PhUSE US Connect 2019

Paper SI02

Creating the Virologic Response ADaM Dataset

Manjusha Alapati, Ephicacy Lifescience Analytics, Bangalore, India

ABSTRACT

Virology studies have exclusive characteristics that effect data collection and analysis. In these studies data is collected from host and infecting virus and it is important to associate the data from host to the non-host organism. Considering one of the most chronic viral infection (HIV) and the associated clinical trials. Current treatment available for HIV infection is a combination of antiretroviral drugs with a goal to suppress viral load i.e. HIV-RNA levels. There are different approaches for evaluating the virologic response data. This paper provides insights on different approaches and considerations to be taken while creating the analysis dataset. This paper also provides challenges about analysis of virologic outcome and handling of missing data.

INTRODUCTION

In contrast to other clinical studies, virology studies have unique way of data collection. In these studies data is collected from both the host i.e. human subject as well as infecting virus. In clinical trials of antiviral drugs, viral load measurement within the host is collected over time in order to understand the antiviral activity or efficacy of the drug. The collection and association of this data is associated with genotypic and phenotypic changes in viral population. This helps to identify viral characteristics associated with treatment failure or success.

HIV infection is a chronic viral infection when untreated causes a progressive destruction of immune system resulting in acquired immunodeficiency syndrome (AIDS). Current treatment of HIV consists of a combination of antiretroviral drugs. Most antiviral drugs entered the market are through accelerated approvals. Viral Load is the term used to describe the amount of HIV in a body fluid. Viral load tests measure the amount of HIV in small sample of blood. The result of a viral load test is described as the number of copies of HIV's genetic material RNA per millimeter (copies/mL).There are number of different viral load tests in use, each having a different technique to measure HIV in the blood. However, each test has a limit below which it cannot be reliably detect HIV. This is referred to as viral load being 'Undetectable'.

A collaborative group of pharmaceutical, academic and government scientists investigated the relationships between treatment induced changes in HIV-1 RNA levels and clinical endpoints collected from ongoing and completed trials. Based on this committee data, treatment decrease in HIV-RNA levels were highly predictive of meaningful clinical benefit and could serve as endpoints in clinical trials designed to support both traditional and accelerated approvals. The goal of antiretroviral treatment is to maintain suppression of plasma HIV-RNA levels below the level of detection of sensitive HIV-RNA assays.

This paper provides an overview of the algorithms used for analysis and illustrates an analysis dataset that supports derivations, identification of baseline values and handling of missing data. This paper also enlists how this dataset and variables can be used for reporting.

ALGORITHMS

There are multiple algorithms recommended for evaluating virologic response in clinical trials supporting antiretroviral approvals. FDA Time to loss of virologic response (TLOVR) algorithm has been commonly used to assess virologic response to antiretroviral regimens. The FDA snapshot algorithm has been proposed to replace the TLOVR algorithm, as it is simpler and is expected to yield similar results.

TLOVR ALGORITHM

Time to loss of virologic response algorithm incorporates longitudinal data of HIV RNA to determine the virologic response at every moment of the clinical trial. The algorithm is affected by changes in the levels of HIV, by the frequency of the analysis of viral load and by the way of handling missing data. The virologic failure was defined as the first in occur between the events of death, abandonment of the drug in study or loss during monitoring, HIV RNA concentrations in greater than or equal to 50 copies/mL detected in two consecutive visits or a value of greater than or equal to 50 copies/mL followed by a permanent discontinuation of study drug during the follow-up.

The Division of Antiviral Products (DAVP), statistical and clinical reviewers conducted the project 'Handling uncertainty in Endpoint selection and other Endpoint issues' in order to determine if simplified end points could be used for approval at week 48. The team evaluated 18 trials from NDAs with 8046 patients. Results from TLOVR algorithm i.e. 61 percent of 8046 patients remained in the study for 48 weeks and were virologic responders compared to 61 percent of the patients using the snapshot algorithm. Differences between the two approaches are minimal. Based on the findings from the project and the ease of using snapshot approach, this approach is followed for the pending supplemental NDAs and future NDAs to include virologic outcome results.

