PharmaSUG 2017 - Paper SS07

Overview and Application of the HCV Vertical Resistance Analysis Template

Yan Xie, Abbvie Inc.

ABSTRACT

FDA proposed an updated hepatitis C virus (HCV) vertical resistance draft template on March 2016, which is quite different from the existing guidance for submitting HCV resistance data. The new vertical template is not published yet and is still under review, but FDA encourages the sponsors to submit resistance data using this new template. Abbvie received this guidance directly from FDA. Compared to the previous horizontal template which contained at least five hundred variables, the new vertical template, with only about hundred variables, is more advanced because: 1. it is compatible with SDTM and ADaM standards, 2. it reduces number of variables by applying the streamlined and simplified vertical format, 3. it can hold the Next Generation Sequencing (NGS) data, and 4. it can hold multiple targets and HCV subtypes in one dataset. The variables in the new vertical template include three categories of variables: subject level characteristics, pharmacogenomics results, and phenotypic results. The subject level characteristic variables are derived from the ADaM datasets (e.g. ADSL), and the pharmacogenomics and phenotypic information can be found in the PF and MS domains in SDTM, respectively. Only variants which are different from the prototypic reference are included in the resistance dataset. Also the data structure in the new vertical template is one record per subject, visit sequenced, genetic region of interest, location, and variant. The new vertical template was implemented in Abbvie's recent HCV drug submission. The HCV resistance data derived from the new vertical template for Abbvie's pilot study was reviewed by FDA in March 2016. This paper will discuss the approaches taken at Abbvie to create such analysis dataset.

INTRODUCTION

The resistance to HCV treatment has been summarized in the reference presentations (see Forum HCV Drug Resistance Slideset in the reference section). Currently the sequences for the direct-acting antivirals (DAAs) HCV drugs are targeted at NS3/4A, NS5A and NS5B genes. The analysis of these sequences can provide information on the HCV's resistance to treatment and association of the variants to breakthrough or relapse, which are two major virologic failures in HCV treatment.

Two kinds of the sequence data, population and deep sequence (or NGS) are widely used in Abbvie's studies. The example raw data from the population and NGS are presented in Table 1 and Table 2 below, respectively. NGS data use 2% cutoff threshold, which means any variants with less than 2% frequency will not be included in the data. The variants refer to the differences from the sequence to the genotype-specific reference sequence as shown in Table 3.

Currently the resistance data are sent from Abbvie's Clinical Virology department in CSV format which may not be structured as SDTM PF, MS, and OI datasets. Refer to Table 1 for the structure used by Abbvie. FDA's vertical template guidance, however, uses PF, OI and MS SDTM domains.

ROWID	USUBJID	SUBTYPE	TARGET	LBDT	NOTE	1	2	2.1	2.2	3	4	5
1	1001	1A	NS5A	03/02/2015	BL	S	G/V	-	-	S	W/T	L
2	1001	1A	NS5A	03/30/2015	W4	S	G	P	I/A	S/T	W/T/Y	L
3	1001	1A	NS5A	06/23/2015	PTW4	S	G	P	I/A	S	W/T	L

Table 1. Sample Population Sequence Data

ROWID	USUBJID	SUBTYPE	TARGET	LBDT	NOTE	1	2	2.1	2.2	3
1	1001	1A	NS5A	03/02/2015		S	G/V	-	-	S
2	1001	1A	NS5A	03/02/2015	VCOV		70628//796	-	-	
3	1001	1A	NS5A	03/02/2015	TCOV		71424	-	-	

ROWID	USUBJID	SUBTYPE	TARGET	LBDT	NOTE	1	2	2.1	2.2	3
4	1001	1A	NS5A	03/30/2015		S	G	P/-	I/-	S/T
5	1001	1A	NS5A	03/30/2015	VCOV			25201	25312	39160//580
6	1001	1A	NS5A	03/30/2015	TCOV			26354	26354	39740
7	1001	1A	NS5A	06/23/2015		S	G	P/-	I/-	S
8	1001	1A	NS5A	06/23/2015	VCOV			24920	24738	
9	1001	1A	NS5A	06/23/2015	TCOV			27549	27549	

ROWID	4	5
1	W/T	L
2	86024//2548	
3	88572	
4	W/T/Y	L
5	43516//615//512	
6	44643	
7	W/T	L
8	42263//1023	
9	43286	

Table 2. Sample NGS Data

ROWID	SUBTYPE	STRAIN	ACCESSION	TARGET	1	2	3	4	5
1	1A	H77	NC_004102	NS5A	S	G	S	W	L
2	1B	Con1	AJ238799	NS5A	S	G	S	W	L

Table 3. Reference Sequence Data

HCV VERITCAL TEMPLATE

FDA's updated HCV vertical template includes three categories of variables: subject level characteristics, pharmacogenomics results, and phenotypic results.

