Research Article

Statistical and Bioanalytical Considerations for Establishing a Depletion Criterion for Specificity Testing During Immunogenicity Assessment of a Biotherapeutic

R. O. Driscoll,¹ L. Zhou,² M. Moxness,¹ D. Mytych,¹ N. Chirmule,¹ and V. Jawa^{1,3}

Received 22 January 2013; accepted 5 August 2013; published online 30 August 2013

Abstract. Immunogenicity assessment of fully human monoclonal antibody-based biotherapeutics requires sensitive and specific ligand binding assays. One of the components of specificity is the depletion of signal by a relevant biotherapeutic that is commonly based on an arbitrary depletion criterion of inhibition of the original response or reduction of the signal below the screening assay cut point (ACP). Hence, there is a need to develop a statistically derived physiologically relevant specificity criterion. We illustrate an optimization approach to determine the concentration of biotherapeutic required for the specificity evaluation. Naïve donor sample sets with and without circulating drug and antitherapeutic/drug antibody (ADA) were prepared. Next, a depletion cut point (DCP) using naïve and ADA-containing donor sets with the optimized biotherapeutic concentration was evaluated. A statistically derived design of experiment was used to establish a validated DCP. A reliable DCP requires naïve (no ADA) donors treated only with an optimized concentration of biotherapeutic. The additional DCPs generated using two distinct concentrations of ADA-spiked sample sets led to a physiologically irrelevant criterion that was not necessarily representative of real-time samples. This increased the risk of false positives or negatives. In this study, well-defined bioanalytical and statistical methods were employed to validate a DCP to confirm the presence of biotherapeutic specific ADA in human serum samples. A physiologically relevant and effective strategy to confirm specificity in immune reactive samples, especially those that are close to the ACP, is proposed through this study.

KEY WORDS: antidrug antibodies; confirmation; depletion cut point; immunogenicity; specificity.

INTRODUCTION

Immunogenicity assessment strategies for fully human monoclonal antibodies (mAb), referred to as biotherapeutics in this manuscript, have been well characterized in the last decade (1–4). Currently, most of the antidrug antibody (ADA) assays follow a robust development and validation plan as recommended by the industry white papers (5). The immunogenicity assessment for binding antibodies comprises a screening step followed by a confirmation and characterization step. The screening assay detects the presence of binding ADAs. Samples with a signal-to-noise (*S/N*) value greater than the assay cut point (ACP) are then tested to

confirm specificity of the response. Samples are analyzed in the specificity assay where depletion of signal is evaluated in the presence of excess soluble biotherapeutic. The samples that deplete to a certain extent are reported as positive for binding ADAs and are further characterized for neutralizing activity in a biological assay. The challenge of immunogenicity assessment has been the lack of statistically defined specificity criteria following the initial screening. Most of the specificity criteria rely on depletion of the initial assay signal with excess amounts of biotherapeutic. However, there is a gap in defining what the extent of depletion should be. Hence, in some instances, it follows an arbitrary criterion or threshold based on 50% depletion or reduction below the ACP, which could lead to inaccurate results (5–7).

Several methods have been discussed for establishing a depletion cut point (DCP), also known as specificity cut point or confirmatory cut point. For the sake of clarity, we have used the term DCP for this manuscript. The specificity evaluation in immunoassays has several components. One aspect is to confirm if the initial response is due to ADA and not due to ligand/receptor interference. One such approach where specificity can be confirmed is by signal reduction using antibody depletion reagents like protein G (8). Another facet of specificity is to confirm that the signal observed during

Electronic supplementary material The online version of this article (doi:10.1208/s12248-013-9523-1) contains supplementary material, which is available to authorized users.

¹ Clinical Immunology, Amgen Inc., One Amgen Center Drive, 30E-3-C, Thousand Oaks, California 91320, USA.

² Biostatistics, Medical Sciences, Amgen Inc., One Amgen Center Drive, Thousand Oaks, California 91320, USA.

³ To whom correspondence should be addressed. (e-mail: vjawa01@amgen.com)

screening is specific to the drug and can be inhibited when excess biotherapeutic or drug is added (5,6).

