dBETS: An online software package for diffusion test breakpoint determination

Glen DePalma Ph.D.
Exponent
Data Sciences Practice
149 Commonwealth Drive
Menlo Park, CA 94025
gdepalma@exponent.com
650-688-7331

John Turnidge Ph.D.
Women's and Children's Hospital
SA Pathology
72 King William Rd
North Adelaide, SA, 5006
Australia

Bruce A. Craig Ph.D.
Department of Statistics
Purdue University
150 N. University Street
West Lafayette, IN 47907

Abstract

dBETS (**d**iffusion **B**reakpoint **E**stimation **T**esting **S**oftware) is a free online software package for the determination of diffusion test breakpoints. It provides both numerical summaries and visual displays of the error-rate bounded (ERB) method and two more recent model-based determination procedures. All functions within the package have user-specified options. This package provides a standard platform to be used for all diffusion breakpoint determination analyses.

Keywords: Minimum Inhibitory Concentration, Disk Diffusion, Breakpoint Estimation, Software Package, Error-Rate Bounded, Susceptibility, Antibiotics, Statistical Modelling

Running Title: dBETS

1 Objectives

The minimum inhibitory concentration (MIC) assay determines the lowest concentration of an antimicrobial agent that completely or near completely inhibits the growth of a microorganism. In the testing of antimicrobial agents, MICs are accepted as the reference to which all other testing methods are compared. One of the most widely-used alternative methods is the disk diffusion test, which determines the diameter (DIA) of complete growth inhibition (the clear zone) around an antimicrobial disk of known potency placed on agar that has been inoculated with the microorganism immediately prior to disk placement. Because there is not a straightforward conversion between clear zone diameter and MIC, paired test results for a large collection of pathogens are generated to help determine the breakpoints for the diffusion test.

Traditionally, the error-rate bounded (ERB) method has been used to do this [1]. This procedure involves selecting DIA breakpoints, which minimize the observed classification discrepancies. While simple and intuitive, this approach is very sample dependent, not precise, and does not fully account for the test characteristics such as rounding and measurement error [2, 3].

Model-based approaches were proposed to better account for these issues (Craig, 2000; Qi 2008; DePalma 2013) [4, 5, 6]. These methods propose a true underlying relationship between the DIA and MIC and link it to the observed results using a probability model based on test behaviors. As a result, these model-based procedures focus on the probabilities of various discrepancies rather than the proportion of observed discrepancies. This leads to more precise and less sample dependent breakpoints.

To date, the model-based approaches have not been readily accessible to users. `dBETS` (**d**iffusion **B**reakpoint **E**stimation **T**esting **S**oftware) is a free online software package that provides access to the model-based procedures as well as the ERB method. In this article, we provide an overview of the methods that are available in this package and run through an example to demonstrate its functionality and ease of use.

2 Methods

dBETS was created using the Shiny application from RStudio [7, 8]. R, along with several R packages, make up the underlying code [9, 10, 11]. However, all that is required to use it is internet access and a modern web browser. The package is located at the following address: https://dbets.shinyapps.io/dBETS/. Source code is available upon request.

Each page in dBETS consists of a navigational panel near the top left corner to jump between the three sections of dBETS: 1) Data Entry, 2) Error-Rate Bounded Method, and 3) Model-Based Approaches. Also on each page is a left-side panel of user-specified inputs and options. Once these inputs are set, a click of a run button (located on the panel below the inputs) activates the procedure. The output will appear in the main section on the right once the procedure completes. The user can always change the input and/or options and re-run the procedure. The user just needs to choose the new options and reclick the run button. Many of the procedures have the option of displaying a graph or visual representation of the results. In these situations, additional options will appear that adjust the visualization. You can also save this visualization to your local computer as a .png file.

Data can be uploaded to dBETS via the local hard drive or from a URL. A description of the proper format of the data is provided in the manual. Various input parameters are available such as setting MIC breakpoints and a description of the data. Once the "Input Options / Load Data / Produce Descriptives" button is activated the data are loaded, descriptive statistics are produced, and graphical options appear in the side panel. Figure 1 displays the home page after activating buttons using the default data set.