SUBJECT LEVEL CHARACTERISTICS

You can populate most of the subject level characteristics variables from ADSL, including STUDYID, USUBJID, ARM, ARMCD and TRTDUR. Other subject level variables can be populated from the efficacy analysis dataset ADEFFOUT or by simple derivations. For example, in the vertical resistance dataset, we have used the following variables:

Stratification variables STRATA and STRATAV. STRATA is the randomization strata, and STRATAV is the verified strata. If the randomization factor, such as HCV genotype subtype, changes from screening period to treatment period, then the values in STRATA and STRATAV may not be consistent. Each individual randomization stratum is stored in the stratum variable STRAT*y*, description of the randomization stratum variable STRAT*y*NM, and verified value of stratum variable STRAT*y*V, in which *y* represents the y^{th} randomization stratum.

Description of comorbidity and comorbidity flag variables COMOR*y*/COMOR*y*FL. Since our trials only have one comorbidity, we have used compensation cirrhosis as the comorbidity variable derived from the efficacy analysis dataset.

Description of co-infection and co-infection flag variables COINF*y*/COINF*y*FL. In our dataset, COINF1='HIV' and COINF1FL='N' for the subjects without HIV co-infection.

Description and value of subgroup variables SBDSCGR*y*/SBVALGR*y*. All subgroup variables, including sex, race, baseline fibrosis score, baseline HCV RNA level, prior HCV treatment history, etc. are pulled from the efficacy dataset ADEFFOUT.

Efficacy endpoint variables include primary efficacy responder flag EFFPRSFL, reason for not achieving efficacy response NEFFREA, efficacy responder flag EFFR*y*FL, exposure duration EXDUR, and exposure duration unit EXDURU. You can pull all these variables from the efficacy dataset ADEFFOUT.

PHARMACOGENOMICS VARIABLES

Visit name, planned study day of visit, study day of specimen collection, and sequence number are named as VISIT, VISITDY, PFDY, and PFSEQ, respectively. These variables should come from the pharmacogenomic domain (PF) in SDTM. Here we derive these variables based on the visit definition in the Statistical Analysis Plan (SAP) because our raw data are not in SDTM format. Similarly, you can find the predecessors in the PF and OI domains in SDTM for the non-host species variable NHSPCES, sponsor non-host organism identifier variable NHOID, and specimen type variable PFSPEC.

PFBLFL is the sequence baseline flag variable. PFGENRI represents the generic region of interest, which is the gene of either NS3/4A, NS5A or NS5B target. PFGENRNG is the range of the sequence, and the values vary on HCV subtype and PFGENRI. For example PFGENRNG='1-448' for genotype 1A and PFGENRNG='1-447' for genotype 1B for NS5A sequence. PFMETHOD shows the test method, which equals to 'POPULATION SEQUENCING' and 'NEXT GENERATION SEQUENCING' as in the example data in Table 1 and Table 2, respectively.

You will see the derivations of PFREF, PFGENLOC, PFNUMLOC, PFORREF, PFORRES, PFSTRESC, PRRESALL, PFRESA15, BLPMFL, and PFRESCAT variables in the example Table 4 below (using the data from ROWID=2). We pulled the reference amino acid into variable PFREF from Table 3, and the variable PFORREF takes the value of STRAIN variable in Table 3. The positions are listed vertically in the variables PFGENLOC and PFNUMLOC, where PFNUMLOC is the sub-numbering system for the insertions or deletions in the sequences. PFRESALL represents all of the amino acid differences from the reference for a given location. If you observe a mixture of the amino acid, the mixture is saved in the variable PFRESALL and each individual amino acid is separated and recorded vertically in variable PFSTRESC. Similarly, PFRESA15 indicates all variants detected at a >=15% frequency, and the result is blank if only reference sequence is detected >=15%. Variable BLPMFL represents whether the variant in PFSTRESC is detected at the baseline sequence. Variable PFRESCAT shows the variant is among substitution, insertion, or deletion.

