Feb 2024

Establishing Appropriate Immunogenicity Assay Cut Points in Oncology Disease Indications

Michael Partridge, Bioanalytical Sciences

REGENERON®

Setting Appropriate ADA Assay Cut Points

- Cut points are critical thresholds to establish ADA positivity

- Poorly established CPs that are too high could potentially miss treatment emergent ADA or, when set too low, result in detection of responses that may have no clinical relevance.

Cut Points can be Challenging for BA Scientists

- Establishing a threshold based on statistical analysis of background assay responses is theoretically straightforward.
- However, key decisions (population type/size, statistical approaches) can substantially impact the cut point.
- Selecting the most appropriate statistical strategies is critical to establishing suitable CPs.
- Bioanalytical scientists need to understand the process to address questions from Health Authorities about high baseline or placebo positivity or poor correlation with clinical outcomes.

Relevant Guidance

Immunogenicity of therapeutic protein products: current considerations for anti-drug antibody assay in Japan

Akiko Ishii-Watabe*,[1], Hiroko Shibata[1], Kazuko Nishimura[1], Jun Hosogi[2], Muneo Aoyama[3], Kazuhiro Nishimiya[4] & Yoshiro Saito[1]

[1]National Institute of Health Sciences, 3-25-26 Tonomachi, Kawasaki-ku,Kawasaki-shi, Kanagawa 210-9501, Japan
[2]Kyowa Hakko Kirin Co., Ltd. 1188 Shimotogari, Nagaizumi-cho, Sunto-gun, Shizuoka 411-8731, Japan
[3]Eisai Co., Ltd. 5-1-3 Tokodai, Tsukuba-shi, Ibaraki 300-2635, Japan
[4]Chugai Pharmaceutical Co., Ltd. 200 Kajiwara, Kamakura-shi, Kanagawa 247-8530, Japan
* Author for correspondence: Tel.: +81-44-270-6512; Fax: +81-44-270-6517; watabe@nihs.go.jp

EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

18 May 2017
EMEA/CHMP/BMWP/14327/2006 Rev 1
Committee for Medicinal Products for Human Use (CHMP)

Guideline on Immunogenicity assessment of therapeutic proteins

Immunogenicity Testing of Therapeutic Protein Products — Developing and Validating Assays for Anti-Drug Antibody Detection

Guidance for Industry

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)

January 2019
Pharmaceutical Quality/CMC

REGENERON

Recent Publications

The AAPS Journal (2023) 25:37
https://doi.org/10.1208/s12248-023-00806-5

RESEARCH ARTICLE

Statistical Approaches for Establishing Appropriate Immunogenicity Assay Cut Points: Impact of Sample Distribution, Sample Size, and Outlier Removal

John Garlits[1] · Sean McAfee[1] · Jessica-Ann Taylor[1] · Enoch Shum[1] · Qi Yang[1,2] · Emily Nunez[1] · Kristina Kameron[1] · Keilah Fenech[1] · Jacqueline Rodriguez[1] · Albert Torri[1] · Jihua Chen[1] · Giane Sumner[1] · Michael A. Partridge[1]

Received: 14 December 2022 / Accepted: 23 March 2023

Immunogenicity of Cemiplimab: Low Incidence of Antidrug Antibodies and Cut-Point Suitability Across Tumor Types