2.1 Error-Rate Bounded Method

The ERB method in dBETS finds the DIA breakpoints that minimize a discrepancy index that is based on user-specified percentages. Two sets of percentages must be specified, the first set refers to the allowable error rates for isolates observed within one dilution of the MIC indeterminant range. The second set specifies the error rate for isolates observed outside this range. These percentages are converted into weights, which are in turn used in the overall discrepancy index. The default values in `dBETS` are the percentages from the Clinical and Laboratory Standards Institute's (CLSI) M23 document [12].

These weights are inversely related to the allowable error rates, so those discrepancies with low error rates will be the most heavily weighted. The index is simply a weighted sum of the discrepancies and can be expressed as:

$$index = w_{VM1} \times \#VM1 + w_{M1} \times \#M1 + w_{m1} \times \#m1 + w_{VM2} \times \#VM2 + w_{M2} \times \#M2 + w_{m2} \times \#m2 \quad (1)$$

where w_* represents the weight associated with a discrepancy * and #* represents the observed count of discrepancy *.

In addition, a bootstrap procedure is available to assess the uncertainty in the DIA breakpoint estimates. This method involves resampling the data (with replacement) and computing the DIA breakpoints for each resample. Both a numerical and graphical summary of this uncertainty are displayed showing the distribution of selected DIA breakpoints. The more disperse the choice of the selected DIA breakpoints in these resamples, the less confidence one can have for using a particular set of DIA breakpoints.

2.2 Model-Based Approaches

Model-based procedures describe the observed pattern in the scatterplot using a probability model. This model accounts for the inherent variability of each test (e.g., a 3-fold dilution range for the MIC test) by estimating the true continuous (but unobservable) MIC and DIA values for each test pathogen (Craig, 2000). These true values are linked based on the assumed monotonically decreasing relationship between the MIC and DIA [13].

2.2.1 True Relationships

The two model-based procedures in this software either assume

1. The true DIA and MIC values follow a 4-parameter logistic curve. This is a generalization of the 3-parameter logistic relationship described in Craig 2000.

2. The true DIA and MIC values follow a nonparametric curve based on I-splines. This method is more flexible and data-driven.

The four-parameter logistic curve can be written as:

$$g(m_i) = \beta_1 \frac{fm_i \exp(-\beta_3(\beta_2 - m_i)) + (1 - fm_i)\exp(-\beta_4(\beta_2 - m_i))}{1 + fm_i \exp(-\beta_3(\beta_2 - m_i)) + (1 - fm_i)\exp(-\beta_4(\beta_2 - m_i))} \tag{2}$$

where

$$fm_i = 1/(1 + \exp(-mb \times (\beta_2 - m_i))) \text{ and } mb = (2 \times \beta_3 \times \beta_4)/(\beta_3 + \beta_4)$$

This curve allows for asymmetry around the inflection point β_2 and reduces to the three-parameter model when $\beta_3 = \beta_4$. It is a weighted average of two logisitic curves with rates β_3 and β_4 and weights fm_i and $1 - fm_i$. This parameterization is sometimes referred to as a five-parameter logistic model in the literature, because mb can be considered an additional parameter [14]. In addition to the coefficients all being positive, the formulation of mb here guarantees a monotonically decreasing function.

The spline model uses I-splines [15] to model the underlying relationship between the MIC and DIA test. Given a knot sequence and bases ($\boldsymbol{I}$), the true DIA test results are connected to the true MIC test results by

$$d_i = g(m_i) = \sum_{j=1}^{j=B} \beta_j I_j(m_i) \tag{3}$$

where $\boldsymbol{\beta}$ are the unknown spline coefficients and B is the number of bases. Restricting the coefficients to be positive results in a monotonic relationship.

Figure 2 displays the fits of both the I-spline and logistic curves to two simulated data sets. The top figure shows both relationships fitting well while in the bottom figure only the I-spline curve fits the true relationship well. Note the MIC is on the x-axis here because the true relationships model d_i given m_i.

