Analysis Data Reviewer's Guide

ADaM Metadata Submissions Guidelines Team

Study CDISCPILOT01

Analysis Data Reviewer's Guide

Contents

1. Introduction

1.1 Purpose

This document provides context for the analysis datasets and terminology that benefit from additional explanation beyond the Data Definition document (define.xml) for an individual study. In addition, this document provides a summary of ADaM conformance findings.

1.2 Acronyms

Acronym	Translation
aCRF	Annotated Case Report Form
ADaM	Analysis Dataset Model
ADRG	Analysis Data Reviewer's Guide
eCRF	Electronic Case Report Form
eDT	Electronic Data Transfer (e.g. central lab data, ECG vendor data, PK data, etc.)
IG	Implementation Guide
NA	Not Applicable
SDTM	Study Data Tabulation Model
TAUG	Therapeutic Area User Guide

1.3 Study Data Standards and Dictionary Inventory

Standard or Dictionary	Versions Used
SDTM	v1.7 / IG 3.3
SDTM Controlled Terminology	2020-12-18
ADaM	v2.1 / IG 1.1
ADaM Controlled Terminology	2020-12-18
Data Definitions	Define.xml v2.0
TAUG (if applicable)	NA
Medications Dictionary	NA
Medical Events Dictionary	MedDRA 8.0

Standard or Dictionary	Versions Used
Other standards (optional)	ADaM Structure for Occurrence Data (OCCDS) v1.0

1.4 Source Data Used for Analysis Dataset Creation

The analysis files for this study were derived from the submitted SDTM files. SDTM files were prepared from CRF data according to version 3.3 of the SDTM IG. No non-CRF or non-SDTM data were used to create the ADaM data.

The datasets include only subjects who were enrolled in the study. Subjects who failed screening criteria were not included in analysis datasets.

2. Protocol Description

2.1 Protocol Number and Title

Protocol Number: CDISCPILOT01

Protocol Title: Safety and Efficacy of Zanomaline in Patients with Mild to Moderate Alzheimer's Disease.

Protocol Versions:

- Original, October 4, 2012.
- Amendment 1, December 23, 2012.

  Amendment 1 changed the vital sign collections of heart rate and blood pressure to be collected in triplicate instead of single measurements.

  Exclusion criteria 12 and 31 were also updated in Amendment 1.

2.2 Protocol Design in Relation to ADaM Concepts

This is a Phase II, multicenter, randomized, double-blind, placebo-controlled, parallel group study. After a two-week screening Epoch, subjects were randomized to placebo, low dose or high dose. The treatment Epoch was 26 weeks long. The LOW and PLACEBO treatment Elements were 26 weeks long, while the HIGH treatment Element was 24 weeks long, with week-long titration Elements at

the beginning and the end. Subjects who discontinued early returned for a visit on Week 24. Those subjects remained in the treatment Epoch.

There are two primary endpoints, the Alzheimer's Disease Assessment Scale - Cognitive Subscale, total of 11 items [ADAS-Cog (11)] at Week 24 and the Video-referenced Clinician's Interview-based Impression of Change (CIBIC+) at Week 24.

3. Analysis Considerations Related to Multiple Analysis Datasets

3.1 Core Variables

Core variables are those that are represented across all/most analysis datasets.

Variable Name	Variable Description
STUDYID	Study identifier used for this protocol
USUBJID	Unique subject identifier
SITEID	Site identifier
TRTSDT	Date of first dose of study treatment
TRTEDT	Date of last dose of study treatment
AGE	Age at first screening visit
AGEGR1	Pooled age group: <65, 65-80, >80
AGEGR1N	Numeric code for pooled age group: 1 = <65, 2 = 65-80, 3 = >80
SEX	Sex
RACE	Race
RACEN	Numeric code for race

3.2 Treatment Variables

ARM versus TRTxxP

Are the values of ARM equivalent in meaning to values of TRTxxP?

Yes. Numeric treatment variable values are based on the numeric dose level (i.e., 0, 54, 81) to avoid the need for a separate variable that contains the treatment dose.

ACTARM versus TRTxxA

If TRTxxA is used, then are the values of ACTARM equivalent in meaning to values of TRTxxA?

The sponsor decided not to store ACTARM in ADSL because it was identical to ARM for all patients.

Use of ADaM Treatment Variables in Analysis

Are both planned and actual treatment variables used in analysis?

Yes. Both planned and actual treatment variables are included in the datasets, though there are no cases where their values are different.

Use of ADaM Treatment Grouping Variables in Analysis

Are both planned and actual treatment grouping variables used in analysis?

No, there are no treatment grouping variables used in analysis.

3.3 Subject Issues that Require Special Analysis Rules

There are no known subject issues requiring special analysis rules.