VARIABLE NAME	USUBJID	PFREF	PFGENLOC	PFNUMLOC	PFORREF	PFORRES
VARIABLE LABEL		Reference Strain for Genotype Analysis	Genetic Location	Position Within Numbering System	Reference Result in Original Units	Result or Finding in Original Format
	1001		2	1	H77	P
	1001		2	2	H77	I
	1001		2	2	H77	A
	1001	S	3		H77	T
	1001	W	4		H77	T
	1001	W	4		H77	Y

VARIABLE NAME	USUBJID	PFSTRESC	PFRESALL	PFRESA15	BLPMFL	PFRESCAT
VARIABLE LABEL		Result or Finding in Standard Format	Accumulated Standardized Results	Accumulated Results at 15% Cutoff	Substitution Present at Baseline Flag	Result Category
	1001	-2.1P	-2.1P	-2.1P	N	INSERTION
	1001	-2.2I	-2.2I/A	-2.2I/A	N	INSERTION
	1001	-2.2A	-2.2I/A	-2.2I/A	N	INSERTION
	1001	S3T	S3S/T	S3S/T	N	SUBSTITUTION

VARIABLE NAME	USUBJID	PFSTRESC	PFRESALL	PFRESA15	BLPMFL	PFRESCAT
	1001	W4T	W4W/T/Y	W4W/T/Y	Y	SUBSTITUTION
	1001	W4Y	W4W/T/Y	W4W/T/Y	N	SUBSTITUTION

Table 4. Sample Population Resistance Dataset in the Vertical Template

For NGS data, variables TCOV, VCOV, AAFREQ, NGSALGO, and NGSCOFF demonstrate total coverage, coverage of variant, frequency of substitution, NGS algorism of method, and NGS frequency cutoff, respectively. If the variant is detected at baseline and is enriched at post-baseline visit, you can flag the enrichment in the variable ENRCHFL, and the enrich criteria in the variable ENRCHFLC.

PHENOTYPIC VARIABLES

We do not collect phenotypic information in our trials. So the derivations of phenotypic variables are not discussed here. You can find the IC50 and EC50 variables, as well as study day, method, reference strain for phenotype analysis readily from MS domain in SDTM.

COMPARESION TO PREVIOUS HORIZONTAL TEMPLATE

The updated, draft, HCV resistance vertical template provides the flexibility and traceability compared to the previous horizontal template. In the horizontal template, the variants are listed horizontally, as shown in Table 5. If a trial has the subjects with different HCV genotype subtype, multiple resistance datasets need to be created for each sequence type (population or clonal), target (NS3/4A, NS5A, NS5B) and subtype (GT 1A, 1B, 2A, 2B, etc.). This is due to difference in sequence length for different HCV genotype subtype, especially in NS5A target. On the other hand, it is structured in a natural way to hold all variants across targets and HCV genotype subtypes into one vertical resistance dataset. In addition, the vertical template has the variables PFSTRESC and PFRESALL for the variants in standard format and mixture format, which provides the valuable information for the resistance analysis. These analysis ready variables are not available in the previous horizontal resistance dataset.

ROWID	USUBJID	SUBTYPE	LBDT	VISIT	N5A0001	N5A0002	N5A0002A	N5A0002B
1					S	G		
2	1001	1A	03/02/2015	BL		V		
3	1001	1A	03/30/2015	W4			P	I/A
4	1001	1A	06/23/2015	PTW4			P	I/A
5	1001	1A	06/23/2015	POST BL ALL		V	P	I/A

ROWID	N5A0003	N5A0004	N5A0005
1	S	W	L
2		T	
3	T	T/Y	
4		T	
5	T	T/Y	

Table 5. Sample Population Resistance Dataset in the Horizontal Template

CONCLUSION

The updated HCV vertical resistance template has more advantages compared to the previous horizontal template, as discussed above. You can create the resistance tables and listings for the clinical virology analysis from the vertical resistance dataset much more easily than from the previous horizontal datasets. Currently the HCV vertical resistance template has not been finalized by FDA, but it is encouraged to

submit the HCV resistance data following the draft vertical template. Abbvie has submitted the new vertical template resistant dataset and it has been accepted by FDA.

REFERENCES

Forum HCV Drug Resistance Slideset (ResisSS 2013 V.1). http://www.hivforum.org/projects/hcv-drug-development-advisory-group/hcv-drug-resistance-slide-set

Food and Drug Administration, 2013. “Attachment to Guidance on Antiviral Product Development — Conducting and Submitting Virology Studies to the Agency – Guidance for Submitting HCV Resistance Data.” https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM340712.pdf

ACKNOWLEDGMENTS

The author thanks Sathasivaharan Thillaiampalam for his contributions and assistance with this project and manuscript.

CONTACT INFORMATION

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

Yan Xie
Abbvie Inc
847-935-6220
yan.xie@abbvie.com
www.abbvie.com

The support of this manuscript was provided by AbbVie. AbbVie participated in the review and approval of the content.

Yan Xie is an employee of AbbVie, Inc.

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.