CDISCPILOT02 – Updated for SDTMIG v3.2

Test Data Factory Project Team

July 2017

[1 Introduction 2](#_Toc486605406)

[2 Specific Comments 3](#_Toc486605407)

[2.1 Split datasets 3](#_Toc486605408)

[3 Data Conformance Summary – Explanation on Remaining Pinnacle 21 Findings 4](#_Toc486605409)

# Introduction

The PhUSE organization initiated the Test Data Factory (TDF) project to make test data for CDISC-based processes and programs publicly available. The CDISCPILOT02 datasets from 2013 were updated to conform to newer standards. The TDF project team used the following approach:

* The team ran the datasets through Pinnacle 21 Community v2.2 using SDTM 3.2 as the Configuration and 2016-06-24 as the CDISC CT.
* Based on the reported findings each dataset was updated so that no more unexplained findings (see section 3) are reported.
* Pinnacle 21 Community v2.2 was again used on all updated datasets to find additional issues resulting from inconsistencies between the datasets. The datasets were updated again.
* For updating the datasets, the team used SAS and R programs.

The team used Pinnacle 21 as a conformance checking tool because it is a commonly used tool to evaluate conformance and is openly available to everyone. The team is very well aware of shortcomings and knows that is it not perfect and comprehensive but decided that for lack of better alternatives, this would be the right choice. In order to resolve some of the findings, the team needed to make some decisions on how the datasets should be updated. These changes should not be construed as definitive approaches for resolving conformance issues in general. Users of the updated CDISCPILOT datasets should be aware that depending on the situation, several options can and need to be considered.

Note that the rather large LB and QS domain datasets were split (see section 2.1) to accommodate limitations of the repository that was used during the work. This also demonstrates splitting of domains. This was not done because of any guidance or other requirement to split these domains.

Section 3 of this document provides explanations for the Pinnacle 21 findings that are still seen in the report. Some of these findings cannot be fully addressed because information on how the data was originally collected was not included in the pilot materials.

Note that this document is not intended to represent a full-fledged Study Data Reviewer’s Guide (SDRG). Instead, its purpose is to serve as documentation for the updated test data, and to explain some of the decisions that were made during the update process. In turn, the test data is intended to be used for the development and testing of standard reporting and analysis scripts, and is not meant to represent a complete regulatory submission package. The original CDISCPILOT02 data may be downloaded from <https://www.cdisc.org/sdtmadam-pilot-project>, if desired.

# Specific Comments

## Split datasets

The LB, SUPPLB, and QS domain datasets were split based on the values in the –CAT variable as recommended in the SDTMIG. Note that this was splitting not done because of any guidance or other requirement to split these domains.

|  |  |  |  |
| --- | --- | --- | --- |
| Domain | Split Domain | Value of - - CAT Variable | Meaning of - - CAT Variable |
| QS | qsco.xpt | ADAS-COG | Alzheimer's Disease Assessment Scale-Cognitive CDISC Version Questionnaire |
| QS | qsda.xpt | DAD | Disability Assessment for Dementia Questionnaire |
| QS | qsgi.xpt | ADCS-CGIC | Alzheimer's Disease Cooperative Study-Clinical Global Impression of Change Questionnaire |
| QS | qshi.xpt | MHIS-NACC | Modified Hachinski Ischemic Scale-NACC Version Questionnaire |
| QS | qsmm.xpt | MMSE | Mini-Mental Status Exam |
| QS | qsni.xpt | NPI | Neuropsychiatric Inventory Questionnaire |
| LB | lbch.xpt | CHEMISTRY | Chemistry |
| LB | lbhe.xpt | HEMATOLOGY | Hematology |
| LB | lbur.xpt | URINALYSIS | Urinalysis |
| SUPPLB | supplbch.xpt | Split based on IDVARVAL matching LBSEQ in lbch.xpt | |
| SUPPLB | supplbhe.xpt | Split based on IDVARVAL matching LBSEQ in lbhe.xpt | |
| SUPPLB | supplbur.xpt | Split based on IDVARVAL matching LBSEQ in lbur.xpt | |