FDA SNAPSHOT ALGORITHM

This algorithm is for analysis of virologic outcome at a given time point or a window period. This approach is used for the determination of primary efficacy end point. The algorithm provides guidance of handling missing data and the associated imputation required for the analysis.

The main principle of snapshot analysis is to follow Virology First Hierarchy i.e. the Primary efficacy endpoint should be primarily a virologic endpoint and not a clinical endpoint. Based on the classification of either virologic success or failure, the algorithm also classifies the virologic failure into two categories: virologic failure or No virologic data at Week 48 window.

Depending on the study design and the endpoints the rules for determining the outcome can be implemented. Based on the above primary endpoint the following rules were considered for determination of virologic outcome during the window period,

1. On treatment plasma HIV-1 RNA data will be utilized to determine the outcome of a subject. Data beyond the last study drug administration date will not be considered in the analysis.
2. If multiple HIV-1 RNA values exist within the same visit window, the last value will be considered for analysis.
3. If multiple outcomes were observed in the same window, then virology first principle will be considered. If subject discontinues study drug for any reason or dies with in the visit window without HIV-1 RNA assessment but on treatment HIV-1 RNA data is available in the visit window, the last assessment value prior the event will be considered for determining the virologic outcome.

ILLUSTRATIVE STUDY EXAMPLE

The primary endpoint for this study is the proportion of subjects with plasma HIV-1 RNA less than 50 copies/mL at week 48. For this analysis subject status will be assessed according to the MSD=F approach i.e. 'Missing, Switch or Discontinuation =Failure'. Secondary endpoints will be based on the mean change from baseline values using an ANCOVA model adjusting for the baseline and stratification factors.

For determining the virologic outcome at the time point on Week X time point as per the protocol following rules were considered,

1. Subject with Plasma HIV-1 RNA less than 50 copies/mL at any of the defined time point will be considered as virologic success.
2. Subject having HIV-1 RNA greater than 50 copies/mL within the defined visit window or has no plasma HIV-1 RNA data within the defined visit window or discontinued the study not due to an adverse event or death will be considered as virologic failure.
3. Subject having no plasma HIV-1 RNA value available and discontinued the study due to an adverse event or death or any other reason and the last available value is less than 50 copies/mL prior to the lower limit of the visit window will be considered as No virologic data at Week X window

ADVIR

To support the analysis, the content and structure of the primary efficacy analysis dataset will be containing records and variables which would be easy to identify the virologic outcome for each of the subject. However, it will be difficult to include these variables in ADSL dataset as it is often the first analysis dataset to be created. End point data will be included in the BDS data structure and will contain variables capturing the information of HIV-1 RNA at baseline, at different time points and new records as well as variables will be included to handle the derivations or imputations as required for the analysis.

When building the analysis data set the below points to be considered,

1. Ordering of the variables to be followed as per the standards in case of the existing standard variables.
2. When defining the source of the variable the immediate predecessor dataset to be mentioned.
3. Datasets should be analysis ready i.e. containing all the variables needed for the specific analysis. In addition to the required variables all the other variables required as per the project like baseline values, stratification or grouping variables, response variables and other supportive variables to facilitate traceability.

All the subject level and treatment variables this dataset also includes the end point data and below are the variables,

1. HIV-1 RNA values at baseline
2. HIV-1 RNA at predefined time points
3. HIV-1 RNA values at other times i.e. either the loss of virologic failure or discontinuation
4. Mean log change of HIV-1 RNA values
5. Virologic outcome for the subject
6. Flag variables to identify the subject had achieved the end points

ADVIR DATASET METADATA

The dataset contains one record per subject per parameter per analysis visit. This dataset follows the format of the ADaM basic data structure (BDS). All the variables present in the source data or the SDTM will be included to ensure the traceability of the derivations while creating this dataset.