3.4 Use of Visit Windowing, Unscheduled Visits, and Record Selection

Was windowing used in one or more analysis datasets?

Visit windowing is used in ADADAS (ADAS-Cog Analysis), ADCIBC (CIBIC+ Analysis) and ADNPIX (NPI-X Item Analysis Data). Visits are placed into analysis visits (represented by AVISIT and AVISITN) based on the date of the visit and the visit windows. If multiple visits fall into the same visit window, then the one closest to the target date is chosen for analysis. Records where ANL01FL='Y' are the ones that were used for analysis.

Were unscheduled visits used in any analyses?

Unscheduled visits were not used in any analyses.

Additional Content of Interest

Protocol deviations were not reported in the source database and have not been incorporated into the SDTM data. As a result, such data do not exist among the analysis data.

3.5 Imputation/Derivation Methods

If date imputation was performed, were there rules that were used in multiple analysis datasets?

Yes, date imputation is performed in ADAE, and will be described in Section 5.2.x below.

Additional Content of Interest

In both ADADAS and ADCIBC, missing values were imputed using last observation carried forward (LOCF) methodology. The LOCF algorithm only considered records used for analysis as candidates to carry forward. Records where DTYPE='LOCF' signify those where AVAL was imputed using the LOCF algorithm.

4. Analysis Data Creation and Processing Issues

4.1 Split Datasets

There are no split datasets

4.2 Data Dependencies

All datasets get core variable values from ADSL. ADTTE (AE Time To 1st Derm. Event Analysis) is derived directly from ADAE, and ADLBHY (Laboratory Results Hy's Law Analysis Data) is derived from ADLBC (Laboratory Results Chemistry Analysis Data). There are no other processing dependencies.

4.3 Intermediate Datasets

No intermediate analysis datasets were created for this study.

5. Analysis Dataset Descriptions

5.1 Overview

Are data for screen failures, including data for run-in screening (for example, SDTM values of ARMCD='SCRNFAIL', or 'NOTASSGN') included in ADaM datasets?

No. Subjects who failed screening criteria were not included in analysis datasets.

Are data taken from an ongoing study?

No. This is a final locked database.

Do the analysis datasets support all protocol- and statistical analysis plan-specified objectives?

Yes.

Additional Content of Interest

All ADAS-Cog data, including the first primary endpoint, can be found in the dataset ADADAS. All CIBIC+ data, including the second primary endpoint, can be found in the dataset ADCIBC.

Key safety data are found in the datasets ADAE (Adverse Events Analysis Data), ADLBC, ADLBH (Laboratory Results Hematology Analysis Data), ADLBHY, and ADVS (Vital Signs Analysis Dataset). In addition, ADTTE is specifically for safety analyses of the time to the first dermatologic AE. Dermatologic AEs are considered an adverse event of special interest).

5.2 Analysis Datasets

Dataset Dataset Label	Class	Efficacy	Safety	Baseline or other subject characteristics	PK/PD	Primary Objective	Structure
ADSL Subject Level Analysis Dataset	ADSL			X			One observation per subject
ADADAS ADAS-Cog Analysis	BDS	X				X	One observation per subject per parameter per analysis visit per analysis date
ADAE Adverse Events Analysis Dataset	OCCDS		X				One observation per subject per adverse event
ADCIBC CIBIC+ Analysis	BDS	X				X	One observation per subject per parameter per analysis visit per analysis date
ADLBC Analysis Dataset Lab Blood Chemistry	BDS		X				One observation per subject per parameter per analysis visit
ADLBCPV Analysis Dataset Lab Blood Chemistry (Previous Visit)	BDS		X				One observation per subject per parameter per analysis visit

Dataset Dataset Label	Class	Efficacy	Safety	Baseline or other subject characteristics	PK/PD	Primary Objective	Structure
ADLBH Analysis Dataset Lab Hematology	BDS		X				One observation per subject per parameter per analysis visit
ADLBHPV Analysis Dataset Lab Hematology (Previous Visit)	BDS		X				One observation per subject per parameter per analysis visit
ADLBHY Analysis Dataset Lab Hy's Law	BDS		X				One observation per subject per parameter per analysis visit
ADNPIX NPI-X Item Analysis Dataset	BDS	X					One observation per subject per parameter per analysis visit per analysis date
ADTTE AE Time to 1st Derm. Event Analysis	BDS		X				One observation per subject per parameter
ADVS Vital Signs Analysis Dataset	BDS		X				One observation per subject per parameter per analysis visit per analysis timepoint

5.2.1 ADSL – Subject Level Analysis Dataset

The ADSL (Subject Level Analysis Data) dataset contains all subject-level variables for demographics, subject characteristics, and population flags.