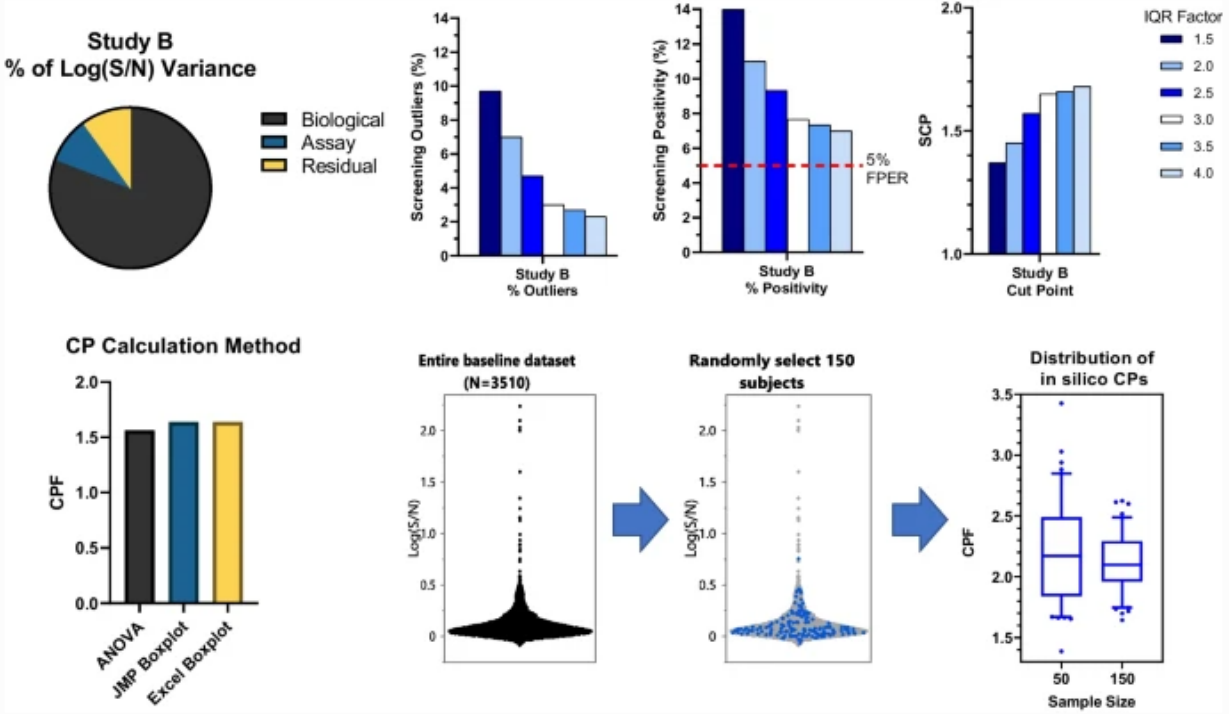
The Journal of Clinical Pharmacology
2023, 0(0) 1–12

DOI: 10.1002/jcph.2340

Jenny L. Valentine, PhD, Andrew Dengler, BSc, An Zhao, PhD, Tiffany Truong, BSc, Sean McAfee, MSc, Mohamed Hassanein, PhD, Susan C. Irvin, PhD, Jihua Chen, PhD, Xiao Meng, MSc, Hong Yan, PhD, Albert Torri, PhD, Giane Sumner, PhD, Matthew D. Andisik, BSc, Anne Paccaly, PhD, and Michael A. Partridge, PhD

REGENERON

Key Factors That Impact Cut Point



- Survey of 16 cut point datasets
- Majority of variability is biological
- Most data sets are non-normal/right-skewed
- Positivity = FPER + outliers
- Higher IQR leads to less outlier removal, higher CPs and positivity closer to target FPER
- Similar CP values with simple Excel box plot vs. more complex ANOVA
- Larger sample sizes better capture variability in the study population

Garlits et al 2023

REGENERON

ADA Assay Cut Point Datasets: Sample Population Characteristics and Sources of Variability

Table I ADA Assay Cut Point Datasets: Sample Population Characteristics and Sources of Variability

Study	Pop'n	No. samples	No. Obs./sample	log(S/N) ANOVA % of variance (random effects)				Shapiro-Wilk (p value)	Skewness
				Sample (biological)	Analyst	Assay	Residual[a]		
A	Diseased	160	2	94.7	ND	2.4	2.9	<0.001	1.51
B	Diseased	150	2	80.8	ND	9.2	10.0	<0.001	1.56
C	Diseased	150	2	93.5	ND	3.8	2.7	<0.001	0.46
D	Diseased	150	2	96.2	ND	1.0	2.8	<0.001	1.59
E	Diseased	200	1	ND	3.4	3.2	93.4	<0.001	1.29
F	Diseased	324	1	ND	0.0	11.1	88.9	<0.001	1.17
G	Healthy	162	2	96.0	ND	0.2	3.9	<0.001	0.95
H	Diseased	150	2	81.9	ND	12.3	5.9	<0.001	0.58
I	Healthy	150	2	78.2	ND	4.2	17.6	0.005	0.47
J	Healthy	150	2	87.4	ND	2.9	9.8	0.250	0.15
K	Healthy	54	6	79.8	10.8	3.2	6.2	<0.001	1.03
L	Diseased	50	6	96.4	0.4	1.0	2.1	<0.001	1.13
M	Healthy	54	6	97.0	0.1	0.6	2.4	<0.001	0.99
N	Diseased	90	4	96.5	ND	0.1	3.4	<0.001	0.66
O	Healthy	150	2	84.6	ND	9.7	5.7	<0.001	1.00
P	Healthy	150	2	92.6	ND	3.3	4.1	<0.001	0.53

