RESEARCH ARTICLE

Implementation of a Three-Way Comparability Assessment for a Bioanalytical Anti-Drug Antibody Method

Rosanna S. Kwok[1] · Ihsan Nijem[1] · Ann Brady[1] · Robert Hendricks[1]

Abstract

Immunogenicity evaluation is a critical part of drug development. Regulatory guidelines from multiple health agencies provide recommendations for the development and validation of anti-drug antibody (ADA) assays to assess immunogenicity in clinical trials. These recommendations primarily describe an ADA method run in one bioanalytical laboratory supporting a biotherapeutic molecule; however, there are increasing instances that may necessitate the support of the ADA method being run in more than one laboratory. A program can rapidly expand into multiple clinical studies within one or multiple countries, where the most appropriate way to support the program is by having multiple laboratories perform ADA sample analysis. In addition, there may be certain country-specific challenges that may make it infeasible to transport samples outside of the country for analysis. China for example has a lengthy sample exportation process that has potential to negatively impact study timelines. If multiple laboratories analyze samples using the same ADA method, comparable method performance should be established. Here, we describe a three-way assessment of ADA assay comparability between two US-based bioanalytical laboratories and one based in China.

Keywords ADA · comparability · immunogenicity · method validation · tiragolumab

Introduction

Immunogenicity assessment, a key component in the process of biotherapeutic development, is done by measuring the body's production of anti-drug antibodies (ADAs) in response to treatment with a biologic medicine. It is imperative to understand the immunogenicity of a drug during research and development since immunogenic responses could have unwanted impacts on efficacy, safety, and pharmacokinetics (PK). ADA assays are developed to test either animal samples before first-in-human studies or human samples during clinical trials. Since the immunogenic response is polyclonal in nature and varies across individuals (1), it is impossible to develop a quantitative assay with a true-positive control that reflects the entire study population. Rather, ADA assays are qualitative and developed using surrogate positive controls that are spiked into a matrix that represents those used for the study population. An ADA method for immunogenicity testing typically follows a tiered testing strategy consisting of screening, confirmatory tests, and further characterization by titer measurements. For a sample to be considered positive, it must be positive in both the screening and confirmation assays. The threshold for positive/negative determination (cut point [CP]) is statistically determined using samples from treatment-naïve subjects and is set to a level that is designed to produce false positives. For screening assays, CPs are typically set to produce 5% false positives; for confirmatory assays, the CP is set to produce 1% false positives. Before analyzing clinical samples, ADA methods are validated based on guidance from health authorities such as the United States (US) Food and Drug Administration (FDA), China's National Medical Products Administration (NMPA), and the European Medicines Agency (EMA) (2–4).

When a program is in early clinical stages, there is usually only a single laboratory supporting the clinical ADA bioanalysis for the study. However, if the program expands into multiple clinical studies, and multiple countries, the bioanalysis needs may exceed the support capacity of a single laboratory. Therefore, it could be necessary to establish

✉ Robert Hendricks
  Hendricks.robert@gene.com

1 Department of BioAnalytical Sciences, Genentech Inc., 1 DNA Way, South San Francisco, California 94080-4990, USA

∠ Springer

the ADA method in additional laboratories, including in other countries, to mitigate challenges with country-specific requirements.

For the late-stage clinical development for our tiragolumab program, it was necessary to evaluate more than one laboratory to support clinical sample testing. The program was originally supported at one bioanalytical laboratory (BioA Lab 1), and was to expand to two additional bioanalytical laboratories (BioA Lab 2 and BioA Lab 3) to (1) mitigate the risk of exceeding the capacity of a single US laboratory to run the increased number of samples from numerous studies and (2) better support new clinical trials with patient populations predominantly in China and avoid the logistical challenges of exporting samples.

The two US-based contract research organizations (CROs) would be able to support the sample analysis from studies conducted in the USA, European Union, and non-China Asia–Pacific (APAC) countries. The CRO in China would support analysis of Chinese and APAC patients, mitigating the challenges involved in sending samples for analysis in the USA. Samples from APAC countries can be shipped into China, which is a more straightforward process. In support of this program, our bioanalytical comparison plan was intended to build confidence in the ADA method being run simultaneously by three separate laboratories with replicable results.

Currently, there is limited guidance from health authorities regarding multiple laboratories using the same methods to support a program; thus, we describe our approach and the techniques used to address satisfactory reproducibility for our methodology. The FDA guidance (2019) gives a brief recommendation on this topic — "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," and mentions that "comparable assay performance, including sensitivity, drug tolerance, and precision, should be established between laboratories" (2). China's guidance, "Technical Guidance for Immunogenicity Studies of Drugs," states similarly that comparable assay performance parameters (sensitivity, drug tolerance, and precision) should be established when using more than one laboratory to generate data using the same ADA method (3).

Previous experience with method comparability evaluations in literature are similarly limited, with the main focus being on the initial method transfer or assay platform comparison rather than comparability evaluations for the purpose of utilizing multiple laboratories for study support (5–7). In one study, a two-way method comparability was conducted between one laboratory in the USA and one in China (6) and incurred samples were not used during the comparability evaluation as it was not feasible to do. Instead, samples spiked with ADA positive control were used for the

evaluation. In our study we opted to use incurred clinical samples (samples selected from a tiragolumab clinical study) for comparability, since ADA methods use surrogate positive controls that may not accurately represent ADA responses in patients.

Here, we describe and show results for our plan to evaluate ADA method comparability between three separate CRO laboratories (Fig. 1). The strategy included independently fully validating the same ADA method at each of the three bioanalytical laboratories (with each establishing their own CP factors), with emphasis on assessing whether assay performance of sensitivity, drug tolerance, and precision parameters were similar as recommended by health authority guidance (2, 3). We also assessed the overall comparability using incurred clinical samples, following a tiered analysis approach to interpret the data from three different laboratories. Incurred samples allowed us to test the comparability of the methods using patient samples and to determine whether the ADA results generated at different laboratories would lead to varying interpretations of the ADA impacts on clinical outcomes, such as efficacy, safety, and PK.

Methods

ADA Assay Format

The enzyme-linked immunosorbent assay (ELISA) method for the detection of ADA uses a bridging format. The ADAs are incubated with biotinylated and the DIGylated tiragolumab to form complexes and are then immobilized to a streptavidin-coated 96-well plate. A horseradish peroxidase (HRP)-conjugated mouse anti-DIG antibody is added and followed by color development with 3,3′,5,5′-tetramethylbenzidine (TMB).