5.2.2 ADADAS – ADAS-Cog Analysis

ADADAS contains analysis data from the ADAS-Cog questionnaire, one of the primary efficacy endpoints. It contains one record per subject per parameter (ADAS-Cog questionnaire item) per visit. Visits are placed into analysis visits (represented by AVISIT and AVISITN) based on the date of the visit and the visit windows. If multiple visits fall into the same visit window, then the one closest to the target date is chosen for analysis. Records where ANL01FL='Y' are the ones that were used for analysis. The last observation carried forward (LOCF) algorithm only considered records used for analysis as candidates to carry forward. Records where DTYPE='LOCF' signify those where AVAL was imputed using the LOCF algorithm. Source data can be traced back to the SDTM.QS domain using USUBJID and QSSEQ. Details on how to derive the primary efficacy result based on ADAS-Cog data can be found in the analysis results metadata in the Define.xml.

5.2.3 ADAE – Adverse Events Analysis Dataset

ADAE contains one record per reported event per subject. Subjects who did not report any Adverse Events are not represented in this dataset. The data reference for ADAE is the SDTM AE (Adverse Events) domain and there is a 1-1 correspondence between records in the source and this analysis dataset. These records can be linked uniquely by STUDYID, USUBJID, and AESEQ.

As with the SDTM AE data set, all MedDRA code variables (i.e., those variables that end in CD) have missing values and dummy terms have been applied to the MedDRA High Level Term (HLT) and High Level Group Term (HLGT). This is due to the proprietary nature of the MedDRA dictionary and the fact that the data with this project will be made available to the public. In a standard submission, these codes and terms should be non-missing and properly populated.

Events of particular interest (dermatologic) are captured in the customized query variable (CQ01NAM) in this dataset. Since ADAE is a source for ADTTE, the first chronological occurrence based on the start dates (and sequence numbers) of the treatment emergent dermatological events are flagged (AOCC01FL) to facilitate traceability between these two analysis datasets.

ADAE also contains additional Occurrence Flags to facilitate traceability, reviewability, and ease of reporting between the analysis dataset and the unique counts in the summary tables. For treatment emergent adverse events, refer to Define.xml documentation for the following variables: AOCCFL, AOCCSFL, and AOCCPFL for summarization at the subject, System Organ Class, and Preferred Term levels, respectively. Similarly, refer to Define.xml documentation for AOCC02FL, AOCC03FL, and AOCC04FL for summarization of serious adverse events at the subject, System Organ Class, and Preferred Term levels.

The three deaths reported during the conduct of this study are captured in the Results in Death Flag (AESDTH='Y') and Outcome of Adverse Event (AEOUT='FATAL'). The Start Date of the Adverse Event in ADAE is imputed to the first of the month if the day is missing. The Study Day of Event Start (ASTDY) and the Treatment Emergent Analysis Flag (TRTEMFL) are derived based on this imputation and may differ from their corresponding SDTM AE/SUPPAE variables Study Day of Start of Adverse Event (AESTDY) and Treatment Emergent Flag (AETRTEM).

5.2.4 ADCIBC – CIBIC+ Analysis

ADCIBC contains analysis data from the CIBIC+ questionnaire, one of the primary efficacy endpoints. It contains one record per subject per visit. Note that for all records, PARAM='CIBIC Score'. Visits are placed into analysis visits (represented by AVISIT and AVISITN) based on the date of the visit and the visit windows. If multiple visits fall into the same visit window, then the one closest to the target date is chosen for analysis. Records where ANL01FL='Y' are the ones that were used for analysis. The last observation carried forward (LOCF) algorithm only considered records used for analysis as candidates to carry forward. Records where DTYPE='LOCF' signify those where AVAL was imputed using the LOCF algorithm. Source data can be traced back to the SDTM.QS domain using USUBJID and QSSEQ. Details on how to derive the primary efficacy result based on CIBIC+ data can be found in the analysis results metadata in the Define.xml.

5.2.5 ADLBC – Analysis Dataset Lab Blood Chemistry

ADLBC contains one record per lab analysis parameter, per time point, per subject. ADLBC contains lab chemistry parameters, and these data are derived from the SDTM LB (Laboratory Tests) domain. Two sets of lab parameters exist in ADLBC. One set contains the standardized lab value from the LB domain and the second set contains change from previous visit relative to normal range values. In some of the summaries, the derived end-of-treatment visit (AVISITN=99) is also presented.

