DRAFT INTERLABORATORY STATISTICAL ANALYSIS REPORT

on

PLACENTAL AROMATASE VALIDATION STUDY
4-OH ASDN POSITIVE CONTROL INHIBITOR STUDY

EPA CONTRACT NUMBER 68-W-01-023
WORK ASSIGNMENT 4-16, TASK 4

October 27, 2005

Prepared for

U.S. ENVIRONMENTAL PROTECTION AGENCY
ENDOCRINE DISRUPTOR SCREENING PROGRAM
WASHINGTON, D.C.

Prepared by

BATTELLE
505 King Avenue
Columbus, Ohio 43201

Summary and Conclusions

The principal results of the inter-laboratory analysis are summarized below.

1. Laboratory C had the highest estimated $\log_{10}IC_{50}$ among the three laboratories; Laboratory B had the lowest. The variance among laboratories was at least 6 times higher than the unweighted average within laboratory variance for $\log_{10}IC_{50}$. The coefficient of variation among laboratories was 10% when replicate 1 in Laboratory C was included and 11% when replicate 1 in Laboratory C was excluded.

2. Results for the slope were consistent for three laboratories. The estimated variance among the laboratories was zero or near zero. The coefficient of variation among laboratories was 3.7% when replicate 1 in Laboratory C was included and 3.2% when replicate 1 in Laboratory C was excluded.

3. No significant differences existed between background activity control at the end and at the beginning for each laboratory or across three laboratories. The estimated variance among the laboratories was zero or near zero.

4. No significant differences existed between full enzyme activity control at the end and at the beginning of the replicates, across the three laboratories or for Laboratories B and C. Laboratory A had a significantly higher value at the beginning when an outlier was excluded but not a significant difference when the outlier was included. The estimated variances among the laboratories were smaller than the unweighted average within laboratory variance whether the outlying value in Laboratory C was included or excluded. The unweighted average within laboratory variance was inflated by the within laboratory variance in Laboratory B.

Introduction and Background

Task 4 of Work Assignment 4-16, the Placental Aromatase Validation Study involved three individual laboratories (labeled as A, B, and C) independently carrying out the placental aromatase assay with positive control inhibitor 4-OH ASDN and centrally prepared microsomes according to a common protocol. An "intra-laboratory" statistical analysis was carried out for each individual laboratory's test data according to a common statistical analysis plan. The "inter-laboratory" statistical analysis discussed in this report combines summary values developed in each intra-laboratory analysis and assesses the relationships among them, the extent of inter-laboratory variation, and overall consensus estimates. This report discusses the methods used and the results obtained from combining the intra-laboratory statistical analysis results.

The intra-laboratory analyses were performed on the "percent of control" responses for placental aromatase assay, at each of the laboratories. The inter-laboratory analysis is based on the $\log_{10}IC_{50}$ and slope parameters of the concentration response curve fits determined in the intra-laboratory analyses. The inter-laboratory analysis also compared across the three laboratories the full enzyme activity control and the background activity control tube responses at the beginnings and the ends of the replicates.

Test Organization

Placental assay aromatase activity levels were determined for the full enzyme activity control, the background activity control, and for six graded concentrations of positive control inhibitor 4-OH ASDN. Three replicates of the positive control inhibitor study were carried out for each laboratory, and an additional replicate was carried out for Laboratory C. Within each replicate three repetitions were run at each of the 4-OH ASDN log (base 10) concentrations -6, -7, -7.3, -7.6, -8, and -9. In addition two repeat tubes of the full enzyme activity and background activity controls were run prior to the 4-OH ASDN runs and two repeat tubes of the full enzyme activity and background activity controls were run following the 4-OH ASDN runs.

Intra-laboratory statistical analyses were carried out on the "percent of control" responses. Percent of control is defined as the ratio of the background adjusted aromatase activity in the tube under consideration to the average background adjusted aromatase activity among the four full enzyme activity control tubes within the replicate, times 100. The average percent of control among the four full enzyme activity control tubes is necessarily 100 percent within each replicate. The average percent of control among the four background activity control tubes is necessarily 0 percent.

Nominally for an inhibitor the percent of control activity values vary between approximately 0% near the high inhibition concentrations and approximately 100% near the low inhibition concentrations, but this may vary with the inhibitor.

Intra-laboratory statistical analyses were performed based on a common analysis plan. The following results were reported for each intra-laboratory analysis.

1. Concentration response curve fits within each replicate to describe the trend in the percent of control activity across varying inhibitor concentrations of test substance 4-OH ASDN.

2. Estimates of the $\log_{10}IC_{50}$ concentration, slope, and associated standard errors within each replicate.

3. Average $\log_{10}IC_{50}$ concentration, average slope, and associated standard errors across replicates.

4. Comparisons between the full enzyme activity and background activity controls obtained at the beginning and those obtained at the end of each replicate.

Results for Laboratories A and B were reported based on three replicates, while Laboratory C provided results based on replicates 1 to 4, as well as results based on replicates 2 to 4. There was an outlying value among the full enzyme activity controls for Laboratory A. The results for Laboratory A were reported both including and excluding this data point. The reported standard error of the average results across replicates for Laboratories A and C incorporated the among replicate component of variation, while that for Laboratory B did not.

The "inter-laboratory" statistical analysis combines summary values developed in each intra-laboratory analysis to assess relationships among the laboratory results, the extent of laboratory-to-laboratory variation, and overall consensus estimates among the laboratories with associated variability estimates (incorporating laboratory-to-laboratory variability). The inter-laboratory analysis is based on the average $\log_{10}IC_{50}$ and slope parameters of the concentration response curve fits determined by the test laboratories in the intra-laboratory analyses. The inter-laboratory analysis also compares among laboratories the average differences of the full enzyme activity and the background activity control results obtained at the end of each replicate with those obtained at the beginning.

The objectives of the inter-laboratory statistical analysis are to:

- Determine the average values and variabilities among laboratories for the parameters mentioned above.
- Determine the coefficients of variation among laboratories for the $\log_{10}IC_{50}$ and slope parameters.
- Estimate the ratios of the among laboratory variation to the within laboratory variation for the parameters mentioned above.

The inter-laboratory analyses were performed on two versions of the data:

- Including all the data
- Excluding replicate 1 for Laboratory C and excluding an outlier for full enzyme activity in Laboratory A.

Statistical Analysis Methods

Statistical analyses were carried out for each of the four endpoints discussed above in the Test Organization section: $\log_{10}IC_{50}$, slope, portion effect (i.e. beginning minus end) for background activity control, and portion effect for full enzyme activity control.

