PharmaSUG 2024 - Paper DS- 367

Handling of Humoral and Cellular Immunogenicity Data in SDTM

Wei Duan, Moderna Therapeutics, Cambridge, MA

ABSTRACT

The immunogenicity endpoints have been broadly used and examined in clinical vaccine studies as the key assessment of immune response to the viral infection. On one hand, humoral immunogenicity, including serum neutralizing and binding antibodies, has been migrated to IS domain per SDTMIG v3.4, as this domain is primarily used to host molecules targeting to antibody recognition, binding, and testing. Along the way of migrating, we faced some challenges when integrating various data sources, from vendor tested immunogenicity results, EDC sample collection data, to central lab data, with added complexity in data flows, discrepancies, and out-of-bound values. On the other hand, there is a trend that clinical vaccine studies pay more and more attention to cellular response, based on reading of CMI (Cell Mediated Immunity) data. Cytokines produced by monocytes, T cells, and lymphokines, play important roles in regulating immune functions and developing antiviral immune responses. This paper will discuss the proper domains for both types of immunogenicity data, along with some mapping challenges into CDISC-compliant and submission-ready SDTM datasets per our experience.

INTRODUCTION

Vaccine immunogenicity data provides strong proof of the magnitude and the duration of specific type of immune responses being induced by the study vaccine. The most common humoral immunogenicity includes the binding and neutralizing antibody assays. One of the main challenges during the mapping to IS domain is how to handle the reconciliations of various data sources and multiple data transfers within a single clinical study. The other challenges include the way that specialty lab handles out-of-bound assay results, the coded terms of various commonly used immunogenicity assays to be enforced at program level, as well as finding the appropriate SDTM host for cellular immunogenicity data including the CMI and flow cytometry datasets, driven by different analysis purposes. This paper may not be able to provide definite answers to the challenge calls but should be able to share some insights in SDTM mapping and considerations in trial management for both humoral and cellular immunogenicity data.

IMMUNOGENICITY DATA FLOW

In a clinical study, the third-party lab data might become available on the later end as it requires extra turnaround for the transportation and processing of the samples. Figure 1 illustrates the biomarker sample flow at high level: the samples were collected in multiple tubes at the site labelled for different analyses, then shipped to the central lab. Part of sample tubes stay in the central lab for the regular tests, and the rest are shipped further to specialty labs for different specialty tests.

At the moment of sample collection at the site, it is required to enter the sample collection date and time into EDC (see Figure 2 for an example of CRF page of Blood Collection of Humoral Immunogenicity), as well as the answer to the question whether this specific sample has been collected at the scheduled visit. A requisition header file containing basic sample information including subject ID, visit, lab name, test code, sample status, sample reference ID, specimen type, and sample collection date and time is also generated and attached to the sample shipment.

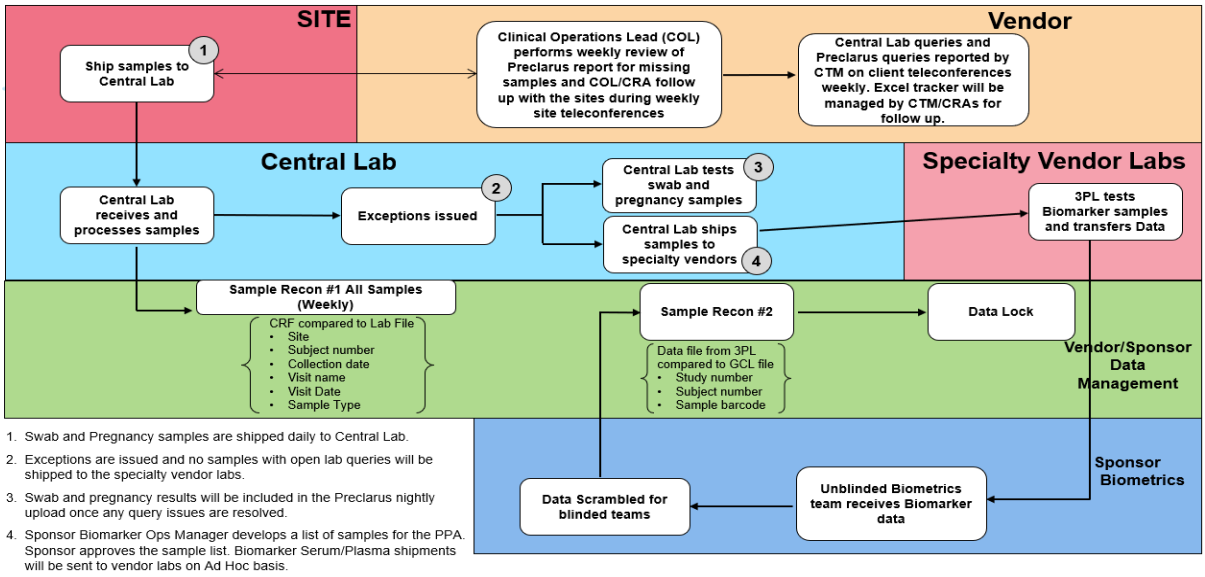

Figure 1. Data Flow for Immunogenicity Data

Form: Blood Collection for Humoral Immunogenicity

Generated On: 24 Aug 2023 20:20:40

Was the Immunogenicity sample collected?	Yes ☐ No ☐
Collection Date (dd MMM yyyy)	_____
Collection Time (24 hour clock)	_____
Collection Date/Time (derived)	_____

