PharmaSUG 2025 - Paper DS-129

ISI in Biopharmaceutical Development: Structure, Challenges, and Experience

Tong Songlin, Sanofi

ABSTRACT

With the widespread use of biopharmaceuticals such as therapeutic proteins and monoclonal antibodies, the immunogenicity risk has become a central concern in drug development. As a key document for systematically assessing the immunogenicity risk of biologics, the Integrated Summary of Immunogenicity (ISI) plays an irreplaceable role in the whole life cycle of drug development. Immunogenicity may not only lead to the production of antidrug antibodies (ADA), which may interfere with the pharmacokinetics (PK) and pharmacodynamics (PD) of drugs but also cause serious allergic reactions and adverse events. Taking the FDA guidelines as a framework and combining with practical project experience, this paper systematically combs through the core structure of ISI documentation, key data integration logic and typical challenges in implementation. This paper provides suggestions for ISI pooling project, optimize the efficiency of multiple studies pooled, and ultimately enhance the practical value of ISI in meeting regulatory requirements and ensuring drug safety.

INTRODUCTION

With the rapid development and widespread clinical application of therapeutic proteins and monoclonal antibodies, various diseases (including cancer and autoimmune diseases) have been better treated. However, along with the progress, people are increasingly aware of the potential immunogenicity risks associated with these biological products. Immunogenicity is defined as the ability of a substance that contains antigens to cause the body to make an immune response against that substance, which may result in the development of anti-drug antibodies (ADAs) [1]. These anti-drug antibodies will cause the body to get an immune response, which may affect the pharmacokinetic (PK) and pharmacodynamic (PD) characteristics of the drug in the body, thereby reducing efficacy or causing unpredictable adverse reactions (such as allergic reactions, cytokine release syndrome, etc.) [1][2].

To address these concerns, regulatory agencies such as the Food and Drug Administration (FDA) point out that conducting a systematic immunogenicity evaluation is crucial throughout the drug development process. The Immunogenicity Summary of Investigations (ISI) has become an important document in the submission package to regulatory agencies during drug approval. The ISI comprehensively integrates the immunogenicity data collected during clinical trials, describing the immunogenicity results of the biological pharmaceutical and its potential impact on safety and efficacy.

Although the ISI report is an important document in biopharmaceuticals, there are still many difficulties and challenges in the data preparation and integration process. For example, the ADA testing protocols may differ across multiple individual studies, inconsistent data collection methods lead to complex pooling of data, and poor communication between departments causes misunderstandings. This paper is based on the FDA and EMA guidelines and summarizes practical project experience to outline the key information required for the ISI document in the submission package. It shares some common challenges biologist-programmers encounter during ISI programming projects and offers some improvement ideas for the design and execution of ISI-related work. This work effort has three goals: first, to improve consistency in integrating data from different studies; second, to improve work efficiency during the ISI pooling process; and finally, to improve the quality of ISI submissions.

IMMUNOGENICITY DATASET OVERVIEW

Immunogenicity data are important for evaluating the safety and efficacy of biologics such as therapeutic proteins and monoclonal antibodies. Generally, immunogenicity datasets include the following basic types of data:

1. **Antibody Status:** Anti-drug antibodies (ADAs) are immunoglobulins generated in response to an

immune reaction against a therapeutic protein-based biologic. ADA status includes POSITIVE-sample and NEGATIVE-sample. In some special cases, there will be samples of Unevaluable-sample or Inconclusive-sample [3]. (Figure 1)

2. **Antibody Titer:** The dilution level of ADA positive sample which determined by serial dilution assays. It is defined as the reciprocal of the highest dilution of the sample that produces a positive result. The Titer is usually expressed as a power of 2, reflecting a serial two-fold dilution process. (Figure 1)
3. **Neutralizing Antibody Status:** This is a subclass of ADA that specifically recognizes and binds to the drug of interest, blocking its interaction with the target. This can significantly affect the drug's efficacy or induce immune-related adverse reactions. Nab status includes POSITIV-sample and NEGATIVE-sample.