There are various bioanalytical aspects that must be investigated closely prior to the validation of a DCP. They include (1) the determination of an appropriate amount of biotherapeutic needed for depletion (in the absence or presence of circulating drug) and (2) determining whether naïve *vs.* ADA-spiked donors should be used to establish such specificity cut point. Other factors like the presence of soluble receptors, targets, serum proteins, or autoantibodies can also interfere in specificity assessments and should be evaluated wherever applicable (9,10) because they can confound the outcome of the screening assay and the specificity assay.

Furthermore, study samples for ADA assessments may contain a high level of circulating drug that can interfere with the assay and reduce the ADA signals (5,9–12). Hence, it is important to validate a cut point that addresses matrix-associated interference to support an accurate specificity threshold (DCP) in ADA immunoassays. In this study, we have established and validated a depletion cut point by keeping bioanalytical and statistical methods in consideration and by mimicking sample sets with and without ADA as well as circulating drug.

MATERIALS AND METHODS

Serum Samples

Pooled naïve normal human serum (PNHS) and serum from individual drug/biotherapeutic naïve human donors (healthy and disease specific) were obtained for use as assay matrix from Bioreclamation, Inc., Hicksville, NY.

Study samples were obtained from an Amgen-sponsored phase 1 study where serum was collected from healthy subjects according to the study protocol and in adherence with informed consent and review of the protocol by relevant ethics committees. None of the subjects had preexisting antitherapeutic antibody reactivity.

Reagent Preparation

General Serum Collection Process

Blood was collected in red-top Vacutainer® tubes (containing no additives or anticoagulants) at protocol-driven time points and maintained at room temperature after collection. Following a 30-min clotting period, samples were centrifuged at room temperature at 1,500×g for approximately 15 min. The collected serum was then transferred into a prelabeled, polypropylene cryovial and stored at –60 to –80°C for future analysis. Collection, centrifugation, and freezing of sample were performed within 60 min.

Antibody Drug (Biotherapeutic Candidate)

A fully human IgG2 mAb generated at Amgen, Inc., was used as test reagent.

Positive Control Antibody

An affinity purified rabbit polyclonal antibody was produced by immunizing rabbits with the biotherapeutic (Amgen Inc.). Purification of the rabbit positive control antibodies specific to the idiotypic region of the biotherapeutic was performed by using a column with biotherapeutic covalently bound to cyanogen bromide Sepharose 4B gel (GE Healthcare). Subsequent to the affinity column purification, immunoabsorption against human immunoglobulin bound to Sepharose 4B was performed (13). This enables generation of a positive control antibody that is predominantly reactive to the F (ab) region of the biotherapeutic and helps eliminate Fc reactivity.

Electrochemiluminescent Bridging Immunoassay

An acid-dissociation electrochemiluminescence (ECL) assay was performed to detect ADA. In brief, samples were treated with 300 mM acetic acid to dissociate antibody complexes in serum samples (dilution factor 1:10) (10,14,15). The biotherapeutic was labeled with Sulfo-TAG-NHS ester ruthenium (MSD) and Sulfo-NHS-LC-Biotin (Thermo Scientific). The conjugation was performed following the vendor procedure as described elsewhere (15). Samples were incubated with a conjugate/neutralization buffer consisting of 1 µg/mL biotinylated biotherapeutic, 1 µg/mL ruthenylated biotherapeutic, and 1 M Tris pH 9.5 in 1% bovine serum albumin (BSA) in 1× phosphate-buffered saline (PBS) (screening assay only). Soluble biotherapeutic was added to the conjugate/neutralization mixture for specificity assays. The concentration of soluble excess biotherapeutic was optimized to be 100 µg/mL. Serum samples were then added to an avidin high bind MSD plate (previously blocked with 1% BSA in 1× PBS). The biotinylated biotherapeutic molecule binds to the avidin-coated surface resulting in the immobilization of the bridged complex. In the specificity assays, excess biotherapeutic competes for and blocks binding to the ADA. Plates were then washed using 1× PBS with 0.05% Tween 20 and 0.25 mM thimerosal to remove unbound complexes. Read buffer containing tripropylamine (MSD) was added to all wells. Plates were read using the MSD SECTOR® Imager 6000 (14,15). Each sample was plated in duplicates. The *S/N* ratio was determined by taking the mean ECL signal of each sample and dividing by the mean ECL signal of the negative control depending on the assay treatment (screening or specificity). Total percent or *S/N* depletion was calculated for each point on the curve pre- and postdrug treatment as follows:

Percent Depletion

$$= (\text{Untreated sample} - \text{Treated sample}) / (\text{Untreated sample}) \\ \times 100$$

Alternative methods for calculating depletion using net ECL or mean ECL are provided in the [supplementary text](#).