Tiragolumab drug materials are conjugated to biotin (Pierce™ Sulfo-NHS-LC-biotin; Cat. No. 21327) and digoxigenin (Invitrogen 3-amino-3-deoxydigoxigenin hemisuccinamide succinimidyl ester; Cat. No. A2952) at a 10:1 challenge ratio. A master mix was prepared as a 1:1 mixture of biotinylated tiragolumab and DIGylated tiragolumab conjugates, each at 3 μg/mL, in assay diluent (phosphate-buffered saline with 0.5% bovine serum albumin (BSA), 0.01% ProClin™ 300, and 0.05% Tween 20). Equal volumes of master mix and diluted sample were mixed in a polypropylene plate and incubated overnight.

For the screening assay, controls and samples were diluted to a minimal required dilution (MRD) of 1/20 into assay diluent and then combined with master mix. For the confirmatory assay, the plate controls/samples were diluted tenfold and then combined 1:1 with a 200 μg/mL tiragolumab working solution. The effective MRD was thus 1/20, and the final tiragolumab concentration 100 μg/mL. This reaction
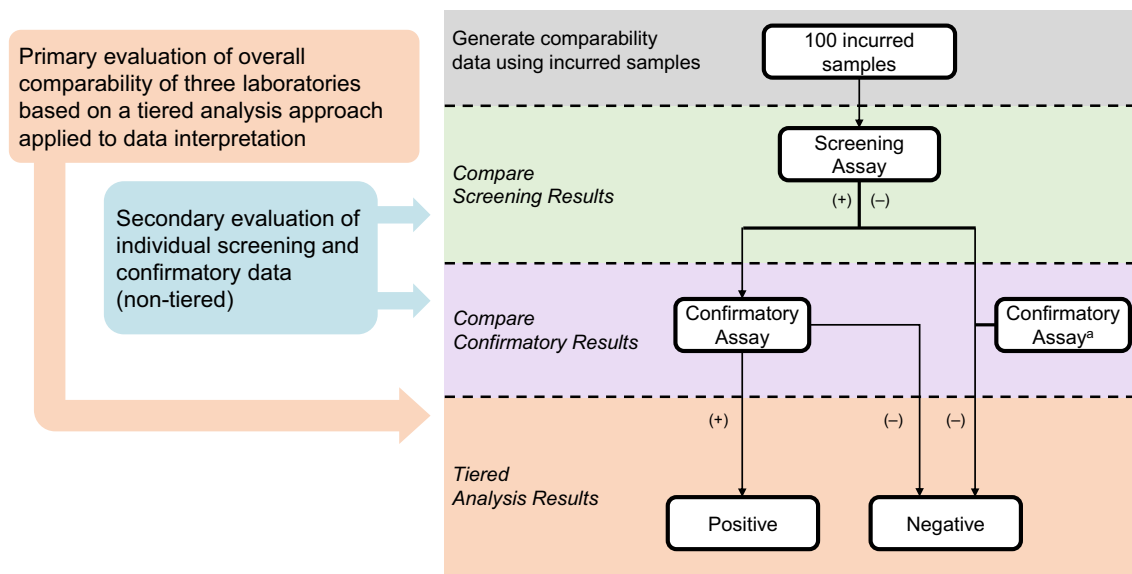
Fig. 1 Schematic of three-way comparability strategy. Each of three BioA Lab generated comparability datasets from 100 incurred samples. Primary and secondary evaluations were conducted using comparability datasets. Each BioA Lab independently validated ADA method prior to running incurred samples. [a]Data for confirmatory assay evaluation only. ADA anti-drug antibody; BioA bioanalytical

mixture was then incubated in a washed high-binding streptavidin plate (StreptaWell™ High Bind; Roche, Cat. No. 11989685001). After washing, bound antibody–conjugate complexes were detected by adding 360 ng/mL mouse monoclonal anti-DIG HRP conjugate (Jackson ImmunoResearch, West Grove, PA, USA; Cat. No. 200–0320156/147166) to the streptavidin plate. After washing, a signal was generated by adding TMB Microwell Peroxidase Substrate System (KPL; SeraCare, Milford, MA, USA; Cat No. 50–76-00) and the reaction was stopped with 1 M phosphoric acid. Absorbance was measured at 450/630 nm on a spectrophotometer.

Assay Controls

A goat anti-tiragolumab complementarity-determining region (CDR) antibody was used for the positive control. Plate control concentrations were as follows: screening low-positive control (LPC) was 15 ng/mL at BioA Lab 1 and 10 ng/mL at BioA Labs 2 and 3. The difference in LPC concentrations is due to an update in our practice for setting the LPC level for ADA assays between the time of the assay validation in BioA Lab 1 (2017) and BioA Labs 2 and 3 (2021 and 2022, respectively). Since ADA assays are qualitative and the LPC is used for plate acceptability instead of positive/negative determination, these small differences should not have an impact on assay results. The three assays have LPC/normalization control (NC) ratio acceptance limits that overlap at the three laboratories; therefore, any concentration differences should not impact assay acceptance.

Confirmatory LPC was 5.0 ng/mL at BioA Lab 1, 3.0 ng/mL at BioA Lab 2, and 2.4 ng/mL at BioA Lab 3. High-positive control (HPC) for both screening and confirmatory was 100 ng/mL in all three laboratories. A human serum pool was used for the negative control or normalization control (NC). Individual and pooled human sera were purchased from BioIVT.

Method Transfer/Validation

The tiragolumab ADA method was developed at Genentech, Inc. and subsequently transferred to and validated at a US-based CRO (BioA Lab 1) in 2017. At the time, validation criteria were based on industry white papers and health authority draft guidance (8–11). Given the increasing bioanalytical support needed for multiple clinical trials, the assay was validated at another US-based CRO (BioA Lab 2) in 2021 according to updated FDA guidance (2). Finally, to accommodate bioanalysis of tiragolumab ADA samples originating from Chinese sites, the method was validated at a third CRO in China (BioA Lab 3) in 2022 following China's NMPA guidance document (3). The following parameters were to be validated: screening CPs (sCPs), confirmatory CPs (cCPs), relative sensitivity, relative drug tolerance, hook effect, specificity (interference/cross-reactivity), matrix interference (hemolysis and lipemia interference/cross-reactivity), selectivity, precision, robustness (minimum and maximum incubation times), and stability.