ND Not determined

[a]Residual effect for Studies E and F includes biological variability

Garlits et al 2023

REGENERON

ADA Assays in Japan: Outlier Removal

- IQR outlier removal assumes a symmetrical distribution
- Most cut point datasets are non-normal (Zhang 2017, Garlits 2023)

> In terms of the outlier exclusion, various approaches (e.g., excluding values higher than × 1.5 or 3 interquartile range by box plot) would be acceptable if scientifically justified and predefined in the standard operating procedure. The approach that all data are included in the cut-point calculation without excluding outliers is also acceptable if it is justified. However, it should be considered that the assigned cut point often becomes higher with outliers included in the calculation, which may increase the risk of obtaining false negative results.

Immunogenicity of therapeutic protein products: current considerations for anti-drug antibody assay in Japan

Akiko Ishii-Watabe*,[1], Hiroko Shibata[1], Kazuko Nishimura[1], Jun Hosogi[2], Muneo Aoyama[3], Kazuhiro Nishimiya[4] & Yoshiro Saito[1]

- 5%/1% false positive concept, tiered testing etc, designed to be conservative
- Outlier removal is an additional layer of stringency, not a requirement
- Excessive outlier removal (e.g., >5-10%) may indicate a potential assay issue

Population Specific Cut Points in Oncology:

Does Every Tumor Type Need a Different Cut Point?

REGENERON®

Verify the appropriateness of the validated cut point with new populations

C. Confirmation of Cut-Point in the Target Population

Samples from different populations can have different background activity in ADA assays. Similarly, the background activity can change when 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. Therefore, it is necessary to confirm that the cut-point determined during assay validation is suitable for the population being studied. 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.

FDA. 2019. Guidance for Industry: Immunogenicity Testing of Therapeutic Protein Products — Developing and Validating Assays for Anti-Drug Antibody Detection.

ADA Assays in Japan

Cut point for the target subject samples

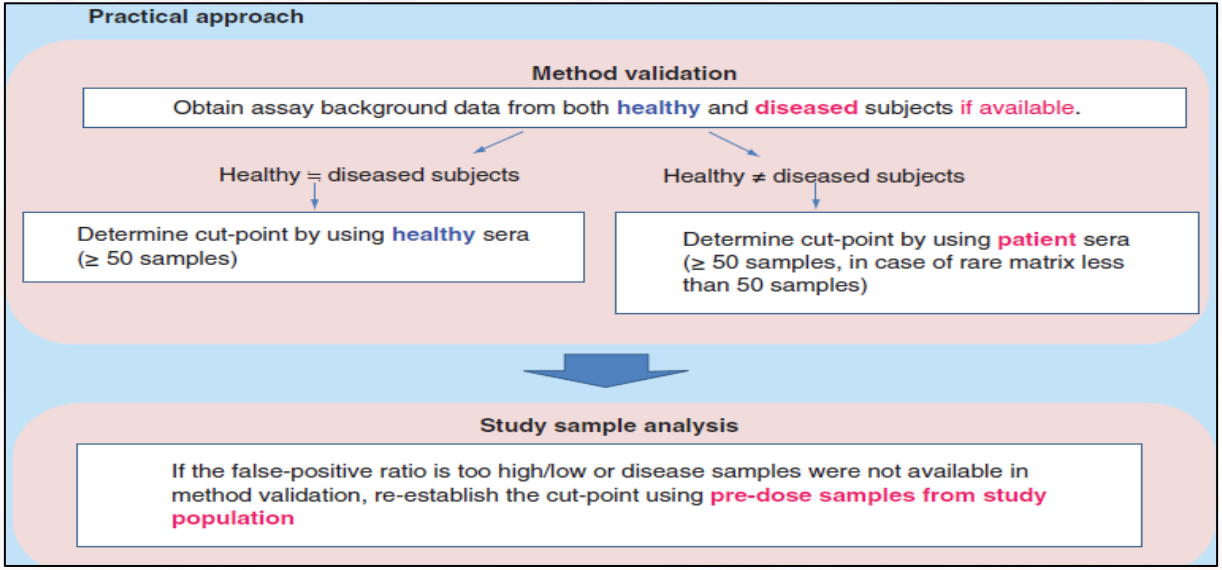
The background response in the ADA assay can be different between healthy and disease subjects. It is also possible that the responses differ among target diseases (e.g., the types of tumor). Therefore, it should be pointed out that using the samples from the target disease subjects should be considered for cut-point setting. However, samples from subjects with target disease are not always available, especially in the early development phase. Therefore, we propose a practical approach for cut-point determination using samples from diseased subjects (Figure 4).