Optimization of Soluble Biotherapeutic Concentration in Specificity Assays

Recovery Phase Sample Set

A set of five 8-point dose–response curves were created in PNHS consisting of increasing amounts of ADA: 0.01, 0.04, 0.16, 0.5, 1.0, 20, 80, and 160 µg/mL. Each curve was treated with one of five soluble biotherapeutic concentrations (1, 7.5, 25, 100, and 400 µg/mL). *S/N* percent depletion was calculated for each point on the curve.

Dosing Phase Sample Set

Twelve 8-point dose–response curves in PNHS were created with the following biotherapeutic concentrations: 0, 4.7, 9.4, 18.8, 37.5, 75.0, 150, and 300 µg/mL (plate maps provided in Table I). They were divided into three panels consisting of four curves apiece. Each panel was spiked with one of three positive control or ADA concentrations (panel 1: high, 100 µg/mL; panel 2: moderate, 0.5 µg/mL; and panel 3: low, 0.05 µg/mL) (Table I). Each curve, regardless of ADA concentration, was depleted with a given concentration of soluble biotherapeutic (curve 1, 7.5 µg/mL; curve 2, 25 µg/mL; curve 3, 100 µg/mL; curve 4, 400 µg/mL). A total of 12 different combinations were achieved. *S/N* percent depletion was calculated for each point on the curve as previously described.

Validation of the DCP

Preliminary Evaluation of DCP Using Naïve vs. ADA-Positive Serum Sets

Serum from 18 naïve human donors (nine males and nine females) was spiked using three ADA concentrations: (1) 0 ng/mL [representing naïve or ADA-negative subjects], (2) 20 ng/mL, and (3) 160 ng/mL [representing ADA-positive subjects] (Table II). Serum samples were treated with

Table II. Preliminary Evaluation of Three DCP Criteria

Donor number designation	<i>S/N</i> % depletion		
	ADA concentration		
	0 ng/mL	20 ng/mL	160 ng/mL
1	1	19	53
2	0	20	52
3	3	18	49
4	11	21	53
5	−4	27	58
6	−3	33	54
7	−1	31	51
8	−4	25	60
9	8	14	53
10	13	3	51
11	11	11	53
12	12	8	56
13	14	25	57
14	4	16	57
15	4	19	54
16	−3	15	52
17	−4	6	48
18	−1	10	48
Mean	3	18	53
STD	6.50	8.40	3.36
DCP	23% ^a	−7% ^b	43% ^b

Serum from 18 individual human donors was spiked with three ADA concentrations (0 ng/mL [naïve donors], 20 ng/mL, and 160 ng/mL). Serum samples were treated with and without excess soluble biotherapeutic (100 µg/mL) as determined in an optimization assay. Potential DCPs were evaluated using *S/N* depletion

^a DCP=Mean+3×STD

^b DCP=Mean−3×STD

(100 µg/mL) and without excess soluble biotherapeutic as determined in the above-mentioned optimization experiments. The *S/N* percent depletion and the standard deviation (STD) were calculated for each of the three serum sets. Potential DCPs were evaluated using one of two formulas

Table I. Optimization of Excess Soluble Biotherapeutic Concentration: Experiment Setup

	ADA (µg/mL)		
	Panel 1, 100	Panel 2, 0.5	Panel 3, 0.05
Circulating drug (µg/mL)	0	0	0
	4.7	4.7	4.7
	9.4	9.4	9.4
	18.8	18.8	18.8
	37.5	37.5	37.5
	75.0	75.0	75.0
	150.0	150.0	150.0
	300.0	300.0	300.0

Twelve-eight-point curves were created ranging from 0 to 300 µg/mL of biotherapeutic mAb in PNHS. Subsequently, they were divided into three panels of four curves each. Each panel was spiked with one of three ADA concentrations (100, 0.5, and 0.05 µg/mL). Each curve per panel was tested in a specificity assay by treating it with one of four soluble biotherapeutic concentrations in the reaction buffer (buffer A, 7.5; buffer B, 25; buffer C, 100; and buffer D, 400 µg/mL). The optimal soluble biotherapeutic concentration produced by this permutation was 100 µg/mL. This concentration is able to deplete serum with low and high ADA levels as well as low and high circulating drug (from dosing) amounts

depending on whether ADA was spiked into the donor serum sample. Formula 1 was used for naïve donors ($DCP = \text{mean } S/N \text{ percent depletion} + 3 \text{ STD}$), and formula 2 was utilized for donors spiked with ADA regardless of concentration ($DCP = \text{mean } S/N \text{ percent depletion} - 3 \text{ STD}$) (5).