Before validation could begin at BioA Labs 2 and 3, a new lot of NC needed to be identified. To minimize potential impacts on the assay, we aimed to identify a new lot that produced similar instrument responses to the original NC. BioA Lab 2 identified a new lot that had optical density (OD) values that were similar (within ± 10%) to the one used by BioA Lab 1 and were therefore shared between both methods at BioA Labs 2 and 3. Lots of BSA were different between the three laboratories, but were also evaluated to produce similar signals in each laboratory.

All three laboratories used the same assay plates (StreptaWell™ High Bind Transparent 96-well Plate, Roche, Product No. 11989685001). The positive control source material was the same for all three laboratories. The capture and detection source material lots were the same in BioA Labs 2 and 3 but differed from the lot used by BioA Lab 1. The new lots used in BioA Labs 2 and 3 were evaluated to ensure that HPC titer and LPC/NC ratios fell within control acceptance ranges set by BioA Lab 1.

Laboratory-specific CPs were established at the three bioanalytical laboratories using commercially purchased samples with disease states matching the expected indications for tiragolumab clinical trials. It is important to note that the 120 commercially purchased CP samples (BioIVT, New York, USA) used in 2017 during the initial validation at BioA Lab 1 were no longer available when we started method validation at BioA Lab 2 in 2021. Therefore, a different set of 100 commercially purchased samples (BioIVT, New York, USA and ProteoGenex, California, USA) were used for determining the CP at BioA Lab 2. The indications of these commercial disease state samples consisted of the same ones used to establish the CP at BioA Lab 1, with the addition of several indications that tiragolumab clinical trials had expanded to by 2021. We attempted to use the same 100 CP samples when performing the validation at BioA Lab 3 in 2022; however, we were only able to use 56 overlapping samples because the other 44 samples did not have adequate documents required for importation into China. Because of this, 44 new CP samples were sourced by BioA Lab 3 to ensure appropriate licensure (BioIVT, New York, USA). Although we had to source 44 new CP samples, they consisted of the same indications. BioA Lab 2 and BioA Lab 3 used the same NC pool.

The sCP and cCP factors were statistically derived based on a target untreated positive rate of approximately 5% and 1%, respectively. The sCP is calculated for each plate using the NC multiplied by the screen CP factor. The cCP factor is a fixed inhibition percentage threshold used to confirm that the ADA detected is positive against tiragolumab. A sample with an inhibition percentage greater than or equal to the fixed inhibition threshold percentage from validation is determined to be positive for ADAs to tiragolumab.

The assessment of assay screen sensitivity was done by running serial dilutions of the HPC and evaluating the lowest concentration that consistently produced a result above the plate-specific CP. Confirmatory sensitivity was determined by spiking tiragolumab drug into HPC curves. The confirmatory sensitivity was determined as the most diluted sample that confirmed as ADA positive (%inhibition ≥ cCP). Final relative sensitivity for both assays was calculated from the mean sensitivity of all qualified curves (Online Resource 1).

Expectations for Sensitivity, Drug Tolerance, and Precision

Sensitivity, drug tolerance, and precision were parameters suggested by the FDA and NMPA guidelines ([2], [3]) to evaluate when using multiple laboratories to support sample analysis in a clinical study. Our expectations for these parameters were as follows: in order for these parameters to be considered similar, all three laboratories should (1) demonstrate sensitivity equal to or better than the FDA recommended level of 100 ng/mL, (2) be able to detect 100 ng/mL of ADA positive control in the presence of drug levels higher than expected trough drug levels (expected range, ~ 30–100 μg/mL), and (3) demonstrate inter- and intra-run precision with coefficient of variation (CV) within 20%.

Approach to Assessing Comparability

As this was the first ADA comparability study conducted at Genentech between three laboratories and there is currently limited health authority guidance regarding comparing ADA assays, an exploratory approach was taken to assess comparability rather than setting strict comparability acceptance criteria. Due to the qualitative nature of ADA assays, and as it was anticipated that BioA Labs 2 and 3 would be used significantly to support sample testing for this program in the future, we opted to not set a reference laboratory in our comparison. Our criteria may evolve as more data is gathered to reflect the intrinsic assay variability between laboratories for ELISA-based ADA assays.

After validation was completed in the three laboratories, comparability datasets were generated using 100 incurred samples from patients dosed with tiragolumab from a phase 1b study ([12], [13]). BioA Lab 1 originally performed the clinical sample testing for these samples. In this phase 1b study, a low incidence of ADAs toward tiragolumab was observed (~ 1%) with only six samples confirmed positive after tiered analysis. From this study, we selected a total of 50 screen negatives and 50 screen positives, which included the six confirmed positive samples (1 baseline and 5 post-baseline), for the comparability evaluation. The sample OD spanned the entire OD range found during the phase 1b study. The selection of 50 screen positives and 50 screen negatives is an

oversampling of screen positives, since at the time of sample selection the clinical study had 111 screen positives out of the over 1000 post-treatment ADA samples analyzed. This was done to allow the comparison of both screen positives and screen negatives. Given that ADA impact assessments often compare ADA negative (ADA−) *vs*. ADA positive (ADA+) patients, it was critical to assess if the newer laboratories were identifying an increased number of ADA false positives that could potentially affect the ADA impact analysis.

All 100 samples were subsequently analyzed in the screening assay and confirmatory assay (regardless of screen result) in all three laboratories (Fig. 1). Samples that screened negative were confirmed so that if a discrepancy was found between the laboratories, insight into if this discrepancy was driven by the screening or confirmation assay steps could be gained. Samples were determined to be positive or negative using the laboratory-specific CPs generated during validation. Each laboratory applied the same plate acceptance criteria set for study phase bioanalysis to generate their comparability dataset. The few ADA+ samples reported in the clinical study (six positives) all had low titers (four with a titer of < 1.3, one with titer of 1.34, and one with a titer of 2.28) (13); therefore, a titer comparison was not included in the analysis.

Primary and Secondary Evaluations of Comparability Datasets

The primary evaluation of this comparison was to determine clinical comparability: whether the results following a tiered analysis approach (applied during data interpretation) would yield similar clinical sample results (Fig. 1). To be considered clinically comparable, we expected that the majority of samples to have the same classification in all three labs without any singular laboratory detecting a significant additional number of positive samples. The secondary evaluation was on the screening and confirmatory steps individually. Our comparability strategy is more extensive than the typical tiered sample testing approach used for clinical sample analysis. This additional approach allows for a comprehensive examination of overall tiered analysis results as well as an understanding of the individual screening and comparability steps. The overall ADA+/− status generated by the three laboratories and the independent results of screening and confirmatory samples were analyzed for concordance.