In cases where samples from subjects with target disease are available, the background response is compared with those of healthy subjects. If these background responses differ, the cut point should be determined using the target disease subjects. When samples from the target disease subjects are not available, or the positive rate of the in-study samples is too high/low, it is recommended to re-evaluate the cut point using predose samples from the study subjects.

Immunogenicity of therapeutic protein products: current considerations for anti-drug antibody assay in Japan

Akiko Ishii-Watabe*,[1], Hiroko Shibata[1], Kazuko Nishimura[1], Jun Hosogi[2], Muneo Aoyama[3], Kazuhiro Nishimiya[4] & Yoshiro Saito[1]

[1]National Institute of Health Sciences, 3-25-26 Tonomachi, Kawasaki-ku,Kawasaki-shi, Kanagawa 210-9501, Japan
[2]Kyowa Hakko Kirin Co., Ltd. 1188 Shimotogari, Nagaizumi-cho, Sunto-gun, Shizuoka 411-8731, Japan
[3]Eisai Co., Ltd. 5-1-3 Tokodai, Tsukuba-shi, Ibaraki 300-2635, Japan
[4]Chugai Pharmaceutical Co., Ltd. 200 Kajiwara, Kamakura-shi, Kanagawa 247-8530, Japan
* Author for correspondence: Tel.: +81-44-270-6512; Fax: +81-44-270-6517; watabe@nihs.go.jp



REGENERON

Oncology Drug: Questions from the Regulators

Bioanalytical Validation Report [redacted] states that the cut points for the ADA screening, confirmatory and titer assays were obtained from the analysis of baseline serum samples from oncology patients participating in [redacted] clinical study. It is unclear which oncology patient population (i.e., [redacted]) was selected for the determination of assay cut points used during the analysis of clinical samples. Due to the variability in baseline serum factors (e.g., pre-existing antibodies and rheumatoid factors) that can exist among different patient populations, serum samples collected from the target patient population should be considered to determine an appropriate population-specific cut point. Clarify which oncology patient population was used in the determination of the ADA screening, confirmatory and titer assay cut points. If the same assay cut point was used to analyze clinical serum samples from the [redacted], then data should be provided to demonstrate that baseline serum factors between the different patient populations are equivalent with regards to assay performance (e.g., signal-to-noise and % inhibition values).

For the validation of the Bioanalytical Analytical Methods used for the detection of [redacted] (ADA assay) and [redacted] (NAb assay) in human serum, baseline serum from different oncology patients (i.e., not limited to [redacted]) were used to establish the screening and confirmation cut points. This raises concerns as samples from different target populations and disease states may have different components that can cause the background signal from the assay to vary *(see US FDA Draft Guidance on Assay Development and Validation for immunogenicity Testing of Therapeutic Protein Products (2016))*. **Provide an adequate rationale/discussion to address the noted concern.**

Biology and Categories of Cancer

Carcinomas
Begin in epithelial tissue

e.g., lung cancer, breast cancer

Sarcomas
Begin in tissues that support and connect the body

e.g., liposarcoma, angiosarcoma

Melanomas
Begin in melanocytes

e.g., nodular melanoma, superficial spreading melanoma

Lymphomas
Begin in the lymphatic system

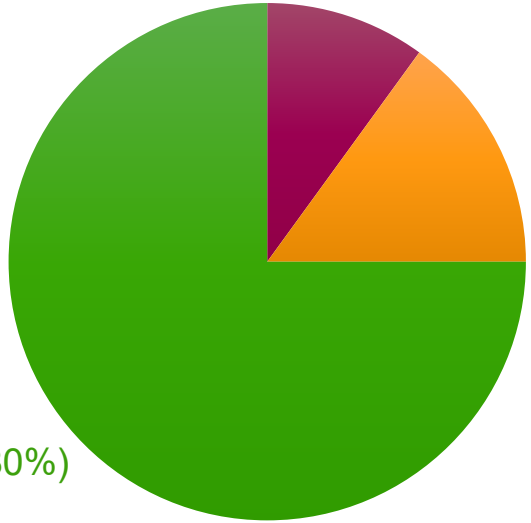
e.g., diffuse large B-cell lymphoma, NHL