For each endpoint a one-way random effects analysis of variance with heterogeneous variances among the participating laboratories was fitted to the summary responses within laboratories. Laboratory was treated as a random effect. The within laboratory variations were based on the squares of the standard errors associated with the endpoint estimates, as determined by each intra-laboratory analysis. The analysis of variance provided an estimated weighted average across all laboratories and its associated standard error as well as an estimate of the laboratory-to-laboratory component of variation. The weights entering into the weighted averages incorporated both laboratory-to-laboratory variations and within laboratory variations. The degrees of freedom associated with the overall average effect was calculated as

$$df = 2*[((1/K)*\sum(S_L^2 + S_i^2))^2]/[(var(S_L^2)+(2/K^2)*\sum(S_i^4/df_i))]$$

where S_L^2 is the random laboratory to laboratory variance, S_i^2 and df_i are the reported within laboratory variance and degrees of freedom for the i^{th} laboratory, $var(S_L^2)$ is the variance of S_L^2, and K is the number of laboratories (Hartung and Makambi, 2001).

For each endpoint, the estimated overall average and its associated standard error (incorporating within-laboratory variation and laboratory to laboratory variation) and degrees of freedom were used to construct a 95% confidence interval. The individual effect and associated 95% confidence interval (based on the within laboratory standard error) for each laboratory were also determined. These were plotted side-by-side to provide a graphical comparison among the laboratories.

It should be noted that when calculating the mean $\log_{10}IC_{50}$ and slope and associated standard errors across replicates, Laboratories A and C incorporated the replicate-to-replicate component of variation in the standard errors of the averages, while Laboratory B did not. Also Laboratories A and C calculated the differences between beginning and end and associated standard errors when comparing the full enzyme activity and background activity controls obtained at the beginning and those obtained at the end of each replicate, while Laboratory B reported only beginning and end values. The sums of the beginning and end values must be equal to 0 for background activity control and 200 for full enzyme activity control. Therefore, for Laboratory B the differences were calculated as −2 x (End value) for background activity control and 200 − 2 x (End value) for full enzyme activity controls. The associated standard errors for these differences are 2 x (standard error associated with the end values).

To describe the variability among the laboratories relative to the average value, coefficients of variation (CV) and their associated 95% confidence intervals (95% CI) were calculated for the $\log_{10}IC_{50}$ and slope parameters. The coefficient of variation is defined as the standard deviation of the effect response divided by its mean. The methods for calculating the CV and the associated 95% CI were different depending on the underlying assumption about the distributions of the endpoint parameter.

For $\log_{10}IC_{50}$, the measurements are assumed to be approximately log normally distributed. The CV therefore is expressed as

$$CV=[\exp(S^2) -1])^{½} \times 100\%$$

where S^2 is the total variance among the three laboratories. S^2 is approximated by $3(se)^2$ where se is the standard error of the pooled mean estimate. This would be exact if the within laboratory variances were equal across laboratories.

The 95% CI is based on the chi square distribution and is calculated as

$$[(\exp(df*S^2/(\chi^2_{df,\,0.975}))- 1)^{½}, (\exp(df*S^2/(\chi^2_{df,\,0.025}))- 1)^{½}]$$

where df is the estimated degrees of freedom among the three laboratories.

For slope (β), the measurements are assumed to be approximately normal. The CV therefore is expressed as

$$CV=S/\beta_{avg}$$

where S^2 is the total variance among the three laboratories, defined as above and $S \equiv \sqrt{S^2}$. The endpoints of the confidence interval for CV are based on the non-central t distribution (Lehmann, 1986).

To describe the variability among laboratories relative to variability within laboratories, the ratio of the variance between laboratories to the average variance within laboratories is calculated as

$$R=S_{lab}^2/[⅓(s_1^2 + s_2^2 + s_3^2)]$$

where S_{lab}^2 is the random component of variance among the three laboratories and (s_1^2, s_2^2, s_3^2) are the squares of the within laboratory standard errors at the three laboratories. A confidence interval for this ratio is based on the F-distribution with (ν_{lab}, ν_{wi}) degrees of freedom,

$$[R/F^{-1}(0.975), R/F^{-1}(0.025)]$$

where $\nu_{lab}=2$ and ν_{wi} is based on Satterthwaite's approximation

$$\nu_{wi} \approx [(s_1^2 + s_2^2 + s_3^2)^2]/[s_1^4/\nu_1 + s_2^4/\nu_2 + s_3^4/\nu_3].$$

This ratio is calculated for each of the four endpoint parameters.

In several places entries in the tables in the interlaboratory analysis report tables differ from corresponding entries in the intralaboratory analysis reports tables by one or a small number of trailing digits in the last decimal place. This is due to differences in rounding in intermediate calculations between the intralaboratory analyses and the interlaboratory analysis.

Statistical Analysis Results

Table 1 displays the estimated parameter values and associated within laboratory 95% confidence intervals about these values. It also displays the overall mean values across laboratories and their associated 95% confidence intervals, incorporating among laboratory variation based on the random effects analysis of variance. The overall mean was calculated with and without replicate 1 for Laboratory C and with and without the full enzyme activity control outlying value for laboratory A. These means and confidence intervals are shown in Figures 1 through 8. Each figure includes reference lines corresponding to the overall average. The estimated CVs and their associated 95% confidence intervals for overall means for $\log_{10}IC_{50}$ and for the slope are also presented in Table 1.

Table 2 displays the within laboratory variances and associated degrees of freedom for each laboratory. These are the squares of the within laboratory standard errors associated with the estimated parameter values. Table 2 also displays the random laboratory-to-laboratory variations and the squares of the standard errors of the overall mean values, as well as their associated degrees of freedom. The ratios of the random among laboratory variances to the unweighted average within laboratory variances are also displayed, with their associated 95% confidence intervals.

Laboratory C had a higher value for $\log_{10}IC_{50}$ than Laboratories A and B (Table 1), which contributed to a relatively high random laboratory variation (more than 6 times higher than the unweighted average within laboratory variation), regardless of whether replicate 1 in laboratory C was included or excluded (Table 2). The coefficient of variation for $\log_{10}IC_{50}$ was 10% when replicate 1 in Laboratory C was included and 11% when replicate 1 in Laboratory C was excluded.

The results for the slope estimates were consistent among the three laboratories Table 1). The estimated variance among the laboratories was zero or near zero (Table 2). The coefficients of variation among laboratories were 3.7% when replicate 1 in Laboratory C was included and 3.2% when replicate 1 in Laboratory C was excluded.

No significant differences (beginning minus end) existed between background activity controls for any laboratory or across the three laboratories (Table 1). The estimated variance among the laboratories was negligible (Table 2). No significant differences (beginning minus end) existed between full enzyme activity controls across the three laboratories or for Laboratories B and C. Laboratory A had a significantly higher full enzyme activity control at the beginning when an outlying value was excluded but not a significant difference when the outlying value was included (Table 1). The estimated variance among the laboratories for the background activity controls was near 0 (Table 2). The estimated variance among the laboratories for the full enzyme activity controls was near 0 when the outlying value was included and was less than the unweighted average within laboratory variance (which is inflated by the within laboratory variance in Laboratory B) when the outlying value was excluded.

References

Hartung, J. and Makambi, K.H. *Simple non-iterative t-distribution based tests for meta-analysis*. South African Statistical Journal, 2001, Vol. 35, p. 1-17.

