***Method Article***

***ShinyBA.plot*: An R Shiny Application for Bland-Altman Plots to Visualize Agreement Between Two Quantitative Measures in Medical Research**

**Abstract**

In medical research, Bland-Altman analysis is a popular technique for determining the boundaries of agreement between two quantitative values. This article presents ‘*ShinyBA.plot*’ an interactive and intuitive application designed to automate the creation of Bland-Altman plots using the Shiny R platform. Researchers and practitioners that need to effectively visualize the agreement between several measuring techniques may find this tool useful. *ShinyBA.plot* can be accessed here: <https://kmk-puranik.shinyapps.io/shinybaplot/>

**Keywords**

Bland-Altman plot, Shiny R, Reproducibility, Interactive tool, Limits of agreement

**Introduction**

In many research domains, determining whether measurements from two (or more) observers or methods yield similar results is crucial. This concept, referred to as agreement, concordance, or reproducibility, is vital in assessing the reliability of different measurement techniques. In clinical and experimental research, discrepancies in measurements can lead to erroneous conclusions, impacting decision-making and outcomes. Therefore, assessing measurement agreement is essential to ensure the accuracy and consistency of results, particularly when comparing a new method to a reference method or evaluating the reliability of repeated measurements.

The Bland-Altman analysis, introduced by Bland and Altman in 19831,2, is one of the most widely used statistical methods for evaluating the agreement between two quantitative measurement methods. It is especially useful for evaluating the reproducibility of measurements made with the same method and for comparing a novel measurement technique to an existing reference or gold standard. It does not reveal the extent of differences between two approaches, in contrast to correlation analysis, which assesses the degree of association between them and shows if they follow similar trends.

Correlation alone is insufficient when comparing measurement methods because it does not imply agreement. Two methods can be highly correlated yet yield significantly different measurements; for example, if one consistently reports values twice as large as the other, correlation remains perfect, but agreement is poor. Additionally, systematic bias may exist, where one method consistently overestimates or underestimates compared to another, potentially affecting clinical decisions—something correlation cannot detect. Assessing variability is also crucial, as Bland-Altman analysis evaluates measurement differences and their clinical significance, determining whether methods can be used interchangeably. By focusing on the differences between measurements rather than just their association, Bland-Altman analysis provides a more comprehensive and reliable approach to evaluating agreement3.

Bland-Altman plots are widely used in biomedical research; however, many popular statistical software programs do not provide menu-driven choices for creating them, making them inaccessible to non-technical users. This work aims to simplify this procedure by introducing ‘*ShinyBA.plot*’ an application built on R Shiny.

The *ShinyBA.plot* program expands the capabilities of the *BA.plot* R function4 through a web-based interface. In order to improve accessibility and usability for a wider audience, the application provides an easy-to-use interface that enables researchers to upload their data, choose variables for comparison, and create Bland-Altman plots.

**Methodology**

Bland-Altman analysis is a statistical approach used to assess the agreement between two quantitative measurement methods by constructing limits of agreement5. This method is widely applied to detect systematic differences between measurement methods and identify potential outliers in the data.

***Steps to construct a Bland-Altman plot:***

* Calculate the difference between the measurements obtained from Method 1 and Method 2 for each data point:

***Difference = Method 1−Method 2***

* Calculate the average of each pair of measurements:

***Average = (Method 1+Method 2)/2***

* Calculate the mean of the differences and standard deviation of the differences
* Determine the limits of agreement as follows,
  + ***Lower Limit* (*LL*) = *Mean of difference* − (1*.*96 × *Standard deviation of difference*)**
  + ***Upper Limit* (*UL*) = *Mean of difference* + (1*.*96 × *Standard deviation of difference*)**
* Create a scatter plot with the mean values of the two methods (Average) on the X-axis and the difference between the measurements on the Y-axis.
* Mark the Lower Limit (LL) and Upper Limit (UL) along with the mean of differences.

***Interpretation of the plot:***

* Data points that are close to the zero line indicate a high level of agreement between the two methods.
* Data points that deviate significantly from the zero line suggest weaker agreement.
* Trends in differences:
* If the differences increase with the mean → Proportional bias.
* If differences are spread out unevenly → Heteroscedasticity (variance depends on the measurement magnitude).
* Outliers:
* Any points beyond the limits of agreement suggest unacceptable variation.
* Investigate these points for potential errors or systematic issues.

While the Bland-Altman plot is a valuable tool, it is important to recognize its limitations. The method assumes normally distributed differences, which may not always hold true. In such cases the data may be logarithmically transformed. Additionally, sample size plays a crucial role as inadequate sample sizes can lead to misleading conclusions about method agreement. To address this, researchers should conduct power analyses to determine appropriate sample sizes before data collection3,7.

**Example Dataset:**

To illustrate the Bland-Altman analysis, we used a dataset provided in Table 1 comparing blood pressure readings from two different devices: a manual sphygmomanometer (Method 1) and a digital blood pressure monitor (Method 2). The table also contains the calculated differences of the blood pressure measurements between two methods and average of the measurements.