# Data Conformance Summary – Explanation on Remaining Pinnacle 21 Findings

| Domain | Variables | Values | Message | Severity | Explanation |
| --- | --- | --- | --- | --- | --- |
| AE | VARIABLE | AEDY | Model permissible variable added into standard domain | Warning | AEDY was included because AEDTC is included. |
| AE | VARIABLE | EPOCH | Model permissible variable added into standard domain | Warning | Per FDA Study Data Technical Conformance Guide, the variable EPOCH should be included for clinical subject-level observations |
| AE | VARIABLE | AEDTC | Model permissible variable added into standard domain | Warning | AEDTC is included because AE information was collected by visit. |
| AE | AEENDTC | null | Missing End Time-Point value | Warning | AEENDTC is missing for some AEs and no value was collected to state that AEs were ongoing or not |
| AE | AEDECOD, AETERM, AESEV, AESTDTC, USUBJID |  | Duplicate records in AE domain | Warning | AE information was collected by visit, with a new record entered each time. Where possible, the records have been collapsed, but there are still cases where the actual event could not be identified from the information provided with the pilot data. |
| CM | VARIABLE | EPOCH | Model permissible variable added into standard domain | Warning | Per FDA Study Data Technical Conformance Guide, the variable EPOCH should be included for clinical subject-level observations |
| CM | VARIABLE | CMDTC | Model permissible variable added into standard domain | Warning | ??? |
| CM | VARIABLE | VISITDY | Model permissible variable added into standard domain | Warning | ??? |
| CM | VARIABLE | VISIT | Model permissible variable added into standard domain | Warning | ??? |
| CM | VARIABLE | CMDY | Model permissible variable added into standard domain | Warning | ??? |
| CM | VARIABLE | VISITNUM | Model permissible variable added into standard domain | Warning | ??? |
| DM | ARMCD, ACTARMCD, USUBJID | Xan\_Lo, Xan\_Lo, 01-703-1119 | No baseline result in LB for subject | Warning | ??? |
| DM | ARMCD, ACTARMCD | Xan\_Hi, Xan\_Lo | ACTARMCD does not equal ARMCD | Warning | ??? |
| DS | VARIABLE | VISIT | Model permissible variable added into standard domain | Warning | VISIT was added to DS because protocol specified events were collected at particular visits. |
| DS | VARIABLE | VISITNUM | Model permissible variable added into standard domain | Warning | VISITNUM was added to DS because protocol specified events were collected at particular visits. |
| DS | VARIABLE | DSDY | Model permissible variable added into standard domain | Warning | DSDY was added because DSDTC is captured. |
| EX | VARIABLE | VISIT | Model permissible variable added into standard domain | Warning | VISIT was added to EX was collected by visits. |
| EX | VARIABLE | VISITNUM | Model permissible variable added into standard domain | Warning | VISITNUM was added to EX was collected by visits. |
| EX | VARIABLE | VISITDY | Model permissible variable added into standard domain | Warning | VISITDY was added to EX was collected by visits. |
| LB | VARIABLE | EPOCH | Model permissible variable added into standard domain | Warning | Per FDA Study Data Technical Conformance Guide, the variable EPOCH should be included for clinical subject-level observations |
| LB (LBHE) | LBORRES, LBTEST, LBORRESU, LBTESTCD |  | Missing value for LBORRESU, when LBORRES is provided | Warning | There are lab tests that do not have units. This is acceptable. |
| LB (LBHE) | LBTEST, LBSTRESN, LBTESTCD, LBSTRESC, LBSTRESU |  | Missing value for LBSTRESU, when LBSTRESC is provided | Warning | There are lab tests that do not have units. This is acceptable. |
| LB (LBUR) | LBORRES, LBTEST, LBORRESU, LBTESTCD |  | Missing value for LBORRESU, when LBORRES is provided | Warning | There are lab tests that do not have units. This is acceptable. |
| MH | VARIABLE | VISITDY | Model permissible variable added into standard domain | Warning | ??? |
| MH | VARIABLE | VISITNUM | Model permissible variable added into standard domain | Warning | ??? |
| MH | VARIABLE | VISIT | Model permissible variable added into standard domain | Warning | ??? |
| MH | VARIABLE | MHSEV | Model permissible variable added into standard domain | Warning | ??? |
| MH | VARIABLE | MHHLGT | Model permissible variable added into standard domain | Warning | ??? |
| MH | VARIABLE | MHLLT | Model permissible variable added into standard domain | Warning | ??? |
| MH | VARIABLE | MHHLT | Model permissible variable added into standard domain | Warning | ??? |
| QS | VARIABLE | EPOCH | Model permissible variable added into standard domain | Warning | Per FDA Study Data Technical Conformance Guide, the variable EPOCH should be included for clinical subject-level observations |
| QS (QSCO) | QSSTRESC, QSBLFL | null, Y | Missing QSSTRESC value for Baseline record | Warning | ??? |
| SE | VARIABLE | SESTDY | Model permissible variable added into standard domain | Warning | ??? |
| SE | VARIABLE | SEENDY | Model permissible variable added into standard domain | Warning | ??? |
| SV |  | EPOCH | Model permissible variable added into standard domain | Warning | Per FDA Study Data Technical Conformance Guide, the variable EPOCH should be included for clinical subject-level observations |
| TS | TSPARM | Age Group | TSPARM value not found in 'Trial Summary Parameter Test Name' extensible codelist | Warning | ??? |
| TS | TSPARMCD | AGESPAN | TSPARMCD value not found in 'Trial Summary Parameter Test Code' extensible codelist | Warning | ??? |
| VS |  | EPOCH | Model permissible variable added into standard domain | Warning | Per FDA Study Data Technical Conformance Guide, the variable EPOCH should be included for clinical subject-level observations |

1. Text for AE issues