Figure 2. Sample Collection CRF Page

In the next step, central lab ships the samples to the corresponding specialty labs, where immunogenicity testing take place. Sometimes for certain specific assays, only a handful of labs in the world have the capability to process and analyze. To be noted, that not every lab has the capability to provide the actual result for out-of-bound values. It also happens quite often that different labs might use different approach when testing the immunogenicity samples, like egg-based testing versus cell-based testing. Essentially it has to do with how the virus is generated for use in an assay. For instance, some types of testing were performed in an egg-based assay where the lab would utilize the chicken eggs as the carrier to grow virus. On the other hand, a cell-based assay would utilize cell lines to grow virus.

As study programmers, ironing out these come-and-go in data flows is critical for correct SDTM dataset mapping. It helps to ensure each data source are integrated into the suggested domain and provide clues or the data checking points from programming perspective, which indicates what type of discrepancies would impact downstream analysis to a non-negligible extent.

For Figure 2, there are two layers of reconciliations to be performed by data management team (sometimes programmers will get involved to help perform these two checks as well). The first one is the reconciliation between the central lab header file and the EDC sample collection data. To do that, we start with filtering out the corresponding immunogenicity testing from the header file, then check through the missing records from one of the two sources, by inner joining the central lab header file and EDC by subject ID and collection date(time). We also compare the visits by merging subject ID and collection date(time). The second layer of reconciliation is to compare the central lab header file with the vendor lab results file. These two files are connected by using the sample barcode, which we also call accession

number. This is the key variable besides subject ID that programmers would use for merging the two files and find any discrepancies, such as the following scenario:

- Sample was collected but with missing results.
- Multiple non-missing results (ISORRES) per subject per visit per test
- Any result with both original testing result (ISORRES) and reason not done (ISREASND) missing.
- What are the non-numeric result values, e.g., > ULOQ, < LLOQ?

Another consideration on utilization of the result file, lies in the design and stage of the study. If programmers are blinded from viewing the actual data, they will generate dummy results file per the data structure outlined in the data transfer agreement. As for how to present out-of-bound values, and what are the validated ULOQ and LLOQ, these need advance planning and discussion with biomarker team and vendor labs. At certain milestone of the study, unblinded biostatisticians and programmers will receive transfer containing actual results from the vendor labs. At this point, the dummy file will no longer be in use for data refreshing. Instead, unblinded programmers will create the scrambled results data files based upon the actual results file, and then pass to blinded programmers to create SDTM on the blinded side.

HUMORAL IMMUNOGENICITY

Humoral immunogenicity, like B-cells and Antibody data, has been broadly used in vaccine clinical trials for comparing serum antibody responses in ways of geometrics mean titer, geometrics mean fold rise, seroconversion, and sero-response, with certain observational period after receiving the injection. Besides the similar challenges mentioned above in the data flow involving several layers of data checks for any potential discrepancies, there are other considerations in SDTM mapping stage, when integrating these multiple sources into Immunogenicity Specimen Assessments (IS) dataset.

Back to SDTM Implementation Guide (IG) v3.1.4, there was new domain Immunogenicity Specimen Assessments (IS) for representing the data pertaining to therapy-induced subject immune response. Since then, IS has been justified for the host of the immunogenicity data based on scientific definition. Per SDTM IG v3.4, it is newly suggested that IS being designed and reserved for representing data regarding specimen-based assessment that measures “presence, magnitude and scale of the immune response upon an antigen stimulation or encounter.” Hence, the scope of representing immunogenicity data has been expanded such that it is not only limited to humoral immunogenicity data. This will be iterated further in the next session.

Like lab test coding system, at program level, it is suggested having consistent code and decode list for ISTESTCD and ISTEST. Besides the header file, EDC blood sample collection of humoral immunogenicity data, and the lab results file, Subject Visits (SV) is also needed for mapping of IS data, especially when unscheduled visits are present. Sometimes for studies that have samples retested by other labs, we usually maintain same ISTESTCD/ISTEST across labs, but reserve Vendor Name (ISNAM) for lab name as the differentiation marker (Table 1). Also in this scenario, the sample barcode, which is mapped into reference ID (ISREFID), will differ as the new lab would use the back-up blood samples different from the original samples.

