Analysis Data Reviewer Guide for BLA Sequencing Data Analysis

BioNTech SE and PFIZER INC. Study C4591001

ANALYSIS DATA REVIEWER GUIDE
REVISION HISTORY

Version	Summary of Major Change(s) and Impact	Version Date
1.0	First approved version of Analysis Data Reviewer Guide for Sequencing data Analysis	03-Jun-2021

Analysis Data Reviewer 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. This ADRG is a supplement to the primary ADRG already submitted for the BLA and covers the efficacy sequencing data analyses in blinded placebo-controlled follow-up period for the COVID-19 cases data (cutoff date:13Mar2021).

1.2 Acronyms

Acronym	Translation
ADaM	Analysis Dataset Model
ADRG	Analysis Data Reviewer's Guide
BLA	Biologics License Application
COVID-19	Coronavirus Disease 2019
IG	Implementation Guide
NA	Not Applicable
SAP	Statistical Analysis Plan
SDTM	Study Data Tabulation Model

1.3 Study Data Standards and Dictionary Inventory

Standard or Dictionary	Versions Used
SDTM	•SDTM v1.4 •SDTM-IG v3.2
SDTM Controlled Terminology	CDISC SDTM Controlled Terminology, 2020-03-27
ADaM	•ADaM v2.1 •ADaM-IG v1.1
ADaM Controlled Terminology	CDISC ADaM Controlled Terminology, 2020-03-27
Data Definitions	Define-XML v2.0
Pinnacle 21	Pinnacle 21 Enterprise 4.2.0

1.4 Source Data Used for Analysis Dataset Creation

The ADaM datasets were derived from the SDTM datasets.

2. Protocol Description

2.1 Protocol Number and Title
Refer to the primary ADRG.

2.2 Protocol Design in Relation to ADaM Concepts
Refer to the primary ADRG.

3. Analysis Considerations Related to Multiple Analysis Datasets

3.1 Study Populations and Core Variables
Refer to the primary ADRG.

3.2 Treatment Variable
ARM versus TRTxxP

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

No, TRT01P is null when ARM equals to "NOT ASSIGNED" or "SCREEN FAILURE".
ARM represents the planned arm for the blinded placebo-controlled period based on
randomization file. TRT01P has the planned treatment for the blinded placebo-controlled
period. TRT02P is only available in ADSL, and was not included and used in the
sequencing data analysis. See details in below table.

PHASE	ARM	TRT01P	TRT02P
Phase 1	BNT162b1 Phase 1 (10 mcg)	BNT162b1 Phase 1 (10 mcg)	-
	BNT162b1 Phase 1 (20 mcg)	BNT162b1 Phase 1 (20 mcg)	-
	BNT162b1 Phase 1 (30 mcg)	BNT162b1 Phase 1 (30 mcg)	-
	BNT162b1 Phase 1 (100/10 mcg)	BNT162b1 Phase 1 (100/10 mcg)	-

PHASE	ARM	TRT01P	TRT02P
	BNT162b2 Phase 1 (10 mcg)	BNT162b2 Phase 1 (10 mcg)	-
	BNT162b2 Phase 1 (20 mcg)	BNT162b2 Phase 1 (20 mcg)	-
	BNT162b2 Phase 1 (30 mcg)	BNT162b2 Phase 1 (30 mcg)	-
	Placebo	Placebo	-
	Placebo	Placebo	BNT162b2 Phase 1 (30 mcg)
	NOT ASSIGNED	-	-
	SCREEN FAILURE	-	-
Phase 2/3	BNT162b2 Phase 2/3 (30 mcg)	BNT162b2 Phase 2/3 (30 mcg)	-
	Placebo	Placebo	-
	Placebo	Placebo	BNT162b2 Phase 2/3 (30 mcg)
	NOT ASSIGNED	-	-
	SCREEN FAILURE	-	-

Note: Unit of dose 'mcg' was displayed as 'μg' in all of outputs.

ACTARM versus TRTxxA

If TRTxxA is used, then are the values of ACTARM equivalent in meaning to values of TRT01A?

No. TRT01A is not used in sequencing analysis data. TRT02A is only available in ADSL, and was not included and used in the sequencing data analysis.