Leukemias
Begin in the blood

e.g., acute lymphocytic leukemia, acute myeloid leukemia

Cancer as a Genetic Disease

Hereditary (5-15%)
- Inherited genetic mutation



Familial (5-20%)
- Occurs more frequently in families than chance alone
- Not linked to a specific mutation

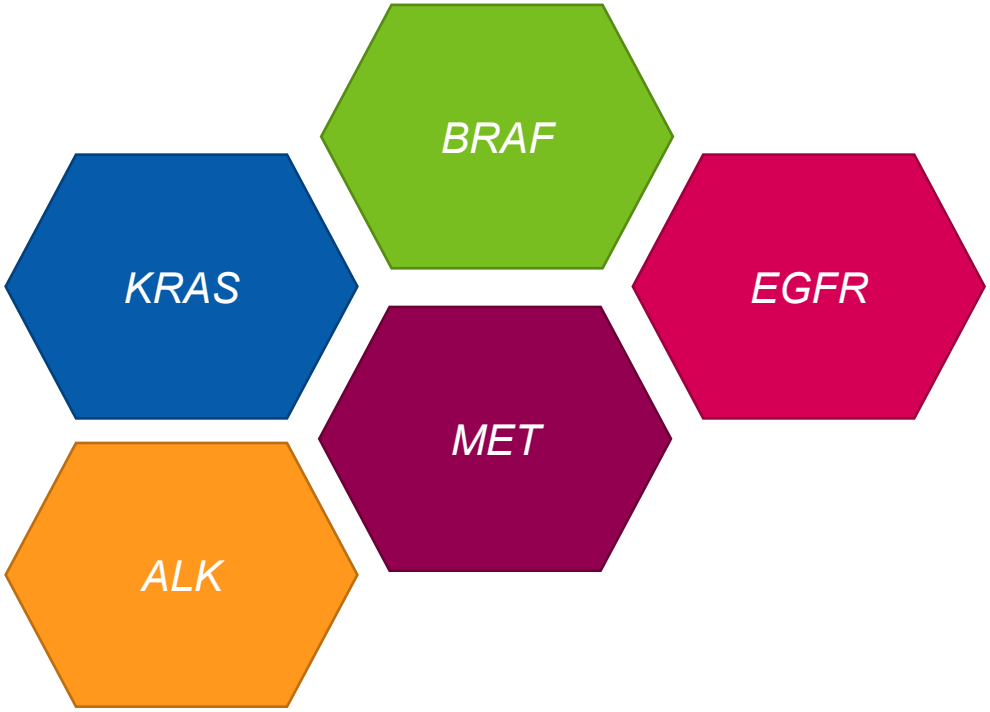
Acquired (70-80%)
- Age
- Tobacco
- UV radiation

BRCA
- Breast
- Stomach
- Prostate

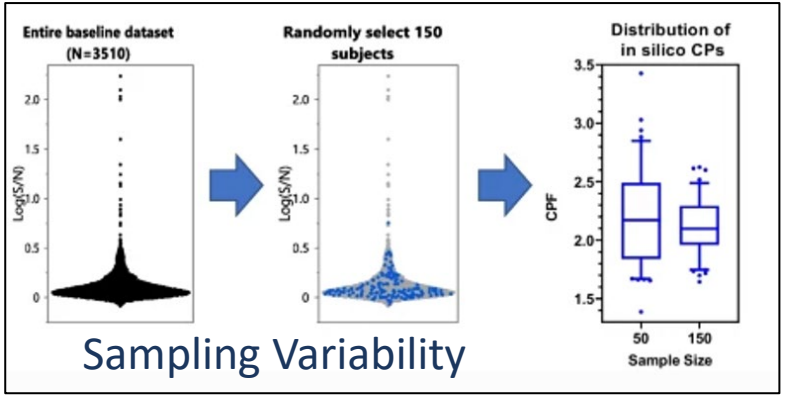
DNA mismatch repair system
- Endometrial
- Colorectal
- Melanoma
- Sarcoma