Validation of DCP Using Biotherapeutic Naïve Samples with No ADA

Forty-eight healthy individual donor samples (24 males and 24 females) and 48 disease-specific individual donor samples (24 males and 24 females) were analyzed by two analysts using two plate lots, two MSD instruments, and two conjugate lots (biotinylated and ruthenylated biotherapeutic) as summarized in the design of experiment approach in Table III. The 96 donors were allocated to six plates with gender and population (normal or disease specific) as stratification factors. Each sample was either treated or left untreated with excess biotherapeutic (100 $\mu\text{g/mL}$) and then assayed in parallel in duplicate wells on the same plate. The S/N was derived for both treated and untreated samples. Percent depletion was determined as described before. Data from 96 donors were then used to derive both screening ACP and specificity DCP. The ACP was established as the upper threshold of a one-sided 95% prediction interval of the S/N of untreated samples, and the DCP was established as 100% minus the lower threshold of one-sided 99% prediction interval of the ratio between treated and untreated conditions (T/U). A mixed effect model was employed for the purpose of deriving these prediction thresholds (refer to the [Supplementary](#) section for detailed statistical definitions and calculations). The two populations were evaluated against each other to decide whether a common DCP could be warranted. Data were analyzed using SAS version 9.1.3 on VISTA Professional operating system.

RESULTS

Evaluation and Optimization of Soluble Biotherapeutic Concentration in Specificity Assays in the Presence and Absence of ADA and Circulating Drug

For definition purposes, the term circulating drug refers to the biotherapeutic mAb inherent to the sample often observed during dosing. The term biotherapeutic denotes the

therapeutic mAb used to confirm specificity. To evaluate the impact of circulating ADA concentration on the depletion efficiency, dose–response curves were created by spiking increasing quantities of ADA into PNHS. These concentrations (0.01–160 $\mu\text{g/mL}$) represent a wide range of ADA levels normally observed during the recovery phase of a study where circulating drug is no longer present. Each curve was depleted using one of five soluble biotherapeutic doses (1–400 $\mu\text{g/mL}$). Such biotherapeutic amounts have been used historically to deplete ADA signal in a specificity assay. There was no impact on the background signal when the negative control was tested in the specificity assay (data not shown). An adequate percent depletion of reactivity measured as S/N was achieved by all concentrations of biotherapeutic except for 1 $\mu\text{g/mL}$ (Fig. 1) which was therefore not evaluated in the subsequent optimization experiment.

To evaluate the optimal amount of biotherapeutic needed to deplete samples containing both ADA and circulating drug, mock samples consisting of biotherapeutic in the presence of ADA were prepared using a statistical design of experiment (Table I). The selected concentrations were based on the circulating drug levels present in samples during the dosing phase of a study as estimated by prior pharmacokinetic (PK) measurements. The three ADA concentrations chosen for each panel were representative of high (panel 1, 100 $\mu\text{g/mL}$), moderate (panel 2, 0.5 $\mu\text{g/mL}$), and low (panel 3, 0.05 $\mu\text{g/mL}$) amounts as seen in panels a, b, and c of Fig. 2, respectively. For samples with high ADA (panel 1), the biotherapeutic concentrations of 400 and 100 $\mu\text{g/mL}$ were effective in presence of circulating drug concentrations up to 300 $\mu\text{g/mL}$. For the samples with moderate ADA concentration (panel 2), the effectiveness in depletion could only be observed when circulating drug concentrations were low (10 $\mu\text{g/mL}$), and for samples with low ADA concentration (panel 3), the reduction of