Table 1: Comparison of blood pressure measurements (mmHg) from two devices

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Subject** | **Method 1**  **(mmHg)** | **Method 2**  **(mmHg)** | **Difference**  **(M1 - M2)** | **Average of**  **M1 and M2** |
| 1 | 120 | 122 | -2 | 121 |
| 2 | 130 | 128 | 2 | 129 |
| 3 | 125 | 127 | -2 | 126 |
| 4 | 140 | 138 | 2 | 139 |
| 5 | 135 | 136 | -1 | 135.5 |
| 6 | 128 | 130 | -2 | 129 |
| 7 | 132 | 131 | 1 | 131.5 |
| 8 | 138 | 140 | -2 | 139 |
| 9 | 142 | 141 | 1 | 141.5 |
| 10 | 136 | 135 | 1 | 135.5 |
| 11 | 127 | 129 | -2 | 128 |
| 12 | 139 | 137 | 2 | 138 |
| 13 | 134 | 133 | 1 | 133.5 |
| 14 | 145 | 144 | 1 | 144.5 |
| 15 | 148 | 147 | 1 | 147.5 |
| 16 | 150 | 149 | 1 | 149.5 |
| 17 | 132 | 130 | 2 | 131 |
| 18 | 137 | 136 | 1 | 136.5 |
| 19 | 140 | 138 | 2 | 139 |
| 20 | 143 | 142 | 1 | 142.5 |

The mean of differences (bias) and standard deviation of differences are calculated as follows:

Mean of differences = 0.40 (rounded)

Standard deviation of differences = 1.57 (rounded)

The limits of agreement are calculated as:

*Lower Limit* (*LL*)= *0.40* − (1*.*96 × *1.57*) =0.40 – 3.08 = –2.68 (rounded)

*Upper Limit* (*UL*)= *0.40* + (1*.*96 × *1.57*) = 0.40 + 3.08 = 3.48 (rounded)

A graph with numbers and lines

AI-generated content may be incorrect.

Fig. 1. Bland-Altman plot for the example data

The Bland-Altman analysis of blood pressure measurements from a manual sphygmomanometer and a digital blood pressure monitor yielded limits of agreement ranging from –2.68 to 3.48 mmHg. This narrow range is crucial in assessing the interchangeability of these two measurement methods. The limits of agreement represent the range within which 95% of the differences between measurements from the two devices are expected to fall. This close agreement is a positive indicator of the digital monitor's accuracy compared to the traditional manual method. In summary, if the estimated limits [-2.67, 3.47] are clinically acceptable, Methods 1 and Method 2 can be considered interchangeable for blood pressure measurement.

***ShinyBA.plot*** **Overview and User Interface**

*ShinyBA.plot* is an interactive application built using the R Shiny framework6, designed for generating Bland-Altman plots to assess the agreement between two measurement methods. The app allows users to upload data in Excel formats (.xls, .xlsx), select the variables of interest, and generate the Bland-Altman plot. It automatically computes key statistical values such as the mean difference (bias), standard deviation of the mean difference, and 95% limits of agreement (lower and upper limits). These values are displayed on the plot for easy interpretation. All statistical calculations are performed using the existing *BA.plot* R function, ensuring reliability and reproducibility of results.

Users are encouraged to use the TrialData provided in the supplementary material and can gain practical experience in interpreting Bland-Altman plots, understanding limits of agreement, and identifying potential systematic differences or outliers in measurement comparisons (Fig. 2). *ShinyBA.plot* is designed for ease of use, requiring minimal setup and allowing researchers to quickly generate accurate results for method comparison studies.

*ShinyBA.plot* can be accessed here: <https://kmk-puranik.shinyapps.io/shinybaplot/>

A screenshot of a computer application

AI-generated content may be incorrect.Fig. 2. *ShinyBA.plot* interface generating a Bland-Altman plot using TrialData (Karun and Puranik, 2021)

**Conclusion**

*ShinyBA.plot* enhances the accessibility of Bland-Altman analysis by transforming the *BA.plot* R function into an interactive web-based tool. It automates the analysis, allowing researchers to generate plots, compute key statistics, and assess method agreement without programming expertise. By simplifying the process, *ShinyBA.plot* makes statistical analysis more user-friendly and is especially valuable in medical and clinical research for quickly evaluating measurement method comparisons.

**Disclaimer (Artificial intelligence)**

Author(s) hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc.) and text-to-image generators have been used during the writing or editing of this manuscript.

**References**

1. Bland JM, Altman DG. Measuring agreement in method comparison studies. Statistical Methods in Medical Research. 1999 Apr;8(2):135–160.
2. Sedgwick P. Limits of agreement (Bland-Altman method). BMJ. 2013 Mar 15;346:f1630.
3. Kaur P, Stoltzfus JC. Bland–Altman plot: A brief overview. International Journal of Academic Medicine. 2017 Jan 1;3(1):110-1.
4. Karun KM, Puranik A. BA. plot: An R function for Bland-Altman analysis. Clinical Epidemiology and Global Health. 2021 Oct 1;12:100831.
5. Giavarina D. Understanding Bland Altman analysis. Biochemia Medica. 2015;25(2):141-151
6. Potter G, Wong J, Alcaraz I, Chi P. Web application teaching tools for statistics using R and shiny. Technology Innovations in Statistics Education. 2016;9(1).
7. Doğan NÖ. Bland-Altman analysis: A paradigm to understand correlation and agreement. Turk J Emerg Med. 2018;18(4):139-41.