Cancer as a Genetic Disease

A single tumor classification like NSCLC can have a variety of genetic origins



- Multiple tumor types have the same origin
- A single tumor type has heterogeneous origins
- **Why would each cancer types have unique assay responses?**

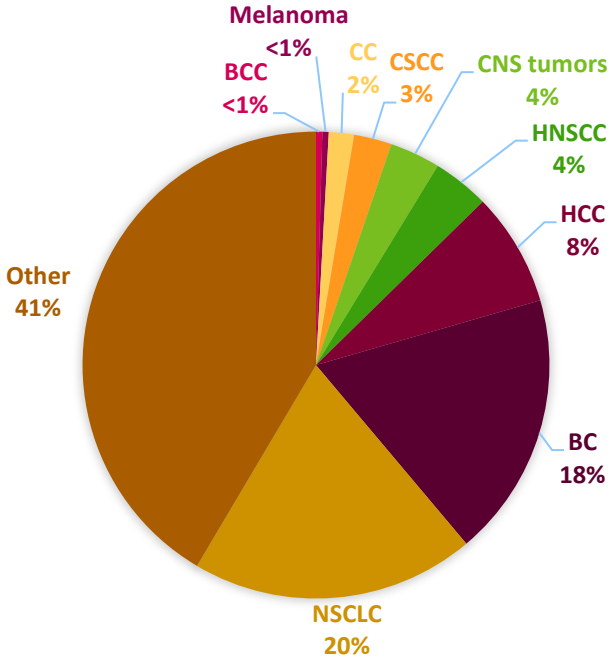


Hypothesis for Validation Cut Point Generation

By using samples from a varied patient population to set the sCPF, we could capture the biological variability observed in serum samples from multiple disease types

- Biological variability - the most important source of variability in cut points
- Used 229 random samples from a solid tumor, first-in-human study population
- Sample collection and handling was the same between this and subsequent studies

Composition of the Validation Population



BC: breast cancer
BCC: basal cell carcinoma
CC: cervical cancer
CNS tumors: central nervous system tumors
CSCC: cutaneous squamous cell carcinoma
HCC: hepatocellular carcinoma
HNSCC: head and neck squamous cell carcinoma
NSCLC: non-small cell lung cancer

*Other category includes glioblastoma multiforme, colorectal cancer, ovarian cancer, pancreatic cancer, endometrial cancer, and small cell lung cancer, among others.

Valentine et al 2023

REGENERON

How We Assessed the Appropriateness of Our Cut Point

1. Examined in-study baseline positive rates
2. Statistically compared baseline responses from validation population and population of interest
3. Examined treatment-emergent ADA rates

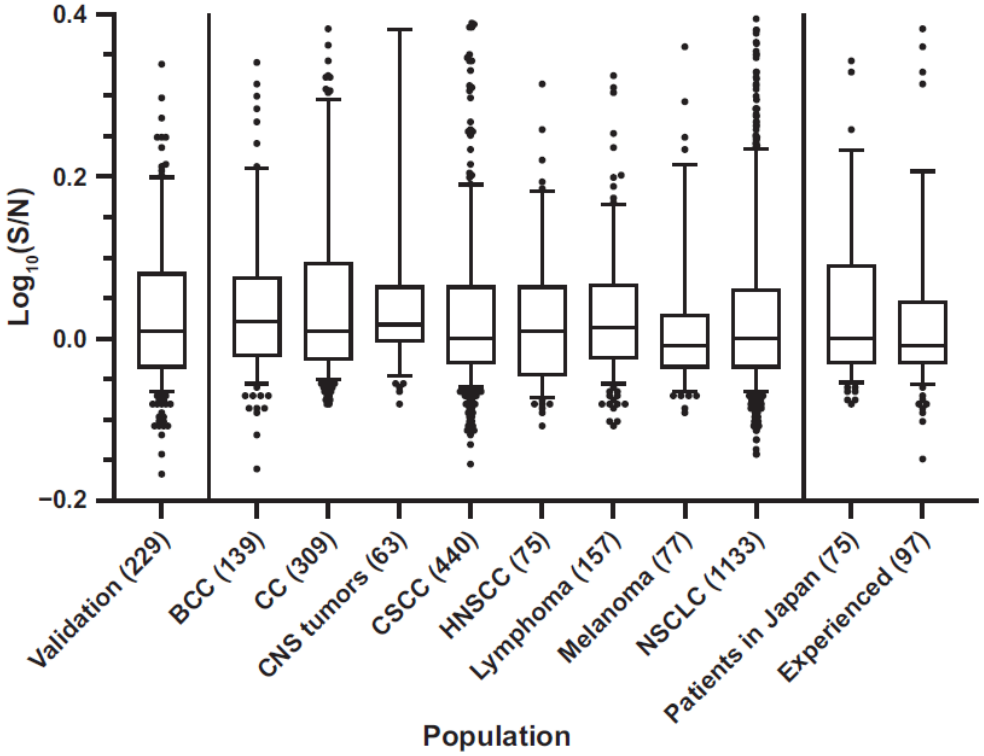
Positive Rates (PR) for Disease Populations