2.2.2 Determination of DIA Breakpoints

Given a true MIC value and the probability model that describes the distribution of the observed MIC results, the probability that an observed MIC will fall in any of the three classification regions (resistant, susceptible, and indeterminant) can be computed. Traditionally, these regions have been defined by the MIC test breakpoints. Given our modelling of the underlying truth, a distinction is made between test MIC breakpoints and true MIC breakpoints. Due to the upwards rounding of the MIC test, we assume the true MIC are shifted down by a mean value of 0.5 ($\log_2$). Thus if the test MIC breakpoints are -1 and 1 the true MIC breakpoints are -1.5 and 0.5. The test MIC breakpoints are referred to as M_L and M_U while the true MIC breakpoints are referred to as M_L^* and M_U^*.

Similarly, given a set of DIA breakpoints, D_U and D_L, and the true underlying relationship $g(m)$, the probability of an observed DIA falling in each region can also be calculated. The model-based approaches focus on the probability of correct classification. These probabilities for given MIC m are

$$p_{MIC}(m) = \begin{cases} \Pr(x \leq M_L) = \Phi\left(\frac{M_L - m}{\sigma_m}\right) & m \leq M_L^* \\ \Pr(M_L < x < M_U) = \Phi\left(\frac{M_U - 1 - m}{\sigma_m}\right) - \Phi\left(\frac{M_L - m}{\sigma_m}\right) & M_L^* < m < M_U^* \\ \Pr(x \geq M_U) = 1 - \Phi\left(\frac{M_U - 1 - m}{\sigma_m}\right) & m \geq M_U^* \end{cases} \tag{4}$$

$$p_{DIA}(m) = \begin{cases} \Pr(y \geq D_U) = 1 - \Phi\left(\frac{D_U - .5 - g(m)}{\sigma_d}\right) & m \leq M_{L^*} \\ \Pr(D_L < y < D_U) = \Phi\left(\frac{D_U - .5 - g(m)}{\sigma_d}\right) - \Phi\left(\frac{D_L + .5 - g(m)}{\sigma_d}\right) & M_{L^*} < m < M_{U^*} \\ \Pr(y \leq D_L) = \Phi\left(\frac{D_L + .5 + g(m)}{\sigma_d}\right) & m \geq M_{U^*} \end{cases} \tag{5}$$

where Φ is the standard normal CDF.

The optimal DIA breakpoints are the set that minimizes a weighted loss function. The loss function is the accumulated squared difference in the probability of correct classification when the DIA test performs worse than the MIC test. The underlying MIC density is used as the weighting function. This can be expressed

$$L = \int_{-\infty}^{\infty} \min\left(0, p_{\text{DIA}}(g(u)) - p_{\text{MIC}}(u)\right)^2 w(u)\, du \tag{6}$$

where $w(.)$ are the weights. Other loss functions, such as absolute loss, are being investigated and will be implemented in the future.

3 Results

To provide more insight into the usability and features of the package, we will now run through an example using the sample data set. We discuss the available input options for each of the methods and show some of the outputs produced. We start with uploading the data, followed by the error-rate bounded method, and finish using the model-based methods. The dataset used throughout this example can be downloaded at: `https://dbets.shinyapps.io/dBETS/data1.csv`.

The left-side panel of Figure 1 displays the inputs for uploading data, setting MIC breakpoints and graphical options. Data may be uploaded from a local hard drive or via the internet. Lower and upper MIC breakpoints are set here, along with an option for only using one MIC breakpoint. Once the data have been uploaded, graphical options appear in the side panel, and descriptive statistics are produced. See Figure 1 for a display of this output given the default inputs.

3.1 ERB Results

The inputs for the ERB method are the min and max widths between DIA breakpoints, and the percentages associated with various discrepancies. Setting min width equal to 3 and max width equal to 7 results in searching for DIA breakpoints at least three mm apart but no further than seven mm. These options are not relevant when only using one breakpoint. The specified discrepancy percentages are used to determine the weights of the index score (1).