USUBJID	ISREFID	ISTESTCD	ISTEST	ISCAT	ISSCAT	ISORRES	ISORRE	ISNAM
1234-PXXX-ABXXX-0001	32NXXXXXX	TESTA	Influenza A Antibody	IMMUNOGENICITY ASSESSMENT - HUMORAL	IMMUNOGENICITY - INFLUENZA A	xxx	titer	VENDOR LAB B
1234-PXXX-ABXXX-0001	32NXXXXXX	TESTB	Influenza B Antibody	IMMUNOGENICITY ASSESSMENT - HUMORAL	IMMUNOGENICITY - INFLUENZA B	xxx	titer	VENDOR LAB A
1234-PXXX-ABXXX-0001	32NXXXXXX	TESTB	Influenza B Antibody	IMMUNOGENICITY ASSESSMENT - HUMORAL	IMMUNOGENICITY - INFLUENZA B	xxx	titer	VENDOR LAB B
1234-PXXX-ABXXX-0001	32NXXXXXX	TESTC	Influenza C Antibody	IMMUNOGENICITY ASSESSMENT - HUMORAL	IMMUNOGENICITY - INFLUENZA C	xxx	titer	VENDOR LAB A
1234-PXXX-ABXXX-0001	32NXXXXXX	TESTC	Influenza C Antibody	IMMUNOGENICITY ASSESSMENT - HUMORAL	IMMUNOGENICITY - INFLUENZA C	xxx	titer	VENDOR LAB B
1234-PXXX-ABXXX-0001	32NXXXXXX	TESTD	Influenza D Antibody	IMMUNOGENICITY ASSESSMENT - HUMORAL	IMMUNOGENICITY - INFLUENZA D	xxx	titer	VENDOR LAB B
1234-PXXX-ABXXX-0001	32NXXXXXX	TESTE	E Neutralizing Antibody	IMMUNOGENICITY ASSESSMENT - HUMORAL	SERUM XXX NEUTRALIZING ANTIBODIES	xxx	AU/mL	VENDOR LAB A

Table 1. Sample IS Data for Humoral Immunogenicity

It is also worth noting that there has been an ongoing debate on how to handle the testing results beyond the limits of quantifications. For instance, PsVNT, which is a mature assay and has been extensively used in many studies and has passed the validation, at analysis level, it is suggested to be capped if the value is beyond the upper limits of quantification (ULOQ). For the results going below the lower limits of quantification (LLOQ), at analysis stage if using antibody titer/concentration that would yield mean/median or similar summary statistics, values < LLOQ should be imputed by LLOQ/2; if percentage of achieving a threshold is needed, such as >=2 Fold-Rise, LLOQ value should be used, e.g. baseline/reference <LLOQ, then >=2 fold rise should count those with post-baseline >=2*LLOQ. For the details regarding how to handle out-of-bound values for immunogenicity testing results, this needs to be aligned among medical leads, biomarker team, and biostats/programming when developing the SAP for the study. It is also strongly driven by the testing capability from the vendor labs, as for whether they can provide the actual testing results when the values go beyond ULOQ and LLOQ.

The difference between the approach of imputation versus using a result beyond the ULOQ is not due to the origin of the virus, but due to differences in testing procedures between the CROs that conducted the testing for each of the assays. For some studies, testing will be conducted in-house, and we would make the imputation dependent on the way the assay will be set up. Some labs might have the capability on if the testing will allow for reporting of values above the ULOQ. The scientific justification for cell-based testing is that cell-grown viruses do better representing viruses relevant to human infection compared to egg-grown viruses.

CELLURLAR IMMUNOGENICITY

CELL-MEDIATED IMMUNITY

Cell-mediated immune responses play an important role in fighting against viral infections. For vaccine studies, it is often defined in some secondary or exploratory endpoints. Thus, this data point is no longer solely for research purpose and outside of the scope of SDTM mapping, and the summary of this data will be part of Clinical Study Report. CMI data includes various T-cell responses. The current IS domain from IG v3.4 is designed for representing specimen-based assessments that measure the presence, magnitude, and scale of the immune response upon any antigen stimulation or encounter. It extends the coverage to also include the detection and quantification of cellular immune response tests, such as antigen-stimulated and activated immune cells and their secreted products. With this new expansion in scope of this domain, the purpose of definition of IS domain is more in line with the scientific justification of the immunogenicity assessments, which also include cell-mediated immunity.

Cytokine-secreting cells and cytokine are main type of cellular immunogenicity data. They fundamentally differ from antibody (humoral) responses in the way they bring about infection control [1]. The immunity requires physical presence of reactive T-cells. Cytokine are proteins playing important roles in normal T-cell-mediated immunity. It is one of the major cellular immunogenicity data that uses non-flow cytometry techniques. Very limited labs have the capability measuring the cytokine production based on blood serum samples. Due to the lack of knowledge of SDTM standards from vendor labs, inputs are needed from sponsor programmers when working on the data transfer agreements, where the column headers of transferred outcome data are defined in detail. It's been in limbo that whether cellular immunogenicity data, like cytokine, or flow cytometry, is supposed to be mapped into CP, IS or LB. For cytokine measurement by specimen of Peripheral Blood Mononuclear Cell (PBMC), IS domain could house the data but with a separate category (ISCAT) reserved for 'CELL-MEDIATED IMMUNOGENICITY', for the analysis purpose of measuring the magnitude of immune response from antigen stimulation.

FLOW CYTOMETRY TESTING

Flow Cytometry testing is a laser-based technique used to detect and analyze the chemical and physical characteristics of cells or particles, by passing each cell through the laser beam. The properties measured by this technique include a cell's particle's relative size, relative granularity or internal complexity, and fluorescence intensity. Using this technique, Clinical trials could detect and compare the

receptor activity through fluorescently conjugated antibodies [2]. In clinical vaccine studies, it is broadly used for testing cellular immunogenicity, such as PBMC, with selected color panel and study population.

During the study, vendor labs will provide the flow cytometry data. As for whether the header file will also come along with the results file, it comes down to the study data transfer agreement. Sometimes the basic sample information, which separately collected outside of EDC, might be integrated into the central lab header file. For Flow Cytometry data, Cell Phenotyping (CP) domain would be the proper place presenting the data, which is for marker-based phenotyping data per IG v3.4. The main difference between the cytokine and flow cytometry mapping, for instance with the same specimen of PBMC, is how PBMC viability results are tested and reported. If that falls into the scope of marker-based cell phenotyping with analytes information provided, CP would be the appropriate place. For some proposed values to show how the specific analytes in the flow assay would map to the various test descriptor fields in the CP domain. CPMETHOD would be 'FLOW CYTOMETRY' and CPCAT would be 'CELL FUNCTION'.

The combination of values in CPTEST (Test Name), CPSBMRKS (Sublineage Marker String), CPCELSTA (Cell State) and CPCSMRKS (Cell State Marker String) are utilized to uniquely identify a test [3]. CPTEST is populated with the name of the cell type being measured, not with the set of markers used to define the cell type. CPMRKSTR (Marker String) is expected to be populated, since it contains the complete set of markers defining a test. Some sub-strings from CPMRKSTR are also present in CPSBMRKS and/or CPCSMRKS [3]. This order in CPMRKSTR depends on how a test is identified using the ordered combination of CPTEST, CPSMRKS, and CPCELSTA [3]. CPGATE (Gate) and CPGATDEF (Gate Definition) conveys gating information used in data collection and analysis. Moreover, if viability is explicitly stated in CPGATE, markers used to designate viability are included in CPGATDEF [3]. The order of markers within a string is generally proceeded in the order that defines the cell hierarchy from highest to lowest, followed by additional non-lineage-defining markers, and ending with cell state and viability markers.

CONCLUSION

As both humoral and cellular immunogenicity data have been brought to the front stage as clinical analysis endpoints, more questions and challenges can emerge along the mapping of different type of immunogenicity data to SDTM. The best approach is always sorting out the data flow and seeking agreement on both humoral and cellular immunogenicity data mapping with stakeholders way in advance. This paper could shed some light on some possible directions of data checking involving specialty labs, as well as mapping of certain types of biomarker immunogenicity data.

REFERENCES

[1] A.M. Zajac, L.E.Harrington “Immune Response to Viruses: Cell-Mediated Immunity” Encyclopedia of Virology (pp.70-77)

[2] Steven G Smith, Kaatie Smits, Simone A Joosten, Krista E van Meijgaarden, Iman Satti, Helen A Fletcher, Nadia Caccamo, Francesco Dieli, Francoise Mascart, Helen McShane, Hazel M Dockrell, Tom H M Ottenhoff, TBVI TB Biomarker Working Group Sep 14, 2015. “Intracellular Cytokine Staining and Flow Cytometry: Considerations for Application in Clinical Trials of Novel Tuberculosis Vaccines” PLoS ONE 10(9): e0138042. https://doi.org/10.1371/journal.pone.0138042

[3] CP Draft SDTMIG 4.0 – Highlight all proposed changes for tracking - CP Specification. https://wiki.cdisc.org/display/CT/CP+Specification

ACKNOWLEDGMENTS

I’d like to sincerely thank our statistical programming team, Baoyu Ding and Yanping Zhou for their valuable feedbacks and continuous support and guidance.

CONTACT INFORMATION

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

Wei Duan
Moderna Therapeutics
Wei.duan@Modernatx.com