Population	# Baseline Samples	Screening PR (%)	Screening False PR (%)	Confirmation PR (%)
Validation*	229	9.2	6.6	2.6
Basal Cell Carcinoma	139	10.1	7.2	2.9
Cervical Cancer	309	13.6	11.0	2.6
Central Nervous System Tumors	63	14.3	9.5	4.8
Cutaneous Squamous Cell Carcinoma	440	9.3	6.8	2.5
Head & Neck Squamous Cell Carcinoma	75	6.7	5.4	1.3
Non-Small Cell Lung Cancer	1133	11.4	8.9	2.5
Lymphoma	157	6.4	5.1	1.3
Melanoma	77	10.4	7.4	3.0
Patients in Japan	75	13.3	10.6	2.7
Anti-PD-1/PD-L1 Experienced	97	9.3	6.2	3.1

*Validation FPR includes 5% targeted FPR + outliers for screening and 1% targeted FPR + outliers for confirmation

Valentine et al 2023

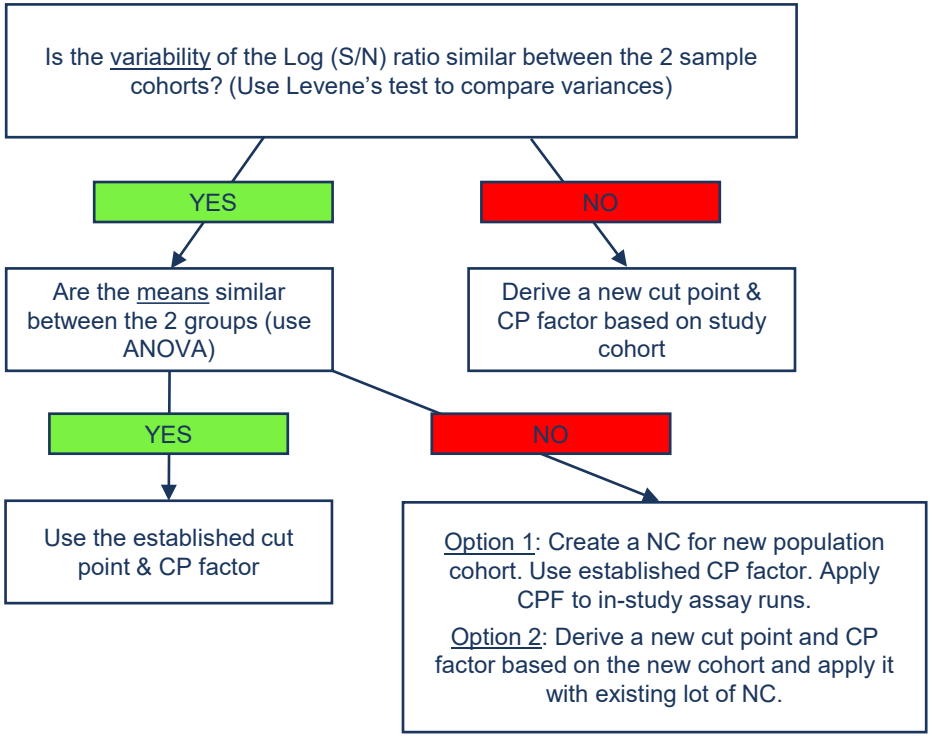
Box Plot of Validation and Disease Populations



Valentine et al 2023

A Statistical Method to Compare Populations

One suggested process for assessing validation sCPF appropriateness in a new population



Adapted from Devanarayan V, *et al*. AAPS J. 2017;19(5):1487-98.