Figure 3 displays the output after the ERB procedure completes. The index score as well as discrepancy statistics are shown. For these data, the optimal DIA breakpoints were (34, 39). There were 64 minor discrepancies within one of the intermediate range. The scatterplot is a visual display of the results where different colors indicate different discrepancies.

3.1.1 Assessing Breakpoint Uncertainty via the Bootstrap

Figure 4 displays the uncertainty in the DIA breakpoints via the bootstrap. For these data, there is a large set of potential breakpoints with the pairs $(34, 39)$, $(33, 39)$, and $(34, 40)$ the most frequent choice among the resamples. Since no set of breakpoints was selected more than 50% of the time, there does not appear to be an obvious choice for a single set of breakpoints using the ERB method.

3.2 Model-Based Results

The inputs for the model-based methods are setting min and max widths of possible DIA breakpoints and specifying the type of loss function. Currently, the only option for the loss function is the squared loss described in (6).

Figure 5 displays the results from the logistic model. At the top is the posterior distribution of DIA breakpoints. There is a high probability that breakpoints (33, 40) are optimal. Below is a visualization of the model fit to the data. Similar output is produced when the spline model is run. For this example, the results are comparable but with (33, 40) selected with 0.95 probability.

Both model-based methods selected breakpoints $(33, 40)$ with high probability. There was no clear choice of breakpoints using the ERB method. This example highlights some of the key differences between the two approaches. First, the model-based methods use the test results to estimate an underlying true relationship between the MIC and DIA. The ERB approach does not assume any particular relationship. As a result, the ERB method focuses on minimizing the observed discrepancies while the model-based procedures use the estimated model to choose breakpoints that maximize the probability that the DIA test gives the correct result. These difference will often result in a different choice of breakpoints. The next section describes an approach to compare different sets of breakpoints.

3.2.1 Probability of Correct Classification given a Fitted Model

Given a fitted model and a set of DIA breakpoints, one can calculate the probability of correct classification profile over any range of true MIC values (see equations 4 and 5). An "overall" probability of correct classification is obtained as a weighted average of this profile using the estimated probability density of the MIC as the weights.

Figure 6 displays the probability of correct classification for the MIC breakpoints and two sets of DIA breakpoints (provided as inputs): (33, 40), and (34, 39). These are the breakpoints selected by the model-based approaches and the ERB method respectively. Unweighted and weighted "overall" probabilities are also provided. The graph shows the probability of correct classification for each MIC in the selected range. The further away from the MIC breakpoints the higher the probability of correct classification as one would expect. We see the model-based DIA breakpoints perform better. The DIA weighted probability of correct classification is 0.82 compared to 0.76. This is primarily due to higher probabilities in the intermediate region. Both DIA test profiles appear superior to that of the MIC test.

4 Conclusions

`dBETS` provides a standard platform for the analysis of data from antimicrobial susceptibility experiments. A standard platform makes reproducible research easier. Given the data set, one can replicate a set of results to see how a particular set of DIA breakpoints were selected. This encourages collaboration among clinicians and agencies.

`dBETS` provides descriptive statistics for the data set, visualization, and the error-rate bounded method. Perhaps most importantly, clinicians now have an easy way to use the model-based procedures that have been proposed in the literature. Comparisons of selected DIA breakpoints between the error-rate bounded method and model-based approaches are now easily produced. In addition, past data sets can be re-analyzed using both approaches.

`dBETS` will continue to evolve to meet the needs of clinicians. For example, the current software assumes known measurement error standard deviations for both the DIA and MIC tests. These were empirically chosen based on QC data. In the future, we plan to relax this and add these standard deviations as additional inputs. We look forward to working with professionals to expand the software's capabilities, fix bugs, improve usability, and introduce new features.