Dataset	Dataset Description	Dataset Structure	Key Variables	Class of Dataset	Documentation
ADVIR	Virologic Response Analysis Dataset	One record per subject per parameter per analysis visit	STUDYID, USUBJID, PARAMCD, AVISIT	Basic Data Structure	advir.sas

ADVIR VARIABLE METADATA

The variables present in this dataset mainly contains information about 1) results for the analysis parameter; 2) representation of multiple findings within the analysis timepoint or visit; 3) derived variables useful for identification of virologic status of the subject.

ADVIR TIMING VARIABLE METADATA

This section represents the variables related to the analysis timepoints. Based on the visit windowing mentioned either in protocol or SAP the analysis visit variables will be derived and used for reporting.

Dataset Name	Variable Name	Variable Label	Variable Type	Derivation
ADVIR	AVISIT	Analysis Visit	Character	Based on the Statistical Analysis Plan use of visit windowing in the reporting.
ADVIR	AVISITN	Analysis Visit (N)	Numeric	Derived based on AVISIT
ADVIR	AWTARGET	Analysis Window Target	Numeric	Based on the lower and upper values defined in the visit windowing section of the SAP
ADVIR	AWTDIFF	Analysis Window Diff from Target	Numeric	Difference between the analysis day and analysis target day
ADVIR	ADT	Analysis Date	Numeric	Numeric date value of XX.XXDTC
ADVIR	ADY	Analysis Day	Numeric	If TRTSDT <= ADT, then ADY=ADT-TRTSDT+1 else ADY=ADT-TRTSDT

SAMPLE CODE

Analysis visit is different from the collected visit or SDTM visit. The visit that is used for analysis and guidance of the derivation is provided in SAP. Below is the example of deriving the analysis visit,

```
data avisit;
  format avisitn 8.;
  format avisit $15.;
  if .< ady<= 1 then do; avisitn = 1; avisit='Baseline'; end;
  else if 2<=ady<=41 then do; avisitn = 4; avisit='Week 4'; end;
  else if 42<=ady<=69 then do; avisitn = 8; avisit='Week 8'; end;
  else if 70<=ady<=97 then do; avisitn = 12; avisit='Week 12'; end;
  else if 98<=ady<=125 then do; avisitn= 16; avisit='Week 16'; end;
  etc...
run;
```

ILLUSTRATION OF ANALYSIS VISIT

STUDYID	USUBJID	TRTP	PARAMCD	VISIT	AVISIT	AVISITN
XXXXX01	101-101	ABC	HIVRNA	Day 1	Baseline	1
XXXXX01	101-101	ABC	HIVRNA	Day 7	Week 4	4
XXXXX01	101-101	ABC	HIVRNA	Day 14	Week 4	4
XXXXX01	101-101	ABC	HIVRNA	Day 21	Week 4	4
XXXXX01	101-101	ABC	HIVRNA	Day 28	Week 4	4
XXXXX01	101-101	ABC	HIVRNA	Day 35	Week 4	4
XXXXX01	101-101	ABC	HIVRNA	Day 42	Week 8	8
XXXXX01	101-101	ABC	HIVRNA	Day 48	Week 8	8
XXXXX01	101-101	ABC	HIVRNA	Day 56	Week 8	8

ADVIR VARIABLE METADATA

This section represents the variables and records associated with the analysis described in the SAP. Missing values to be imputed using the Last Observation Carried Forward (LOCF) method.

Dataset Name	Variable Name	Variable Label	Variable Type	Derivation
ADVIR	PARAM	Parameter	Character	Additional parameters required for the analysis and reporting
ADVIR	PARAMCD	Parameter Code	Character	Assigned based on PARAM
ADVIR	AVAL	Analysis Value	Numeric	Equal to xx.xxSTRESN along with any derived result as per SAP.
ADVIR	DTYPE	Derivation Type	Character	Based on the derivation type of the variable. Will be populated as 'LOCF' if the AVAL is populated based on this approach
ADVIR	PARAMTYP	Parameter Type	Character	Will be assigned DERIVED when the new parameter is derived as per SAP and do not exist in source data.

