

# Analysis Data Reviewer's Guide

22nd Century Group, Inc.

Study CEG-P1-078

# Analysis Data Reviewer's Guide

## Contents

1.	Introduction .....	3
1.1	Purpose .....	3
1.2	Study Data Standards and Dictionary Inventory .....	3
1.3	Source Data Used for Analysis Dataset Creation .....	3
2.	Protocol Description .....	3
2.1	Protocol Number and Title .....	3
2.2	Protocol Design in Relation to ADaM Concepts .....	3
3.	Analysis Considerations Related to Multiple Analysis Datasets .....	4
3.1	Comparison of SDTM and ADaM Content .....	4
3.2	Core Variables .....	5
3.3	Treatment Variables .....	6
3.4	Subject Issues that Require Special Analysis Rules .....	6
3.5	Use of Visit Windowing, Unscheduled Visits, and Record Selection .....	6
3.6	Imputation/Derivation Methods .....	6
4.	Analysis Data Creation and Processing Issues .....	7
4.1	Split Datasets .....	7
4.2	Data Dependencies .....	7
4.3	Intermediate Datasets .....	7
4.4	Variable Conventions .....	7
5.	Analysis Dataset Descriptions .....	8
5.1	Overview .....	8
5.2	Analysis Datasets .....	8
6.	Data Conformance Summary .....	10
6.1	Conformance Inputs .....	10
6.2	Issues Summary .....	10

## 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). In addition, this document provides a summary of ADaM conformance findings.

### 1.2 Study Data Standards and Dictionary Inventory

Standard or Dictionary	Versions Used
SDTM	
ADaM	•ADaM v2.1 •ADaM-IG v1.1
Controlled Terminology	
Data Definitions	Define-XML v2.0
Other standards (optional)	CDISC ADaM Controlled Terminology, 2018-12-21

### 1.3 Source Data Used for Analysis Dataset Creation

The ADaM datasets were derived from SDTM version 3.2.

