Analytical Plan for Goodness of fit of measurements from a bioassay (redacted) between customer-site against laboratory distributions

DOCUMENT: SAP-2022-035-SP-v02

From: Felipe Figueiredo **To:** Undisclosed Medical Equipment Company (US)

2022-12-07

TABLE OF CONTENTS

Analytical Plan for Goodness of fit of measurements from a bioassay (redacted) between customer-site against laboratory distributions

Document version

Version	Alterations
01	Initial version
02	New objective to assess GoF between laboratory and customer-site conditions: changed hypothesis, variables, inclusion criteria and statitical methods.

1 ABBREVIATIONS

- CI: confidence interval
- CV: coefficient of variation
- GoF: goodness of fit
- IQR: interquartile range
- QC: Quality control
- SD: standard deviation

2 CONTEXT

2.1 Objectives

Test the goodness of fit of distribution of measurements from a bioassay (redacted) between laboratory and customer-site conditions.

2.2 Hypotheses

The distribution of QC assay measurements from the customer sites fits the laboratory distribution.

3 DATA

3.1 Raw data

The raw data was received as an Excel table containing measurements from several assays. Two sources of measurements were made available, measurements from lab assays that serve as reference values and measurements from customer sites that will be the object of QC monitoring. The original data base had 5 variables collected on 492648 observations.

3.2 Analytical dataset

Since this analysis will evaluate the distributions of the gold standard defined as the lab assays the measurements from customer assays will be dropped and not used at this moment. Future analyses will test the goodness of fit between the distributions of application assays and lab assays.

After the cleaning process 5 variables were included in the analysis with 492648 observations. The total number of observations excluded due to incompleteness and exclusion criteria will be reported in the analysis. Table 1 shows the structure of the analytical dataset.

Table 1 *Analytical dataset structure after variable selection and cleaning.*

id	sample_date	lot_number	golden	analyte
1				
2				
3				
…				
492648				

All variables in the analytical set were labeled according to the raw data provided and values were labeled according to the data dictionary for the preparation of production-quality results tables and figures.

4 STUDY PARAMETERS

4.1 Study design

Cross-sectional.

4.2 Inclusion and exclusion criteria

Measurements from both reference assay and customer-site will be included in the analysis.

4.3 Exposures

N/A

4.4 Outcomes

Specification of outcome measures (Zarin, 2011):

1. (Domain) Bioassay (redacted)
2. (Specific measurement) Analyte
3. (Specific metric) Density
4. (Method of aggregation) Average

Primary outcome

Average Analyte of the QC assay.

4.5 Covariates

The Analyte distribution on the available lots will be evaluated per lot.

5 STATISTICAL METHODS

5.1 Statistical analyses

5.1.1 Descriptive analyses

The distribution of Analyte measurements will be described as mean (SD), minimum, maximum, and their quartiles will be presented as additional information. The parameters of the overall Analyte distribution will be summarized in tables and visualized in exploratory plots.

In each subset of samples the deviation of Analyte measurements from the Normal distribution will be assessed visually using QQ-plots and findings will be compared with the summary statistics.

5.1.2 Inferential analyses

The goodness of fit of the customer-site operating conditions will be assessed against laboratory conditions with the asymptotic one-sample Kolmogorov-Smirnov test.

5.1.3 Statistical modeling

N/A

5.1.4 Missing data

No missing data imputation will be performed. All evaluations will be performed as complete case analyses.

5.2 Significance and Confidence Intervals

All analyses will be performed using the significance level of 5%. All significance hypothesis tests and confidence intervals computed will be two-tailed.

5.3 Study size and Power

N/A

5.4 Statistical packages

This analysis will be performed using statistical software R version 4.2.1.

6 OBSERVATIONS AND LIMITATIONS

N/A

7 REFERENCES

- **SAR-2022-035-SP-v01** – Goodness of fit of measurements from a bioassay (redacted) between customer-site against laboratory distributions
- Zarin DA, et al. The ClinicalTrials.gov results database – update and key issues. N Engl J Med 2011;364:852-60 (https://doi.org/10.1056/NEJMsa1012065).
- Gamble C, et al. Guidelines for the Content of Statistical Analysis Plans in Clinical Trials. JAMA. 2017;318(23):2337–2343 (https://doi.org/10.1001/jama.2017.18556).

8 APPENDIX

This document was elaborated following recommendations on the structure for Statistical Analysis Plans (Gamble, 2017) for better transparency and clarity.

8.1 Availability

All documents from this consultation were included in the consultant's Portfolio.

The portfolio is available at:

https://philsf-biostat.github.io/SAR-2022-035-SP/