References

1. **Brunden MN, Zurenko GE, Kapik B.** 1992. Modification of the error-rate bounded classification scheme for use with two MIC break points. *Diagnostic Microbiology and Infectious Disease* **15 (2)**:135–140.
2. **Annis D, Craig B.** 2005. Statistical properties and inference of the antimicrobial MIC test. *Statistics in Medicine* **24**:3631–3644.
3. **Craig BA.** 2000. Modeling Approach to Diameter Breakpoint Determination. *Diagnostic Microbiology & Infectious Disease* **36.3**:193–202.
4. **Qi X.** 2008. Nonparametric Calibration of Two Common Susceptibility Tests using Interval-Censored Data with Measurement Error. PhD thesis. Purdue University.
5. **Craig BA.** 2010. *Drug Dilution vs Drug Diffusion: Calibrating the Two Suusceptibilty Tests*. Tech. rep. 99-15. Purdue University, Department of Statistics.
6. **DePalma G.** 2013. Disk Diffusion Breakpoint Determination Using a Bayesian Nonparametric Variation of the Errors-in-Variables Model. PhD thesis. Purdue University.
7. **RStudio, Inc.** 2014. *shiny: Web Application Framework for R*. R package version 0.10.0.
8. **Allaire J.** 2013. *shinyapps: Interface to ShinyApps*. R package version 0.3.54.
9. **R Core Team.** 2014. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria.
10. **Wickham H.** 2009. *ggplot2: elegant graphics for data analysis*. Springer New York.
11. **Venables WN, Ripley BD.** 2002. *Modern Applied Statistics with S*. Fourth. ISBN 0-387-95457-0. New York: Springer.
12. **NCCLS.** 1994. *National Committee for Clinical Laboratory Standards development of in vitro susceptibility testing criteria and quality control parameters; Approved guideline M23-A*. Wayne, PA.

13. **Murray PR, Rosenthal KS, Pfaller MA.** 2009. *Medical Microbiology, 6th Edition.* Saunders.

14. **Gottschalk PG, Dunn JR.** 2005. The five-parameter logistic: A characterization and comparison with the four-parameter logistic. *Analytical Biochemistry* **343(1)**:54–65.

15. **Ramsay J.** 1988. Monotone Regression Splines in Action. *Statistical Science* **3**.4:425–461.

5 Funding

This study was conducted as part of our routine work.

6 Transparency declarations

None to declare.

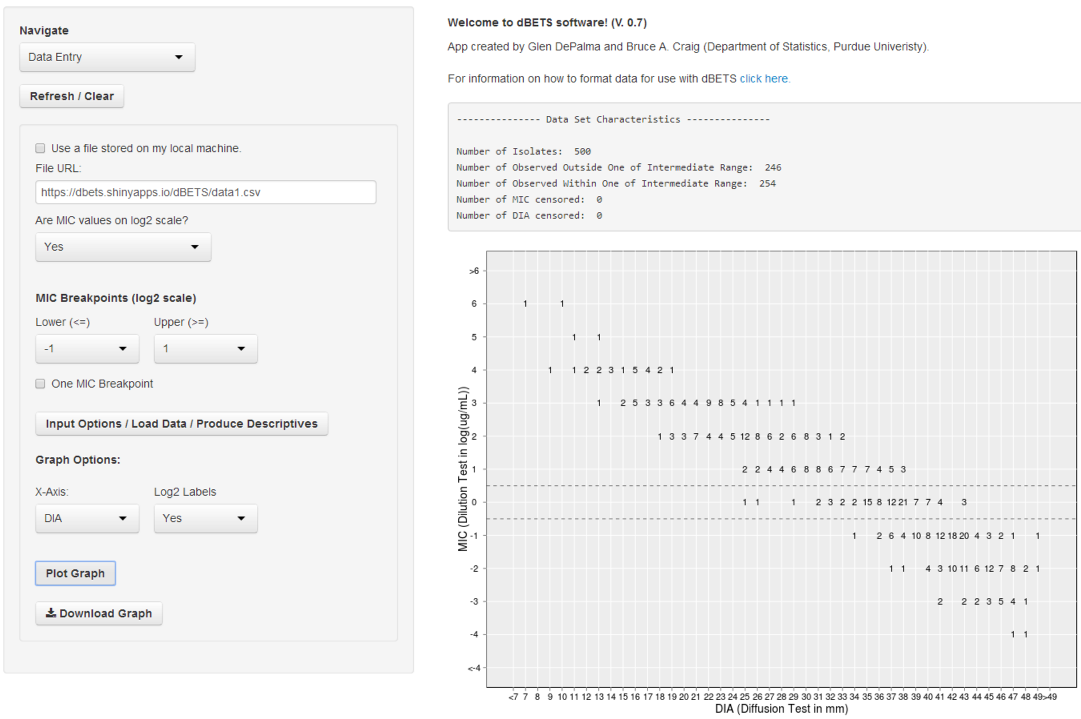

Figure 1: Date Entry page after the default input parameters have been set for both data input and graphical display.