## 2. Protocol Description

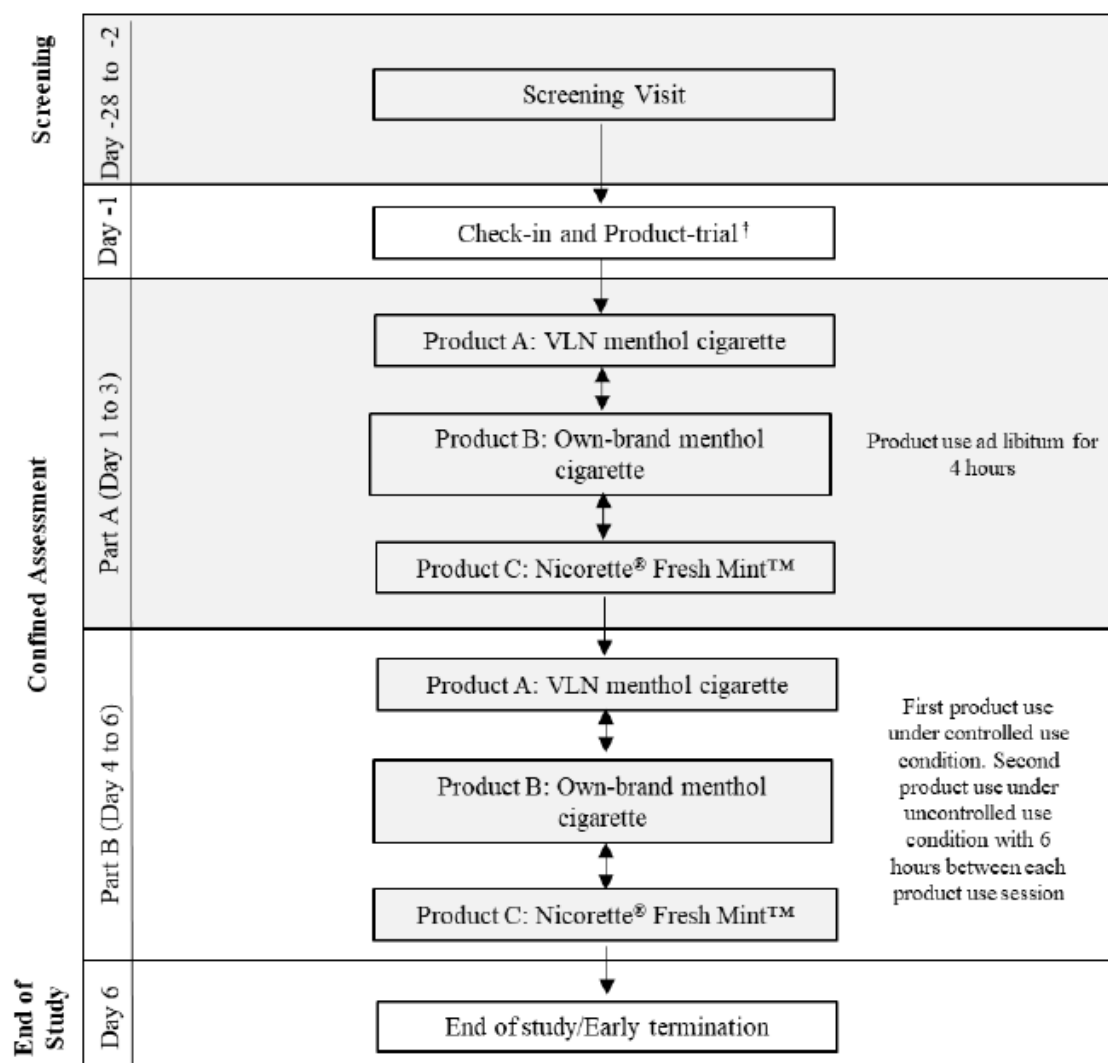
### 2.1 Protocol Number and Title

Protocol Number: CEG-P1-078

Protocol Title: Evaluation of the Abuse Liability of Very Low Nicotine (VLN) Mentholated Cigarettes with Characterization of Nicotine Exposure Profiles in Adult Smokers

Protocol Versions: Version 1.0 / 2018-08-15

### 2.2 Protocol Design in Relation to ADaM Concepts



<sup>†</sup> Ad libitum use of the nicotine gum for 10 minutes. Subjects will be instructed on how to correctly use the nicotine gum using the “chew and park” method.

### 3. Analysis Considerations Related to Multiple Analysis Datasets

#### 3.1 Comparison of SDTM and ADaM Content

- 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**
- Are data taken from an ongoing study? **No**

### 3.2 Core Variables

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

Variable Name	Variable Description
TRTSDTM	Datetime of First Exposure to Treatment
ACTARMCD	Actual Arm Code
RANDBFL	Randomized Population Flag in part B
TRTP	Planned Treatment
TRTA	Actual Treatment
STUDYID	Study Identifier
PDFL	PD Population Flag
ARMCD	Planned Arm Code
SITEID	Study Site Identifier
TRTEDT	Date of Last Exposure to Treatment
TRTEDTM	Datetime of Last Exposure to Treatment
COMPLFL	Completers Population Flag
SAFFL	Safety Population Flag
RANDAFL	Randomized Population Flag in part A
TRTPN	Planned Treatment (N)
PKFL	Pharmacokinetic Population Flag
TRTAN	Actual Treatment (N)
USUBJID	Unique Subject Identifier
TRTSDT	Date of First Exposure to Treatment
ENRLFL	Enrolled Population Flag
SUBJID	Subject Identifier for the Study
ARM	Description of Planned Arm
ACTARM	Description of Actual Arm
ADUR	Substance use Duration
TRT01PN	Planned Treatment for Period 01 (N)
TRT01AN	Actual Treatment for Period 01 (N)
EPOCH	Epoch

Variable Name	Variable Description
TRT01P	Planned Treatment for Period 01
TRT01A	Actual Treatment for Period 01
AESPID	Sponsor-Defined Identifier

### 3.3 Treatment Variables

ARM versus TRTxxP

- *Are the values of ARM equivalent in meaning to values of TRTxxP?* **Yes**

ACTARM versus TRTxxA

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

Use of ADaM Treatment Variables in Analysis

- *Are both planned and actual treatment variables used in analyses?* **Yes**

### 3.4 Subject Issues that Require Special Analysis Rules

No special analysis rules were applied to address subject issues.