Results

Validation Results

The ADA assay was successfully validated at each laboratory. The sCPs were 1.16, 1.75, and 1.56 and the cCPs were 21.6%, 23.5%, and 19.5% inhibition for BioA Labs 1, 2, and 3, respectively (Table I). The relative sensitivity of the screening assay at BioA Labs 1, 2, and 3 were 3.69, 4.98, and 3.35 ng/mL, respectively (Table I). The relative sensitivity of the confirmation assay was 2.90, 1.37, and 1.62 ng/mL at BioA Labs 1, 2, and 3, respectively (Table I). All three laboratories showed a screening sensitivity < 5 ng/mL and confirmatory sensitivity < 3 ng/mL (Table I). All three laboratories had sensitivities well below the health authority recommended 100 ng/mL (2), so were considered acceptable.

For drug tolerance, the screening assay run in all three laboratories was able to detect 100 ng/mL control in the presence of 250 µg/mL tiragolumab. The confirmation assay was able to detect 100 ng/mL control in the presence of 500, 500, and 250 µg/mL of tiragolumab at BioA Labs 1, 2, and 3, respectively (Table I). Thus, the validated drug tolerance for screening and confirmation assays in all three laboratories met expectations of being able to detect 100 ng/mL of control in at least the estimated drug trough level (~ 30–100 ug/mL). The inter-assay precision of the LPC OD, LPC signal to NC ratios, and %inhibition (confirmatory LPC and confirmatory HPC, and HPC titer at the three laboratories) were all within the 20% CV recommended in the FDA guidance (Table I) (2).

Primary Evaluation of Overall Clinical Comparability

The primary evaluation in this study consisted of comparing the three laboratories following a tiered approach. Following this approach, any sample with a screen negative would be given the final result of ADA−; a sample would only be considered ADA+ if both screening and confirmatory assays showed a positive result (Fig. 1). This reflects the standard testing procedure used during clinical immunogenicity assessment and is crucial to understand when investigating how different laboratory results compare. If the three datasets led to similar conclusions of ADA status, we could consider that the assay demonstrated clinical comparability, as the three laboratories would lead to similar clinical interpretations of data.

We first conducted three comparisons, each between two sets of laboratory results, and found high sample concordance in the final ADA result (after tiered analysis). Between BioA Lab 1 and BioA Lab 2, 96/100 sample results were in agreement (94 ADA−, 2 ADA+; Table II). BioA Lab 1 and BioA Lab 3 showed agreement in 95/100 samples (92 ADA−, 3 ADA+; Table III), and BioA Labs 2 and 3 showed agreement in 97/100 samples (95 ADA−, 2 ADA+; Table IV). Altogether, all three laboratories had consensus among 94/100 samples (92 ADA−, 2 ADA+) (Table V).

Table I Cross-Laboratory Validation of the ADA Assay Validation Data

BioA Lab	Cut points		Relative sensitivity (ng/mL)		Drug tolerance (μg/mL)[a]		Inter-assay precision (%)					
	Screening	Confirmation	Screening	Confirmation	Screening	Confirmation	LPC OD[b]	LPC/NC ratio	HPC titer[c]	cLPC[d,e]	cHPC[c,d]	
1	1.16	21.6%	3.69	2.90	250	500	13.7	10.2	2.1	9.7	1.7	
2	1.75	23.5%	4.98	1.37	250	500	15.0	5.1	7.6	14.1	1.8	
3	1.56	19.5%	3.35	1.62	250	250	8.8	5.0	0.2	5.2	0.9	

ADA anti-drug antibody, *cHPC* confirmatory high-positive control, *cLPC* confirmatory low-positive control, *HPC* high-positive control, *LPC* low-positive control, *NC* normalization control, *OD* optical density

[a] 100 ng/mL of positive surrogate control is detected in the presence of the tiragolumab concentration tested

[b] LPC concentration = 15 ng/mL at BioA Lab 1, 10 ng/mL at BioA Labs 2 and 3

[c] HPC/cHPC concentration = 100 ng/mL

[d] Precision of inhibition percentages

[e] cLPC concentration: 5 ng/mL at BioA Lab 1, 3 ng/mL at BioA Lab 2, and 2.4 ng/mL at BioA Lab 3

Table II Comparability of Tiered Analysis Results Between BioA Lab 1 and BioA Lab 2[a]

BioA Lab 1		Tiered approach	
	Positive	4	2
	Negative	94	0
		Negative	Positive
		BioA Lab 2	

[a] Two-way comparisons were conducted between the bioanalytical laboratories. Tables show the number of ADA negative (bottom left quadrants) and positive samples (top right quadrants) that the laboratories have in agreement, and samples with disagreements (top left and bottom right quadrants)

ADA anti-drug antibody, *BioA* bioanalytical

Table III Comparability of Tiered Analysis Results Between BioA Lab 1 and BioA Lab 3[a]

BioA Lab 1		Tiered approach	
	Positive	3	3
	Negative	92	2
		Negative	Positive
		BioA Lab 3	

[a] Two-way comparisons were conducted between the bioanalytical laboratories. Tables show the number of ADA negative (bottom left quadrants) and positive samples (top right quadrants) that the laboratories have in agreement, and samples with disagreements (top left and bottom right quadrants)

ADA anti-drug antibody, *BioA* bioanalytical

Table IV Comparability of Tiered Analysis Results Between BioA Lab 2 and BioA Lab 3[a]

BioA Lab 2		Tiered approach	
	Positive	0	2
	Negative	95	3
		Negative	Positive
		BioA Lab 3	

[a] Two-way comparisons were conducted between the bioanalytical labs. Tables show the number of ADA negative (bottom left quadrants) and positive samples (top right quadrants) that the labs have in agreement, and samples with disagreements (top left and bottom right quadrants)

ADA anti-drug antibody, *BioA* bioanalytical

Table V Primary Evaluation of Sample Agreement Following Tiered Analysis[a]