Use of ADaM Treatment Variables in Analysis

Are both planned and actual treatment variables used in analyses?

No. Only the planned treatment was used in the sequencing data analysis.

Use of ADaM Treatment Grouping Variables in Analysis

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

No. Neither planned nor actual treatment grouping variables are used in analysis

3.3 Subject Issues that Require Special Analysis Rules

No change to this section, please refer to the details from the primary ADRG.

3.4 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?

No.

3.5 Imputation/Derivation Methods

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

No.

Was DTYPE used in one or more analysis datasets?

No.

4. Analysis Data Creation and Processing Issues

4.1 Split Datasets

There are no split datasets.

4.2 Data Dependencies

ADXB dataset pull core variable values from ADSL, and also uses ADC19EF dataset as an input to create efficacy parameter variables.

4.3 Intermediate Datasets

No intermediate analysis datasets were created in this trial.

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?

Yes. Subjects with 'NOTASSGN' 'SCRNFAIL' are only included in ADSL.

Are data taken from an ongoing study?

Yes. The sequencing data contains SARS-CoV-2 lineage data for the COVID-19 cases as of 13Mar2021 cutoff. For the other data, please refer to the primary package.

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

No. The analysis dataset in this supplemental package only supports analysis of SARS-CoV-2 lineage for COVID-19 cases.

Additional Content of Interest

No additional content of Interest.

5.2 Analysis Datasets

Dataset Label	Class	Efficacy	Safety	Baseline or other subject PK/PD	Primary	Structure
ADSL Subject-Level Analysis Dataset	SUBJECT LEVEL ANALYSIS DATASET			X		One record per subject
ADXB Sequencing Analysis Dataset	BASIC DATA STRUCTURE	X				One record per subject per analysis parameter per analysis timepoint

5.2.1   ADSL – Subject-Level Analysis Dataset

The ADSL is identical to the dataset included in the BLA esub package, for Pinnacle 21 validation checks.

5.2.2   ADXB – Sequencing Analysis Dataset

The purpose of this dataset is to get SARS-CoV-2 lineage phylogenetic analysis information for the COVID-19 cases. Key variables in this dataset are

Variable Name	Variable Description
FC19D27	First COVID Onset 7D Post D2 (PD)
FSEVPD27	First Sev COVID Onset 7D Post D2 (PD)
FSEVCD27	First Sev COVID Onset 7D Post D2 (CDC)
FSEVPD1	First Sev COVID Onset Post D1 (PD)
FSEVCD1	First Sev COVID Onset Post D1 (CDC)

CDC= Centers for Disease Control and Prevention defined; D1=Dose 1; D2=Dose 2; PD=Protocol-defined; Sev=Severe.

Table below describes subset condition for total number of subjects with various definition of confirmed COVID-19 cases (N) from ADC19EF and ADSL.

Analysis Population		Total Number of Subjects (N)		Subset Condition for Total N
	Sub-condition	BNT162b2	Placebo	
Evaluable Efficacy	Subjects with first COVID-19 occurrence From 7 Days After Dose 2	81	873	ADSL.EVALEFFL="Y" and ADSL.PHASEN ne 1 and ADSL.MULENRFL ne "Y" and ADSL.HIVFL = 'N' and UPCASE(ADC19EF.PARAMCD) = "C19ONST" and FIND(ADC19EF.AVALC, 'POS' , 'i') and ADC19EF.ILD27FL = "Y" and ADC19EF.FILOCRFL = "Y" and ADC19EF.PDRMUPFL = "N" and ((NOT MISSING(ADC19EF.DVSTDT) and ADC19EF.ADT <=