### 3.5 Use of Visit Windowing, Unscheduled Visits, and Record Selection

- *Was windowing used in one or more analysis datasets?* **No**
- *Were unscheduled visits used in any analyses?* **Yes**

*Additional Content of Interest*

The baseline value will be defined as the last non-missing evaluation prior to the first dose of study medication and this may include values from unscheduled visits. When assessments are repeated for a given time point, only the result which is the closest to the dosing time will be included in summary tables and this may include values from unscheduled visits.

### 3.6 Imputation/Derivation Methods

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

No date imputation was performed.

## **4. Analysis Data Creation and Processing Issues**

### **4.1 Split Datasets**

No datasets were split.

### **4.2 Data Dependencies**

All ADaM datasets (except ADSL) are dependent on ADSL and their corresponding SDTM domains.  
ADSL is dependent on multiple SDTM domains.

### **4.3 Intermediate Datasets**

No intermediate datasets were generated.

### **4.4 Variable Conventions**

All variables generated in the ADaM datasets follow variable conventions specified in ADaM IG v1.1

## 5. Analysis Dataset Descriptions

### 5.1 Overview

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

Yes

### 5.2 Analysis Datasets

Dataset - Dataset Label	Class	Safety	Baseline or other subject characteristics	PK/PD	Structure
ADSL - subject-level analysis	SUBJECT LEVEL ANALYSIS DATASET		Y		One record per subject
ADAE - Adverse Event Analysis Dataset	OCCURRENCE DATA STRUCTURE	Y			One record per subject per adverse event
ADEG - ECG Analysis Dataset	BASIC DATA STRUCTURE	Y			One or more records per subject per analysis parameter per analysis timepoint
ADEX - Analysis Exposure Dataset	ADAM OTHER	Y			One or more records per subject per treatment per analysis timepoint
ADLB - Laboratory Analysis	BASIC DATA STRUCTURE	Y			One or more records per subject per analysis parameter per analysis timepoint
ADPC - Pharmacokinetic Concentrations	BASIC DATA STRUCTURE			Y	One or more records per subject per analysis parameter per analysis timepoint



<b>Dataset - Dataset Label</b>	<b>Class</b>	<b>Safety</b>	<b>Baseline or other subject characteristics</b>	<b>PK/PD</b>	<b>Structure</b>
ADPE - Physical Examinations	BASIC DATA STRUCTURE	Y			One or more records per subject per analysis parameter per analysis timepoint
ADPP - Pharmacokinetic Parameters	BASIC DATA STRUCTURE			Y	One or more records per subject per analysis parameter per analysis timepoint
ADQS - Questionnaires Analysis Dataset	BASIC DATA STRUCTURE	Y			One or more records per subject per analysis parameter per analysis timepoint
ADSU - Substance Use Analysis Dataset	BASIC DATA STRUCTURE		Y		One or more records per subject per analysis parameter per analysis timepoint
ADVS - Vital Signs Analysis Dataset	BASIC DATA STRUCTURE	Y			One or more records per subject per analysis parameter per analysis timepoint

## 6. Data Conformance Summary

### 6.1 Conformance Inputs

*Was Pinnacle21 used to evaluate conformance?*

**Yes**

*If yes, specify the versions of Pinnacle21 and the Pinnacle21 validation rules:*

**Pinnacle 21 Enterprise version 3.4.4**

*Were sponsor-defined validation rules used to evaluate conformance?*

**No**

*If yes, describe any significant sponsor-defined validation rules:*

**n/a**

*Were the ADaM datasets evaluated in relation to define.xml?*

**Yes**

*Was define.xml evaluated?*

**Yes**

### 6.2 Issues Summary

Check ID	Diagnostic Message	Severity	Dataset	Count (Issue Rate)	Explanation
AD0018	Variable label mismatch between dataset and ADaM standard	Error	ADAE	4 (9.09%)	The labels for ASTDT, AENDT, ASTDY, AENDY are missing one space between one of the words and was left as is, as it does not impact analysis.
AD0018	Variable label mismatch between dataset and ADaM standard	Error	ADEG	1 (2.86%)	The label ANL01FL is updated for clarity.
AD0018	Variable label mismatch between dataset and ADaM standard	Error	ADLB	1 (2.78%)	The label ANL01FL is updated for clarity.

