I can see that many changes have been made. However, some important issues still remain.

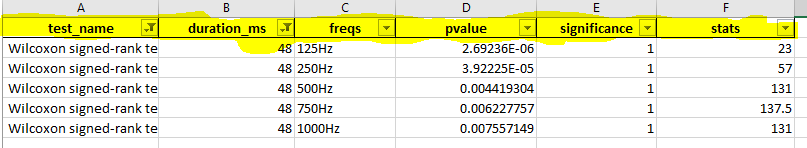
* It does not make sense to test BMLD vs S0N0 vs SpiN0 as the BMLD is the difference of S0N0 and SpiN0. One of the three, most likely the BMLD, will always be different. It would make sense to test S0N0 vs SpiN0 though. I suggest you do this, using appropriate tests.
* **Hence we are using sono vs spino not BMLD**

**Appropriate test**

1. Chi square is used for catergorical data but we have numerical data
2. Numerical data :either parametric or non-parametric
3. sample size: Large sample sizes allow for the use of parametric tests, while small sample sizes require non-parametric tests.
4. Non parametric because sample size is 35, In general, nonparametric tests are considered more robust to small sample sizes than parametric tests
5. Normality: If the data is not normally distributed, use non-parametric tests such as the Wilcoxon rank-sum test and Kruska-Wallis.
6. Paired or non-paired : same subjects and different conditions hence paired
7. A Kruska-Wallis test would assume that all observations are independent, whereas repeat observations on the same subjects are related. The Wilcoxon signed rank test correctly accounts for the fact that observations are paired by subjects by making a pairwise comparisons.

* You essentially have three different types of independent variables: frequency, duration and     in phase or out phase. You have tested for frequency and you should test for S0N0 vs SpiN0 . You might also test for 3 ms vs 18 ms vs 48 ms.

You can use the filters for the comparison of different variables for example in following screen shot I have tried to observe effect of different frequency for 48 ms duration and the value in stat col is explain below higher value mean null hypothesis is rejected strongly .

* 

**Stats for Wilcoxon rank-sum test**

* + - The test statistic of the Wilcoxon signed-rank test is based on the signed ranks of the differences between the paired observations. The signed rank is simply the absolute value of the difference between the two observations, with the sign of the difference (+ or -) indicating which observation was larger.
    - The Wilcoxon signed-rank test calculates the sum of the signed ranks, which is then used to determine whether the median difference between the two groups is zero or not. If the sum of the signed ranks is significantly different from zero, it indicates that there is a significant difference between the two groups. The test statistic is compared to a critical value from the Wilcoxon signed-rank distribution to determine whether the null hypothesis should be rejected.
    - The Wilcoxon rank-sum test compares the ranks of two samples to determine whether they come from the same distribution. The larger the value of the test statistic (wc\_stat), the stronger the evidence against the null hypothesis, which states that the two samples come from the same distribution. Therefore, a larger value of wc\_stat suggests that the two samples are more likely to come from different distributions.

Example

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Patient** | **Difference A-B** | **Absolute Difference** | **Rank** | **Sign** |
| **1** | **4** | **4** | **8** | **+** |
| **2** | **-2** | **2** | **4** | **-** |
| **3** | **1** | **1** | **6** | **+** |
| **4** | **-3** | **3** | **3** | **-** |
| **5** | **0** | **0** | **5.5** | **0** |
| **6** | **2** | **2** | **4** | **+** |
| **7** | **-1** | **1** | **6** | **-** |
| **8** | **5** | **5** | **10** | **+** |
| **9** | **-2** | **2** | **4** | **-** |
| **10** | **3** | **3** | **3** | **+** |

* There is an issue with low frequency signals of short duration. For a 125 Hz signal, the period is 8 ms which is considerably longer than 3 ms. Even the period for the 250 Hz signal is longer than 3 ms and for 500 Hz there are still only one and a half period in 3 ms so it is not clear whether people could really hear the frequency of the tone in such a short signal. This probably explains why lower frequencies do not result in bigger BMLDs for 3 ms. You need to discuss this issue in your paper.
  + - Add this to paper.
* I would not talk about different combinations of datasets as this is only confusing and you have used all data anyway. Instead, simply explain what tests you are doing and what the purpose of it is.
  + - Already explained

* You might also explain the reasons for doing both the Friedman test and the Kruskal Wallis test. What is the difference and why is it useful to do both?
  + Not using these test
* The test statistic for the Shapiro-Wilks test should be called W, nor "statistic", even though it comes out of SPSS that way
  + - Shapiro-Wilks is used to test normality to decided which test we we want to use for hypothesis testing .its not for hypothesis testing here related to binaural hearing
* It is not true that 3 ms signals are only processes in the brainstem whereas longer signals are processed by different parts of the brain. The reason for using a very short signal for the ABR is that the ABR would otherwise overlap with the stimulus.
* It would be good if you could make more clear what the contribution of this paper is.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  | Rank | | | Test Statistic | | | |
|  |  |  | N | Mean Rank | Sum Rank | Chi-Square | N | