**Appendix 2. Directions for using JMP, R, or an R-based application to fit LacZ data to a four-parameter logistic regression curve and calculate estimates of interpolated estradiol equivalents (EEQs) for test samples in ng/ml.**

**A. JMP (version 13.0.0)**

1) Copy and paste LacZ values, log-transformed E2 standard concentrations (log ng/ml), and sample names into JMP (as in Figure 1 below). Ensure that LacZ values and E2 standard concentrations are defined in JMP as continuous and numeric data. Fit the means of LacZ values for E2 standards and the corresponding log-transformed E2 concentrations to a four-parameter logistic regression model by clicking on Analyze > Specialized Modeling > Nonlinear. Select LacZ values as the Y variable and log-transformed E2 concentrations as the X variable.

2) Click on the red “Fit Curve” button above Model Comparison and choose Sigmoid Curves > Logistic Curves > Fit Logistic 4P.

3) Click on the red “Logistic 4P” button above Prediction Model and choose Save Formulas > Save Inverse Prediction Formula. Estimates of interpolated log E2 equivalents (log ng/ml EEQs) will then appear in a new column of the JMP data file.

4) Export log ng/ml EEQ data to spreadsheet software such as Excel. In a new column, calculate the antilog of log ng/ml EEQ (=power(10,log ng/ml EEQ)). Label the new column “EEQ (ng/ml).” The EEQ (ng/ml) values represent the concentration of estrogen equivalent ligands in each well.

5) If any estrogenicity was detected in extraction control wells, be sure to subtract the mean EEQ (ng/ml) for extraction controls (prepared in step 2.2) from the means of sample EEQs (ng/ml).

6) Output from the JMP analysis represents estradiol equivalents (EEQs in ng/ml) in the wells and should be multiplied by 325/5 to yield EEQs per ml or gram of extracted sample (see step 8.1).

**B. R code**

1) Construct an Excel data file using the layout shown in Figure 1 below and save the file as a template.csv file.

****

Fig. 1. Layout for an Excel document saved as a template.csv file. Before running the logistic regression analysis in R, paste calculated lacZ values for estradiol standards and samples into column D (data shown here are for example only).

2) Open R or RStudio and enter the code below. Run this code for each of the data files.

* + Note: The code below is intended for use with a *YESdataalpha.csv* file. To analyze the *YESdatabeta.csv* file, simply replace “YESdataalpha.csv” with “YESdatabeta.csv” in the second and third lines of code.

#Set a working directory and read in the data file. Type the name of the folder containing *YESdataalpha.csv* and *YESdatabeta.csv* files in the place of **folder name here**. For example, John Smith might have a folder on his desktop called *assay data*. The folder name for this folder could be C:/Users/jsmith/assay data. Include the quotation marks around the folder name, and be sure all slash marks are forward slashes (/).

setwd("**folder name here**")

read.csv("YESdataalpha.csv")

YESdata<-read.csv("YESdataalpha.csv")

#Assign column headings to terms X and Y and then set up a data frame with these data.

X<-YESdata$LogE2.ng.ml.[1:7]

Y<-YESdata$LacZ[1:7]

df<-data.frame(X,Y)

#Determine the parameters of the four-parameter logistic regression.

cstart<-(min(Y))

dstart<-(max(Y))

minx<-(min(YESdata$LogE2.ng.ml.[1:7]))

maxx<-(max(YESdata$LogE2.ng.ml.[1:7]))

bstart<-(maxx+minx)/2

astart<-((dstart-cstart)/(maxx-minx))

model<-nls((Y~(lower.asymptote+((upper.asymptote-lower.asymptote)/(1+exp(-slope\*(X-inflection)))))),data=df,start=list(slope=astart,inflection=bstart,lower.asymptote=cstart,upper.asymptote=dstart),trace=TRUE)

model

#NOTE: If R returns an error when running the model, try adjusting the number of iterations by adding a line of code before the model<-nls(…) code above. The new line of code is **nls.control(maxiter=1000)**. This code changes the number of model iterations attempted from the default of 50 to the number specified (1000).