USUBJID	PARAM	PARAMCD	AVAL	AVISIT	AVISITN
Patient-xxx-xxx	Anti-Drug 2 Antibody Titer	AD2AABT	0	VISIT 1	1
Patient-xxx-xxx	Anti-Drug 2 Antibody Titer	AD2AABT	0	VISIT 2	2
Patient-xxx-xxx	Anti-Drug 2 Antibody Titer	AD2AABT	0	VISIT 3	3
Patient-xxx-xxx	Anti-Drug 2 Antibody Titer	AD2AABT	200	VISIT 4	4
Patient-xxx-xxx	Anti-Drug 2 Antibody Titer	AD2AABT	800	VISIT 5	5
Patient-xxx-xxx	Anti-Drug 2 Antibody Titer	AD2AABT	400	VISIT 6	6

USUBJID	PARAM	PARAMCD	AVALC	AVISIT	AVISITN
Patient-xxx-xxx	Anti-Drug 2 Antibody	AD2AAB	NEGATIVE	VISIT 1	1
Patient-xxx-xxx	Anti-Drug 2 Antibody	AD2AAB	NEGATIVE	VISIT 2	2
Patient-xxx-xxx	Anti-Drug 2 Antibody	AD2AAB	NEGATIVE	VISIT 3	3
Patient-xxx-xxx	Anti-Drug 2 Antibody	AD2AAB	POSITIVE	VISIT 4	4
Patient-xxx-xxx	Anti-Drug 2 Antibody	AD2AAB	POSITIVE	VISIT 5	5
Patient-xxx-xxx	Anti-Drug 2 Antibody	AD2AAB	POSITIVE	VISIT 6	6

Figure 1ADA Status and Titer

The evaluation of ADAs follows a tiered approach. A screening test was first performed to detect the presence of drug antibodies in the serum. If the screening result is positive, the investigator conducts a secondary test to confirm the presence of antibodies; in this process, samples initially classified as positive may be reclassified as negative or indeterminate. Subsequently, the confirmed positive samples are subjected to semi-quantitative analysis, and the test result will be recorded as Titer. Finally, a neutralizing antibody assay is performed to assess whether the ADA interferes with the binding of the drug to its target. Please refer to Figure 2.

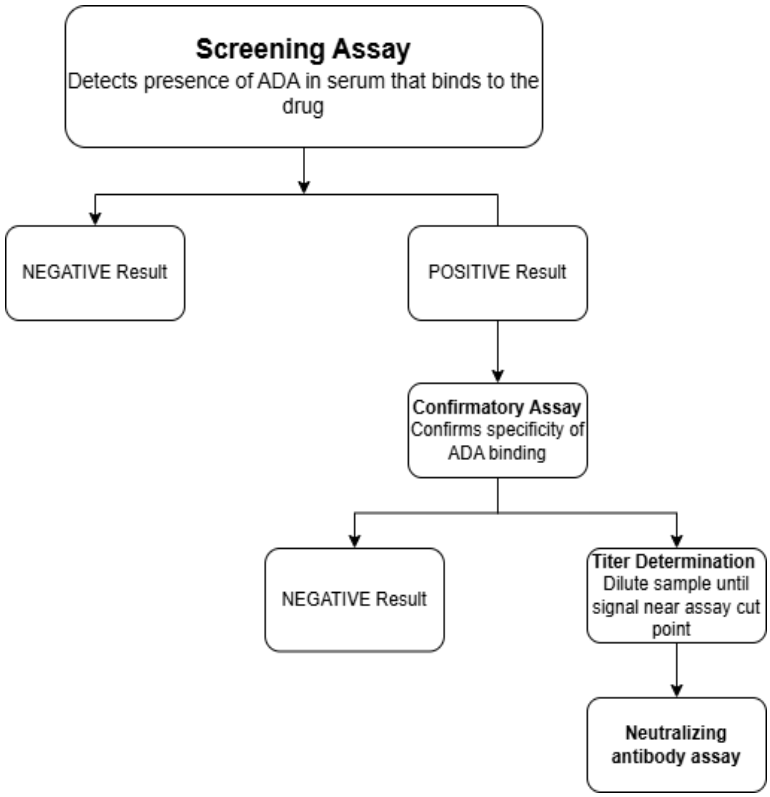

Figure 2 ADA testing workflow

MAIN CONTENT OF ISI DOCUMENTATION

The FDA's *Immunogenicity Assessment for Therapeutic Protein Products* (2014) and the EMA's *Guideline on Immunogenicity Assessment of Therapeutic Proteins* (2017) are guidelines that comprehensively describe the clinical consequences of immune responses to therapeutic protein products and provide recommendations for risk mitigation during clinical development. Both guidelines indicate that treating patients with therapeutic protein products may trigger varying immune responses, ranging from antibody reactions with no clinical manifestations to life-threatening reactions, such as anaphylaxis, cytokine release syndromes, or hypersensitivity reactions. These guidelines do not mandate the preparation of related documents at the time of drug submissions but suggest that pharmaceutical companies consider relevant factors during clinical trials and include key information in the reports when submitting to the FDA. The following information should be included:

Basic Analysis

- **ADA Titer and Status:** Descriptive statistics of ADA titer, including mean, maximum, minimum, peak titer, etc., should be summarized for each scheduled visit. Additionally, the number of participants with ADA status per visit should be reported.
- **ADA Incidence and Prevalence:** Report the frequency and distribution of anti-drug antibodies within each patient population.
- **Persistence and Clinical Impact:** Analyze the persistence of ADA (transient vs. persistent) and its impact on pharmacokinetics (PK), pharmacodynamics (PD), and safety.

Patient-Related Factors

- **Immune Status:** Include patient-specific factors such as HLA genotypic defects, pre-existing antibodies (Pre-Abs), etc.

- **Population Variability:** Evaluate immunogenicity differences among subgroups (e.g., pediatric, elderly, immunocompromised participants).

- **Dose-Response Relationships:** Investigate the association between ADA development and drug dosing regimens, administration routes, or treatment duration.

Impact on Efficacy

- **Pharmacological Inhibition:** Neutralizing antibodies (NAbs) may block drug efficacy by targeting critical functional domains. Therefore, the analytical figures or tables showing the association between NAbs and efficacy endpoints can be included in the report.

Safety Consequences (Adverse Events):

- **Immunogenicity-Related Safety Risks:** The immunogenicity in participants treated with therapeutic protein products can lead to unpredictable and highly variable adverse events. The primary safety concerns associated with anti-drug antibodies (ADAs) include acute and non-acute immune responses. Acute reactions may include anaphylaxis, cytokine release syndrome, and infusion-related events, typically presenting as systemic or localized symptoms during or shortly after drug administration. Non-acute reactions may involve hypersensitivity or immune complex effects, potentially leading to organ-specific inflammation or deposition-related disorders. Additionally, ADAs can cross-react with endogenous proteins, causing dysfunction of normal physiological functions and producing toxic side effects. Because all of these effects are related to ADAs, adverse event summaries should be listed by ADA status to provide information for safety assessment.

CHALLENGES AND RECOMMENDATIONS FOR ISI POOLING STRATEGY

Data pooling is critical to ensuring Immunogenicity Risk Assessment results are accurate and reliable. However, we may face many challenges while preparing ISI data. These challenges include but are not limited: ensuring all studies use the same MedDRA dictionary version, selecting the correct data and combining across different studies, and facilitating seamless collaboration between different teams. For example, inconsistent variable naming conventions for the same type of variables used in different studies may lead to misunderstandings during data pooling. Therefore, ongoing communication and close cooperation with programmers from each study project are very important to pool datasets efficiently. This section discusses key points during actual project implementation and proposes improvement suggestions based on project experience. By selecting suitable data structures and improving data quality at the individual study level, we can make ISI data integration more efficient and consistent, helping us complete high-quality ISI submissions.

POINT 1: USE THE SAME MEDDRA DICTIONARY VERSIONS ACROSS DIFFERENT STUDIES

One of the challenges in ISI analysis is the inconsistent MedDRA (Medical Dictionary for Regulatory Activities) versions used in different studies. Pooling analyses often combine data from multiple studies completed at different time points. The dictionary versions used in these studies will not be updated after the study completed, so inconsistencies in MedDRA versions often occur between different studies when pooling studies.

Different MedDRA versions may cause the same adverse event to be classified into different Preferred Terms (PTs) or System Organ Classes (SOCs). These inconsistencies can affect the accuracy and reliability of safety assessments and may mask important safety signals. The adverse event tables are the core of safety tables, especially when identifying and interpreting immunogenicity-related events. Therefore, we must unify the MedDRA version in individual studies when we combine AE data in the ISI pooling project.

To ensure the MedDRA versions are consistent, once the project team confirms that ISI analysis is required, the programmer should check the MedDRA versions in each study and decide which version will be used in the ISI report. Appropriate mapping strategies should be applied for studies that use a different MedDRA version to align the adverse event (AE) data with the selected dictionary version.

Medical coders must clearly document and verify all mappings to ensure traceability, scientific accuracy, and data integrity.

POINT 2: ENSURING CONSISTENCY OF ANALYSIS VISIT WINDOWS

In addition to using the same MedDRA version, it also needs to check whether the protocol-defined visit or analysis windows of immunogenicity data in different studies are consistent. Each study usually collects immunogenicity-related data based on its protocol, which may lead to differences in the frequency of data collection. Without a unified time window to define the analysis visit, it becomes very difficult to perform meaningful summary analysis.

For example, Study 1 may collect ADA data at baseline, week 2, week 4, and week 6, while Study 2 collects data more frequently, including baseline, week 2, week 3, and week 4. Combining the datasets in these two studies without unifying the analysis visit windows may lead to fluctuations in the number of participants displayed in descriptive statistical tables or graphs. This will affect the accuracy and interpretability of immunogenicity trend analysis.

The primary strategy for dealing with this risk is to define a unified analysis visit window for data pooling. Two primary approaches can be considered:

1. **Redefine a Visit Window Across Studies:**

   Redefine a standard analysis visit window in ISI to remap all individual studies.

2. **Select Primary Visits for Analysis by the Medical Team:**

   The medical or statistical team can select a subset of clinically relevant time points for the pooled analysis.

In practice, statistical teams can collaborate with medical teams to establish unified analysis visit windows, ensuring that the selected time windows are appropriate for trend analysis. All decisions regarding analysis visit window selection and definition should be documented, including their rationale, to support transparency and regulatory review.

POINT 3: PAY ATTENTION TO THE ADA DATA COLLECTION METHOD UPDATE

The duration of a clinical study may span four to eight years or even longer. Therefore, with the protocol version updated, the method for collecting and recording anti-drug antibody data may change at specific time points. Suppose programmers do not notice the method update or forget to update the programs on time. In that case, it may lead to errors in deriving some ADaM variables and subsequent outputs.

For example, some studies may record an ADA sample status as two separate records based on the initial protocol version, such as: Record 1 = Positive-Screen sample; Record 2 = Negative/Positive-Confirm sample. The new protocol may only require recording the confirmed status, such as Record 1 = Negative. The data recording process can also impact ADA data. For example, in the old data spec, if a patient’s ADA status is NEGATIVE at a time point, the corresponding ADA titer value does not need to be recorded in the dataset. Some years later, the new data spec may require that even for the samples with a NEGATIVE ADA status, the ADA Titer should still be recorded as 0 in the dataset.

This inconsistency may cause challenges during the ISI pooling process, as different programmers may have different habits in selecting data when summarizing some results. If the data collection method changes and programmers do not update the programs on time, it may cause errors in the final summarized results. Therefore, it is crucial to closely monitor whether data collection methods change during the individual study. Ideally, individual studies' data presentation in the ADaM datasets should be consistent before being included in the ISI pooling effort.

If data standardization cannot be achieved across different studies for operational or technical reasons, individual study programmers must proactively inform the ISI team of these differences. Providing detailed documentation regarding the data structure changes, including justifications, mapping logic, etc., will help ISI programmers accurately use the data during pooling and analysis.

POINT 4: UTILIZING THE ADADA DATASET IN INDIVIDUAL STUDIES FOR EFFICIENT ISI POOLING

In addition to defining consistency analysis windows and using the same dictionary versions, selecting suitable data structures at the study level is important for facilitating efficient ISI pooling. It is strongly recommended that the ADIS (Anti-Drug Antibody Analysis Data – Observed) (Figure 3) and ADADA (Anti-Drug Antibody Analysis Data - Derived) (Figure 4) datasets be created at the individual study level to organize immunogenicity data better. In general, the immunogenicity data can be categorized into two main types:

- **Observed Values Collected by Visit:**

  These records include raw assay results obtained directly from the laboratory at the Protocol-defined VISIT, such as ADA status or titer values collected at each VISIT. These records are best stored in the ADIS dataset, one record per participants per parameter per analysis visit per analysis timepoint.

- **Derived or Summary Data:**

  This type includes statistical results derived from lab observations, such as peak titer, mean titer, time to first positive result, or ADA response results for each patient. These results are more appropriately saved in the ADADA dataset, one record per participants per parameter per period.

Figure 3 Summary Variables in ADIS

USUBJID	PARAMCD	AAPERIDC	AVISIT	ADAPKG	ADPPKG
Patient-xxx-xxx	AD2AABT	PAP	VISIT 0	Peak titer 1600-6400	Peak titer 400-800
Patient-xxx-xxx	AD2AABT	PAP	VISIT 1	Peak titer 1600-6400	Peak titer 400-800
Patient-xxx-xxx	AD2AABT	PAP	VISIT 2	Peak titer 1600-6400	Peak titer 400-800
Patient-xxx-xxx	AD2AABT	PAP	VISIT 3	Peak titer 1600-6400	Peak titer 400-800
Patient-xxx-xxx	AD2AABT	ETP	VISIT 4	Peak titer 1600-6400	Peak titer 1600-6400
Patient-xxx-xxx	AD2AABT	ETP	VISIT 5	Peak titer 1600-6400	Peak titer 1600-6400
Patient-xxx-xxx	AD2AABT	ETP	VISIT 6	Peak titer 1600-6400	Peak titer 1600-6400

Figure 4 Summary Variables in ADADA

USUBJID	AAPERIDC	PARAMCD	PARAM	AVALC
Patient-xxx-xxx	ETP	ADPPKG	Peak titer range group by period	Peak titer 1600-6400
Patient-xxx-xxx	Overall	ADAPKG	Peak titer range group by Overall	Peak titer 1600-6400
Patient-xxx-xxx	PAP	ADPPKG	Peak titer range group by period	Peak titer 400-800

Risk with Only Using the ADIS Dataset

Some individual studies select complex trial designs, such as a multiple-period or crossover trial design, that may include multiple treatment periods in a study. If both VISIT-observed results and the summarized results are placed in the ADIS dataset, we may face the following problems:

- **Increased Number of Variables:** In multiple-period studies, new variables need to be created for each summary result corresponding to each period, leading to a times increase in the number of variables in the ADIS dataset.
- **Data Redundancy:** Summary results for the same participant and period may be saved in multiple records, causing data duplication and making the presentation unfriendly.
- **Difficulty in Variable Interpretation:** Even within an individual study, if the trial spans several years,

different programmers may be responsible for the ADIS programming. Differences in programming styles and habits may lead to inconsistent variable naming and representation, requiring updating relevant information time by time during program handover.

- **Challenges in Multiple-Study Pooling:** The ISI report may only select specific phases from each study (e.g., only the ETP phase of Study 1 and all phases of Study 2). If the variable design in ADIS is too complex, ISI programmers—especially those unfamiliar with the individual study design—may face difficulties when combining data, potentially leading to increased workload or errors.

Advantages of Using the ADADA Dataset

To address these challenges, placing subject-level and period-level derived results in the ADADA dataset and only retaining ADA observations for each VISIT in the ADIS dataset is a more reasonable and practical approach. This separation ensures:

- **Clearer Data Structure and Purpose:** Each dataset has a well-defined role, the ADIS for observed data and ADADA for derived summaries data. This facilitates easier interpretation and use.
- **Improved Interpretation for Programmers and Statisticians:** Separating the observed value and derived value makes the datasets and variables easier to understand and use, which minimizes the risk of misinterpretation during the programming and summarization process.
- **Greater and Flexibility in ISI Pooling Strategies:** Programmers can easily select and combine data from specific periods or overall summaries without duplication or confusion.

Fortunately, a paper introducing the concept of the ADADA dataset was also presented at the PharmaSUG US 2025. This paper provides valuable guidance and practical reference for data scientists, statistical programmers, and regulatory submission teams involved in immunogenicity assessment. For more detailed information on designing and implementing an ADADA dataset, please refer to the referenced paper in the PharmaSUG 2025 conference proceedings (PharmaSUG 2025 - Paper PO-052 [6]).

POINT 5: STANDARDIZATION OF PARAM / PARAMCD ACROSS STUDIES

ISI data pooling also faces challenges: not all studies use the same PARAM and PARAMCD for the same laboratory test. For example, one study might use the "ADA Status" as PARAM and "ADST" as PARAMCD to represent the assay result. While another study might use "Anti-Drug Antibody Result" as PARAM and "ADRES" as PARAMCD.

In ISI pooled analyses, immunogenicity data come from multiple studies to evaluate trends such as ADA incidence, titer development, or neutralizing activity. Inconsistent PARAMCD values or parameter labels across studies for the same scientific endpoint can cause data pooling, interpretation, and use difficulties by both ADaM and output programmers. So, standardization of PARAM/PARAMCD prior to the dataset is essential to support efficient and accurate analysis.

- Standardizing PARAMCD and PARAM values across all studies in ISI pooling is recommended, not only for ADA data, but also for other parameters used in the ISI analysis plan. If inconsistent PARAMCD values were used in historical studies, the programmer should create a transparent and traceable mapping specification document. This document should include:
- Identification of parameters across studies
- Define the mapping rules between old PARAMCD values and standardized PARAMCD values
- Reviewed and approved by the project team (such as Biostatistics, Programming, Medical team)
- The unit conversion rules should also be documented if different units are used for parameters in the study.

POINT 6: COMMUNICATION AND DOCUMENT CRITICAL INFORMATION ACROSS TEAMS

Finally, beyond technical considerations, a successful ISI pooling project depends on effective collaboration and communication among all relevant study teams. Effective communication is critical to successfully pooling immunogenicity data in ISI work. Immunogenicity analysis involves multiple teams, including the clinical team, data management, statistical programming, a Biostatistician, and regulatory teams. Timely communication ensures:

Standardized Pooling Strategies

Immunogenicity data from different studies often vary in structure, format, terminology, etc. Clear communication ensures all team members understand these differences and agree on pooling strategies.

Accurate Notification of Team Decisions

Decisions include which periods to include in ISI reporting, how to define response criteria, how to handle missing data, and others. These decisions can impact programming, analysis, and regulatory reporting, and they must be communicated clearly and documented.

Timely Issue Resolution

During the data pooling or review process, misunderstandings or inconsistencies may arise due to documentation gaps or historical reasons. Timely communication among programmers, statisticians, and the medical team is important to resolve issues quickly.

Document traceability

Regulatory agencies expect transparency and traceability in data and outputs. Any significant decisions impacting the data or output should be appropriately documented.

CONLUSION

Throughout the entire lifecycle of biopharmaceutical development, the Immunogenicity Summary and Interpretation (ISI) plays an important role in comprehensively assessing and monitoring immunogenicity risks. With therapeutic proteins and monoclonal antibodies becoming increasingly important in modern medicine, a systematic assessment of their potential to elicit an immune response and the possible consequences is essential to ensure drug efficacy and patient safety.

This paper briefly introduces the ISI report's main components and discusses some common challenges encountered during the project process. It aims to improve the efficiency and quality of ISI submissions. Inconsistencies in MedDRA versions across studies, differences in analysis visit windows, the lack of structured ADADA datasets, and non-standardized use of PARAMCD can significantly impact the reliability of pooling data. To address these issues, early in the development process, MedDRA versions should be coordinated and unified on time, and standard analysis visit windows should be selected. At the same time, using ADIS and ADADA datasets helps to clearly distinguish between observed ADA results and derived ADA results, reduce redundancy, and improve data clarity. Finally, standardizing parameter naming conventions across studies enables more effective data integration and improves the quality of the pooled datasets.

Based on FDA and EMA guidelines and project experience, proactive planning and unified strategies are essential in the ISI project submission process. Biopharmaceutical developers can significantly improve the quality of ISI submissions by adopting practices such as structured dataset design. These efforts increase the transparency and scientific rigor of the data and help expedite the regulatory review process.

REFERENCES

[1] FDA. "Guidance for Industry: Immunogenicity Assessment for Therapeutic Protein Products." U.S. Food and Drug Administration. 2014. Available at https://www.fda.gov/media/85017/download

[2] EMA. "Guideline on Immunogenicity assessment of therapeutic proteins." European Medicines Agency. 2017. Available at https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-immunogenicity-assessment-therapeutic-proteins-revision-1_en.pdf

[3] Immunogenicity Testing of Therapeutic Protein Products - Developing and Validating Assays for Anti-Drug Antibody Detection. FDA, 2019, Available at https://www.fda.gov/media/119788/download

[4] Assessment and Reporting of the Clinical Immunogenicity of Therapeutic Proteins and Peptides - Harmonized Terminology and Tactical Recommendations, G Shankar, S Arkin, et al. 2014, Available at https://pubmed.ncbi.nlm.nih.gov/24764037/

[5] Anti-drug Antibody Sample Testing and Reporting Harmonization, Darshana Jani, et al. 2022, https://pubmed.ncbi.nlm.nih.gov/36307592/

[6] ADaM Implementation for ADA Data, Jiannan Kang, Merck & Co., Inc., Rahway, NJ, USA; Luke Reinbolt, Navitas Data Sciences Inc., 2025, Available at https://www.lexjansen.com/pharmasug/2025/PO/PharmaSUG-2025-PO-052.pdf

ACKNOWLEDGMENTS

The author would like to thank Fan Zhang (Sanofi) and Fangrun He (Sanofi) for their advice and support.

CONTACT INFORMATION

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

Author Name: Tong Songlin
Company: Sanofi
E-Mail: Songlin.Tong@Sanofi.com