Check ID	Diagnostic Message	Severity	Dataset	Count (Issue Rate)	Explanation
AD0018	Variable label mismatch between dataset and ADaM standard	Error	ADPE	1 (3.33%)	The label ANL01FL is updated for clarity.
AD0018	Variable label mismatch between dataset and ADaM standard	Error	ADVS	1 (3.03%)	The label ANL01FL is updated for clarity.
AD0092B	Inconsistent value for TRTPN	Error	ADQS	181 (0.58%)	Due to study design (two cross-overs 3x3 with the same treatments), each treatment is taken twice: once during Day 1 to 3 with a corresponding TRTPN (1, 2 or 3) and once during Day 4 to 6 with a different corresponding TRTPN (4, 5 or 6)
AD0095B	Inconsistent value for TRTAN	Error	ADQS	273 (0.88%)	Due to study design (two cross-overs 3x3 with the same treatments), each treatment is taken twice: once during Day 1 to 3 with a corresponding TRTAN (1, 2 or 3) and once during Day 4 to 6 with a different corresponding TRTAN (4, 5 or 6)
AD0124	Inconsistent value for PARCAT1 within a unique PARAMCD	Error	ADPE	24 (4.15%)	Same tests were collected in two different CRF pages which are identified by PARCAT1.
AD0124	Inconsistent value for PARCAT2 within a unique PARAMCD	Error	ADQS	28860 (89.86%)	Same tests were collected in three different CRF pages which are identified by PARCAT2.

Check ID	Diagnostic Message	Severity	Dataset	Count (Issue Rate)	Explanation
AD0141	Inconsistent value for PARAM within a unique PARAMCD	Error	ADSU	5 (6.17%)	More than one value is mapped to the same PARAMCD as they were collected differently.
AD0146	Inconsistent value for PARAM	Error	ADSU	5 (6.17%)	PARAM values are kept as collected as was needed for analysis.
AD0146B	Inconsistent value for PARAM	Error	ADSU	5 (6.17%)	PARAM values are kept as collected as was needed for analysis.
AD0154	Multiple baseline records exist for a unique USUBJID,PARAMCD,BASETYPE	Error	ADQS	3640 (82.60%)	The baseline is defined as the pre-dose value for each product use session. Therefore there is a unique baseline record each USUBJID, PARCAT2, VISIT, PARAMCD
AD0242B	Inconsistent value for TRT06AN	Error	ADSL	1 (1.79%)	Due to study design (two cross-overs 3x3 with the same treatments), each treatment is taken twice: once during Day 1 to 3 with a corresponding TRTPN (1, 2 or 3) and once during Day 4 to 6 with a different corresponding TRTPN (4, 5 or 6)
AD1012	Secondary variable is present but its primary variable is not present	Error	ADEG	1 (25.00%)	EGPREC is not a secondary variable.
AD1012	Secondary variable is present but its primary variable is not present	Error	ADPC	1 (20.00%)	NBPREC is not a secondary variable.
AD1012	Secondary variable is present but its primary variable is not present	Error	ADPP	2 (33.33%)	NBPREC, PPSPEC are not secondary variables.

Check ID	Diagnostic Message	Severity	Dataset	Count (Issue Rate)	Explanation
AD1012	Secondary variable is present but its primary variable is not present	Error	ADSL	1 (25.00%)	ETHNIC is not a secondary variable.
DD0007	Invalid content was found starting with element 'ItemRef'.	Error	DEFINE	11 (100.00%)	This issue is occurring because Pinnacle21's system is incorrectly identifying "ADXX" domains as split datasets. This does not affect the xml content and the information displayed in the define file.