Analysis Population	Sub-condition	Total Number of Subjects (N)		Subset Condition for Total N
		BNT162b2	Placebo	
				ADC19EF.DVSTDT) or MISSING(ADC19EF.DVSTDT))
	Subjects with first severe COVID-19 occurrence Based on FDA-Definition From 7 Days After Dose 2	1	21	ADSL.EVALEFFL="Y" and ADSL.PHASEN NE 1 and ADSL.MULENRFL NE "Y" and ADSL.HIVFL = 'N' and UPCASE(ADC19EF.PARAMCD) = "SEVCONST" and FIND(ADC19EF.AVALC, 'POS' , 'i') and ADC19EF.ILD27FL = "Y" and ADC19EF.FILOCRFL = "Y" and ADC19EF.PDRMUPFL = "N" and ((NOT MISSING(ADC19EF.DVSTDT) and ADC19EF.ADT <= ADC19EF.DVSTDT) or MISSING(ADC19EF.DVSTDT))
	Subjects with first severe COVID-19 occurrence Based on CDC-Definition From 7 Days After Dose 2	0	32	ADSL.EVALEFFL="Y" and ADSL.PHASEN NE 1 and ADSL.MULENRFL NE "Y" and ADSL.HIVFL = 'N' and UPCASE(ADC19EF.PARAMCD) = "CDCSONST" and FIND(ADC19EF.AVALC, 'POS' , 'i') and ADC19EF.ILD27FL = "Y" and ADC19EF.FILOCRFL = "Y" and ADC19EF.CDRMUPFL= "N" and ((NOT MISSING(ADC19EF.DVSTDT) and ADC19EF.ADT <= ADC19EF.DVSTDT) or MISSING(ADC19EF.DVSTDT))
Dose 1 All-Available Efficacy	Subjects with first severe COVID-19 occurrence Based on FDA-Definition After Dose 1	1	30	ADSL.AAI1EFFL="Y" and ADSL.PHASEN ne 1 and ADSL.MULENRFL ne "Y" and ADSL.HIVFL = 'N' and UPCASE(ADC19EF.PARAMCD) = "SEVCONST" and FIND(ADC19EF.AVALC, 'POS' , 'i') and ADC19EF.ADT >= ADC19EF.VAX101DT and ADC19EF.FILOCRFL= "Y" and ADC19EF.PDRMUPFL = "N"
	Subjects with first severe COVID-19 occurrence Based on	1	45	ADSL.AAI1EFFL="Y" and ADSL.PHASEN ne 1 and ADSL.MULENRFL ne "Y" and

| Analysis Population | | Total Number of Subjects (N) | | Subset Condition for Total N |
	Sub-condition	BNT162b2	Placebo	
	CDC-Definition After Dose 1			ADSL.HIVFL = 'N' and UPCASE(ADC19EF.PARAMCD) = "CDCSONST" and FIND(ADC19EF.AVALC, 'POS' , 'i') and ADC19EF.ADT >= ADC19EF.VAX101DT and ADC19EF.FILOCRFL= "Y" and ADC19EF.CDRMUPFL= "N"

6. Data Conformance Summary

6.1 Conformance Inputs
Specify the software name and version for the analysis datasets
Pinnacle 21 Enterprise 4.2.0., Validation Engine version FDA 2010.1

Specify the version of the validation rules (i.e. CDISC, FDA) for the analysis datasets
CDISC ADaM-CT 2020-03-27

Specify the software name and version for the define.xml
Pinnacle 21 Enterprise 4.2.0.

Specify the version of the validation rules (i.e. CDISC, FDA) for the define.xml
CDISC ADaM CT 2020-03-27

6.2 Issues Summary (Pinnacle 21 Enterprise Validation Report)
Note: The issues associated with ADSL are not included as they are covered in the BLA esub package.

Check ID	Diagnostic Message	FDA Severity	Dataset	Count (Issue Rate)	Explanation
CT2002	RACE value not found in 'Race' extensible codelist	Warning	ADXB	36 (3.14%)	New terms were added to extensible codelist RACE for the study protocol needs: Multiple

7. Submission of Programs
All programs for analysis datasets are submitted as shown below. All programs were created on a SAS platform using 9.4. ADSL.sas (adsl-sas.txt) must be run first before any other ADaM datasets; ADXB program is dependent on ADSL and ADC19EF. Please be aware ADC19EF and ADSYMPT (ADSYMPT was used to populate efficacy listing) have been included in the primary package so will not be submitted again.

7.1 ADaM Programs

Program Name	Output	Input	Macro Used
adxb-sas.txt	adxb.xpt	xb adsl adc19ef	NA

7.2 Analysis Output Programs

NA

8. Appendix

NA