Table III. Experiment DCP Validation

Assay run	Analyst	Plate lot	Conjugate lot	Plate
1	1	1	1	1, 2, 3
			2	4, 5, 6
2	1	2	1	1, 2, 3
			2	4, 5, 6
3	2	1	1	1, 2, 3
			2	4, 5, 6
4	2	2	1	1, 2, 3
			2	4, 5, 6

Forty-eight healthy individual donor samples (24 males and 24 females) and 48 disease-specific individual donor samples (24 males and 24 females) were analyzed by two analysts using two plate lots and two conjugate lots (biotinylated biotherapeutic and ruthenylated biotherapeutic)

Fig. 1. Evaluation of biotherapeutic concentrations in depletion buffer at various ADA levels. Five dose–response curves were created consisting of increasing levels of ADA (0.01, 0.04, 0.16, 0.50, 1.0, 20, 80, 160 $\mu\text{g/mL}$). Each curve was depleted with one out of five biotherapeutic concentrations (1, 7.5, 25, 100, 400 $\mu\text{g/mL}$). The lowest concentration of biotherapeutic (1 $\mu\text{g/mL}$) was not able to reduce the signal at the highest concentration of ADA (last point on the curve) which indicates the possibility of reporting false negatives in a study where high amounts of ADA could be present. Biotherapeutic performance was comparable at 25, 100, and 400 $\mu\text{g/mL}$ in depletion buffers. Note: the ADA concentration was transformed prior to plotting on the X-axis

signal was not effective. The samples with minimal level of detectable ADA or low ADA concentration (0.05 $\mu\text{g/mL}$) did not deplete in a similar pattern as observed for panel 1 and 2. The samples in panel 3 treated with 400 $\mu\text{g/mL}$ of soluble biotherapeutic specifically showed inconsistency in the depletion profiles. This could be attributed to saturation of the sample due to high biotherapeutic amounts. The final optimized concentration of biotherapeutic that could adequately deplete the different concentrations of ADA-induced signal was established at 100 $\mu\text{g/mL}$ for both dosing and recovery sample sets.

Establishing the DCP

Preliminary Evaluation of DCP Using Naïve vs. ADA-Positive Serum Sets

Prior to the validation of the DCP using a large sample set, a preliminary evaluation was performed on a subset of healthy donor serum using various ADA concentrations. The purpose of this pilot evaluation was to understand whether naïve or ADA-containing serum could impact DCP determination. This preliminary assessment was performed using a small number of donors ($n=18$) to improve manageability, but still gain an understanding of the effects of ADA levels in a confirmatory assay. As previously mentioned, the amounts of ADA chosen were 0, 20, and 160 ng/mL. The 20-ng/mL concentration was selected to represent a sample set at assay sensitivity as determined during validation (data not shown). The concentration at 160 ng/mL was chosen to reflect samples with an ADA level higher than assay sensitivity.

The DCPs generated by the preliminary evaluation for 0, 20, and 160 ng/mL ADA were 23%, -7%, and 43%, respectively. The sample set spiked with 20 ng/mL of ADA yields an extremely low DCP (negative in some cases) which can generate false positives. On the other hand, mock samples spiked at 160 ng/mL result in a high DCP (43%) that can potentially produce false negatives when used in a specificity assay. Only samples with strong S/N values (high ADA and low or no circulating drug levels) will be considered positive using this high DCP. The most suitable and yet conservative DCP (23%) was generated utilizing naïve donors (0 ng/mL sample set). It provided a moderate risk in producing false positives or negatives. This DCP offers the most optimal approach to effectively detect true-positive samples regardless of ADA levels and/or affinity. In addition, it addresses samples close to the ACP (low positives) in a highly sensitive validated assay. Hence, following the findings in this preliminary assessment, an extensive validation with a larger set of donors was performed using the naïve donor set.

Validation of DCP Using Naïve Samples with No ADA

The validation of DCP involves inclusion of assay variables that are expected to contribute to the interassay precision leading to the development of a robust DCP. Table III illustrates the design of experiment used to validate and establish a DCP using naïve donors. The difference between the normal donor and diseased donor populations was found to be insignificant at -0.29% with a 90% confidence interval (-1.61%, 1.03%). Therefore, it was

statistically acceptable to support a common DCP for the two populations and was hence combined for the statistical analysis. A mixed effect model was intentionally designed into the experiment to test assay variability due to reagents, analysts, instrumentation, and donor samples. None of the effects except for analyst was statistically significant (P value <0.05). The distribution of the residuals from this model was normal indicating that the normality assumption was met. The analysis showed that the contribution to the total assay variability was 4% from donor-to-donor variability, 40% from plate-to-plate variability, 13% from run-to-run (a combination of analyst and plate lot) variability, and 43% from residual variance (unexplained variability by the statistical model). A final DCP of 19% was established.

