Assay Development and Validation for Immunogenicity Testing of Therapeutic Protein Products

Guidance for Industry

DRAFT GUIDANCE

This guidance document is being distributed for comment purposes only.

Comments and suggestions regarding this draft document should be submitted within 60 days of publication in the *Federal Register* of the notice announcing the availability of the draft guidance. Submit electronic comments to http://www.regulations.gov. Submit written comments to the Division of Dockets Management (HFA-305), Food and Drug Administration, 5630 Fishers Lane, rm. 1061, Rockville, MD 20852. All comments should be identified with the docket number listed in the notice of availability that publishes in the *Federal Register*.

For questions regarding this draft document, contact (CDER) Susan Kirshner at 301-827-1731; (CBER) Office of Communication, Outreach and Development, 800-835-4709 or 240-402-8010; or (CDRH) Office of Communication and Education, 800-638-2041 or 301-796-7100.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)
Center for Devices and Radiological Health (CDRH)

April 2016
Pharmaceutical Quality/CMC
Revision 1

Assay Development and Validation for Immunogenicity Testing of Therapeutic Protein Products

Guidance for Industry

Additional copies are available from:

Office of Communications, Division of Drug Information
Center for Drug Evaluation and Research
Food and Drug Administration
10001 New Hampshire Ave., Hillandale Bldg., 4th Floor
Silver Spring, MD 20993-0002
Phone: 855-543-3784 or 301-796-3400; Fax: 301-431-6353
Email: druginfo@fda.hhs.gov
http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm
and/or
Office of Communication, Outreach and Development
Center for Biologics Evaluation and Research
Food and Drug Administration
10903 New Hampshire Ave., Bldg. 71, Room 3128
Silver Spring, MD 20993-0002
Phone: 800-835-4709 or 240-402-8010
Email: ocod@fda.hhs.gov
http://www.fda.gov/BiologicsBloodVaccines/GuidanceComplianceRegulatoryInformation/Guidances/default.htm
and/or
Office of Communication and Education
CDRH-Division of Industry and Consumer Education
Center for Devices and Radiological Health
Food and Drug Administration
10903 New Hampshire Ave., Bldg. 66, Room 4621
Silver Spring, MD 20993-0002
Phone: 800-638-2041 or 301-796-7100; Fax: 301-847-8149
Email: DICE@fda.hhs.gov
http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/default.htm

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)
Center for Devices and Radiological Health (CDRH)

April 2016
Pharmaceutical Quality/CMC
Revision 1

Contains Nonbinding Recommendations

Draft — Not for Implementation

TABLE OF CONTENTS

Assay Development and Validation for Immunogenicity Testing of Therapeutic Protein Products Guidance for Industry[1]

This draft guidance, when finalized, will represent the current thinking of the Food and Drug Administration (FDA or Agency) on this topic. It does not establish any rights for any person and is not binding on FDA or the public. You can use an alternative approach if it satisfies the requirements of the applicable statutes and regulations. To discuss an alternative approach, contact the FDA office responsible for this guidance as listed on the title page.

I. INTRODUCTION

This guidance provides recommendations to facilitate industry's development and validation of immune assays for assessment of the immunogenicity of therapeutic protein products during clinical trials. Specifically, this document includes guidance regarding the development and validation of screening assays, confirmatory assays, titering assays, and neutralization assays.[2,3] For the purposes of this guidance, immunogenicity is defined as the propensity of the therapeutic protein product to generate immune responses to itself and to related proteins or to induce immunologically related adverse clinical events. The recommendations for assay development and validation provided in this document apply to assays for detection of anti-drug antibody(ies) (ADA).[4] This guidance may also apply to some combination products on a case-by-case basis.[5]

[1] This guidance has been prepared by the Office of Medical Policy in the Center for Drug Evaluation and Research in cooperation with the Center for Biologics Evaluation and Research and the Center for Devices and Radiological Health at the Food and Drug Administration.

[2] This document specifically does not discuss the development or validation of anti-drug antibody(ies) (ADA) assays for animal studies; however, some concepts discussed are relevant to the design of ADA studies for nonclinical testing. Refer to the International Conference on Harmonisation (ICH) guidance for industry *S6(R1) Preclinical Safety Evaluation of Biotechnology-Derived Pharmaceuticals* for more information regarding immunogenicity assessments in animal toxicology studies. Also see the guidance for industry *Immunogenicity Assessment for Therapeutic Protein Products*, where the topic "Utility of Animal Studies" is covered in more detail. We update guidances periodically. To make sure you have the most recent version of a guidance, check the FDA guidance Web page at http://www.fda.gov/RegulatoryInformation/Guidances/default.htm.

[3] For information on clinical immunogenicity assessment of proposed biosimilar biological products, see the guidance for industry *Scientific Considerations in Demonstrating Biosimilarity to a Reference Product*.

[4] This guidance does not pertain to immunogenicity assays for assessment of immune response to preventative and therapeutic vaccines for infectious disease indications.

[5] General information on combination products is available at http://www.fda.gov/CombinationProducts/default.htm.

This document does not discuss the product and patient risk factors that may contribute to immunogenicity.[6] This guidance, including any discussions of terminology used in this guidance, does not apply to in vitro diagnostic products.[7] This guidance revises the draft guidance for industry *Assay Development for Immunogenicity Testing of Therapeutic Proteins* issued in December 2009. The information in this guidance has been reorganized for clarity and includes new information on titering and confirmatory assays.

In general, FDA's guidance documents do not establish legally enforceable responsibilities. Instead, guidances describe the Agency's current thinking on a topic and should be viewed only as recommendations, unless specific regulatory or statutory requirements are cited. The use of the word *should* in Agency guidances means that something is suggested or recommended, but not required.

II. BACKGROUND

Patient immune responses to therapeutic protein products have the potential to affect product safety and efficacy.[8] The clinical effects of patient immune responses are highly variable, ranging from no effect at all to extremely harmful effects to patient health. Detection and analysis of ADA formation is a helpful tool in understanding potential patient immune responses. Information on immune responses observed during clinical trials, particularly the incidence of ADA induction and the implications of ADA responses for therapeutic protein product safety and efficacy, is crucial for any therapeutic protein product development program. Accordingly, such information, if applicable, should be included in the prescribing information as a subsection of the ADVERSE REACTIONS section entitled "Immunogenicity." Therefore, the development of valid, sensitive, specific, and selective assays to measure ADA responses is a key aspect of therapeutic protein product development.

III. GENERAL PRINCIPLES

The risk to patients of mounting an immune response to a therapeutic protein product will vary with the product. FDA recommends adoption of a risk-based approach to evaluating and mitigating immune responses to or immunologically related adverse clinical events associated

---

[6] See the guidance for industry *Immunogenicity Assessment for Therapeutic Protein Products*, where these topics are covered in more detail.

[7] Per 21 CFR 809.3(a), "in vitro diagnostic products are those reagents, instruments, and systems intended for use in the diagnosis of disease or other conditions, including a determination of the state of health, in order to cure, mitigate, treat, or prevent disease or its sequelae. Such products are intended for use in the collection, preparation, and examination of specimens taken from the human body. These products are devices as defined in section 201(h) of the Federal Food, Drug, and Cosmetic Act (the act), and may also be biological products subject to section 351 of the Public Health Service Act."

[8] See the guidance for industry *Immunogenicity Assessment for Therapeutic Protein Products*.

with therapeutic protein products that affect their safety and efficacy.[9] Immune responses may have multiple effects, including neutralizing activity and the ability to induce hypersensitivity responses. Immunogenicity tests should be designed to detect ADA that could mediate unwanted biological or physiological consequences.

Screening assays, also known as binding antibody (BAb) assays, are used to detect all antibodies that bind to the therapeutic protein product. The specificity of BAb for the therapeutic protein product is established using confirmatory assays. ADA are further characterized using titering and neutralization assays. Titering assays are used to characterize the magnitude of the ADA response. It is important to characterize this magnitude with titering assays because the impact of ADA on safety and efficacy may correlate with ADA titer and persistence rather than incidence (Cohen and Rivera 2010). Neutralization assays assess the ability of ADA to interfere with the therapeutic protein product-target interactions. Therefore, neutralizing antibodies (NAb) are a subset of BAb. It is important to characterize neutralizing activity of ADA with neutralization assays because the impact of ADA on safety and efficacy may correlate with NAb activity rather than ADA incidence (Calabresi, Giovannoni, et al. 2007; Goodin, Frohman, et al. 2007; Cohen and Rivera 2010). Similarly, it may be important in some cases to establish NAb titers. Additional characterization assays, such as isotyping, epitope mapping, and assessing cross-reactivity, e.g., to endogenous counterparts or to other products, may be useful.

The optimal time to design, develop, and validate ADA assays during therapeutic protein product development depends on the risk assessment of the product (Mire-Sluis, Barrett, et al. 2004; Gupta, Indelicato, et al. 2007; Shankar, Devanarayan, et al. 2008; Gupta, Devanarayan, et al. 2011). The sponsor should provide a rationale for the immunogenicity testing paradigm, preferably at the investigational new drug application (IND) stage, during phase 1. Because ADA assays are critical when immunogenicity poses a high clinical risk (e.g., assessment of a therapeutic protein product with a non-redundant endogenous counterpart) and real-time data concerning patient responses are needed, the sponsor should implement preliminary validated assays early, before and during phase 1, and obtain data in real time. Real-time assessments entail analyses of the samples as soon as possible after sampling, before banking of the samples, and prior to additional dosing when the dosing regimen allows. In lower risk situations, the sponsor may bank patient samples so they can be tested when suitable assays are available. FDA encourages sponsors to test samples during phase 1 and phase 2 studies using suitable assays. Samples derived from pivotal studies should be tested with fully validated assays. At the time of license application, the sponsor should provide data supporting full validation of the assays. Recommendations regarding the timing of ADA sample collection can be found in section VII.A.[10]

[9] See the guidance for industry *Immunogenicity Assessment for Therapeutic Protein Products*.