Statistical Comparison of Baseline Assay Responses

Population	# Baseline Samples	Mean Log_{10}(S/N)	Stdev Log_{10}(S/N)	Levene *p* value	ANOVA *p* value
Validation	229	0.055	0.175	N/A	N/A
BCC	139	0.066	0.180	0.8561	0.5775
Cervical	309	0.081	0.205	0.0930	0.1284
CNS Tumors	63	0.092	0.224	0.1284	0.1658
CSCC	440	0.059	0.219	0.5465	0.8048
HNSCC	75	0.049	0.229	0.8217	0.8003
NSCLC	1133	0.064	0.219	0.1515	0.5810
Lymphoma	157	0.041	0.119	0.0124	N/A
Melanoma	77	0.035	0.161	0.3698	0.3780
Patients in Japan	75	0.070	0.190	0.4358	0.5492
Anti-PD-1/PD-L1 Experienced	97	0.038	0148	0.4932	0.4014

- The validation sCPF is appropriate for all carcinoma populations and melanoma population.
- The lymphoma population variance is statistically different from the validation population as shown by Levene's test, though the mean is within the range of the other seven populations.

Valentine et al 2023

REGENERON

Lymphoma Data Set Evaluation

- Screening FPR of 5.1% is within the recommended 2-11% range
- Lymphoma data set is not normal
  - Shapiro-Wilk test is <0.05 for S/N and Log(S/N) data
  - Skewness coefficient is >1 for S/N and Log(S/N) data
- Non-parametric cut point estimated for this population was very similar to the established validation cut point
- Levene's test may not always be suitable for comparison of variance in the non-normal datasets commonly observed in cut point populations

➢ Detemined validation CP was appropriate for lymphoma population

Observed Immunogenicity

- A low treatment-emergent ADA rate observed across all populations
- No effect on clinical response

	BCC	CSCC	NSCLC	CC	Other	Overall
N	129	311	240	213	45	938
Pre-existing (%)	4 (3.1)	7 (2.3)	6 (2.5)	6 (2.8)	2 (4.4)	25 (2.7)
Treatment Emergent (%)	4 (3.1)	6 (1.9)	5 (2.1)	4 (1.9)	3 (6.7)	22 (2.3)

Conclusions I

1. Data show that 8/8 disease populations are within the 2-11% FPR range.
2. Different disease indications had Log_{10}(S/N), from 0.035 ± 0.161 to 0.092 ± 0.224.
   - We saw 0.055 ± 0.175 in our validation population.
3. Levene's test showed 7/8 disease populations had similar variance; only lymphoma was different.
   - Lymphoma data set was non-parametric. A different statistical test may be more appropriate.
4. Using ANOVA, 7/7 disease populations had a similar mean log_{10}(S/N).
5. A low treatment-emergent ADA rate was observed across all populations with no effect on clinical response, suggesting that the validation sCPF was suitable for monitoring immunogenicity.

Conclusions II

- Data is consistent samples representing a mixture of different biological phenotypes in serum, as opposed to groupings like “NSCLC” or “BCC,” since cancer is a genetic disease.
- Other therapeutic areas (e.g., inflammatory diseases) may have subpopulations that require a specific cut point.
  - Rheumatoid arthritis has a type of autoantibody, rheumatoid factors, that binds to IgGs.
- Populations known to have potential impact on cut point calculations, disease matrix is typically used during validation.
- Sampling variability may contribute to differences in observed positivity (vs. population differences)

The biological differences of indications should be considered when sCPFs are evaluated for and applied to different populations.

For a low-risk mAb, we found that new cut points were not required for each oncology population.

Acknowledgements

Jenny Valentine
An Zhao
Susan Irvin
John Garlits
Jihua Chen
Sean McAfee
Jessica-Ann Taylor
Andrew Dengler
Tiffany Truong
Jackie Rodriguez
Qi Yang
Kristina Kameron
Mohamed Hassanein
Xiao Meng
Hong Yan
Anne Paccaly
Christine Grimaldi
Boris Gorovits
Matthew Andisik
Giane Sumner
Albert Torri

Questions?

Notes