5.2.6 ADLBH – Analysis Dataset Lab Hematology

ADLBH contains one record per lab analysis parameter, per time point, per subject. ADLBH contains lab hematology parameters, and these data are derived from the SDTM LB (Laboratory Tests) domain. Two sets of lab parameters exist in ADLBC. One set contains the standardized lab value from the LB domain and the second set contains change from previous visit relative to normal range values. In some of the summaries, the derived end-of-treatment visit (AVISITN=99) is also presented.

5.2.7 ADLBHY – Analysis Dataset Lab Hy's Law

ADLBHY contains one record per lab test code per sample, per subject for the Hy's Law based analysis parameters. ADLBHY is derived from the ADLBC (Laboratory Results Chemistry Analysis Data) analysis dataset. It contains derived parameters based on Hy's law.

5.2.8 ADNPIX – NPI-X Item Analysis Dataset

ADNPIX contains one record per subject per parameter (NPI-X questionnaire item, total score, and mean total score from Week 4 through Week 24) per visit. The analysis visits (represented by AVISIT and AVISITN) are derived from days between assessment date and randomization date and based on the visit windows that were specified in the statistical analysis plan (SAP). If multiple assessments fall into the same visit window, then the one closest to the target day is chosen for analysis. Records where analysis flag (ANL01FL) = 'Y' are the ones that were used for analysis. The last observation carried forward (LOCF) algorithm was not used for these data. Source data can be traced back to the SDTM.QS domain using USUBJID and QSSEQ. All the NPI-X parameters, except for the mean total score from Week 4 through Week 24 (NPTOTMN), are from SDTM.QS domain. The value of parameter, NPTOTMN,

contains the mean total score for each patient who had any assessments from Week 4 through Week 24. The baseline value of the parameter, NPTOTMN, is the same as the baseline value of total score. The baseline value is a covariate in the analysis of covariance (ANCOVA) model.

6. Data Conformance Summary

6.1 Conformance Inputs

Specify the software name and version for the analysis datasets

Pinnacle 21 Community version 3.1.2

Specify the version of the validation rules (i.e. CDISC, FDA) for the analysis datasets

FDA validation engine 2010.1

Specify the software name and version for the define.xml

Pinnacle 21 Community version 3.1.2, Define.xml version 2.0

Specify the version of the validation rules (i.e. CDISC, FDA) for the define.xml

FDA validation engine 2010.1

6.2 Issues Summary

Dataset	Diagnostic Message	Severity	Count	Explanation
ADLBC	BR2A1LO is greater than BR2A1HI		36772	Not an error- BR2A1LO and BR2A1HI are independent variables representing the ratio of the baseline value to the lower limit of the normal range and the ratio of the baseline value to the upper limit of the normal range, respectively.
ADLBH	BR2A1LO is greater than BR2A1HI		19848	Not an error- BR2A1LO and BR2A1HI are independent variables representing the ratio of the baseline value to the lower limit of the normal range and the ratio of the baseline value to the upper limit of the normal range, respectively.

Dataset	Diagnostic Message	Severity	Count	Explanation
ADLBHY	BR2A1LO is greater than BR2A1HI		4938	Not an error- BR2A1LO and BR2A1HI are independent variables representing the ratio of the baseline value to the lower limit of the normal range and the ratio of the baseline value to the upper limit of the normal range, respectively.
ADSL	Secondary custom variable is present but its primary variable is not present		1	VISNUMEN contains the number of the end of treatment visit. There is no corresponding character variable.

7. Submission of Programs

All programs for analysis datasets and primary and secondary efficacy results would normally be submitted. However, for the CDISCPILOT01, on which this submission package is based, only 1 table program and no analysis data or macro programs were available. The ADaM datasets were all created using SAS version 9.4. The internal reference date used to create dates in ADaM datasets is January 1, 1960.

7.1 ADaM Programs

Program Name	Output	Macro Used
adsl.sas	ADSL	n/a
adadas.sas	ADADAS	n/a
adae.sas	ADAE	n/a
adcibc.sas	ADCIBC	n/a
adlbc.sas	ADLBC	n/a
adlbcpv.sas	ADLBCPV	n/a
adlbh.sas	ADLBH	n/a
adlbhpv.sas	ADLBHPV	n/a
adlbyh.sas	ADLBHY	n/a
adnpix.sas	ADNPIX	n/a

Program Name	Output	Macro Used
adtte.sas	ADTTE	n/a
advs.sas	ADVX	n/a

7.2 Analysis Output Programs

Program Name	Output Number	Title	Input
at14-5-02.sas	14.5.02	Serious Adverse Events by SOC and Preferred Term	ADAE

7.3 Macro Programs

No macro programs are included in this package.

8. Appendix

Note that there is incomplete information in the define.xml file. More specifically, Source/Derivation/Comment information is not included for all datasets and variables. This information was not provided for the datasets in the CDISCPILOT01, and so was not available for inclusion in this package.