[10] See the guidance for industry *Immunogenicity Assessment for Therapeutic Protein Products*, where immunogenicity risk assessment and mitigation considerations are covered in more detail. Guidance on appropriate assay development and validation for immunogenicity testing is also available in the ICH guidances for industry *Q2A Text on Validation of Analytical Procedures* and *Q2B Validation of Analytical Procedures: Methodology*.

Assays for detection of ADA facilitate understanding of the immunogenicity, safety, and efficacy of therapeutic protein products. However, the detection of ADA is dependent on key operating parameters of the assays (e.g., sensitivity, specificity), which vary between assays.[11] Although information on ADA incidence is typically included in the prescribing information under an "Immunogenicity" subsection of the ADVERSE REACTIONS section, FDA cautions that comparison of ADA incidence among products, even for products that share sequence or structural homology, can be misleading. This is because detection of ADA formation is highly dependent on the sensitivity and specificity of the assay. Additionally, the observed incidence of ADA (including NAb) positivity in an assay may be influenced by factors such as method, sample handling, timing of sample collection, concomitant medications, and disease condition. Therefore, comparing immunogenicity rates among therapeutic protein products with structural homology for the same indication is unsound, even though fully validated assays are employed. When a true comparison of immunogenicity across different therapeutic protein products that have homology is needed, it should be obtained by conducting a head-to-head clinical study using a standardized assay under the same conditions that has equivalent sensitivity and specificity for both therapeutic protein products.[12]

The recommendations on assay development and validation provided in this guidance are based on common issues encountered by the Agency upon review of immunogenicity submissions. Sponsors should contact FDA for any product-specific guidance. Isotyping and cross-reactivity assay designs should be discussed with FDA. Other publications may also be consulted for additional insight (see Mire-Sluis, Barrett, et al. 2004; Gupta, Indelicato, et al. 2007; Shankar, Devanarayan, et al. 2008; Gupta, Devanarayan, et al. 2011). In general, FDA recommends that sponsors develop assays that are optimized for sensitivity, specificity, selectivity, precision, reproducibility, and robustness (see sections IV.C through G).

IV. ASSAY DESIGN ELEMENTS

This section applies to all types of assays for detection of ADA, unless specified otherwise.

A. Testing Strategy

1. Multi-Tiered Testing Approach

FDA recommends a multi-tiered ADA testing approach because of the size of some clinical trials and the necessity of testing patient samples at several time points. In this paradigm, a rapid, sensitive screening assay is initially used to assess clinical samples. The initial screening assay should be sensitive to low levels of low- and high-affinity ADA (see section V.A). Samples testing positive in the screening assay are then subjected to a confirmatory assay to demonstrate

[11] See the United States Pharmacopeia (USP) General Chapter 1106 *Immunogenicity Assays – Design and Validation of Immunoassays to Detect Anti-Drug Antibodies* for a broader discussion of various assay types.

[12] For information on proposed biosimilar products, see the guidance for industry *Scientific Considerations in Demonstrating Biosimilarity to a Reference Product*.

that ADA are specific for the therapeutic protein product. For example, a competition assay could confirm that antibody is specifically binding to the therapeutic protein product and that the positive finding in the screening assay is not a result of non-specific interactions of the test serum or detection reagent with other materials in the assay milieu such as plastic or other proteins.

Samples identified as positive in the confirmatory assay should be further characterized in other assays, such as titering and neutralization assays. In some cases, assays to detect cross-reactivity to other proteins with homology, such as the corresponding endogenous protein, may be needed. Further, tests to assess the isotype of the antibodies and their epitope specificity may also be recommended once samples containing antibodies are confirmed as positive.

2. Immunoglobulin Isotypes

The initial screening assay should be able to detect all relevant immunoglobulin (Ig) isotypes. For non-mucosal routes of administration, and in the absence of anaphylaxis, the expected ADA isotypes are IgM and IgG. For mucosal routes of administration, IgA isotype ADA are also expected. Although FDA expects that all relevant isotypes be detected in screening assays, it is not necessary that the screening assay establish which isotypes are being detected. For example, assays using the bridging format may provide no information on which isotypes are being detected. Bridging assay format can theoretically detect antibodies of most isotypes, but may not detect IgG4 isotypes. In some circumstances the sponsor should develop assays that discriminate between antibody isotypes. For example, for therapeutic protein products where the risk for anaphylaxis is a concern, antigen-specific IgE assays should be developed. In addition, the generation of IgG4 antibodies has been associated with immune responses generated under conditions of chronic antigen exposure, such as with factor VIII treatment, and in erythropoietin-treated patients with pure red cell aplasia (Matsumoto, Shima, et al. 2001; Aalberse and Schuurman 2002). Consequently, depending on the clinical concern, assessing for specific isotypes may be needed.

3. Epitope Specificity

FDA recommends that the sponsor direct initial screening tests against the whole therapeutic protein product and, when relevant, its endogenous counterpart. For some therapeutic protein products, the sponsor may need to investigate the ADA to specific epitopes to which immune responses are specifically generated. For example, determination of epitope specificity is recommended for some fusion molecules because the region where the two molecules join may form a neoantigen, and immune responses to this region may arise. Because of epitope spreading, immune responses to other parts of the molecule may ensue, leading to the generation of antibodies to the therapeutic protein product or its endogenous counterpart (Prummer 1997; Miller, Korn, et al. 1999; Disis, Goodell, et al. 2004; Thrasyvoulides, Liakata, et al. 2007; van der Woude, Rantapaa-Dahlqvist, et al. 2010; Hintermann, Holdener, et al. 2011). For these therapeutic protein products, FDA encourages sponsors to investigate the initiating event in the immune cascade. This knowledge may allow for modification to the protein to reduce its potential immunogenicity. Similarly, for therapeutic protein products with modifications, such as PEGylation, sponsors should develop assays to determine the specificity of ADA for the

protein component as well as the modification to the therapeutic protein product. Also see sections IV.K.4 and 5.

B. Assay Cut Point

The cut point of the assay is the level of response of the assay that defines the sample response as positive or negative. Information specific to establishing the cut point for the respective assay types is provided in sections V and VI. Establishing the appropriate cut point is critical to ensuring acceptable assay sensitivity.

The cut point of the assay can be influenced by a myriad of interfering factors, such as pre-existing antibodies, rheumatoid factor (RF), human anti-mouse antibodies, and the levels of product-related material or homologous proteins in the matrix. These factors should be considered early on in assay development when defining the cut point. Because samples from different target populations and disease states may have components that can cause the background signal from the assay to vary, different cut points may be needed for discrete populations being studied.

The cut point should be statistically determined using samples from treatment-naïve subjects.[13] By performing replicate assay runs with these samples, the variability of the assay can be estimated. During assay development, a small number of samples may be used to estimate the cut point. This may be done with as few as 5–10 samples from treatment-naïve subjects.

The specific approach employed to determine the cut point will depend on various factors. Specifically, because the cut point should identify any samples that produce a signal beyond that of the variability of the assay, the sponsor should consider the impact of statistically determined outlier values as well as true-positive samples when establishing the cut point. The sponsor should provide justification for the removal of any data points, along with the respective method used to determine their status as outliers. Positive values and samples may derive from non-specific serum factors or the presence of pre-existing antibodies in patient samples (Ross, Hansen, et al. 1990; Turano, Balsari, et al. 1992; Coutinho, Kazatchkine, et al. 1995; Caruso and Turano 1997; van der Meide and Schellekens 1997; Boes 2000). Although pre-existing antibodies to a variety of endogenous proteins are present in healthy individuals, these can be much higher in some disease states. The sponsor should identify those samples with pre-existing antibodies, for example, through immunodepletion approaches, and remove them from the cut point analysis. If the presence of pre-existing antibodies is a confounding factor, it may be necessary to assign positive responses or a cut point based on the difference between individual patient results before and after exposure. It is possible to arrive at a reasonable value to define assay cut point through careful design consideration, such as utilizing the minimal required dilution (MRD) of the sample, removing statistical outliers from analyses, minimizing the impact

[13] Treatment-naïve subjects could be healthy individuals or a patient population not exposed to therapeutic protein product, depending on the stage of assay development or validation and on the availability of samples. Sponsors should provide justification for the appropriateness of the samples used.

of interfering factors, improving assay drug tolerance, and using an approach to account for pre-existing antibodies.

C. Sensitivity

1. Assay Sensitivity

The sponsor should determine the sensitivity of the assay to have confidence when reporting immunogenicity rates. Assay sensitivity represents the lowest concentration at which the antibody preparation consistently produces either a positive result or readout equal to the cut point determined for that particular assay.[14] FDA recommends that screening and confirmatory ADA assays achieve a sensitivity of at least 100 nanograms per milliliter (ng/mL). Although traditionally FDA has recommended sensitivity of at least 250–500 ng/mL, recent data suggest that concentrations as low as 100 ng/mL may be associated with clinical events (Plotkin 2010; Zhou, Hoofring, et al. 2013). However, it is understood that neutralization assays may not always achieve that level of sensitivity.

The assays should have sufficient sensitivity to enable detection of low levels of ADA before the amount of ADA reaches levels that can be associated with altered pharmacokinetic, pharmacodynamic, safety, or efficacy profiles. Because assessment of patient antibody levels will occur in the presence of biological matrix, testing of assay sensitivity should be performed with the relevant dilution of the same biological matrix (e.g., serum or plasma, with the same anticoagulant as the diluent, from the target population). The final sensitivity should be expressed as mass of antibody detectable/mL of undiluted matrix. Therefore, assay sensitivity should be reported after factoring in the MRD. Assay sensitivity should not be reported as titer. During development, sensitivity should be assessed using both individual as well as pooled samples from treatment-naïve subjects so that the suitability of the negative control can be established.

Assay sensitivity should be determined by testing serial dilutions of a positive control antibody of known concentration in pooled negative control matrix. The dilution series should be no greater than two- or threefold, and a minimum of five dilutions should be tested. Alternatively, sensitivity can be calculated by interpolating the linear portion of the dilution curve to the assay cut point. As noted previously, assay sensitivity should be reported in mass units per volume of undiluted matrix.

A purified preparation of antibodies specific to the therapeutic protein product should be used to determine the sensitivity of the assay so that assay sensitivity can be reported in mass units/mL of matrix. Antibodies used to assess sensitivity can take the form of affinity purified polyclonal preparations or monoclonal antibodies (mAb).

A low positive system suitability control containing a concentration of ADA slightly above the sensitivity of the assay should be used to ensure that the sensitivity of the assay is consistent

[14] See the USP General Chapter 1106 *Immunogenicity Assays – Design and Validation of Immunoassays to Detect Anti-Drug Antibodies* for a discussion on *Relative Sensitivity*.

across assay runs. The low positive system suitability control should be designed to fail in 1% of the runs (see section IV.I.1).

2. Drug Tolerance

Therapeutic protein product or the endogenous counterpart present in the serum may interfere with the sensitivity of the assay. Specifically, complexes formed between ADA and the therapeutic protein product, also called ADA-drug complexes, that prevent detection of ADA in the test format can form if product-related materials are present in the test sample. This is because ADA assays are generally designed to detect uncomplexed ADA. The assessment of assay sensitivity in the presence of the expected levels of interfering therapeutic protein product, also known as the assay's drug tolerance, is critical to understanding the suitability of the method for detecting ADA in dosed patients.[15] FDA recommends that the sponsor examine assay drug tolerance early in assay development. The sponsor may examine drug tolerance by deliberately adding different known amounts of purified ADA into individual ADA-negative control samples in the absence or presence of different quantities of the therapeutic protein product under consideration and determining quantitatively whether the therapeutic protein product interferes with ADA detection. Results obtained in the absence and presence of different quantities of the therapeutic protein product under consideration should be compared. There should be a relationship between the quantity of antibody and the amount of therapeutic protein product required for a specified degree of inhibition. Data from pharmacokinetic studies may be useful in establishing optimal sample collection times. Acid dissociation pretreatment or other approaches may be used to disrupt circulating ADA-drug complexes, which may lead to increased assay drug tolerance. Interference from the therapeutic protein product can be minimized if the sponsor collects patient samples at a time when the therapeutic protein product has decayed to a level where it does not interfere with assay results.

D. Specificity and Selectivity

Demonstrating assay specificity and selectivity is critical to the interpretation of immunogenicity assay results. Specificity refers to the ability of a method to detect ADA that bind the therapeutic protein product but not assay components such as surfaces or reagents. The assays should exclusively detect the target analyte, in this case the ADA.[16] The selectivity of an ADA assay is its ability to identify therapeutic protein product-specific ADA in a matrix such as serum or plasma that may contain potential interfering substances. Assay results may be affected by interference from the matrix or from on-board therapeutic protein product.[17] Lack of assay specificity or selectivity can lead to false-positive results, which could obscure relationships between ADA response and clinical safety and efficacy measures. Demonstrating the specificity

[15] See the USP General Chapter 1106 *Immunogenicity Assays – Design and Validation of Immunoassays to Detect Anti-Drug Antibodies*.

[16] See the USP General Chapter 1106 *Immunogenicity Assays – Design and Validation of Immunoassays to Detect Anti-Drug Antibodies*.

[17] See the USP General Chapter 1106 *Immunogenicity Assays – Design and Validation of Immunoassays to Detect Anti-Drug Antibodies*.

and selectivity of antibody responses to mAb, Fc-fusion protein, and Ig-fusion proteins poses particular challenges because of the high concentration of Ig in human serum. The sponsor should clearly demonstrate that the assay method specifically detects anti-mAb and not the mAb product itself, non-specific endogenous antibodies, or antibody reagents used in the assay. Similarly, for patient populations with a high incidence of RF, the sponsor should demonstrate that RF does not interfere with the detection method. Host cell proteins and other product-related impurities may interfere with demonstrating the assay specificity and selectivity as well.

A straightforward approach to addressing specificity and selectivity is to demonstrate that binding can be blocked by soluble or unlabeled purified therapeutic protein product. One approach is to incubate positive and negative control antibody samples with the purified therapeutic protein product or its components under consideration. Inhibition of signal in the presence of the relevant therapeutic protein product or its components demonstrates that the response is specific and selective. For responses to mAb products, inclusion of another mAb with the same Fc but different variable region can be critical. For responses to other proteins, an unrelated protein of similar size and charge can be used. If the assay is specific and selective for the protein in question, generally the addition of that protein in solution should reduce the response to background or the cut point, whereas the addition of an unrelated protein of similar size and charge should have no effect. Conversely, addition of the protein in question should have little effect on antibodies specific to an unrelated protein. Selectivity should further be evaluated by performing recovery studies, in which positive control antibodies are spiked into matrix at defined concentrations, and the positive control antibody signal is compared to that obtained from antibody spiked into assay buffer alone.

1. Matrix Interference

An important consideration is how interference from the assay matrix, which is composed of the sample and the diluent, can affect assay performance. Components in the matrix other than therapeutic protein product can interfere with assay results. For example, different anticoagulants used during sample collection may have different effects in the assay, potentially affecting the assay sensitivity and linearity. Sponsors should evaluate different salt anticoagulant sample collection solutions for their effect on assay results.

Endogenous and exogenous components in serum or plasma may influence assay results, and it is usually necessary to dilute patient samples for testing to minimize such effects. The sponsor should examine the effect of such interferents by performing spike-and-recovery studies. The sponsor should define the dilution factor that will be used for preparation of patient samples before performing validation studies assessing potential interference of this matrix on assay results (see section IV.D.2 on MRD).

Buffer components that are chemically related to the therapeutic protein product may also interfere in the assay. For example, polysorbate is chemically similar to polyethylene glycol (PEG) and therefore may interfere in the detection of anti-PEG antibodies. The chemical composition of the buffer should be carefully considered during assay development.

The sponsor may examine matrix interference by spiking different known amounts of purified ADA into the assay buffer in the absence or presence of different matrix components. Comparing the recovery of ADA in buffer alone with that in the matrix can provide input on the degree of interference from matrix components. Furthermore, such analysis may guide decisions on the MRD recommended for sample testing. In addition, the sponsor should examine other parameters affecting patient samples, such as hemolysis, lipemia, presence of bilirubin, and presence of concomitant medications that a patient population may be using. Samples that have very high antibody titers may need additional testing, such as with different dilutions of the competing product in the confirmatory assay, to ensure their identification.

2. Minimal Required Dilution

Matrix components can contribute to non-specific signal if undiluted, thereby obscuring positive results. Therefore, there is frequently a need to dilute patient samples to maintain a reasonable ability to detect ADA (sensitivity). Ideally, the MRD is the sample dilution that yields a signal close to that of the assay diluent and allows for the highest signal-to-noise ratio. MRD typically ranges from 1:5 to 1:100.

FDA recommends that the sponsor determine the MRD from a panel of appropriate number of samples from treatment-naïve subjects. Determination of MRD usually involves serially diluting treatment-naïve ADA-negative samples, as well as testing known amounts of purified antibody (at high, medium, and low concentrations) in serially diluted matrix in comparison to the same amount of antibody in buffer. This ensures a reasonable signal-to-noise ratio throughout the range of the assay. The MRD should be calculated using at least 10 individual serum samples; the appropriate number of samples will depend on various factors, including the variability of the individual samples.

Although the MRD ultimately selected by the sponsor will depend on the assay design and patient population, FDA recommends that dilutions not exceed 1:100. Higher dilution may result in the spurious identification of a negative response when patients may actually possess low levels of therapeutic protein product-specific antibodies, the occurrence of which can be related to significantly altered pharmacokinetics, pharmacodynamics, safety, or efficacy profiles. However, in some instances greater initial dilutions may be required, and the overall effect of such dilutions on assay sensitivity and immunogenicity risk assessment should be considered.

E. Precision

Precision is a measure of the variability in a series of measurements for the same material run in a method. Results should be reproducible within and between assay runs to assure adequate precision.[18] Demonstrating assay precision is critical to the assessment of ADA because assay variability is the basis for determining the cut points and ensuring that low positive samples are

[18] For more information on precision, see the guidance for industry *Bioanalytical Method Validation.* Also see the USP General Chapter 1106 *Immunogenicity Assays – Design and Validation of Immunoassays to Detect Anti-Drug Antibodies.*

detected as positive. To provide reliable estimates, the sponsor should evaluate both intra-assay (repeatability) and inter-assay (intermediate precision) variability of assay responses.

F. Reproducibility

Reproducibility is an important consideration if an assay will be run by two or more independent laboratories during a study, and a sponsor should establish the comparability of the data produced by each laboratory.[19] In addition, the assays should have the same precision between different laboratories under the established assay operating conditions (for example, using the same instrument platform).

G. Robustness and Sample Stability

Assay robustness is an indication of the assay's reliability during normal usage[20] and is assessed by the capacity of the assay to remain unaffected by small but deliberate variations in method and instrument performance that would be expected under relevant, real-life circumstances in routine laboratory practice. For example, changes in temperature, incubation times, or buffer characteristics, such as pH and salt concentration, can all impact assay results. The complexity of bioassays makes them particularly susceptible to variations in assay conditions, and it is essential to evaluate and optimize parameters such as cell passage number, incubation times, and culture media components. The sponsor should examine robustness during the development phase, and if small changes in specific steps in the assay affect results, specific precautions should be taken to control their variability. FDA recommends storing patient samples in a manner that preserves antibody reactivity at the time of testing. FDA recommends that the sponsor avoid freeze-thaw cycles because freezing and thawing patient samples may also affect assay results. However, studies evaluating long-term stability of positive control antibodies may be useful.[21]

H. Selection of Format

A number of different assay formats and instrumentation are available that can be employed for detection of ADA. These include, but are not limited to, direct binding assays, bridging assays, and equilibrium binding assays. Each assay format has advantages and disadvantages, including rapidity of throughput, sensitivity, selectivity, dynamic range, ability to detect various Ig isotypes, ability to detect rapidly dissociating antibodies, and availability of reagents. One of the major differences between each of these assay formats is the number and vigor of washes, which

[19] For more information on reproducibility, see the guidance for industry *Bioanalytical Method Validation*. Also see the USP General Chapter 1106 *Immunogenicity Assays – Design and Validation of Immunoassays to Detect Anti-Drug Antibodies,* the USP General Chapter 1225 *Validation of Compendial Procedures*, and the ICH guidance for industry *Q2B Validation of Analytical Procedures: Methodology*.

[20] For more information on robustness, see the ICH guidance for industry *Q2B Validation of Analytical Procedures: Methodology*. Also see the USP General Chapter 1106 *Immunogenicity Assays – Design and Validation of Immunoassays to Detect Anti-Drug Antibodies*.

[21] For more information on stability studies, see the guidance for industry *Bioanalytical Method Validation*.

can have an effect on assay sensitivity. All assays should be evaluated for their ability to detect rapidly dissociating antibodies such as IgM, which are common in early immune responses. Failure to detect such antibodies in early immune responses to therapeutic protein products may result in under-detection of true-positive antibody samples. Epitope exposure is also important to consider because binding to plastic or coupling to other agents, such as reporters (i.e., fluorochromes, enzymes, or biotin), can result in conformational changes of the antigen that can obscure, expose, modify, or destroy relevant antibody binding sites on the therapeutic protein product in question.

I. Selection of Reagents

Many components of the assays for ADA detection may be standard or obtained from commercial sources, for example, commercially available reagents such as Protein A/G coated resins used in the depletion approach for confirmatory assays. Other components, however, including positive control antibodies, negative controls, and system suitability controls, may need to be generated specifically for the particular assay.

1. Development of Positive Control Antibodies

Sponsors may use different or the same positive control antibodies to establish and monitor system suitability during routine assessment of assay performance, as well as to determine that the assay employed is fit for purpose. For system suitability controls, a positive control antibody, either mono- or polyclonal, used at concentrations adjusted to control the cut point and dynamic range levels, may be suitable.

Positive control antibodies frequently are generated by immunizing animals in the absence or presence of adjuvants. FDA recommends that positive control antibodies generated by immunizing animals be affinity purified using the therapeutic protein product. This approach enriches the polyclonal antibody preparation for ADA, which enables a more accurate interpretation of sensitivity assessment results. The selection of animal species when generating positive control antibodies should be carefully considered. For example, if an anti-human Ig reagent will be used as a secondary reagent to detect patient antibodies, the positive control antibodies and quality control (QC) samples should be detectable by that same reagent. When the positive control antibody is not detectable by that same reagent, an additional secondary reagent to detect the positive control antibody may be needed. In those cases, an additional positive control antibody for the secondary reagent used to detect human antibodies should be implemented to ensure that the reagent performs as expected. In some instances, the sponsor may be able to generate a positive control antibody from patient samples.[22] Although such antibodies can be very valuable, such samples are generally not available in early trials. Alternatively, individual mAb or panels of mAb may be used for positive control antibodies. Sponsors should discuss with FDA alternative approaches to assay development and validation in the rare event that a sponsor is not able to generate a positive control antibody.

[22] Proper informed consent from patients is needed and should be planned ahead of time.

Ideally, the positive control antibody used to determine assay applicability for the purpose of the respective assay should reflect the anticipated immune response that will occur in humans. For therapeutic mAb, the sponsor should give special consideration to the selection of a positive control antibody for the assay. When animals are immunized with a chimeric, humanized, or human mAb to develop a positive control antibody, the humoral response may be against the human Fc and not the variable region of the molecule. Such positive control antibodies may not be relevant for the anticipated immune response in patients where the response is primarily directed to the antigen-binding regions.

Once a source of a positive control antibody has been identified, the sponsor should use that source to assess assay performance characteristics such as sensitivity, selectivity, specificity, and reproducibility. FDA recommends that sponsors generate and reserve positive control antibody solution for use as a quality or system suitability control. For assay development and validation, dilutions should be representative of a high, medium, and low value in the assay. This is needed even for qualitative assays to understand whether assay performance is acceptable across a broad range of antibody concentrations. Although high- and low-value QC samples should be used, medium-value QC samples for detection of ADA are generally not needed for monitoring system suitability during routine assessment of assay performance.

2. Development of Negative Controls

For negative control samples, it is recommended that when possible, the control population should have the same disease condition. The control samples should represent a similar gender, age, and concomitant medications so that the sample matrix is representative of the study population. Similarly, control samples should be collected and handled in the same manner as study samples with respect to, for example, type of anticoagulant used, sample volume, and sample preparation and storage, because these pre-analytical variables can impact the performance of control samples in the assay. It is frequently the case that such control samples are not available for use during development or pre-study validation exercises. In those situations, it is acceptable to use purchased samples or samples from healthy donors, but important parameters of assay performance such as cut point, sensitivity, and selectivity should be confirmed when samples from treatment-naïve subjects from the appropriate target population become available.

FDA recommends that the sponsor establish a negative control for validation studies and patient sample testing. In this regard, a pool of sera from an appropriate number of treatment-naïve subjects can serve as a useful negative control. Importantly, the value obtained for the negative control should be below but close to the cut point determined for the assay in the patient population being tested. Negative controls that yield values far below the mean value derived from individual serum samples used to establish the cut point may not be useful in ensuring proper assay performance.

3. Detection Reagent Consideration

The selection of a suitable detection reagent (i.e., reporter) depends on the assay format chosen. It is critical to minimize the non-specific signal from the detection reagent. The detection

reagent chosen should have the adequate sensitivity required for the particular assay. These factors should be taken into consideration when deciding on the detection reagent.

4. Controlling Non-Specific Binding

Every reagent, from the plastic of the microtiter plates to the developing agent, can affect assay sensitivity and non-specific binding. One of the most critical elements is the selection of the proper assay buffer and blocking reagents used to prevent non-specific binding to the solid surface. The sponsor should carefully consider the number and timing of wash steps as well as the detergents added to the assay buffer (i.e., blocking or wash buffer) to reduce background noise, but still maintain sensitivity. A variety of proteins can be used as blocking reagents to provide acceptable signal-to-noise ratio. However, these proteins may not all perform equivalently in specific immunoassays. For example, they may not bind well to all types of solid phases or may show unexpected cross-reactivity with the detecting reagent. Therefore, the sponsor may need to test several blocking agents to optimize assay performance. Moreover, including uncoated wells is insufficient to assess non-specific binding. Rather, determining the capacity of ADA to bind to an unrelated protein of similar size and charge that may be present in the sample may prove to be a better test of binding specificity.

J. Reporting Results for Qualitative and Semi-Quantitative Assays

Several approaches may be used to report positive antibody responses, and the appropriateness of the approach used should be evaluated on a case-by-case basis. The most common approach is qualitative, with patients reported as having a positive or negative antibody response.

For patients who are confirmed to be ADA positive, determining antibody levels can be informative because it allows for the stratified assessment of ADA levels and their impact on safety and efficacy. These relationships may not be elucidated unless ADA levels are determined. Positive antibody responses may be reported as a titer (e.g., the reciprocal of the highest dilution that gives a readout at or just above the cut point of the assay), when appropriate. The MRD should be factored in the calculations of titers and provided when reporting titers. Reporting levels of antibodies in terms of titers is appropriate and generally understood by the medical community. Values may also be reported as amount of mass units of therapeutic protein product neutralized per volume serum with the caveat that these are arbitrary in vitro assay units and cannot be used to directly assess therapeutic protein product availability in vivo.

Unless the assay method used allows for independent determination of mass, antibody levels reported in mass units are generally not acceptable because they are based on interpolation of data from standard curves generated with a positive control antibody, and parallelism between the reference standard and test article cannot be assumed. Thus, FDA does not consider it necessary nor desirable for the sponsor to report patient antibody results in terms of mass units unless (1) the results are determined by quantitative means or (2) a universally accepted and accessible source of validated antibody is available as a control and parallelism between the dilution curves of the control antibody and patient samples has been demonstrated. Furthermore, even if parallelism is demonstrated, because the reference standard and test articles are likely to

contain different populations of antibodies, the absolute mass units cannot be calculated. Therefore, FDA understands that the mass units reported are relative rather than absolute values.