Overall sample status

Sample ID[b]	BioA Lab 1	BioA Lab 2	BioA Lab 3	Overall agreement? (yes/no)
1001	Negative	Negative	Negative	Yes
1002	Negative	Negative	Negative	Yes
1003	Negative	Negative	Negative	Yes
1004	Negative	Negative	Negative	Yes
1005	Negative	Negative	Negative	Yes
1006	Negative	Negative	Negative	Yes
1007	Negative	Negative	Negative	Yes
1008	Negative	Negative	Negative	Yes
1009	Negative	Negative	Negative	Yes
1010	Negative	Negative	Negative	Yes
1011	Negative	Negative	Negative	Yes
1012	Negative	Negative	Negative	Yes
1013	Negative	Negative	Negative	Yes
1014	Negative	Negative	Negative	Yes
1015	Negative	Negative	Negative	Yes
1016	Negative	Negative	Negative	Yes
1017	Negative	Negative	Negative	Yes
1018	Negative	Negative	Negative	Yes
1019	Negative	Negative	Negative	Yes
1020	Negative	Negative	Negative	Yes
1021	Negative	Negative	Negative	Yes
1022	Negative	Negative	Negative	Yes
1023	Negative	Negative	Negative	Yes
1024	Negative	Negative	Negative	Yes
1025	Negative	Negative	Negative	Yes
1026	Negative	Negative	Negative	Yes
1027[c]	Positive	Negative	Negative	No
1028	Negative	Negative	Negative	Yes
1029	Negative	Negative	Negative	Yes
1030	Negative	Negative	Negative	Yes
1031	Negative	Negative	Negative	Yes
1032	Negative	Negative	Negative	Yes
1033	Negative	Negative	Negative	Yes
1034	Negative	Negative	Negative	Yes
1035	Negative	Negative	Negative	Yes
1036	Negative	Negative	Negative	Yes
1037	Negative	Negative	Negative	Yes
1038	Negative	Negative	Negative	Yes
1039	Negative	Negative	Negative	Yes
1040	Negative	Negative	Negative	Yes
1041	Negative	Negative	Negative	Yes
1042[c]	Positive	Negative	Positive	No
1043	Negative	Negative	Negative	Yes
1044	Negative	Negative	Negative	Yes
1045	Negative	Negative	Negative	Yes
1046	Negative	Negative	Negative	Yes

Table V (continued)

Overall sample status

Sample ID[b]	BioA Lab 1	BioA Lab 2	BioA Lab 3	Overall agreement? (yes/no)
1047	Negative	Negative	Negative	Yes
1048	Negative	Negative	Negative	Yes
1049	Negative	Negative	Negative	Yes
1050	Negative	Negative	Negative	Yes
1051	Negative	Negative	Negative	Yes
1052	Negative	Negative	Negative	Yes
1053[c]	Positive	Negative	Negative	No
1054	Negative	Negative	Negative	Yes
1055	Negative	Negative	Negative	Yes
1056	Negative	Negative	Negative	Yes
1057	Negative	Negative	Negative	Yes
1058	Negative	Negative	Negative	Yes
1059	Negative	Negative	Negative	Yes
1060	Negative	Negative	Negative	Yes
1061	Negative	Negative	Positive	No
1062	Negative	Negative	Negative	Yes
1063	Negative	Negative	Negative	Yes
1064	Negative	Negative	Negative	Yes
1065	Negative	Negative	Negative	Yes
1066	Negative	Negative	Negative	Yes
1067	Negative	Negative	Negative	Yes
1068	Negative	Negative	Negative	Yes
1069	Negative	Negative	Negative	Yes
1070	Negative	Negative	Negative	Yes
1071	Negative	Negative	Negative	Yes
1072	Negative	Negative	Negative	Yes
1073	Negative	Negative	Negative	Yes
1074	Negative	Negative	Negative	Yes
1075	Negative	Negative	Negative	Yes
1076	Negative	Negative	Negative	Yes
1077	Negative	Negative	Negative	Yes
1078	Negative	Negative	Negative	Yes
1079	Negative	Negative	Negative	Yes
1080	Negative	Negative	Negative	Yes
1081	Negative	Negative	Negative	Yes
1082	Negative	Negative	Negative	Yes
1083	Negative	Negative	Negative	Yes
1084	Negative	Negative	Negative	Yes
1085	Negative	Negative	Negative	Yes
1086[c,d]	Positive	Positive	Positive	Yes
1087	Negative	Negative	Negative	Yes
1088[c,d]	Positive	Positive	Positive	Yes
1089[c]	Positive	Negative	Negative	No
1090	Negative	Negative	Negative	Yes
1091	Negative	Negative	Negative	Yes
1092	Negative	Negative	Negative	Yes
1093	Negative	Negative	Negative	Yes

Table V (continued)

Overall sample status

Sample ID[b]	BioA Lab 1	BioA Lab 2	BioA Lab 3	Overall agreement? (yes/no)
1094	Negative	Negative	Negative	Yes
1095	Negative	Negative	Positive	No
1096	Negative	Negative	Negative	Yes
1097	Negative	Negative	Negative	Yes
1098	Negative	Negative	Negative	Yes
1099	Negative	Negative	Negative	Yes
1100	Negative	Negative	Negative	Yes
Total samples with overall ADA status agreement				94

[a]The ADA positive/negative determinations by the three bioanalytical laboratories are shown in this table, along with whether there was agreement across all three laboratories

[b]Sample IDs have been anonymized

[c]Based on clinical data generated by BioA Lab1, these samples were reported as ADA+ in the phase 1b study that they were selected from

[d]One of two samples where ADA positivity was detected at all three laboratories

ADA anti-drug antibody, *BioA* bioanalytical

There were 6/100 samples that did not have agreement (Table V). These samples consisted of one baseline and five post-baseline samples from a total of six patients. The screening, confirmatory, and overall ADA result (following tiered analysis of screening and confirmatory data) of these samples were evaluated to identify potential reasons for discrepancies (Online Resource 2). Instances where sample signals hovered close to the sCP or cCP in at least one laboratory were noted (e.g., samples 1042, 1027, 1053, 1095). Proximity to either sCP or cCP may be a reason why the ADA results in these samples were not consistent in all three laboratories. Despite these six discrepancies, we consider 94/100 to be a high degree of agreement, and we feel that the differences in ADA+ samples between laboratories would not be impactful to clinical interpretations of tiragolumab ADA data. Thus, in the primary evaluation we consider the assay to demonstrate clinical comparability.