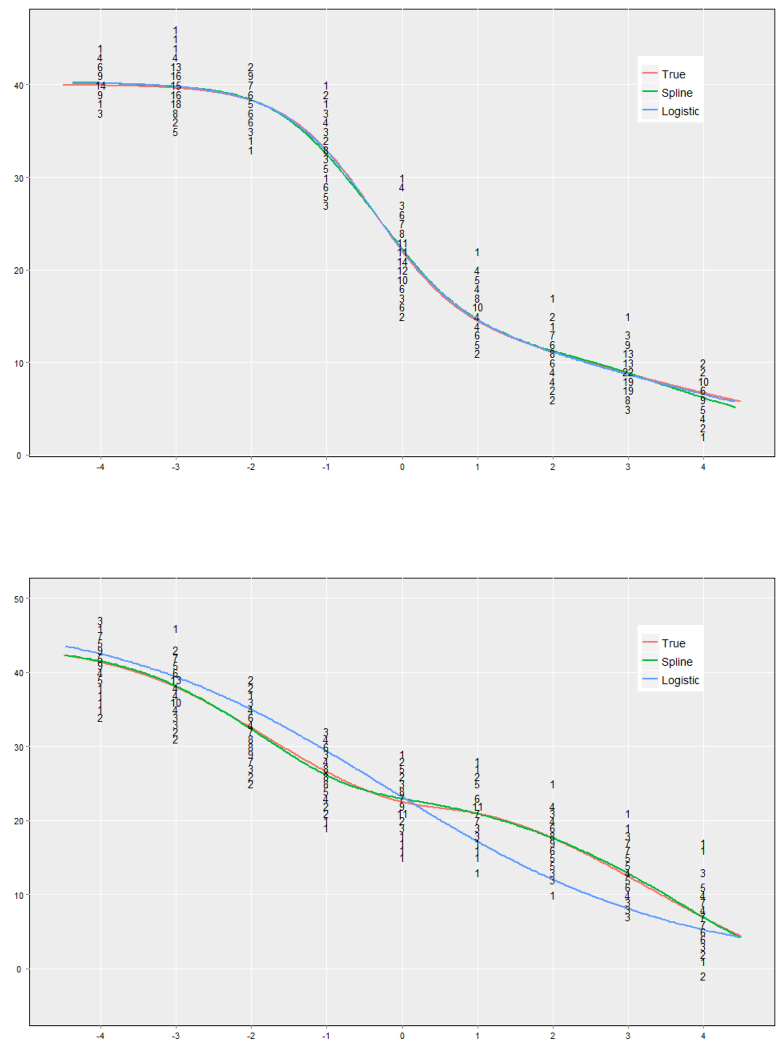

Figure 2: Example I-spline and logistic fits to simulated data. In the top graph, both curves fit the true relationship well. In the bottom graph, the I-spline curve fits well but the logistic curve struggles.

```
Optimal DIA Breakpoints for ERB: 34 39
Number of Isolates:  500
Number Observed Outside One of Intermediate Range:  246
Number Observed Within One of Intermediate Range:  254
Index Score =  0.3327572

Count (%)
   Range       Agree    Very Major   Major      Minor
  Within 1  189 (74.41)   0 (0)     1 (0.39)  64 (25.2)
 Outside 1  244 (99.19)   0 (0)      0 (0)     2 (0.81)
```