K. Other Considerations for Assay Development

A myriad of factors can affect the assessment of antibody levels, such as patient sample variability, therapeutic protein product-dose response of the cells used to generate the standard curve in a cell-based neutralization bioassay, affinity and avidity of the ADA, and concentration of competing product in confirmatory assays. Accounting for such factors is important to understand and analyze assay variability and avoid errors. Common factors that should be considered include the following:

1. Pre-Existing Antibodies

A growing body of evidence in the medical literature suggests that B-cells and T-cells with specificity for a number of self-proteins exist naturally and may even be heightened in some disease states, such as in patients subjected to cytokine therapy or suffering from a variety of immunological or immunoinflammatory diseases (Coutinho, Kazatchkine, et al. 1995; van der Meide and Schellekens 1997; Boes 2000). For example, antibodies to interferon can be found in normal individuals (Ross, Hansen, et al. 1990; Turano, Balsari, et al. 1992; Caruso and Turano 1997). Less surprisingly, subjects may have pre-existing antibodies to foreign antigens, such as bacterial products, most likely as a result of exposure to the organism or cross-reactivity. Pre-existing antibodies may have clinical effects and may affect the efficacy of the therapeutic protein product being tested. An alternative to the qualitative screening assay approach may be needed to assess the quantity and quality of ADA when pre-existing antibodies are present. For example, testing samples for an increase in ADA using a semi-quantitative assay type such as a titering assay (see sections V.C and VI.D) can provide information on the impact of a therapeutic protein product on product immunogenicity that is not provided by a qualitative assay.

2. Rheumatoid Factor

Measuring immune responses to therapeutic protein products that possess Ig tails, such as mAb and Fc-fusion proteins, may be particularly difficult when RF is present in serum or plasma. RF is generally an IgM antibody that recognizes IgG, although other Ig specificities have been noted. Consequently, RF will bind Fc regions, making it appear that specific antibody to the therapeutic protein product exists. Several approaches for minimizing interference from RF have proven useful, including treatment with aspartame (Ramsland, Movafagh, et al. 1999) and careful optimization of reagent concentrations so as to reduce background binding. When examining immune responses to Fc-fusion proteins in clinical settings where RF is present, FDA recommends developing an assay specific for the non-Fc region of the proteins.

3. Monoclonal Antibodies

Some special considerations pertain to the detection of antibodies against mAb. Animal-derived mAb, particularly those of rodent origin, are expected to be immunogenic with the immune response directed against the whole mAb molecule. In the early days of the therapeutic mAb

industry, this was a key reason for the failure of clinical trials (Kuus-Reichel, Grauer, et al. 1994).

Technologies reducing the presence of non-human sequences in mAb, such as chimerization and humanization, have led to a dramatic reduction but not elimination of immunogenicity. In these cases, the immune responses are directed largely against the variable regions of the mAb (Harding, Stickler, et al. 2010; van Schouwenburg, Kruithof, et al. 2014). As immune responses against the variable regions of human mAb are anticipated, FDA does not expect that the use of human mAb will further reduce immunogenicity by a significant margin. The assays that can detect the reactivity against variable regions are considered more appropriate to evaluate the potential impact of antibodies against mAb-based therapeutics in patients. However, engineering of Fc portion (e.g., modification of the levels of afucosylation) in human antibodies may affect immunogenicity. Many of these concerns also pertain to Fc-fusion proteins containing a human Fc region.

4. Conjugated Proteins

Because antibody-drug conjugates (ADCs) are antibodies conjugated with small molecule drugs, they represent a classic hapten-carrier molecule. Therefore, the immunogenicity assays should be able to measure the responses to all components of the ADC therapeutic protein product, including the antibody, linker-drug, and new epitopes that may result from conjugation. When ADCs need to be labeled for immunogenicity assays, the conjugation should be performed carefully because ADCs are already modified. The potential for increased hydrophobicity of the labeled molecules may cause aggregation, and therefore the stability and solubility of these capture reagents should be adequately characterized.

5. Products With Multiple Functional Domains

Some proteins possess multiple domains that function in different ways to mediate clinical efficacy. An immune response to one domain may inhibit a specific function while leaving others intact. Examination of immune responses to therapeutic protein products with multiple functional domains may require development of multiple assays to measure immune responses to different domains of the molecules.

V. ASSAY DEVELOPMENT

Information specific to development of respective assay types is provided in sections A through D below. These sections supplement information relevant to all assay types provided in section IV.

A. Development of Screening Assay

Based on the multi-tiered approach discussed previously in section IV.A, the first assay to be employed for detection of ADA should be a highly sensitive screening assay that detects low- and high-affinity ADA. Approximately 10 individual samples may be used to estimate the cut

point early in assay development; however, this may need to be adjusted when treatment-naïve samples from the target population become available. A low but defined false-positive rate is desirable for the initial screening assay because it maximizes detection of true positives. Subsequent assays can be employed to exclude false-positive results when determining the true incidence of immunogenicity.

B. Development of Confirmatory Assay

Because the screening assay is designed to broadly detect the presence of antibodies that bind product in serum samples with a defined false-positive rate, FDA recommends that the sponsor develop assays to confirm the binding of antibodies that are specific to the therapeutic protein product. Implementation of a suitable confirmatory assay is important to prevent data on ADA false-positive patients from confounding the analyses of the impact of ADA on safety and efficacy.

1. Selection of Format for Confirmatory Assay

It is expected that the selected confirmatory assay will be at least as sensitive as the screening assay but have higher specificity and at least as good selectivity in order to identify any false-positive samples. The method and instrument platform selected may be similar to or different from those used for the screening assay. Frequently, both screening and confirmatory assays use the same method and instrument platform. In such cases, the sensitivity of each assay will need to be determined in mass units and confirmed using system suitability controls to ensure that the assay is sensitive to the presence of binding antibody. When using a binding competition assay, the concentration of competing product should be optimized to confirm the presence of antibodies throughout and above the range of the assay.

2. Cut Point of Confirmatory Assay

If a competitive inhibition format is selected, a recommended approach to determining the cut point uses the data from the binding of antibody-negative treatment-naïve patient samples in the presence of the competitor, which is usually the therapeutic protein product. In this case, the amount of therapeutic protein product used to establish the cut point should be the same as the amount of therapeutic protein product that will be used as a competitive inhibitor in the assay. However, this approach may not be appropriate when dealing with samples where pre-existing antibodies are present in the treatment-naïve population. In those cases, the sponsor should exclude true positives from the cut point assessment. In rare cases when baseline negative samples are not available, sponsors may evaluate changes in titer or use an orthogonal method to confirm samples that screen positive.

C. Development of Titering Assay

1. Titer Determination

Titers are defined as the maximal dilution where a sample gives a value above the screening cut point. Titers are often informative and can be linked to clinical impact of the ADA. Titering

assays can be particularly informative when patients have pre-existing antibodies. Titering assays most often are performed using the same platform as the screening assay. Sera are tested in sequential dilutions. Alternatively, titer may be determined by extrapolating the dilution to the assay cut point using the linear portion of the dose response curve.

2. Cut Point of Titering Assay

When patients have pre-existing ADA, treatment-boosted ADA responses may be identified by post-treatment increases in titer. A cut point for defining the treatment-emergent or boosted responses is needed. Frequently this cut point is determined as a titer that is two dilution steps greater than the pre-treatment titer, when twofold dilutions are used to determine the titer. If titer is established by extrapolating the dilution curve to the assay cut point, treatment-emergent responses may be determined using estimates of assay variability.

D. Development of Neutralization Assay

In vitro neutralization assays provide an indication of the potential of the ADA to inhibit the biological activity of the product. Such NAb can interfere with the clinical activity of a therapeutic protein product by preventing the product from reaching its target or by interfering with receptor-ligand interactions. The testing method selected to assess neutralizing potential for ADA-positive samples should be based on the mechanism of action of the therapeutic protein product.

1. Selection of Format for Neutralization Assay

Two formats of assays have been used to measure NAb activity: cell-based bioassays and non-cell-based competitive ligand-binding assays. Selection of the appropriate assay format depends on various factors. These factors include, but are not limited to, the mechanism of action of the therapeutic protein product, its ability to reflect the in vivo situation most closely, and the selectivity, sensitivity, precision, and robustness of the assay. FDA recommends that neutralization assays use a cell-based bioassay format depending on the therapeutic protein product's mechanism of action because, frequently, cell-based bioassays more closely reflect the in vivo situation and therefore provide more relevant information than ligand-binding assays. Because the cell-based bioassays are often based on the product's potency, historically the format of these assays has been extremely variable. The choice and design of potency bioassays are generally based on a cell line's ability to respond to the product in question and the potency bioassay's relevance to the therapeutic protein product's mechanism of action.

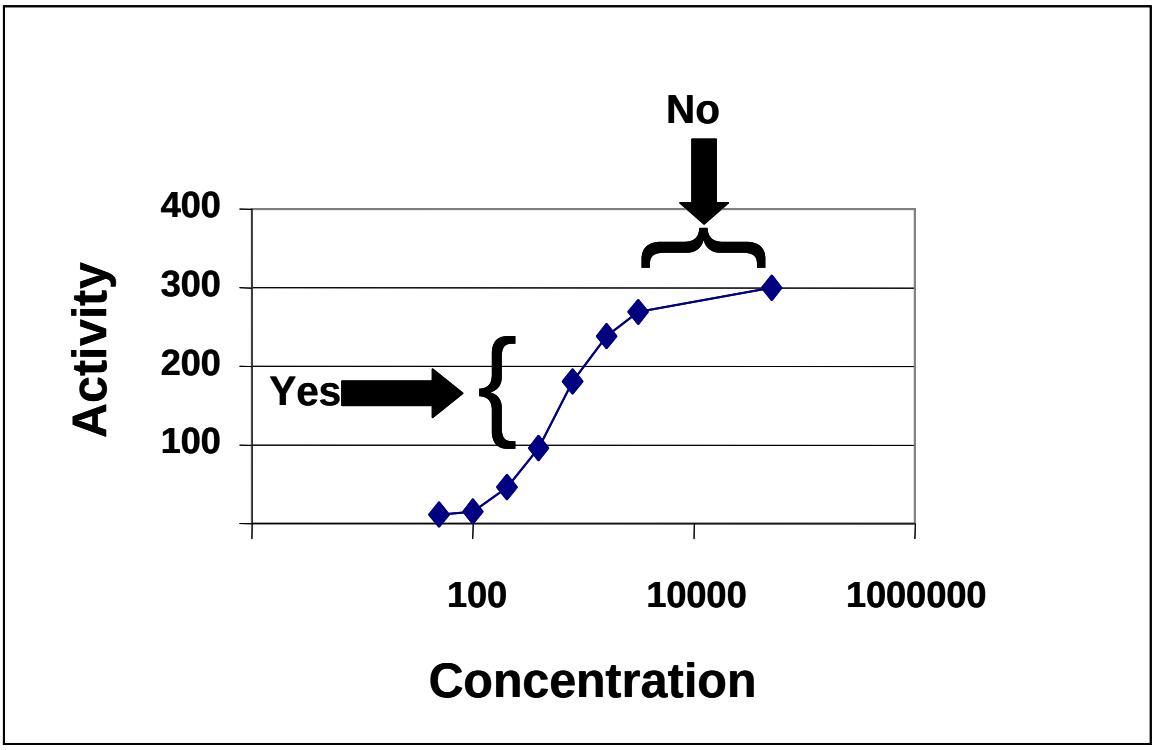
The cellular responses measured in these bioassays are numerous and can include outcomes such as phosphorylation of intracellular substrates, calcium mobilization, proliferation, and cell death. In some cases, sponsors have developed cell lines to express relevant receptors or reporter constructs. When therapeutic protein products directly stimulate a cellular response, the direct effect of NAb on reducing bioactivity in the bioassay can be measured. When therapeutic protein products indirectly impact cellular activity; for example, by blocking a receptor-ligand interaction, the indirect effect of the NAb on restoring bioactivity in a bioassay can be measured. Generally, bioassays have significant variability and a limited dynamic range for their activity curves. Such problems can make development and validation of neutralization assays difficult.

There are cases when ligand-binding assay formats may be used. One such case is when sufficiently sensitive or selective cell-based bioassays cannot be developed. Another case is when the therapeutic protein product does not have a cell-based mechanism of action; for example, enzyme therapeutic protein products that target serum proteins. Ligand-binding assays may also be appropriate for therapeutic protein products that bind serum ligands, preventing them from interacting with their receptor. However, cell-based bioassays may still be more appropriate for such therapeutic protein products to demonstrate that ADA are inhibiting cellular activity. Sponsors should discuss using ligand-binding assays with FDA in such cases.

2. Activity Curve of Neutralization Assay

The sponsor should carefully consider the dose response curve (product concentration versus activity) before examining other elements of neutralization assay validation. Assays with a small dynamic range may not prove useful for determination of neutralizing activity. Generally, the neutralization assay will employ a single concentration of therapeutic protein product with a single dilution of antibody. Consequently, the sponsor should choose a therapeutic protein product concentration whose activity readout is sensitive to inhibition. If the assay is performed at concentrations near the plateau of the dose-response curve (marked “No” in Figure 1, below), it may not be possible to discern samples with low amounts of NAb. FDA recommends that the neutralization assay be performed at therapeutic protein product concentrations that are on the linear range of the curve (marked “Yes” in Figure 1). The assay should also give reproducible results.

Figure 1. Activity Curve for a Representative Therapeutic Protein Product



The x-axis (Concentration) indicates a concentration of the therapeutic protein product, and the y-axis (Activity) indicates resultant activity; for example, the concentration of cytokine secretion of a cell line upon stimulation with the therapeutic protein product. The curve demonstrates a steep response to a therapeutic protein product that plateaus at approximately 300. The "No" arrow indicates a concentration of a therapeutic protein product that would be inappropriate to use in a single dose neutralization assay because it would represent a range of concentrations where the activity induced by the therapeutic protein would be relatively insensitive to inhibition by NAb. The "Yes" arrow represents a range of concentrations on the linear part of the curve where the activity induced by the therapeutic protein product would be sensitive to neutralization by antibody.

3. Considerations for Matrix Interference for Neutralization Assay

The matrix can cause interference with neutralization assays, particularly as serum or plasma components may enhance or inhibit the activity of a therapeutic protein product in bioassays. For example, sera from patients with particular diseases may contain elevated levels of one or more cytokines that might serve to activate cells in the bioassay and obscure the presence of NAb by increasing the response to the original stimulatory factor or therapeutic protein product. Therefore, the sponsor should understand matrix effects in these assays. Approaches such as enriching for ADA from serum or plasma samples may be appropriate for these types of situations. However, this approach may result in the loss of NAb, and consequently will require careful examination and validation by the sponsor.

The concentration of therapeutic protein product employed in the neutralization assay has a critical impact on assay sensitivity. FDA recognizes that although the use of low concentrations of therapeutic protein product may lead to a neutralization assay that is more sensitive to inhibition by antibodies, very low concentrations of therapeutic protein product may result in poor precision of the assay. Also see section IV.D.1 for general information on matrix interference.

4. Cut Point of Neutralization Assay

Determination of assay cut point has historically posed a great challenge for neutralization assays. As with all assays, the cut point should be determined based on the assay variability established using samples from treatment-naïve subjects. If neutralization assays are performed on samples that tested positive in screening and confirmatory assays, a 1% false-positive rate is acceptable. If neutralization assays are used for screening, a 5% false-positive rate should be used (see section VI.B.2). If the degree of sample variation makes it difficult to assess NAb activity, other approaches may be considered but should be discussed with FDA before implementation. Alternatively, exploring other assay formats that lead to less variability and provide a more accurate assignment of cut point may be necessary. Also see section IV.B for general information on assay cut point.

5. Additional Considerations for Neutralization Assay

Because neutralization assays are most commonly performed only on samples that are confirmed to have antigen-specific ADA, confirmatory approaches are not usually necessary. However, because of the complexity of bioassays, confirmation of assay specificity may be useful in determining whether patients have mounted a true NAb response. The sponsor should consider the following approaches:

a. Unrelated inhibitory molecules may cause neutralizing activity, and sometimes it may be unclear whether the observed neutralizing activity is caused by neutralizing antibodies or by other inhibitory molecules. Test results from baseline pre-exposure samples may be informative. When there is concern that there is non-specific inhibition, antibody depletion assays should be performed to evaluate whether the neutralizing activity is truly caused by ADA and not caused by other inhibitory molecules.

b. Cell lines may be responsive to multiple stimuli other than the therapeutic protein product under study. In such cases, the presence of NAb can be examined in the presence of the therapeutic protein product, which should be blocked by a specific NAb response, versus alternative stimuli, which should not be blocked by a specific NAb response.

c. Serum may contain components such as soluble receptors or endogenous product counterparts that may yield false results in the neutralization assay. In such instances,

adding test serum or plasma samples directly to the bioassay in the absence of therapeutic protein product may be useful in understanding assay results.

VI. ASSAY VALIDATION

Assay validation is a process of demonstrating, by the use of specific laboratory investigations, that the performance characteristics of the ADA assay employed are suitable for its intended use.[23] The level of validation depends on the stage of product development and the risks of consequences of immunogenicity to patients associated with the therapeutic protein product. A partial validation involving assessments of assay sensitivity, specificity, and precision requirements with less emphasis on robustness, reproducibility, and stability may be adequate for the earlier stages of clinical development such as phase 1 and phase 2 studies. However, as a scientific matter, as stated in section VI.A, fully validated assays should be used for pivotal and postmarketing studies.

Information specific to validation of respective assay types is provided in sections VI.B through E. These sections supplement information relevant to all assay types provided in sections IV and VI.A.

A. General Considerations for Assay Validation

Samples derived from pivotal studies should be tested with fully validated assays. At the time of license application, the sponsor should provide data supporting full validation of the assays. Validation includes all of the procedures that demonstrate that a particular assay used for quantitative measurement of ADA in a given sample is reliable and reproducible for the intended use. The fundamental parameters for validation include (1) cut point, (2) sensitivity, (3) specificity and selectivity, (4) precision, (5) reproducibility when relevant, and (6) robustness of some assay features and stability of reagents and control samples. The acceptability of clinical data generated by an assay corresponds directly to the criteria used to validate the assay.

Determination of cut point is a fundamental aspect of assay validation. If treatment-naïve samples from the appropriate patient population are not available for the pre-study validation exercise, alternative samples may be used. Frequently these are samples from commercial sources. When alternative samples are used to determine the cut point in the validation exercise, the cut point should be determined again once samples from the appropriate population (e.g. treatment-naïve patients) are available. The cut point validated using the appropriate samples should be used to determine whether samples are positive for ADA.

For validation of the fundamental assay parameters, FDA recommends, at the minimum, that inter-assay precision be evaluated on at least 3 different days with two analysts each preparing a

[23] See the USP General Chapter 1106 *Immunogenicity Assays – Design and Validation of Immunoassays to Detect Anti-Drug Antibodies*. Also see the guidance for industry *Bioanalytical Method Validation*, the USP General Chapter 1225 *Validation of Compendial Procedures*, and the ICH guidance for industry *Q2(R1) Validation of Analytical Procedures: Text and Methodology*.

minimum of six otherwise independent preparations of the same sample using the same instrument platform and model. Intra-assay precision should be evaluated with a minimum of six independent preparations of the same sample per plate independently prepared by the same analyst. In cases where intra-assay or inter-assay precision has a coefficient of variance (%CV) greater than 20%, sponsors should consider the need to refine the assay parameters to optimize the assay precision to the extent possible or provide justification to explain why higher %CV should be acceptable. Alternatively, in assays with low throughput (e.g., titer assay) when it may not be possible to run six independent preparations of the same sample on a plate, intra-assay precision should be evaluated with a minimum of three independent preparations of the same sample per plate and at least nine total independent preparations of the same samples. Samples should include negative controls and positive samples whose testing yields values in the low, medium, and high levels of the assay dynamic range. The sponsor should evaluate inter-instrument and inter-operator precision when relevant. Assays should have comparable precision between different operators under the same operating conditions.

When changes are made to a previously validated method, the sponsor should exercise judgment as to how much additional validation is needed. During the course of a typical product development program, a defined ADA assay may undergo modifications. Occasionally, samples may need to be re-tested with the optimized validated assay; therefore, provisions should be made to preserve sufficient sample volume under conditions that allow for re-testing until the assays have been completely validated and evaluated by the Agency.[24]

Critical method parameters, for example, incubation times and temperatures, should be validated to demonstrate that the assay performs as expected within predetermined ranges for these parameters. Generally, the low, middle, and high values of the allowed range are tested in the validation exercise.

Additional parameters may need to be validated depending on the method (or technology) and instrument platform used for the assay. For example, surface plasmon resonance assays should be validated for surface stability upon regeneration, and criteria should be set for baseline performance of the chip. The efficiency and stability of the labeled[25] reagents should be established. The sponsor should examine robustness during the development phase, and if small changes in specific steps in the assay affect results, specific precautions should be taken to control their variability.

[24] See the guidance for industry *Bioanalytical Method Validation* for different types and levels of validation. Also see the USP General Chapter 1106 *Immunogenicity Assays – Design and Validation of Immunoassays to Detect Anti-Drug Antibodies*.

[25] A reagent is considered *labeled* if it is conjugated or fused to a moiety that will aid in its capture or visualization; for example, conjugation to biotin, streptavidin, or a fluorochrome. *Unlabeled* reagent is a reagent (for example, a drug) that is not *labeled*.

B. Validation of Screening Assay

1. Sensitivity of Screening Assay

All the general considerations for assay validation discussed previously apply to validation of screening assay. As noted earlier, the sensitivity is particularly important in the initial screening assay because these results dictate the further analysis of the sample.

2. Cut Point of Screening Assay

The cut point should be determined statistically with a minimum of 50 samples tested on at least 3 different days by at least two analysts using suitable statistical methods. FDA recommends that the cut point for screening assays be determined by a 90% one-sided lower confidence interval for the 95th percentile of the negative control population (Shen, Dong, et al. 2015). This will assure at least a 5% false-positive rate with a 90% confidence level. This approach improves the probability of the assay identifying all patients who may develop antibodies. The statistical method used to determine the cut point should be based on the statistical distribution of the data. For example, the 95th percentile of the normal distribution is estimated by the mean plus 1.645 standard deviation. Other approaches may be used for estimating 95th percentile, including the use of median and median absolute deviation value instead of mean and standard deviation.

The mean response of negative control samples may be constant or may vary between assays, plates, or analysts. When the mean is constant, a cut point may be established during assay validation that can be applied to the assay in-study. This is frequently called a fixed cut point. When the mean varies between assays, plates, or analysts but the variance around the mean is constant, a normalization factor can be statistically determined and applied in-study. This is also known as a floating cut point. When both the mean and variance vary, a cut point must be established for each assay, plate, or analyst. This is known as a dynamic cut point. One drawback of the dynamic cut point is the need to have more replicates of the negative control in the assay. Dynamic cut points should not be used to compensate for deficient assay optimization.

C. Validation of Confirmatory Assay

Confirmatory assays should be fully validated in a manner similar to screening and neutralization assays because these assays raise some specific issues. As a scientific matter, the studies to validate the assay will depend on the assay format and instrumentation chosen. If these assays are based on competition for antigen binding[26] by the antibodies in patient samples and the measurement is loss of response, it is critical to identify the degree of inhibition or depletion that will be used to ascribe positivity to a sample. In the past, fixed percentages of binding reduction were used, but these numbers were often arbitrary and are unlikely to be relevant for all assays.

[26] *Competition for antigen binding* refers to a competition assay where the ability of antigen-specific antibodies to bind to either labeled or plate-bound antigen is inhibited by unlabeled or soluble antigen.

FDA recommends establishing a cut point based on the assessment of the binding changes observed in samples that are known to lack the antibodies when competing antigen is added. FDA also recommends that the sensitivity of the confirmatory assay be confirmed using a low concentration of the positive control antibody.

For the estimation of the confirmatory assay cut point, an 80% one-sided lower confidence interval for the 99th percentile is recommended. Because the purpose of this assay is to eliminate false-positive samples arising as a result of non-specific binding, it is adequate to use a 1% false-positive rate for the calculation of the confirmatory cut point. The use of tighter false-positive rates such as 0.1% is not recommended because it will lead to an increased risk of false-negative results. See section IV.B for general information on assay cut point.

If the confirmatory assay format is a competiton assay in which a competitor, usually unlabeled therapeutic protein product,[27] will be added to the reaction mixture to inhibit ADA binding to the capture reagent for the cut point assay, the same concentration of unlabeled therapeutic protein product should be added to the samples when determining the confirmatory cut point.

D. Validation of Titering Assay

The principles of assay validation described in section VI.A apply in general to validation of titering assays. The cut point of the titration assay may be the same as or different from that of the screening assay. When the titering assay is not used for screening and the cut point is different than that of the screening assay, the validation of the separate titration method cut point can become necessary; for example, when the signal from the assay diluent or matrix causes higher results than the screening assay cut point because of a blocking effect of serum or if samples at a dilution higher than the MRD do not generate consistently negative results, i.e., when the screening cut point falls on the lower plateau of the positive-control dilution curve.[28]

E. Validation of Neutralization Assay

A minimum of 30 samples tested on at least 3 different days by at least two analysts should be used to determine the cut point, using suitable statistical methods.

FDA recognizes that not all ADA are neutralizing, and it can be difficult to identify positive control antibodies with neutralizing capacity. Further, if an affinity purified polyclonal positive control antibody preparation is used, it is likely that only a portion of the antibodies are neutralizing, which can make the assay appear less sensitive. Therefore, it is important to validate assay sensitivity.

Sponsors should validate assay specificity for cell-based neutralization bioassays. As mentioned, for cells that may be responsive to stimuli other than the specific therapeutic protein product, the

[27] See footnote 25.

[28] See the USP General Chapter 1106 *Immunogenicity Assays – Design and Validation of Immunoassays to Detect Anti-Drug Antibodies*.

ability to demonstrate that NAb only inhibit the response to therapeutic protein product and not the response to other stimuli is a good indication of assay specificity. In such studies, FDA recommends that the other stimuli be employed at a concentration that yields an outcome similar to that of the therapeutic protein product. The sponsor should also confirm the absence of alternative stimuli in patient serum (see sections IV.C and D).

Cell-based neutralization bioassays frequently have reduced precision when compared to ligand-binding assays because biologic responses can be inherently more variable than carefully controlled binding studies. Consequently, the sponsor should perform more replicates for assessment of precision and assessment of patient responses than for the screening assay (see section IV.E).

Additional parameters that should be validated are assay performance when cells at the low, middle, and high range of the allowed passage numbers, cell density, and cell viability are used (see section IV.G).

VII. IMPLEMENTATION OF ASSAY TESTING

A. Obtaining Patient Samples

FDA recommends that the sponsor obtain pre-exposure samples from all patients. Because there is the potential for pre-existing antibodies or confounding components in the matrix, understanding the degree of reactivity before treatment is essential. The sponsor should obtain subsequent samples, with the timing depending on the frequency of dosing. Optimally, samples taken 7 to 14 days after the first exposure can help elucidate an early IgM response. Samples taken at 4 to 6 weeks after the first exposure are generally optimal for determining IgG responses. For individuals receiving a single dose of therapeutic protein product, the above time frame may be adequate. However, for patients receiving a therapeutic protein product at multiple times during the trial, the sponsor should obtain samples at appropriate intervals throughout the trial and also obtain a sample approximately 30 days after the last exposure.

Obtaining samples at a time when there will be minimal interference from the therapeutic protein product present in the serum is essential. A sponsor should consider the therapeutic protein product's half-life to help determine appropriate times for sampling. This is especially important for mAb products because these products can have half-lives of several weeks or more; and depending on the dosing regimen, the therapeutic mAb itself could remain present in the serum for months. Under circumstances when testing for IgE is needed, the timing of sample collection should be discussed with FDA.

The level of therapeutic protein product that interferes with the assay, as determined by immune competition, may also help define meaningful time points for sampling. If therapeutic protein product-free samples cannot be obtained during the treatment phase of the trial, the sponsor should take additional samples after an appropriate washout period (e.g., five half-lives). Obtaining samples to test for meaningful antibody results can also be complicated if the therapeutic protein product in question is itself an immune suppressant. In such instances, the

sponsor should obtain samples from patients who have undergone a washout period either because the treatment phase has ended or because the patient has dropped out of the study.

Samples to determine serum concentrations of therapeutic protein product should be obtained at the same time as immunogenicity samples. Testing such samples can provide information on whether the therapeutic protein product in the samples may be interfering with ADA testing and whether ADA may be altering the therapeutic protein product's pharmacokinetics.

B. Concurrent Positive and Negative Quality Controls

If the sponsor completes the proper validation work and makes the cut point determinations, the immunogenicity status of patients should be straightforward to determine. However, positive control or QC samples are critical and should be run concurrently with patient samples. We recommend that these samples span a level of positivity with QC samples having a known negative, low, and high reactivity in the assay. More important, the QC samples should be diluted in the matrix in which patient samples will be examined; for example, the same percent serum or plasma (specify salt anticoagulant used). In this way, the sponsor ensures that the assay is performing to its optimal degree of accuracy and that patient samples are correctly evaluated. For the low-positive QC sample, we recommend that a concentration be selected that, upon statistical analysis, would lead to the rejection of an assay run 1% of the time. Such an approach would ensure the appropriate sensitivity of the assay when performed on actual patient samples. The concentration of high-positive QC samples should be set to monitor prozone effects.[29]

FDA also recommends that these QC samples be obtained from humans or animals possessing antibodies that are detected by the secondary detecting reagent, to ensure that negative results that might be observed are truly caused by lack of antigen reactivity and not caused by failure of the secondary reagent. This issue is not a problem for antigen bridging assays because labeled antigen is used for detection.

C. Confirmation of Cut Point in the Target Population

Samples from different populations can have different background activity in ADA assays. Therefore, it is necessary to confirm that the cut point determined during assay validation is suitable for the population being studied. Similarly, if samples used to determine the cut point during assay validation were not obtained and handled in a manner that represents how samples will be obtained and handled in-study, the cut point should also be confirmed with appropriate samples in-study. A sufficient number of samples from the target population should be used, and justification for the number used should be provided. If sufficient numbers of samples are not available, agreement with the Agency should be sought for the number of samples to be used.

[29] Prozone effects (also referred to as hook effects) are a reduction in signal that may occur as a result of the presence of a high concentration of a particular analyte or antibody and may cause false-negative results.

VIII. DOCUMENTATION

The rationale and information for the immunogenicity testing paradigm should be provided in module 5.3.1.4 of the electronic common technical document (eCTD) on *Reports of Bioanalytical and Analytical Methods for Human Studies*.[30] The standard operating procedure of the respective assay being used should be provided to the FDA, together with the results of the validation studies and relevant assay development information for parameters that were not validated, such as the MRD, the stimulatory concentration of therapeutic protein product used in the NAb assay, and some robustness parameters that are critical for assay performance (see section VII. Documentation in the draft guidance for industry *Bioanalytical Method Validation*.)[31]

[30] See the FDA Web site for further information on eCTD submissions, available at http://www.fda.gov/Drugs/DevelopmentApprovalProcess/FormsSubmissionRequirements/ElectronicSubmissions/ucm153574.htm. For more information about the agreed-upon common format for the preparation of a well-structured Efficacy section of the CTD for applications that will be submitted to regulatory authorities, see the ICH guidance for industry *M4E: The CTD — Efficacy*. For more information on how sponsors and applicants must organize the content they submit to the Agency electronically for all submission types under section 745A(a) of the FD&C Act, see the guidance for industry (and the technical specification documents it incorporates by reference) *Providing Regulatory Submissions in Electronic Format — Certain Human Pharmaceutical Product Applications and Related Submissions Using the eCTD Specifications*.

[31] When final, this guidance will represent the FDA's current thinking on this topic. To make sure you have the most recent version of a guidance, check the FDA guidance Web page at http://www.fda.gov/RegulatoryInformation/Guidances/default.htm.

REFERENCES

Aalberse, R. C. and J. Schuurman (2002). "IgG4 breaking the rules." Immunology 105(1): 9-19.

Boes, M. (2000). "Role of natural and immune IgM antibodies in immune responses." Mol Immunol 37(18): 1141-1149.

Calabresi, P. A., G. Giovannoni, et al. (2007). "The incidence and significance of anti-natalizumab antibodies: results from AFFIRM and SENTINEL." Neurology 69(14): 1391-1403.

Caruso, A. and A. Turano (1997). "Natural antibodies to interferon-gamma." Biotherapy 10(1): 29-37.

Cohen, B. A. and V. M. Rivera (2010). "PRISMS: the story of a pivotal clinical trial series in multiple sclerosis." Curr Med Res Opin 26(4): 827-838.

Coutinho, A., M. D. Kazatchkine, et al. (1995). "Natural autoantibodies." Curr Opin Immunol 7(6): 812-818.

Disis, M. L., V. Goodell, et al. (2004). "Humoral epitope-spreading following immunization with a HER-2/neu peptide based vaccine in cancer patients." J Clin Immunol 24(5): 571-578.

Goodin, D. S., E. M. Frohman, et al. (2007). "Neutralizing antibodies to interferon beta: assessment of their clinical and radiographic impact: an evidence report: report of the Therapeutics and Technology Assessment Subcommittee of the American Academy of Neurology." Neurology 68(13): 977-984.

Gupta, S., V. Devanarayan, et al. (2011). "Recommendations for the validation of cell-based assays used for the detection of neutralizing antibody immune responses elicited against biological therapeutics." J Pharm Biomed Anal 55(5): 878-888.

Gupta, S., S. R. Indelicato, et al. (2007). "Recommendations for the design, optimization, and qualification of cell-based assays used for the detection of neutralizing antibody responses elicited to biological therapeutics." J Immunol Methods 321(1-2): 1-18.

Harding, F. A., M. M. Stickler, et al. (2010). "The immunogenicity of humanized and fully human antibodies: residual immunogenicity resides in the CDR regions." MAbs 2(3): 256-265.

Hintermann, E., M. Holdener, et al. (2011). "Epitope spreading of the anti-CYP2D6 antibody response in patients with autoimmune hepatitis and in the CYP2D6 mouse model." J Autoimmun 37(3): 242-253.

Kuus-Reichel, K., L. S. Grauer, et al. (1994). "Will immunogenicity limit the use, efficacy, and future development of therapeutic monoclonal antibodies?" Clin Diagn Lab Immunol 1(4): 365-372.

Matsumoto, T., M. Shima, et al. (2001). "Immunological characterization of factor VIII autoantibodies in patients with acquired hemophilia A in the presence or absence of underlying disease." Thromb Res 104(6): 381-388.

Miller, L. L., E. L. Korn, et al. (1999). "Abrogation of the hematological and biological activities of the interleukin-3/granulocyte-macrophage colony-stimulating factor fusion protein PIXY321 by neutralizing anti-PIXY321 antibodies in cancer patients receiving high-dose carboplatin." Blood 93(10): 3250-3258.

Mire-Sluis, A. R., Y. C. Barrett, et al. (2004). "Recommendations for the design and optimization of immunoassays used in the detection of host antibodies against biotechnology products." J Immunol Methods 289(1-2): 1-16.

Plotkin, S. A. (2010). "Correlates of protection induced by vaccination." Clin Vaccine Immunol 17(7): 1055-1065.

Prummer, O. (1997). "Treatment-induced antibodies to interleukin-2." Biotherapy 10(1): 15-24.

Ramsland, P. A., B. F. Movafagh, et al. (1999). "Interference of rheumatoid factor activity by aspartame, a dipeptide methyl ester." J Mol Recognit 12(5): 249-257.

Ross, C., M. B. Hansen, et al. (1990). "Autoantibodies to crude human leucocyte interferon (IFN), native human IFN, recombinant human IFN-alpha 2b and human IFN-gamma in healthy blood donors." Clin Exp Immunol 82(1): 57-62.

Shankar, G., V. Devanarayan, et al. (2008). "Recommendations for the validation of immunoassays used for detection of host antibodies against biotechnology products." J Pharm Biomed Anal 48(5): 1267-1281.

Shen, M., X. Dong, et al. (2015). "Statistical evaluation of several methods for cut-point determination of immunogenicity screening assay." J Biopharm Stat 25(2): 269-279.

Thrasyvoulides, A., E. Liakata, et al. (2007). "Spreading of antibody reactivity to non-thyroid antigens during experimental immunization with human thyroglobulin." Clin Exp Immunol 147(1): 120-127.

Turano, A., A. Balsari, et al. (1992). "Natural human antibodies to gamma interferon interfere with the immunomodulating activity of the lymphokine." Proc Natl Acad Sci U S A 89(10): 4447-4451.

van der Meide, P. H. and H. Schellekens (1997). "Anti-cytokine autoantibodies: epiphenomenon or critical modulators of cytokine action." Biotherapy 10(1): 39-48.

van der Woude, D., S. Rantapaa-Dahlqvist, et al. (2010). "Epitope spreading of the anti-citrullinated protein antibody response occurs before disease onset and is associated with the disease course of early arthritis." Ann Rheum Dis 69(8): 1554-1561.

van Schouwenburg, P. A., S. Kruithof, et al. (2014). "Functional analysis of the anti-adalimumab response using patient-derived monoclonal antibodies." J Biol Chem 289(50): 34482-34488.

Zhou, L., S. A. Hoofring, et al. (2013). "Stratification of antibody-positive subjects by antibody level reveals an impact of immunogenicity on pharmacokinetics." AAPS J 15(1): 30-40.