Secondary Evaluation of Screening and Confirmatory Data

The secondary evaluation in our comparability was to examine the screening and confirmatory steps individually
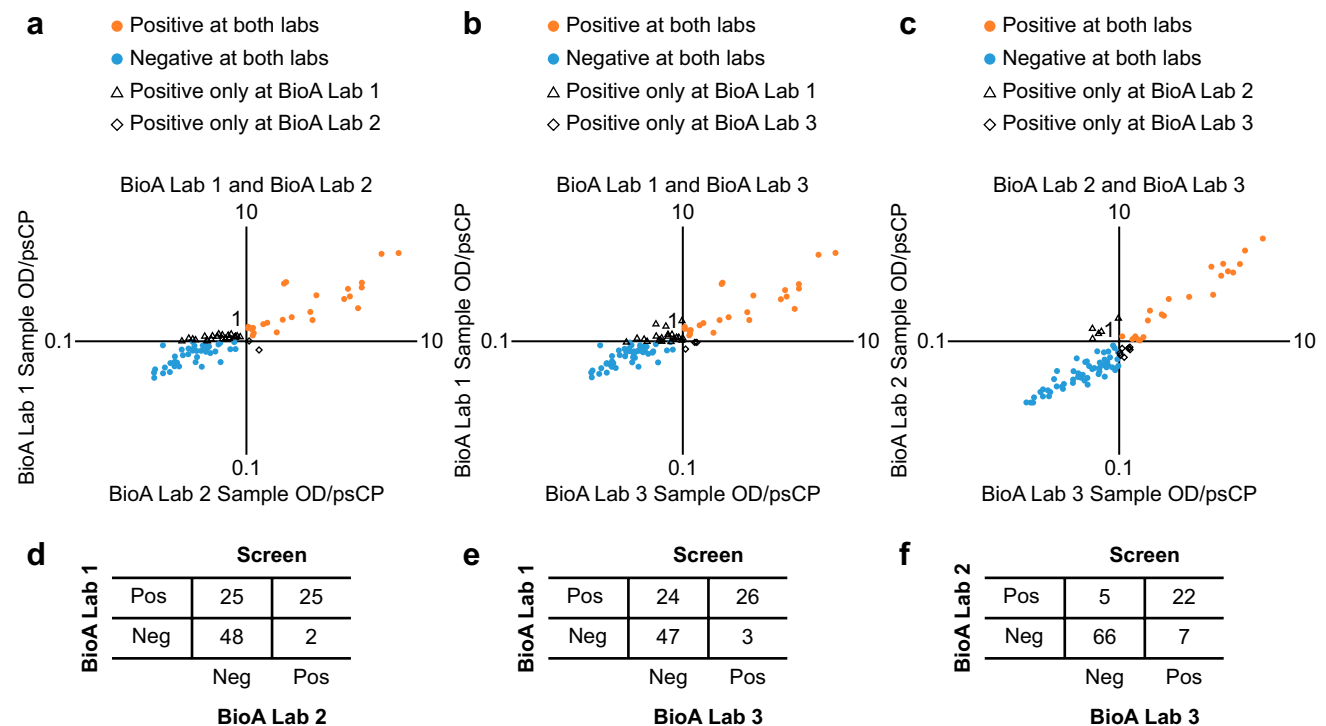
Fig. 2 Secondary evaluation — screening results. One-hundred incurred samples from a tiragolumab phase 1b study were selected for comparability evaluation of the tiragolumab ADA method. All three laboratories independently validated the method with lab-specific CPs. BioA Lab 1 was used to support the initial sample analysis. Fifty screen-negative and 50 screen-positive samples with OD covering the entire OD range were selected for ADA method comparability. Of the 100 samples, the 6 confirmed positive samples from the phase 1b study were also included. Shown are the results plotted as sample OD/psCP (**a**, **b**, and **c**) as well as sample concordance (**d**, **e**, and **f**) on the 100 incurred samples. ADA anti-drug antibody; BioA bioanalytical; CP cut point; OD optical density; psCP plate specific cut point

(Fig. 1). Screening results between BioA Lab 1 and BioA Lab 2 (both in the USA) showed the same screening results for 73/100 samples (48 ADA−, 25 ADA+; Fig. 2a, d). We found the same 73/100 sample concordance when comparing BioA Lab 1 (USA) and BioA Lab 3 (China) (47 ADA−, 26 ADA+; Fig. 2b, e). BioA Labs 2 and 3 showed the highest sample agreement (88/100) (66 ADA−, 22 ADA+; Fig. 2c, f), which is not surprising given that their CPs were set using 56 overlapping commercial samples, while the CP at BioA Lab 1 was established with different commercial samples. This was reflected in the sCP factors of BioA Labs 2 and 3 being more similar (1.75 and 1.56, respectively) compared with the sCP factor of BioA Lab 1 (1.16). Since the CP is the threshold for determining ADA status, the two CPs set with 56 overlapping samples at BioA Lab 2 and BioA Lab 3 and the same NC may have contributed to more similar screening results. When evaluating the sample OD responses normalized by the plate-specific CP OD, samples that were not concordant between different laboratories tended to be very close to the CP (Fig. 2a, b, c). This suggests that the close proximity to the CP may also have been a contributing factor to why samples may not have agreeing results. When evaluating the combined screening results, 67/100 samples had the same results from all three laboratories (Online Resource 3).