Comparing Clinical Study Specific Predoses with the Naïve Donor Validation Data Set

Case Study 1. The DCP validated from biotherapeutic naïve healthy donor-derived samples (with no ADA) described above was tested on a small set of baseline or predose (prior to dosing) samples ($n=16$) from a phase 1 clinical study comprising of healthy subjects. The purpose of this exercise was to ensure that study specific predose/baseline samples were within the range of responses generated using a larger data set during the DCP validation. S/N percent depletion was calculated for the phase 1 study samples and plotted along with validation data (Fig. 3). The S/N percent depletion from the predose study samples was slightly higher than the DCP validation data set ($n=96$). However, predose phase 1 samples were still within the expected parameters justifying the assay performance and assigned criteria.

Comparison of Arbitrary 50% Depletion Threshold vs. DCP Established Using Clinical Study Specific Predoses

Case Study 2. Serum samples ($n=241$, belonging to 56 subjects) from a phase 1 clinical study were analyzed in the screening assay. Twenty-two of these samples were further confirmed in the specificity assay (Fig. 4). Assessment of their S/N value in the screening portion of the confirmatory assay indicated they were low positives as they were close to the ACP. It is noted that the low positivity in the screening assay could also be due to circulating drug interference as well as other matrix-associated soluble factors (10,11). Eighteen of 22 samples tested positive using the arbitrary depletion threshold of 50% (using net ECL depletion as opposed to S/N depletion), and an additional four were positive due to their S/N values falling below ACP postdepletion with drug. However, when the same sample set was analyzed using a DCP of 23% established from naïve donors (see Preliminary evaluation of DCP criteria section [Table II]), only 8 of the original 22 samples were positive. Applying the DCP of 19% (established using 96 naïve donors) did not impact the number of positive results as compared to the DCP from the preliminary exercise where only 18 donors were used. The DCP of -7% (determined using 18 donors spiked with 20 ng/mL ADA) resulted in 22 of 22 samples as positive. On the contrary, the DCP of 43% (estimated by 18 donors spiked with 160 ng/mL ADA) led to all samples being reported

Fig. 2. a–c Optimization of excess soluble biotherapeutic concentration for specificity assay in the presence of circulating biotherapeutic and ADAs. Three ADA panels were prepared (100, 0.5, and 0.05 $\mu\text{g/mL}$) consisting of four serum curves each with increasing concentrations of biotherapeutic (mimicking circulating drug in serum). Each curve per panel was depleted using one of four soluble biotherapeutic concentrations (7.5, 25, 100, and 400 $\mu\text{g/mL}$), giving a final number of 12 curves with 12 different combinations. Panels a–c illustrate that buffer A (7.5 $\mu\text{g/mL}$ of biotherapeutic) does not deplete as robustly as the other three concentrations. Buffers B, C, and D perform comparably in panels a and b, but panel c demonstrates high variability due to low level of ADA (50 ng/mL). This phenomenon may lead to false positives or negatives in response to the already circulating drug. Buffer C offers the most consistent depletion across ADA and circulating drug. Note: the circulating drug concentration was transformed prior to plotting on the X-axis

Fig. 3. Comparing clinical study specific predoses with the naïve donor validation data set. S/N percent depletion was calculated for 16 predose samples from a phase 1 clinical study and for 96 donor samples (48 healthy and 48 disease specific). The plot indicates that although baseline samples depleted a bit higher, they are still within the distribution of naïve donors from the DCP validation experiment

as negative. Once again, the naïve donor-derived DCP demonstrated a more careful approach in a confirmatory assay.

DISCUSSION

Mire-Sluis *et al.* recommend a confirmatory step during immunogenicity assessment of a biotherapeutic that ascertains the specificity of the initial screening response (16,17) to be biotherapeutic specific. The specificity step usually involves the depletion of the antibody signal by adding an excess amount of biotherapeutic (5–9,13,18). The analytical considerations around such assays and statistical methods needed to establish a DCP require further evaluation.