Lehmann, E. *Testing Statistical Hypotheses*. John Wiley & Sons, Inc., 1986, p. 352-356

Table 1. Parameter Estimates and 95% Confidence Intervals for the Percent of Control Responses for Placental Aromatase Assay

Parameter	Estimate and 95% Confidence Interval[1]				CV(%)and 95% CI[4]
	Lab A	Lab B	Lab C	Average[2,3]	
Rep 1-4 for Lab C					
$Log_{10}IC_{50}$	-7.2190 (-7.4543, -6.9837)	-7.3260 (-7.4293, -7.2227)	-7.0940 (-7.1885, -6.9995)	-7.2136 (-7.3881, -7.0392)	10.1621 (5.8912, 34.3980)
Slope	-0.9830 (-1.2685, -0.6975)	-1.0070 (-1.1619, -0.8521)	-0.9662 (-1.0616, -0.8708)	-0.9816 (-1.0403, -0.9228)	3.7072 (2.2125, 10.8498)
Difference Between End and Beginning for Background Activity Control	-0.1416 (-1.8038, 1.5206)	-0.0040 (-0.1933, 0.1853)	0.1340 (-0.0715, 0.3395)	0.0253 (-0.0611, 0.1116)	
Difference Between End and Beginning for Full Enzme Activity Control	0.6019 (-21.8796, 23.0834)	-1.9780 (-37.4060, 33.4500)	2.5365 (-1.8436, 6.9166)	2.2127 (-2.1833, 6.6087)	
Rep 2-4 for Lab C Outlier Deleted for Lab A					
$Log_{10}IC_{50}$	-7.2190 (-7.4543, -6.9837)	-7.3260 (-7.4293, -7.2227)	-7.0720 (-7.1783, -6.9657)	-7.2047 (-7.3959, -7.0135)	11.0910 (6.4149, 38.1054)
Slope	-0.9830 (-1.2685, -0.6975)	-1.0070 (-1.1619, -0.8521)	-0.9852 (-1.0791, -0.8913)	-0.9907 (-1.0432, -0.9381)	3.1878 (1.8770, 9.8890)
Difference (Beginning Minus End) for Background Activity Control	-0.1472 (-1.8078, 1.5134)	-0.0040 (-0.1933, 0.1853)	0.1787 (-0.1270, 0.4844)	0.0207 (-0.0697, 0.1110)	
Difference (Beginning Minus End) for Full Enzme Activity Control	10.5925 (4.0417, 17.1433)	-1.9780 (-37.4060, 33.4500)	0.3623 (-4.3839, 5.1085)	4.2022 (-4.9895, 13.3939)	

1. The estimates and 95% CI were as reported in the intra-laboratory analyses based on the data tested by the three participating laboratories. Laboratory C provided results separately for replicates 1 to 4 and for replicates 2 to 4. Laboratory A had results with and without an outlier for full enzyme activity controls.
2. The overall effects and standard errors were estimated using a one-way ANOVA mixed model assuming the variances differed among the three laboratories, where the variances for each laboratory were fixed to be the reported variances.
3. The averages were calculated as the following:
   - including all three replicates for Laboratories A and B and all four replicates for Laboratory C;
   - including all three replicates for Laboratory B, all three replicates for Laboratory A but excluding an outlier for full enzyme activity control, and replicates 2 to 4 for Laboratory C.
4. CV is calculated for the average results for $Log_{10}IC_{50}$ and slope parameters.

Table 2. Variance Components and Ratio of Between Laboratories and Within Laboratories. Percent of Control Responses for the Placental Aromate Assay.

Parameter	Within Lab Variance[1]				Random Laboratory Variance and (p-value) (df=2)[3]	Mean Variance[4,5]	Ratio and 95% CI of Random Lab-to-Lab Variation to Average Within Lab Variation[6]
	Lab A	Lab B	Lab C	Pooled Unweighted Simple Average Results[2]			
Rep 1-4 for Lab C							
$Log_{10}IC_{50}$	0.003045 /df=2.019	0.000575 /df=2	0.00082 /df=2.823	0.00148/df=3.95	0.008904 (p=0.1297)	0.00342/df=3.40	6.0149 (0.5560, 236.055)
Slope	0.005089 /df=2.166	0.001296 /df=2	0.000771 /df=2.646	0.002385/df=3.93	0 (p=1.000)	0.000441/df=3.93	0 (-)
Difference (Beginning Minus End) for Background Activity Control	0.5565 /df=10	0.0019 /df=2	0.0071 /df=6	0.1885 /df=10.322	$3.33x10^{-22}$ (p=1.000)	0.001515/df=10.32	$1.7684x10^{-21}$ ($3.2831x10^{-22}$, $6.9678x10^{-20}$)
Difference (Beginning Minus End) for Full Enzme Activity Control	101.80 /df=10	67.7988 /df=2	4.1706 /df=14	57.9245 /df=9.052	$3.4x10^{-22}$ (p=1.000)	3.78291/df=9.05	$5.8692x10^{-24}$ ($1.0298x10^{-24}$, $2.3117x10^{-22}$)

Parameter	Within Lab Variance[1]				Random Laboratory Variance and (p-value) (df=2)[3]	Mean Variance[4,5]	Ratio and 95% CI of Random Lab-to-Lab Variation to Average Within Lab Variation[6]
	Lab A	Lab B	Lab C	Pooled Unweighted Simple Average Results[2]			
	Rep 2-4 for Lab C Outlier Deleted for Lab A						
$Log_{10}IC_{50}$	0.003045 /df=2.019	0.000575 /df=2	0.000548 /df=1.894	0.00139/df=3.535	0.01094 (p=0.1234)	0.00408/df=3.37	7.8730 (0.6308, 308.744)
Slope	0.005089 /df=2.166	0.001296 /df=2	0.00049 /df=2.03	0.002292 /df=3.66	$4.14x10^{-22}$ (p=1.000)	0.000332/df=3.66	$1.8x10^{-19}$ ($1.5x10^{-20}$, $7.1x10^{-18}$)
Difference (Beginning Minus End) for Background Activity Control	0.5556 /df=10	0.0019 /df=2	0.0121 /df=4	0.1898 /df=10.499	$5x10^{-22}$ (p=1.000)	0.001664/df=10.50	$2.6361x10^{-21}$ ($4.927x10^{-22}$, $1.0387x10^{-19}$)
Difference (Beginning Minus End) for Full Enzme Activity Control	8.3857/ df=9	67.7988 /df=2	4.5373 /df=10	26.9072 /df=2.8230	17.7142 (p=0.2214)	10.5319/df=3.81	0.6583 (0.0369, 25.7706)

1. The within laboratory variance for a given laboratory is the square of the standard error associated with the parameter estimate, which was reported in the intra-laboratory analyses based on the data tested by the three participant laboratories. Laboratory C provided results separately for replicates 1 to 4 and for replicates 2 to 4. Laboratory A had results with and without an outlier for the full enzyme activity controls
2. Pooled unweighted average results for within laboratory are the simple averages of the within laboratory variances among the three laboratories, and the associated degree of freedom was calculated using Satterthwaite's approximation.
3. A one-way ANOVA mixed model assuming the variances differed among the three labs, where the within laboratory variance for each laboratory was fixed to be the reported variance, was fitted to estimate the random laboratory variance.

4. Mean Variance is the square of the standard error of the pooled weighted mean, including among laboratory variation.
5. Degrees of freedom for the Mean Variance was estimated as $2*((1/K)*\sum(S_L^2 + S_i^2))^2/(var(S_L^2)+(2/K^2)* \sum(S_i^4/df_i))$, where S_L^2 is random lab variance, S_i^2 and df_i are reported variance and degree of freedom for a given laboratory, $(var(S_L^2)$ is the variance associated with the estimation of S_L^2, and K is the number of laboratories (Hartung and Makambi, 2001).
6. Ratio of random among laboratory variance and unweighted simple average of within laboratory variances

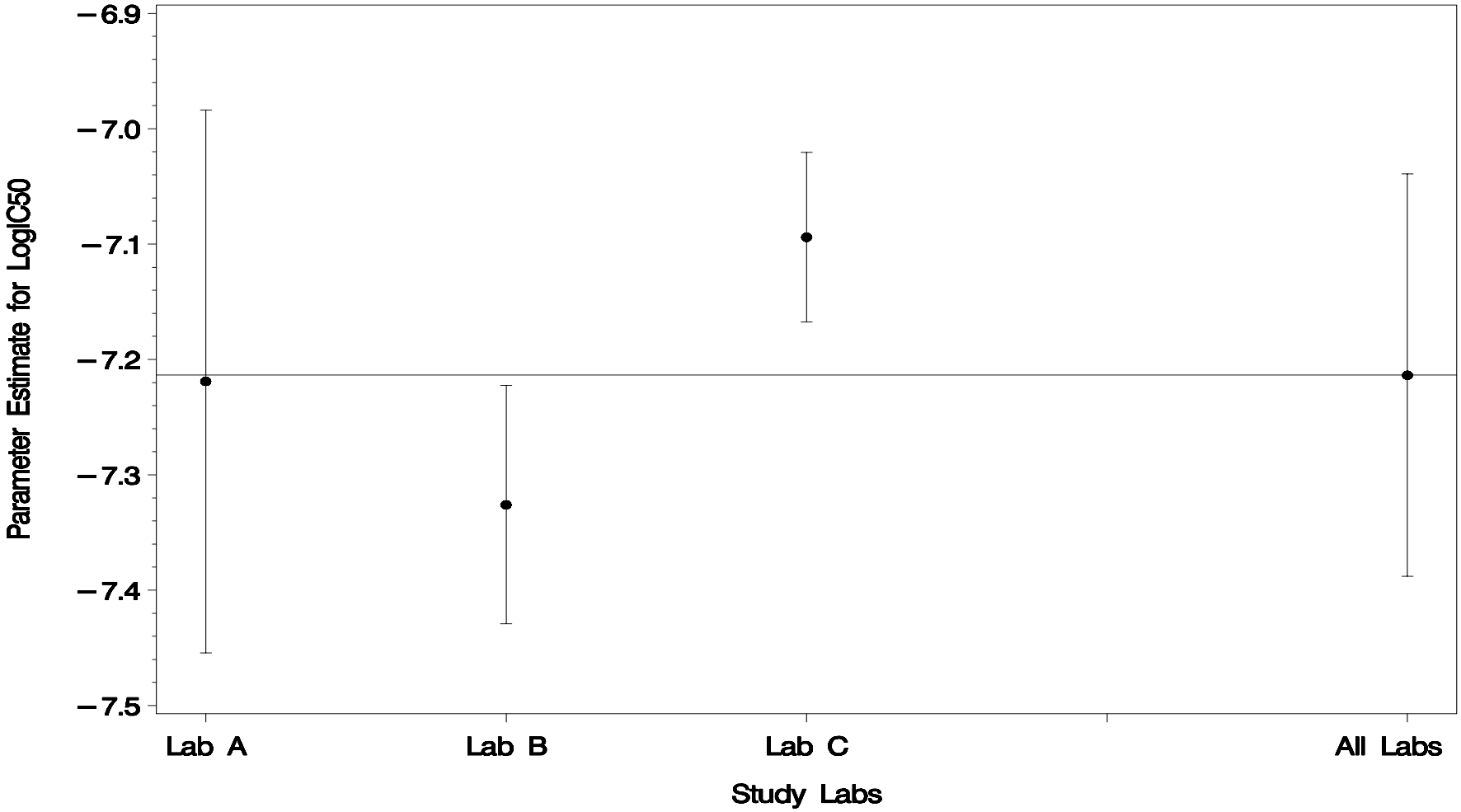

Figure 1. **Parameter Estimates and Their Associated 95% Confidence Intervals for $Log_{10}IC_{50}$ in the Placental Aromatase Assay, Across Laboratories and by Each Laboratory. The Reference Line Corresponds to the Average Across Laboratories. Laboratory C Includes Replicates 1-4.**

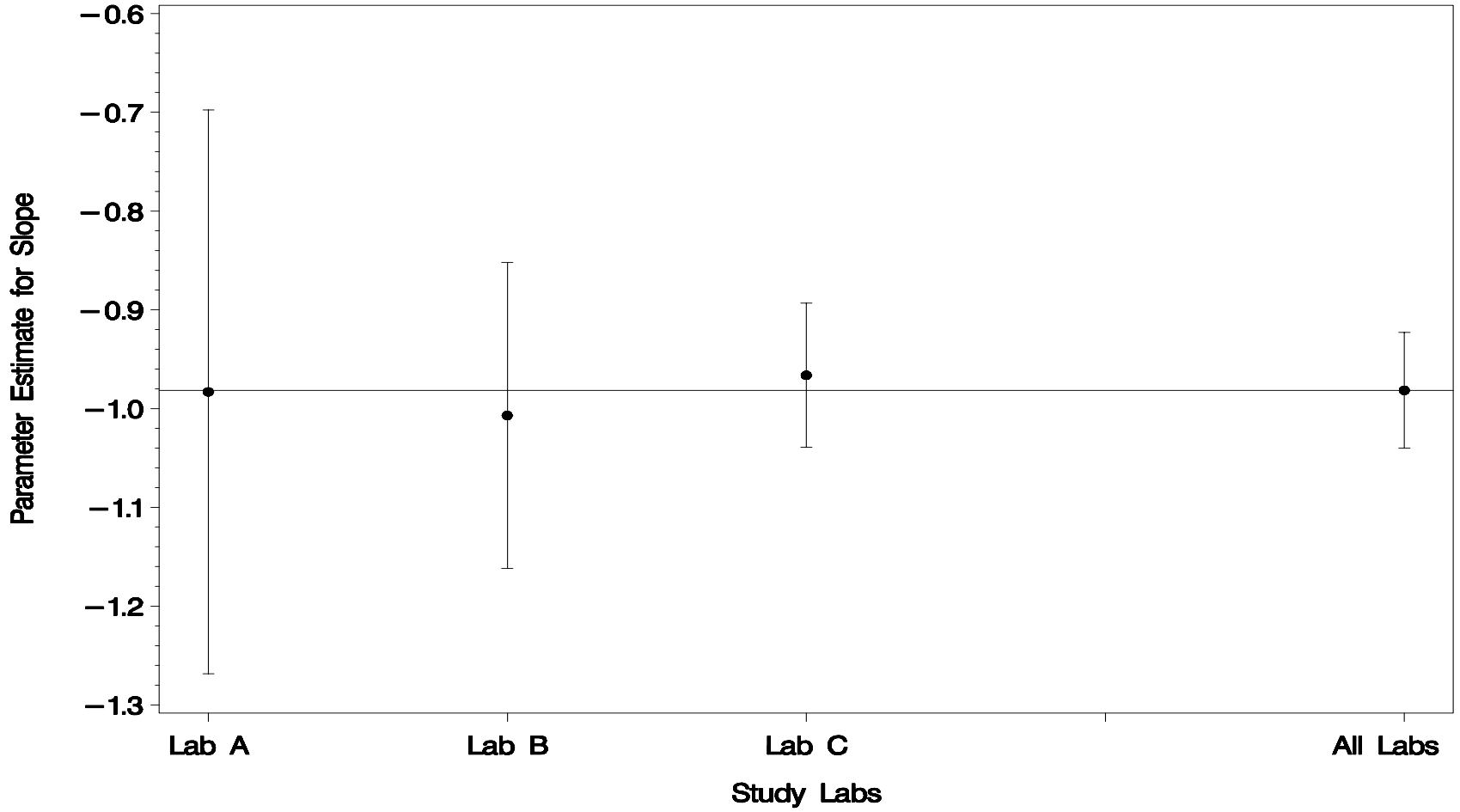

Figure 2. Parameter Estimates and Their Associated 95% Confidence Intervals for Slope in the Placental Aromatase Assay, Across Laboratories and by Each Laboratory. The Reference Line Corresponds to the Average Across Laboratories. Laboratory C Includes Replicates 1-4.

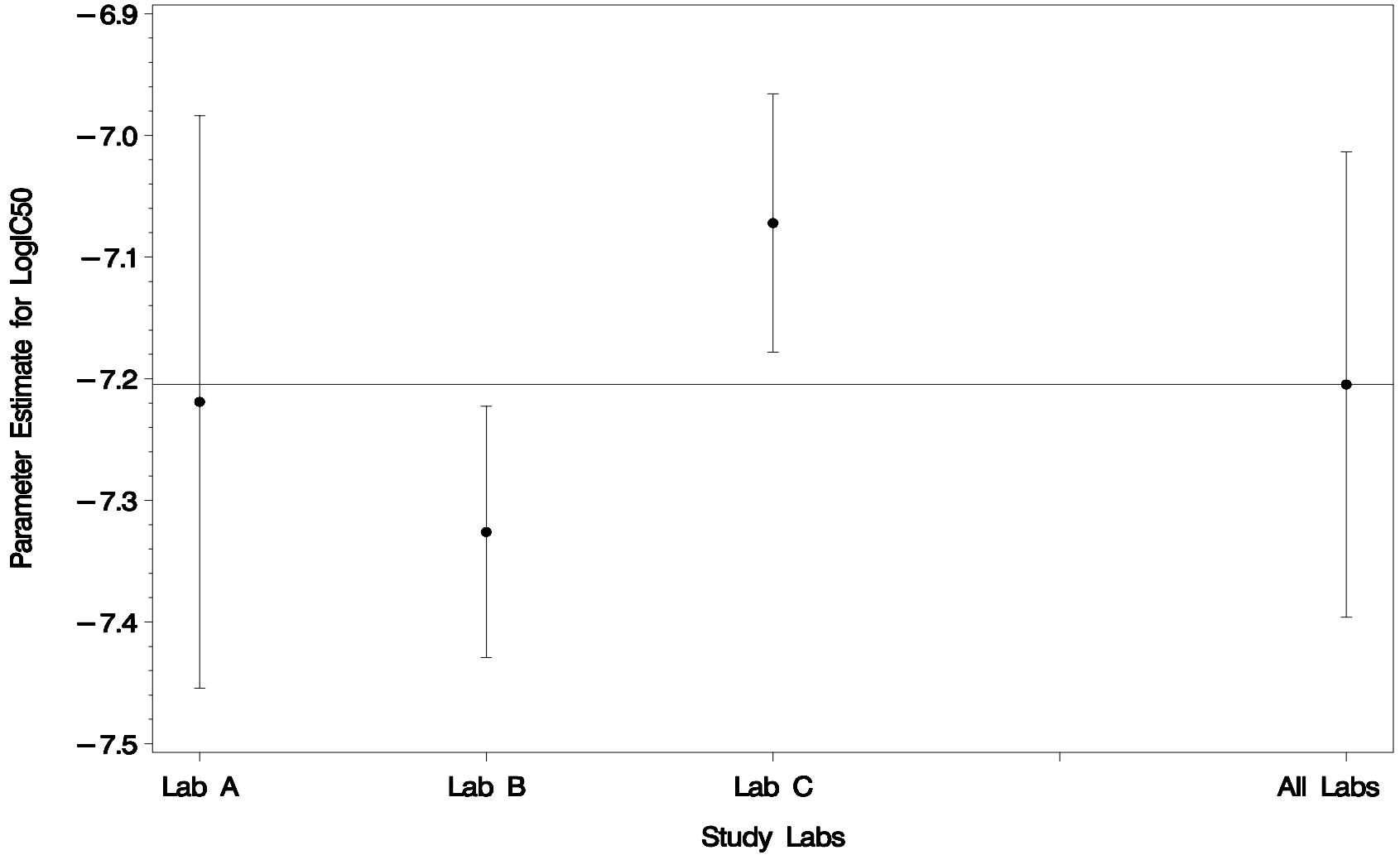

Figure 3. **Parameter Estimates and Their Associated 95% Confidence Intervals for $Log_{10}IC_{50}$ in the Placental Aromatase Assay, Across Laboratories and by Each Laboratory. The Reference Line Corresponds to the Average Across Laboratories. Laboratory C Includes Replicates 2-4.**

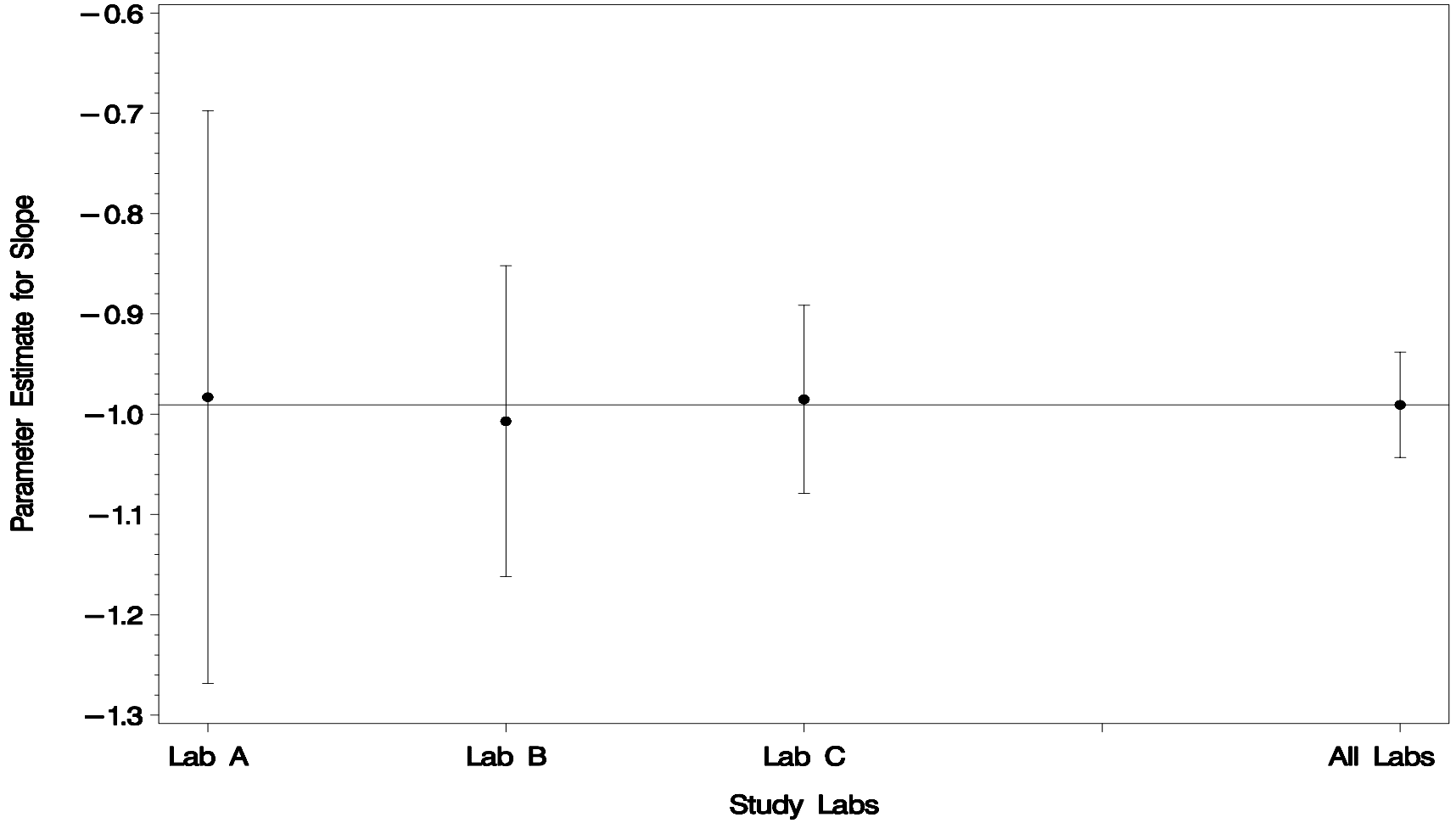

Figure 4. Parameter Estimates and Their Associated 95% Confidence Intervals for Slope in the Placental Aromatase Assay, Across Laboratories and by Each Laboratory. The Reference Line Corresponds to the Average Across Laboratories. Laboratory C Includes Replicates 2-4.

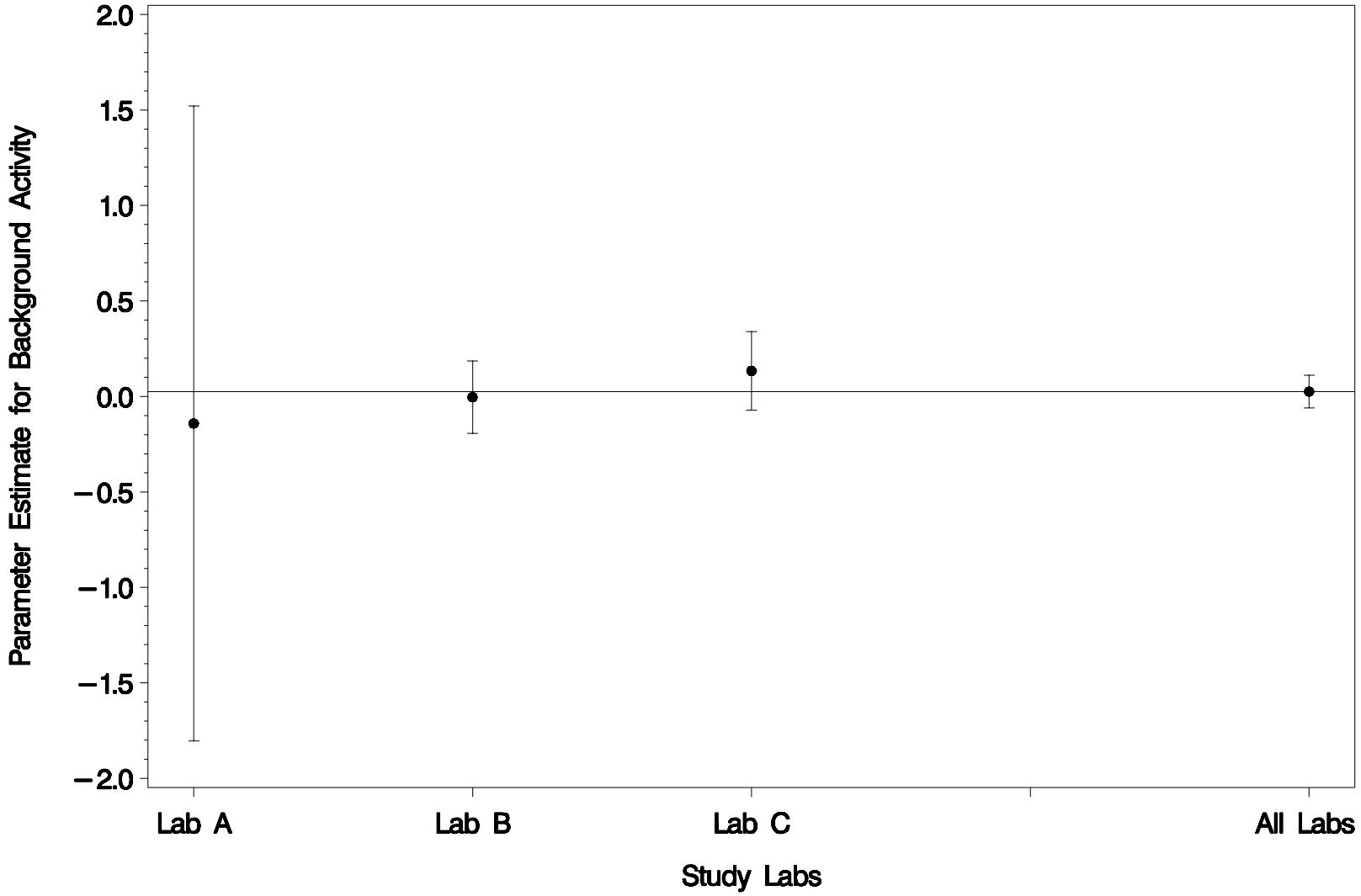

Figure 5. Parameter Estimates and Their Associated 95% Confidence Intervals for Difference (Beginning Minus End) for the Background Activity Controls in the Placental Aromatase Assay, Across Laboratories and for Each Laboratory. The Reference Line Corresponds to the Average Across Laboratories. Laboratory C Includes Replicates 1-4.