When reviewing all the confirmatory data (independent of screen results), the three laboratories show high agreement in sample results. BioA Lab 1 and BioA Lab 2 agreed on 97/100 samples (92 ADA−, 5 ADA+; Fig. 3a, d). Between BioA Lab 1 and BioA Lab 3, 93/100 sample results were in agreement (87 ADA−, 6 ADA+; Fig. 3b, e). Between BioA Lab 2 and BioA Lab 3, 94/100 sample results were in agreement (88 ADA−, 6 ADA+; Fig. 3c, f). We also evaluated the confirmatory results when normalized by the cCP and found that the samples without results in agreement tended to be in close proximity to the cCP (Fig. 3a, b, and c). When evaluating the confirmatory results across all three laboratories, 92/100 samples had the same results from all three laboratories (Online Resource 3). In our assessment, the comparability assay showed higher sample agreement than the screening assay. In the confirmatory assay, individual samples are normalized by the same sample compared with and without spiked drug; therefore, they are not as impacted by individual variability compared with the screening assay which normalizes using a NC. Therefore, the greater agreement observed in the confirmation assay compared with the screening assay was anticipated.
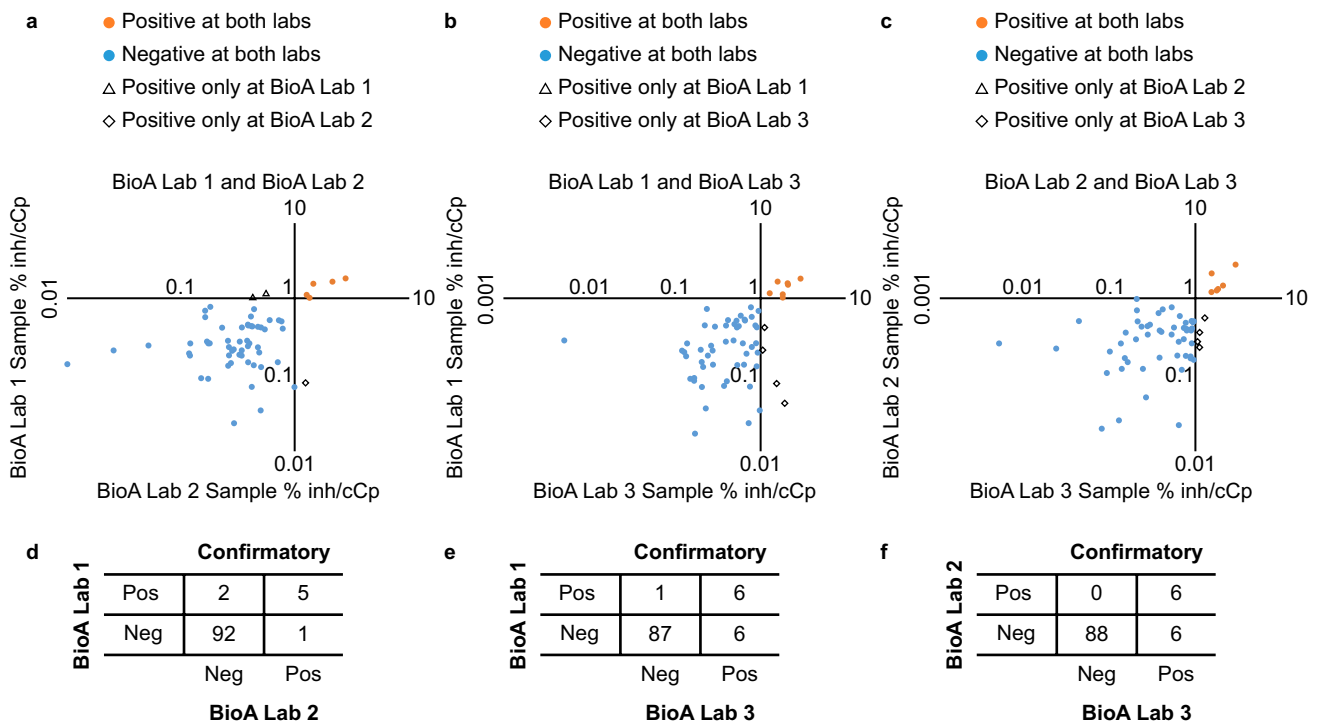
Fig. 3 Secondary evaluation — confirmatory results. All 100 samples selected for the comparability evaluation were run on the confirmatory assay regardless of screen result. Shown are the results plotted as sample % inh/cCP (**a**, **b**, and **c**) as well as sample concordance (**d**, **e**, and **f**), cCP confirmation cut point; BioA bioanalytical; inh inhibition

Discussion

Due to the expansive clinical development plan of tiragolumab which included multiple clinical studies with Chinese patients, it was advantageous to establish the ADA assay at multiple laboratories. In addition to allowing for higher capacity and study support flexibility by using two laboratories in the USA, transferring the method to a bioanalytical laboratory in China also allows us to analyze patient samples in China without the need to export samples to the USA.

A major challenge with a multi-laboratory approach is ensuring the assay is performing comparably between the laboratories, especially when a single-study analysis is split between laboratories. While there is a recommendation from the FDA that comparability should be evaluated, there are no existing acceptance criteria provided (2). There is more guidance for the cross-validation of PK assays (14, 15), but given that ADA assays are semi-quantitative or qualitative tests that are designed to produce false-positive results, such comparisons require a different approach than would be used to test the equivalence of quantitative assays.

A robust comprehensive comparison plan was designed to include separate full validations at each laboratory and comparison of validation data, and use of incurred samples from a previous study. As BioA Labs 2 and 3 used different reagents from the initial laboratory, including the NC and the individual sera used to determine CPs, it was crucial that each laboratory perform its own CP experiments. When using different NCs, different CP factors can be observed in the screen assay, which can still result in similar assay sensitivity. This is shown by the different NCs used between the three laboratories and similar screen assay sensitivities observed (3.35 ng/mL, 3.69 ng/mL, and 4.98 ng/mL). Ideally, the same individual sera would be used in validations at each laboratory, but (as here) this is not always possible. If the same individual sera cannot be used in subsequent laboratories, the disease indication and other factors of the patient samples should be matched as closely as possible to what was used in the initial laboratory. We believe it is important for each laboratory to develop its own CP factors and assay sensitivity when using different reagents.

The validation data between the three laboratories showed similarity when evaluating key validation parameters. Had the assay shown differences that could potentially impact interpretation of comparability results, we would have investigated the source of such discrepancies, examining the screening and confirmation steps in isolation, and taking consideration of any variations in reagents, before generating comparability datasets.

For this comparability assessment, we used incurred samples from a phase 1b tiragolumab study (12, 13) to evaluate comparability in three bioanalytical laboratories. Given the low number of ADA+ samples in the phase 1b study, we selected 50 samples that screened positive, including six samples that were confirmed positive, and 50 samples that screened negative for our comparison. Altogether, the samples represented the entire OD range observed during sample testing. This sample selection was designed to achieve two objectives: ensuring that the two additional laboratories would identify at least some of the positive samples and ensuring that a significant additional number of samples were not identified as positive.

The primary evaluation in this comparison was to determine clinical comparability by examining the overall ADA results through tiered analysis. The secondary evaluation was to independently review the results for the 100 incurred samples on both the screening and confirmatory assays. This included testing samples in the confirmatory assay in all laboratories regardless of whether the samples screened positive. This is more extensive than the approach used during clinical sample analysis. However, comparability data are not for clinical ADA reporting, and we felt this was appropriate for our exploratory evaluation and to build experience with conducting ADA comparability across laboratories. If there were significant discrepancies between the overall results, then data in the individual assay steps of the tiered analysis could be used to determine which of the steps was responsible for the discrepancy and likely identify which laboratory's CPs were functionally different from the others. Due to the low number of positive samples, titers were not included in our comparisons. For molecules with higher numbers of ADA+ results, a comparison involving titer values could be valuable, particularly if ADA titer values were associated with clinical sequelae.

In the primary evaluation of the overall results (tiered analysis of screening and confirmatory), there was a very high degree of agreement in sample results (94/100) (Table V). While there were six discrepancies, we believe that they would not lead to different interpretations of tiragolumab clinical immunogenicity. Thus, with this high degree of agreement, we consider the assay comparable at all three laboratories and would be comfortable pooling data from these laboratories in a regulatory filing.