There are two well-recognized approaches for DCP determination in immunogenicity assays (5). One evaluates the amount of biotherapeutic required for confirmation in the samples containing different levels of ADA as well as circulating drug, while the other assesses the amount of biotherapeutic in serum of ADA-negative samples. The former scenario is especially relevant for samples obtained from nonclinical PK and toxicology studies where both high ADA and excess circulating drug are prevalent (11). In this study, we established a DCP utilizing both approaches.

Before establishing the DCP criterion, we first optimized the biotherapeutic concentration needed for treating samples in a specificity assay. This was evaluated by using mock sample sets in presence of ADA alone or ADA and circulating drug. Once the biotherapeutic concentration was

optimized, we further assessed the DCP estimations in three different scenarios: (1) naïve donors (0 ng/mL ADA), (2) naïve donors spiked with 20 ng/mL ADA, and (3) naïve donors spiked with 160 ng/mL ADA. It should be noted that this preliminary exercise was performed using a small data set population ($n=18$). Scenario 1 resulted in a DCP of 23%, and the DCP from scenario 2 and 3 were –7% and 43% respectively. Hence, scenario 1-associated DCP seemed to be the most relevant and appropriately derived criterion. Finally, we validated a DCP (23%) using a large population ($n=96$) of naïve donors.

Even though the naïve donor set or predose samples from study-specific populations can reveal the impact of biotherapeutic on preexisting drug-reactive antibodies, they can overestimate the DCP due to the limited sample set. The DCP validation is performed on a representative population to help assign the appropriate criteria for study sample analysis. During validation, a purposely large population of samples is desirable in order to gain the statistical power needed to assess normality. The findings from case study I illustrate the advantage of using a sufficiently large naïve donor data set and the importance of incorporating variables that contribute to assay variation for DCP estimation. This is because the DCP is a prediction-bound parameter requiring a large sample size that can be applied to day-to-day sample analysis which is subject to the interassay imprecision due to variables dependent on the analysts, reagents, instruments, *etc.* In addition to the naïve donor set, we also evaluated sample sets spiked with two levels of ADA, low (close to assay sensitivity [20 ng/mL]) and high (close to the lowest limit of reliable detection [160 ng/mL]). Some caveats to using the ADA-containing sample sets to establish a DCP include the variability of ADA attributes such as affinity, avidity, and their monoclonal vs. polyclonal nature which can impact the amount of biotherapeutic used for spiking such samples. The preliminary DCP evaluation illustrated that utilizing a low concentration of ADA (20 ng/mL) produced a very low DCP (–7%) where all samples would falsely confirm as positive. On the other hand, the use of a higher amount of ADA (160 ng/mL) would yield a high DCP (43%) in which only extremely robust responses would verify as positive, leading to higher false negatives. Consequently, a careful consideration of the appropriate biotherapeutic concentration and use of the suitable number of donor serum sample set can help in validating and establishing the most accurate DCP for specificity assessments.

Fig. 4. Comparison of arbitrary 50% depletion threshold vs. DCP established using clinical study specific predoses. Study samples ($N=22$) were tested in the confirmatory assay and analyzed using both available depletion options (*net ECL* and *S/N % depletion*). When using *net ECL* depletion (*left half of the plot*), 18 out of 22 samples confirm as positive because their depletion is above the 50% threshold (Note: the remaining four samples [denoted in red] were considered positive even though they were under the 50% threshold because their treated *S/N* value fell below the untreated *S/N* ACP). Eight samples out of 22 would confirm as positive if *S/N %* depletion and the DCP of 23% were employed (*right half of the plot*). When the DCP or criterion II produced by spiking donors with 20 ng/mL of ADA was applied, all samples would be positive due to the unrealistic DCP of –7%. But on the other hand, if the cut point generated by adding 160 ng/mL of ADA to the donor set would be used, none of the samples would confirm positive due to the artificially elevated DCP (43%)

CONCLUSION

Specificity is an important aspect of any immunogenicity assessment when reporting an ADA-positive sample. The specificity of response can be confirmed by adding an excess amount of biotherapeutic to deplete the ADA-associated signal (19). In this study, we have evaluated a subset of potential influencing factors that can impact the specificity of an ADA assay. We also attempted to optimize the amount of the biotherapeutic required to establish and statistically validate a DCP. This specificity evaluation was performed utilizing PNHS representative of ADA-positive and negative samples at both dosing and recovery phase of the study. A DCP criterion using naïve serum samples with no ADA or

circulating drug was considered the most relevant to establish specificity of an ADA-positive sample. We have also shown that an accurate determination of ADA relies on many other factors, and at high circulating drug concentrations, the optimized depletion method may still have limitations to confirm specificity of ADA.