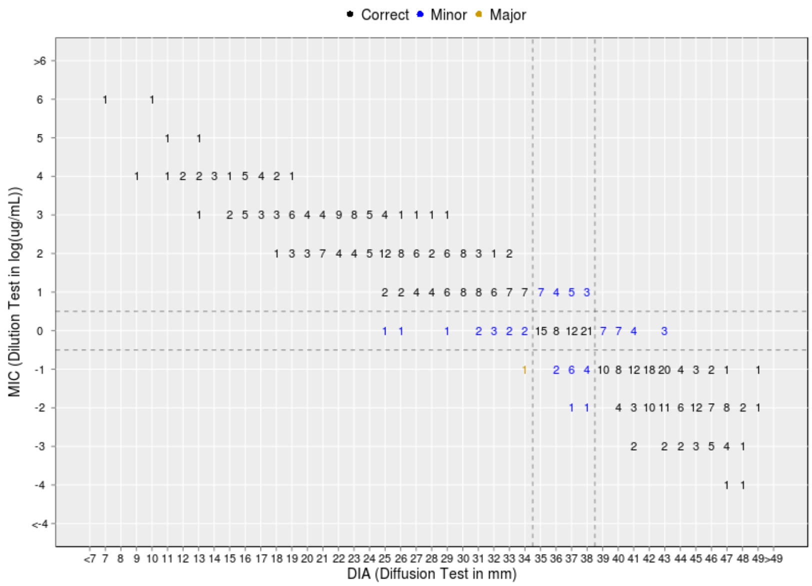

Figure 3: Example output from the error-rate bounded procedure.

Bootstrap samples = 12000

-------DIA Breakpoints by Confidence--------

DIABrkptL	DIABrkptU	Percent	Cumulative
34	39	46.60	46.60
33	39	22.65	69.25
34	40	15.21	84.46
33	40	7.81	92.27
34	41	3.04	95.31
35	39	1.70	97.01
33	41	1.47	98.48
35	40	0.66	99.14
32	39	0.26	99.40
32	40	0.21	99.61
31	39	0.10	99.71
31	40	0.09	99.80
35	41	0.08	99.88
36	40	0.04	99.92
32	41	0.04	99.96
36	41	0.02	99.98
33	38	0.01	99.99
34	38	0.01	100.00

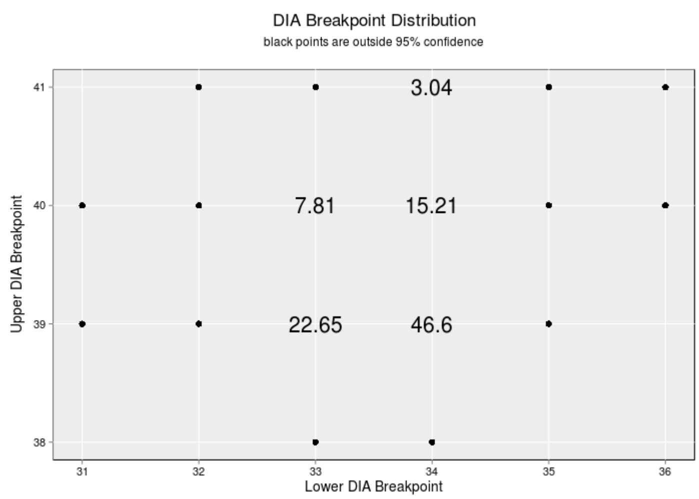

Figure 4: Distribution of possible ERB breakpoints resulting from the bootstrap.

```
-------Optimal DIA Breakpoints--------
 DIABrkptL  DIABrkptU   Percent   Cumulative
     33         40       77.67       77.67
     32         40       20.17       97.83
     33         41        2.17      100.00
```