SAMPLE CODE

As defined in the SAP for the missing results at any timepoint to be taken from the last non-missing value from a prior record. The assigned value has to be identified in the dataset with the help of an additional variable which will indicate the derivation type of the analysis value. In the below example LOCF (Last Observation Carried Forward) approach is followed.

```
proc sort data=forlocf;
      by usubjid paramcd avisitn;
run;
data locf;
      set forlocf;
      by usubjid paramcd avisitn;
      retain retval ;
```

```
      if first.avisitn then retval=.;
      if aval eq . and retval ne . then aval=retval;
      if aval ne . then retval=aval;
run;
```

ILLUSTRATION OF DERIVED RESULT

STUDYID	USUBJID	PARAMCD	VISIT	AVISIT	AVISITN	LBSTRESN	AVAL	DTYPE
XXXXX01	101-101	HIVRNA	Day 1	Baseline	1	46	46	
XXXXX01	101-101	HIVRNA	Day 7	Week 4	4	48	48	
XXXXX01	101-101	HIVRNA	Day 14	Week 4	4	43	43	
XXXXX01	101-101	HIVRNA	Day 21	Week 4	4	44	44	
XXXXX01	101-101	HIVRNA	Day 28	Week 4	4	46	46	
XXXXX01	101-101	HIVRNA	Day 35	Week 4	4		46	LOCF
XXXXX01	101-101	HIVRNA	Day 42	Week 8	8	46	46	
XXXXX01	101-101	HIVRNA	Day 48	Week 8	8	48	48	
XXXXX01	101-101	HIVRNA	Day 56	Week 8	8		48	LOCF

ILLUSTRATION OF ADDITIONAL PARAMETER

STUDYID	USUBJID	PARAMCD	VISIT	AVISIT	AVAL	DTYPE	PARAMTYP
XXXXX01	101-101	HIV1 RNA	Day 1	Baseline	46		
XXXXX01	101-101	HIV1 RNA	Day 7	Week 4	48		
XXXXX01	101-101	HIV1 RNA	Day 14	Week 4	51		
XXXXX01	101-101	HIV1 RNA	Day 21	Week 4	44		
XXXXX01	101-101	HIV1 RNA	Day 28	Week 4	46		
XXXXX01	101-101	HIV1 RNA	Day 35	Week 4	46	LOCF	
XXXXX01	101-101	HIV1 RNA	Endpoint	Week 4	46		DERIVED
XXXXX01	101-101	HIV1 RNA	Day 42	Week 8	54		
XXXXX01	101-101	HIV1 RNA	Day 48	Week 8	48		
XXXXX01	101-101	HIV1 RNA	Day 56	Week 8	48	LOCF	

Based on the analysis and to summarize the data an additional parameter can be added in the dataset. PARAMTYP will be added in the dataset in order to identify the derived parameter. In the above example, additional record is included in the dataset to identify the Endpoint data.

ILLUSTRATION OF DERIVED FLAG VARIABLES

STUDYID	USUBJID	PARAMCD	VISIT	AVISIT	AVAL	DTYPE	CRIT1	CRIT1 FL
XXXXX01	101-101	HIV1 RNA	Day 1	Baseline	46		<50 copies/mL	Y
XXXXX01	101-101	HIV1 RNA	Day 7	Week 4	48		<50 copies/mL	Y
XXXXX01	101-101	HIV1 RNA	Day 14	Week 4	51		<50 copies/mL	
XXXXX01	101-101	HIV1 RNA	Day 21	Week 4	44		<50 copies/mL	Y
XXXXX01	101-101	HIV1 RNA	Day 28	Week 4	46		<50 copies/mL	Y
XXXXX01	101-101	HIV1 RNA	Day 35	Week 4	46	LOCF	<50 copies/mL	Y
XXXXX01	101-101	FINVALUE		Week 4	46		<50 copies/mL	Y
XXXXX01	101-101	HIV1 RNA	Day 42	Week 8	54		<50 copies/mL	
XXXXX01	101-101	HIV1 RNA	Day 48	Week 8	48		<50 copies/mL	Y
XXXXX01	101-101	HIV1 RNA	Day 56	Week 8	48	LOCF	<50 copies/mL	Y