ACKNOWLEDGMENTS

The authors of this work would like to thank the following Amgen team members for their support and contribution: Marta Starcevic, Ami Bautista, and Steven J. Swanson. All work has been funded by Amgen Inc. and performed by employees of Amgen Inc. that may hold company stock.

REFERENCES

1. CBER UFC. Assay development for immunogenicity testing of therapeutic proteins 2009.
2. EMEA. Guideline on immunogenicity assessment of biotechnology derived therapeutic proteins. 2007.
3. Chamberlain P, Mire-Sluis A. An overview of scientific and regulatory issues for the immunogenicity of biological products. *Dev Biol.* 2003;112:3–11.
4. CHMP. Guideline of immunogenicity assessment of biotechnology-derived therapeutic proteins. EMEA Guidance Document. 2008.
5. Shankar G, Devanarayan V, Amaravadi L, Barrett YC, Bowsher R, Finco-Kent D, *et al.* Recommendations for the validation of immunoassays used for detection of host antibodies against biotechnology products. *J Pharm Biomed Anal.* 2008;48(5):1267–81.
6. Swanson S, Chirmule N. Assessing specificity for immunogenicity assays. *Bioanalysis.* 2009;1:611–7.
7. Ponce R *et al.* Immunogenicity of biologically-derived therapeutics: assessment and interpretation of nonclinical safety studies. *Regul Toxicol Pharmacol: RTP.* 2009;54:164–82.
8. Sanchez S *et al.* Strategy to confirm the presence of anti-erythropoietin neutralizing antibodies in human serum. *J Pharm Biomed Anal.* 2011;55:1265–74.
9. Koren E, Smith HW, Shores E, Shankar G, Finco-Kent D, Rup B, *et al.* Recommendations on risk-based strategies for detection and characterization of antibodies against biotechnology products. *J Immunol Methods.* 2008;333(1–2):1–9.
10. Bautista A *et al.* Impact of matrix-associated soluble factors on the specificity of the immunogenicity assessment. *Bioanalysis.* 2010;2(4):721–31.
11. Zhong ZD *et al.* Identification and inhibition of drug target interference in immunogenicity assays. *J Immunol Methods.* 2010;355:21–8.
12. Wang YM, Fang L, Zhou L, Wang J, Ahn HY. A survey of applications of biological products for drug interference of immunogenicity assays. *Pharm Res.* 2012;29:3384–92.
13. Grodzki AC, Berenstein E. Antibody purification: affinity chromatography—protein A and protein G Sepharose. *Methods Mol Biol.* 2010;588:33–41.
14. Moxness M *et al.* Immunogenicity testing by electrochemiluminescent detection for antibodies directed against therapeutic human monoclonal antibodies. *Clin Chem.* 2005;51:1983–85.
15. Tatarewicz S *et al.* Rheumatoid factor interference in immunogenicity assays for human monoclonal antibody therapeutics. *J Immunol Methods.* 2010;357(1–2):10–6.
16. Mire-Sluis AR, Barrett YC, Devanarayan V, Koren E, Liu H, Maia M, *et al.* Recommendations for the design and optimization of immunoassays used in the detection of host antibodies against biotechnology products. *J Immunol Methods.* 2004;289(1–2):1–16.
17. Smith H, Moxness M, Marsden R. Summary of confirmation cut point discussions. *AAPS J.* 2011;13(2):227–9.
18. Peng K *et al.* Clinical immunogenicity specificity assessments: a platform evaluation. *J Pharm Biomed Anal.* 2011;54(3):629–35.
19. Wakshull E, Coleman D. Confirmatory immunogenicity assays. In: Tovey M, editor. *Detection and quantification of antibodies to biopharmaceuticals: practical and applied considerations.* New York: Wiley; 2011.