Document Date:20.04.2023

Statistical Analysis Plan (SAP)

The Shapiro-Wilk goodness-of-fit test was used to determine whether the distributions of numerical variables conformed to a normal distribution. For numerical variables that followed a normal distribution, descriptive statistics such as mean and standard deviation were provided. For numerical variables that did not follow a normal distribution, descriptive statistics such as median and interquartile range were given. Percentages and frequency tables were provided for categorical variables. To compare the patient and control groups, variables that met the normality assumption were analyzed using the parametric test for the significance of the difference between two means, while those that did not meet the normality assumption were analyzed using the non-parametric Mann-Whitney U test. The assumption of homogeneity of variances was examined with the Levene test statistic. Kruskal-Wallis analysis of variance was used to examine whether there were differences between diagnosis groups in terms of age. The relationship between age and numerical parameters was examined with the Spearman rank correlation coefficient since the age variable did not meet the normality assumption. Whether there were differences in numerical measurements between diagnostic groups was examined with one-way analysis of variance, and when the parametric assumptions of one-way analysis of variance were not met, it was examined with the Kruskal Wallis test. The differences between the groups were examined using the Bonferroni test approach, one of the Post-Hoc tests.

The area under the ROC curve was used to evaluate the performance of the diagnostic test. The Youden index (J), which assumes sensitivity and selectivity, or false positive and false negative rates, as equally important, was used to determine the best cutoff point.

Statistical significance level was accepted as $p < 0.05$. Categorical variables are presented as percentages and numbers, and continuous variables are presented as mean (mean), standard deviation (SD) or median, 25th Quarter (Q1) and 75th Quarter (Q3). The suitability of the variables to normal distribution was examined using visual (histogram and probability graphs) and analytical methods (Kolmogorov Smirnov Test). In comparing two independent groups, Mann Whitney U Test was used for those that did not show normal distribution. Spearman correlation analysis was performed to

evaluate the relationships of continuous variables. In comparing three or more measurements of dependent groups, ANOVA test was used for repeated measurements for data that conformed to normal distribution, and the Friedman test was used for data that did not comply with normal distribution. Wilcoxon test was used for pairwise group comparisons in repeated measurements that were statistically significant. The results were evaluated at a 95% confidence interval, with alpha error set at 0.05. Statistical evaluation was made using the Statistical Package for Social Sciences (SPSS) for Windows 26.0 (IBM SPSS Inc., Chicago, IL) program.

Evaluation of Sample Size Calculation for Correlation Analysis between VAS and Tryptophan Metabolite Measurements

The sample size calculation for assessing the correlation between VAS (Visual Analog Scale) and tryptophan metabolite measurements was conducted using G*Power software, employing a Point biserial model test. It was determined that a minimum sample size of 26 participants is required to achieve a statistical power of 80% with a significance level (alpha) of 5% and an effect size of 0.50 (1).

This approach ensures that the study is adequately powered to detect meaningful correlations between pain perception (measured by VAS) and the levels of tryptophan metabolites in the blood, contributing to the robustness of the statistical analysis.

Reference:

1. Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior Research Methods, 39, 175-191.