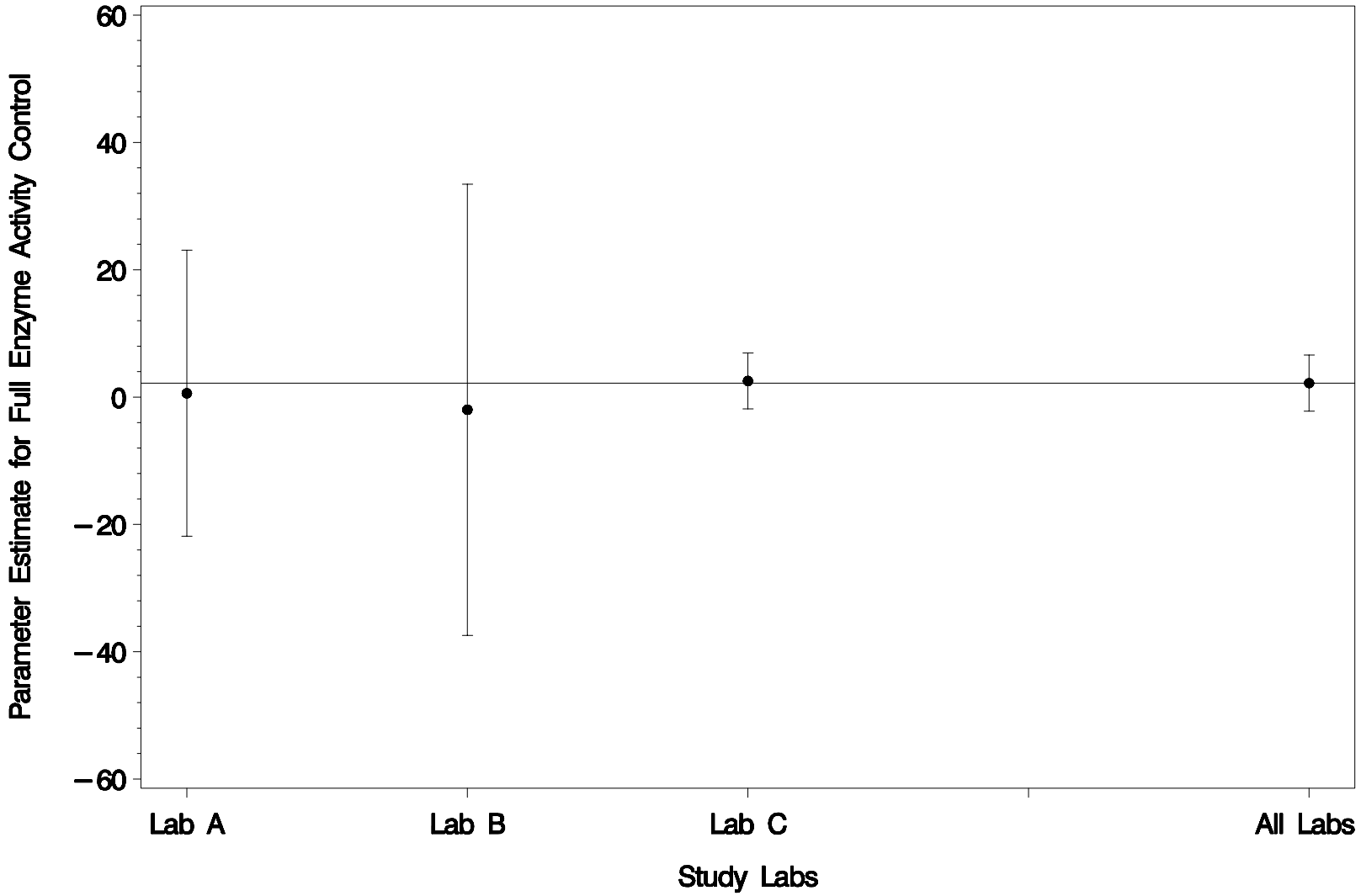

Figure 6. Parameter Estimates and Their Associated 95% Confidence Intervals for Difference (Beginning Minus End) for the Full Enzyme Activity Controls in the Placental Aromatase Assay, Across Laboratories and for Each Laboratory. The Reference Line Corresponds to the Average Across Laboratories. Laboratory C Includes Replicates 1-4.

