |

**Figure 5. Example of final output.**

Here is an example of a report after the comments have been migrated from the previous report with the new issues documented with the curated list of standard explanations. Notice the comment and red cell that are circled. This red cell with a comment and the green cell with a comment are examples of how the macro provides a way to quickly compare the changes in the number of issues without manual inspection. Documentation from the previous report migrates without an interior color, otherwise the interior color is inherited from the curated list of standard explanations.

## Conclusion

Automating CDISC conformance issue management has dramatically reduced time spent reviewing CDISC data and documenting unavoidable issues in the reviewer's guides.

*Presented at PhUSE/CSS 2018*

## Acknowledgment & Contact Information

Special thanks to Steve Kirby, Terek Peterson and Covance for supporting this poster. The authors can be reached at:

Eric Crockett – Eric.Crockett@Chiltern.com
Mario Widel – Mario.Widel@Chiltern.com

COVANCE®
SOLUTIONS MADE REAL®