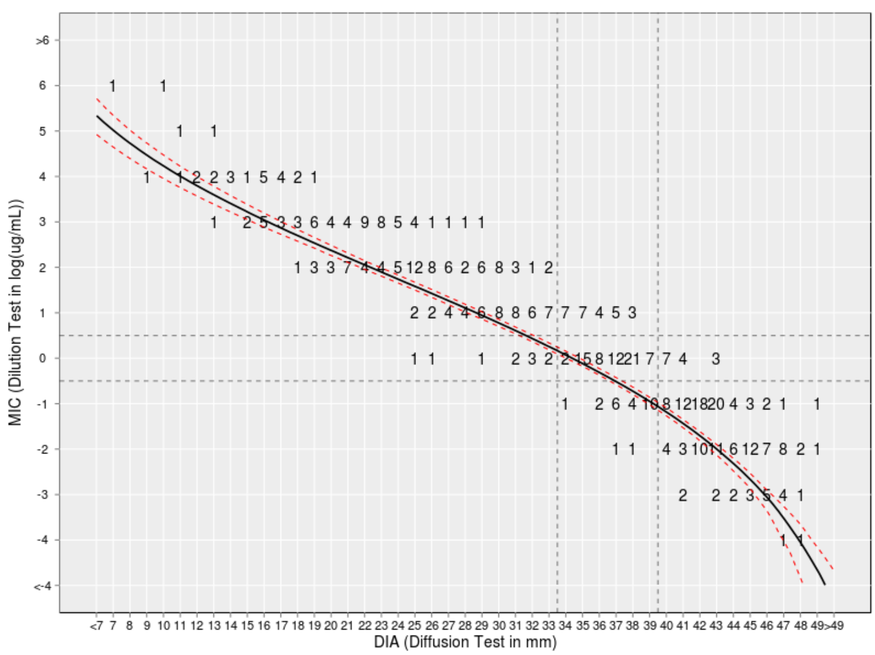

Figure 5: Example output from the logistic model procedure.

```
DIA Breakpoints
Set 1: 33, 40
Set 2: 34, 39

Probability Correct Classification
  MIC Correct  DIA 1 Correct DIA 2 Correct
     0.861         0.904         0.874

Probability Correct Classification Weighted by Isolate Distribution
  MIC Correct  DIA 1 Correct DIA 2 Correct
    0.7438         0.817        0.7597
```

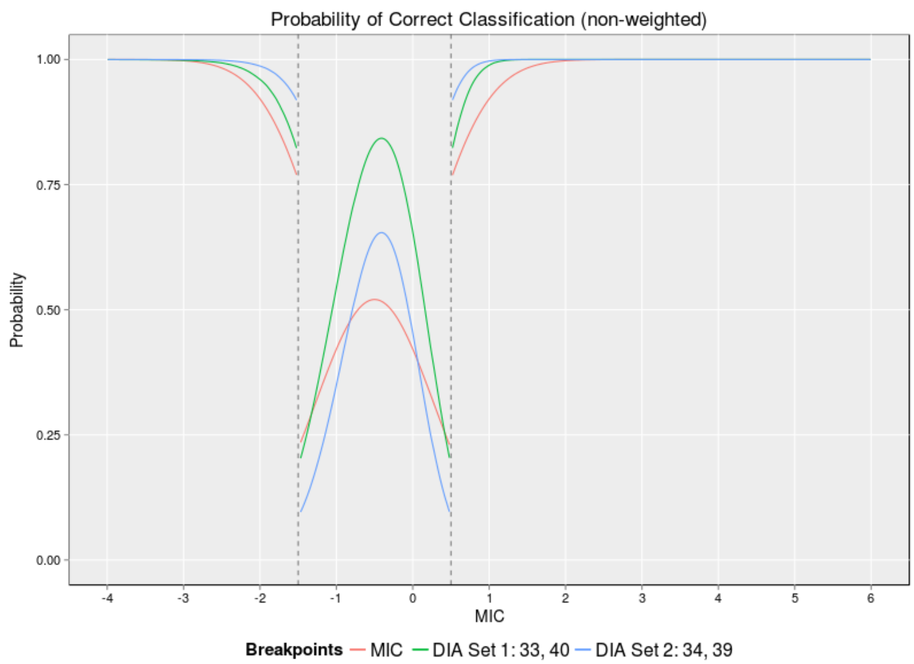

Figure 6: Information on the probability of correct classification for two sets of DIA breakpoints and the MIC breakpoints.