In the secondary evaluation, the most contrasting results were found at the screening assay level (particularly in BioA Lab 1). Upon further investigation we found that most of the samples with differences in screening assay results were close to the CP in all three laboratories (Fig. 2). This close proximity to the CP may be one reason why a sample would be detected by only one laboratory instead of all three. Additionally, the CPs for BioA Labs 2 and 3 were established with a partially overlapping set of CP samples and the same NC, whereas the CP for BioA Lab 1 was established with

entirely different samples and NC. This may also explain why the screening assay data from BioA Lab 1 showed more differences than the other two laboratories.

For the samples close to the CP, we hypothesize that even if the same laboratory were to re-test the samples another time, a low-positive sample may very well be determined as negative, and a sample just below the CP may be determined as positive on a repeat analysis. In addition, some screen positives might be due to antibodies in the samples that bind non-specifically in the ADA assay. During typical immunogenicity testing of clinical samples, specificity of ADAs toward the drug is determined by running screened positive samples on the confirmatory assay. This emphasizes the importance of our primary evaluation using the tiered strategy (screening and confirmation assays together) to understand how the overall results in a study will compare.

As this was our first experience with ADA comparison between three laboratories, we decided to use incurred samples in our tiragolumab ADA assay comparison. The incurred sample comparison results reflected tiragolumab's clinically low incidence of immunogenicity, with ADA positivity found in only 2/94 samples with matching results at all three laboratories (Table V). In future studies, we feel that for low-risk immunogenicity molecules like tiragolumab, it may not be necessary to use incurred samples to compare ADA assays. The use of spiked samples may be sufficient to provide adequate confidence in assay performance across multiple laboratories. The use of incurred samples, if available, are likely only necessary for ADA assay comparison for high immunogenicity risk programs.

Conclusion

It is important to address the comparability of ADA assays between laboratories to ensure that consistent immunogenicity interpretations can be made on a program. This can be critical in programs with higher immunogenicity risk in which ADAs have been associated with clinical sequelae like lower PK concentration, lower efficacy, or adverse safety events. We have described our approach for validation and comparison of an ADA assay in three different laboratories. The assay was fully validated in three laboratories and demonstrated similar validation results in the parameters of relative sensitivity, drug tolerance, and assay precision. We saw a high degree of agreement between the laboratories in our comparison using incurred study samples.

We conducted our comparability using an exploratory approach rather than setting strict acceptance criteria. In the future, as more practical knowledge accumulates in this area, it is our hope that a more quantitative approach for assessing equivalence (e.g., the use of Cohen's kappa or two one-sided tests (TOST)) will be established that could provide benefits to the bioanalytical community. Our experience presented here may help contribute to furthering these necessary discussions in the bioanalytical industry and with regulators to determine the best strategy to use when conducting ADA comparability.

Declarations

References

1. Koren E, Zuckerman LA, Mire-Sluis AR. Immune responses to therapeutic proteins in humans—clinical significance, assessment and prediction. Curr Pharm Biotechnol. 2002;3(4):349–60. https://doi.org/10.2174/1389201023378175.
2. U.S. Food and Drug Administration. Guidance for industry: immunogenicity testing of therapeutic protein products — developing and validating assays for anti-drug antibody detection. February 2019. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/immunogenicity-testing-thera

peutic-protein-products-developing-and-validating-assays-anti-drug. Accessed 13 May 2023.

3. National Medical Products Administration. Technical guidance for immunogenicity studies of drugs. March 2021. https://www.cde.org.cn/main/att/download/7d89a3b3b078e2682e935bd5d85ce7a8. Accessed 13 May 2023.

4. European Medicines Agency. Immunogenicity assessment of biotechnology-derived therapeutic proteins—— scientific guideline EMEA/CHMP/BMWP/14327/2006. May 2017. https://www.ema.europa.eu/documents/scientific-guideline/guideline-immunogenicity-assessment-therapeutic-proteins-revision-1_en.pdf. Accessed 13 May 2023.

5. Tatarewicz S, Moxness M, Weeraratne D, Zhou L, Hale M, Swanson SJ, et al. A step-wise approach for transfer of immunogenicity assays during clinical drug development. AAPS J. 2009;11(3):526–34. https://doi.org/10.1208/s12248-009-9130-3.

6. Mora JR, White JT, Chilewski SD, Qu Q, Stocker D, Luo L, et al. Strategies for method comparison when changes in the immunogenicity method are needed within a clinical program. Bioanalysis. 2020;12(7):431–43. https://doi.org/10.4155/bio-2019-0300.

7. Qiu ZJ, Ying Y, Lewin-Koh S-C, Coleman D, Brignoli S, Hendricks R, et al. Strategies to compare clinical antitherapeutic antibody data when changing assay platforms: a case study. Bioanalysis. 2015;7(14):1775–83. https://doi.org/10.4155/bio.15.91.

8. U.S. Food and Drug Administration. Guidance for industry. Assay development and validation for immunogenicity testing of therapeutic protein products. Draft Guidance. April 2016. https://www.fda.gov/media/77796/download. Accessed 19 February 2024.

9. European Medicines Agency. Guideline on immunogenicity assessment of biotechnology-derived therapeutic proteins. April 2008. https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-immunogenicity-assessment-biotechnology-derived-therapeutic-proteins-first-version_en.pdf. Accessed 19 February 2024.

10. 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. https://pubmed.ncbi.nlm.nih.gov/15251407/.

11. 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):1267–1281. https://pubmed.ncbi.nlm.nih.gov/18993008/.

12. Kim TW, Bedard PL, LoRusso P, Gordon MS, Bendell J, Oh D-Y, et al. Anti-TIGIT antibody tiragolumab alone or with atezolizumab in patients with advanced solid tumors. JAMA Oncol. 2023;9(11):1574–82. https://doi.org/10.1001/jamaoncol.2023.3867.

13. Garralda E, Oh DY, Italiano A, Bedard PL, Delord JP, Calvo E, et al. Pharmacokinetics (PK) of tiragolumab in first-in-human study in patients with mixed solid tumors (GO30103). J Clin Pharmacol. 2023. [Online ahead of print]. https://doi.org/10.1002/jcph.2397

14. U.S. Food and Drug Administration. Guidance for industry: bioanalytical method validation. May 2018. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/bioanalytical-method-validation-guidance-industry. Accessed 13 May 2023.

15. U.S. Food and Drug Administration. M10 bioanalytical method validation and study sample analysis. November 2022. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/m10-bioanalytical-method-validation-and-study-sample-analysis. Accessed 13 May 2023.