#Plot assay data and fit a curve.

equation<-nls((Y~(lower.asymptote+((upper.asymptote-lower.asymptote)/(1+exp(-slope\*(X-inflection)))))),data=df,start=list(slope=astart,inflection=bstart,lower.asymptote=cstart,upper.asymptote=dstart))

xdata<-seq(min(X),max(X),length.out=300)

plot(X,Y,xlab = "Log[E2] (ng/ml)]", ylab = "LacZ")

lines(xdata,predict(equation,data.frame(X=xdata)))

#Calculate the R-squared value for the regression.

r.squared<-(1-((sum(residuals(model)^2))/(sum((Y - mean(Y))^2))))

r.squared

#Use the regression parameters to inverse predict EEQs from LacZ values.

#Note: A warning about NaNs produced means that some samples have EEQs of 0.

coefs<-as.numeric(summary(model)$coefficients)

slope<-coefs[1]

inflection<-coefs[2]

lower.asymptote<-coefs[3]

upper.asymptote<-coefs[4]

LacZ<-YESmodel["LacZ"]

logEEQ<-(inflection-((1/slope)\*log(((upper.asymptote-lower.asymptote)/(LacZ-lower.asymptote))-1)))

EEQ<-10^logEEQ

#Make tables of log(EEQ) and EEQ values.

names<-YESmodel$Sample

dflogEEQ<-data.frame(names,logEEQ)

dflogEEQ[sapply(dflogEEQ,is.na)] = NA

colnames(dflogEEQ)<-c("Sample name","Log EEQ (ng/ml)")

print(dflogEEQ,row.names=FALSE,right=FALSE,digits=5)

dfEEQ<-data.frame(names,EEQ)

dfEEQ[sapply(dfEEQ,is.na)] = NA

colnames(dfEEQ)<-c("Sample name","EEQ (ng/ml)")

print(dfEEQ,row.names=FALSE,right=FALSE,digits=5)

#Repeat all of these steps using the *YESdatabeta.csv* file.

3) If any estrogenicity was detected in extraction control wells, subtract the mean EEQ (ng/ml) for extraction controls (prepared in step 2.2) from the means of sample EEQs (ng/ml).

4) Output from the R analysis represents estradiol equivalents (EEQs in ng/ml) in the wells and should be multiplied by 325/5 to yield EEQs per ml or gram of extracted sample (see step 8.1).

**C. R-based application**

1) Construct a .csv data file as instructed in step B1 above. Upload this file on the first tab of the online YES Assay Data Analysis application (Roark, A.M., 2016, <https://furmanbiology.shinyapps.io/YESapp/>). This tab is labeled “Upload File.”

2) Click on the second tab of the YES Assay Data Analysis application. This tab is labeled “Plot of Your Data.” Using the example figure, estimate the four parameters of the plot that was generated using the uploaded data and enter these estimates in the fields provided.

3) Click on the third tab of the YES Assay Data Analysis application. This tab is labeled “Regression Model.” If any error messages appear in red, return to the second tab and enter different parameter estimates. Repeat until a plot with a regression model and an R2 value is shown on the third tab.

4) Click on the fourth tab of the YES Assay Data Analysis application. This tab is labeled “Log EEQs.” The values on this tab represent the interpolated x-axis values corresponding to the LacZ values of samples.

5) Click on the fifth tab of the YES Assay Data Analysis application. This tab is labeled “EEQs.” The calculated EEQ values on this tab were calculated as the anti-logs of the values on the fourth tab and represent the concentration of estrogen equivalent ligands (EEQs (ng/ml)) in each well of the plate assay.

6) If any estrogenicity was detected in extraction control wells, subtract the mean EEQ (ng/ml) for extraction controls (prepared in step 2.2) from the means of sample EEQs (ng/ml).

7) Output from the online application represents estradiol equivalents (EEQs in ng/ml) in the wells and should be multiplied by 325/5 to yield EEQs per ml or gram of extracted sample (see step 8.1).