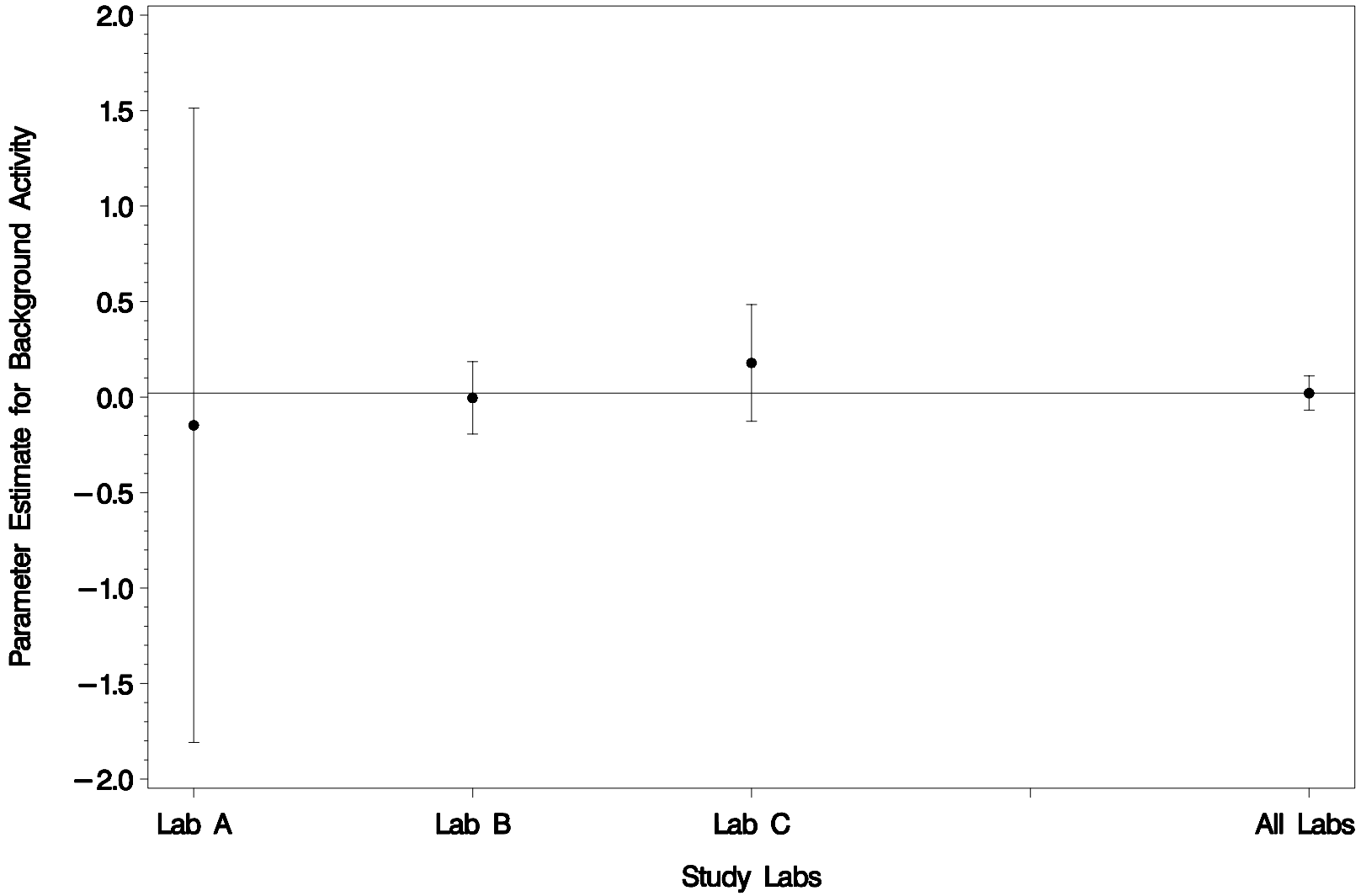

Figure 7. Parameter Estimates and Their Associated 95% Confidence Intervals for Difference (Beginning Minus End) for the Background Activity Controls in the Placental Aromatase Assay, Across Laboratories and for Each Laboratory. The Reference Line Corresponds to the Average Across Laboratories. Laboratory C Includes Replicates 2-4 and Laboratory A Excludes an Outlying Value for Full Enzyme Activity.

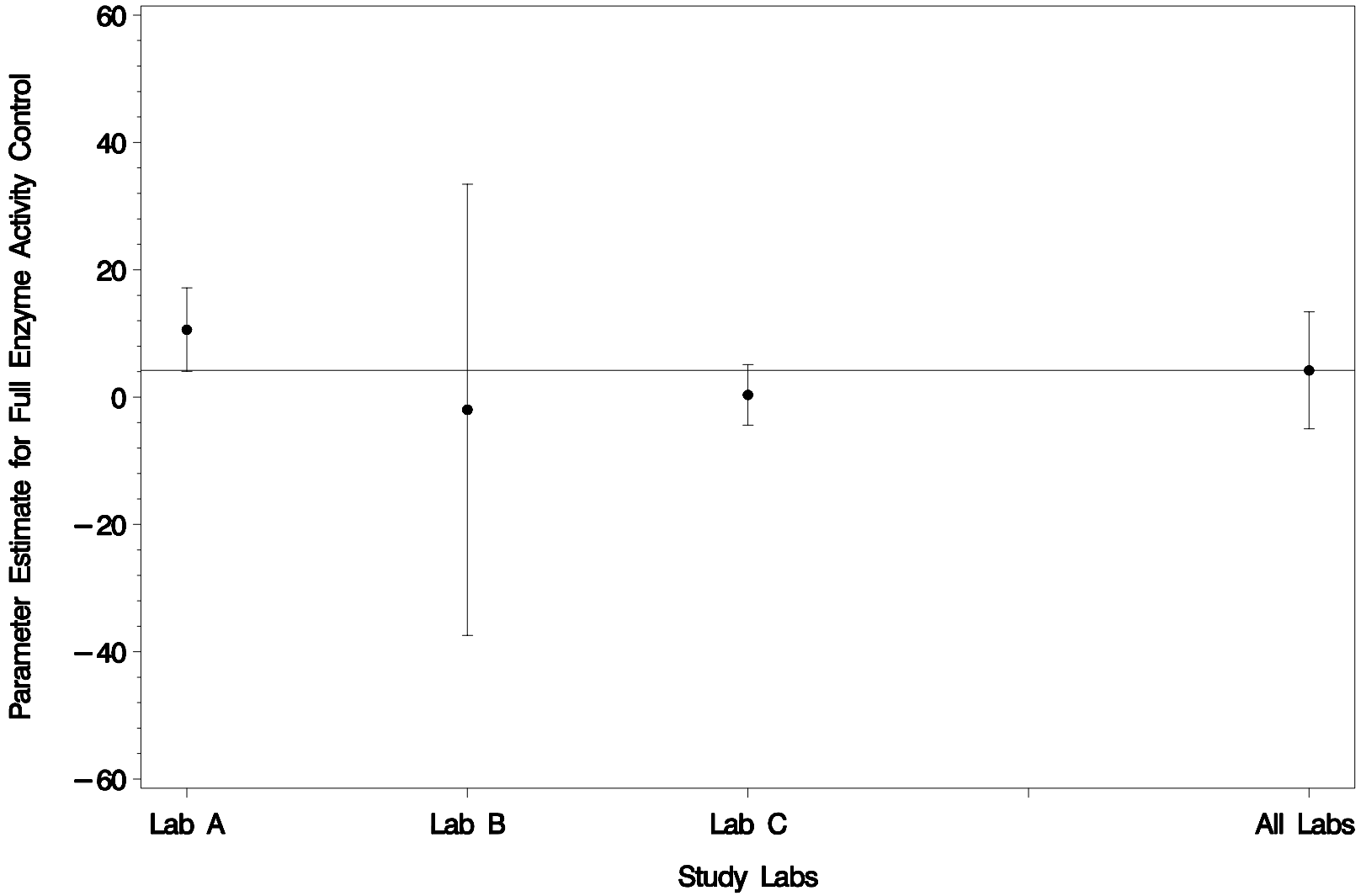

Figure 8. Parameter Estimates and Their Associated 95% Confidence Intervals for Difference (Beginning Minus End) for the Full Enzyme Activity Controls in the Placental Aromatase Assay, Across Laboratories and for Each Laboratory. The Reference Line Corresponds to the Average Across Laboratories. Laboratory C Includes Replicates 2-4 and Laboratory A Excludes an Outlying Value for Full Enzyme Activity.