Based on the reporting criteria with in a parameter the variables can be added in the dataset. In the above illustration, the variables for the parameter HIV-1 RNA when satisfied the condition the flag will be populated as 'Y'. These variables can be used to summarize the data of derived records by using the DTYPE variable and to summarize the data based on the subjects satisfying the criteria.

ANALYSIS RESULTS

The following data display illustrates an ANCOVA analysis performed using ADVIR as described above,

The primary efficacy analysis endpoint is the percent change from baseline at Week48 in the HIV-1 RNA data. The analysis is conducted using the Intent-To-Treat (ITT) population, with LOCF imputation for any missing values. As the primary efficacy analysis method, an ANCOVA model is used, with planned treatment as categorical variable, baseline HIV-1 RNA value as a covariate. Other covariates such as age, race and screening BMI will be included as covariates. The adjusted least squares (LS) means and 95% confidence interval are presented.

HIV-1 RNA Percent Change from Baseline at Week 48 (ITT Population, LOCF Data, ANCOVA Model)						
	%Change from Baseline			**Treatment Difference**		
	n	LSMEAN*	95% CI*	LSMEAN*	95% CI*	p-value*
Drug A (N=)						
Placebo (N=)						
N=ITT population; n=number of subjects with non-missing percentage change from baseline at week48; CI=confidence Interval; LS=Least Squares; *Based on ANCOVA model						

ANALYSIS RESULTS METADATA FOR ANCOVA EXAMPLE

Metadata Field	Metadata
Title	HIV-1 RNA Percentage Change from Baseline at Week 48 (ITT Population)
Result Identifier	LSMEANS, CI, p-value
PARAM	HIV RNA (Copies/mL)
PARAMCD	HIVRNA
Analysis Variable	PCHG
Reason	Primary Efficacy Analysis as per SAP
Dataset	ADVIR
Selection Criteria	ITTFL='Y' and PARAMCD='HIVRNA' and AVISIT='WEEK 48' and ANL01FL='Y'

CONCLUSION

This paper provides basic knowledge of the different approaches followed in the analysis of data in HIV studies. Although the detailed analysis and derivations will always be based on the clinical trial, the above-mentioned illustrations help in understanding few of them. The number of variables required for the analysis may increase based on the complexity of clinical trial. The viral load monitoring and response is preferred approach compared to the immunological and clinical monitoring is to provide an early and more accurate identification of treatment failure in early stages and to switch to second-line drugs to improve clinical outcomes.

REFERENCES

- Hill, AM, R DeMasi, and D Dawson, 1999, Meta-Analysis of Antiretroviral Effects on HIV-1 RNA, CD4 Cell Count and Progression to AIDS or Death, Antiviral Therapy.
- Murray, JS, MR Elashoff, LC Iacono-Connors et al., 1999, The Use of Plasma HIV RNA as a Study Endpoint in Efficacy Trials of Antiretroviral Drugs, AIDS.
- http://www.fda.gov/BiologicsBloodVaccines/GuidanceComplianceRegulatoryInformation/Guidances/Vaccines/ucm074775
- https://www.cdisc.org/system/files/members/standard/foundational/adam/ADaMIG_v1.1.pdf
- Human Immunodeficiency Virus-1 Infection: Developing Antiretroviral Drugs for Treatment.

- Antiviral Product Development--Conducting and Submitting Virology Studies to the Agency: Guidance for Submitting HIV Resistance Data.
- Guidance for Industry Human Immunodeficiency Virus-1 Infection: Developing Antiretroviral Drugs for Treatment

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Manjusha Alapati
Ephicacy Lifescience Analytics
Bangalore,
India
Email: manjusha.alapati@ephicacy.in;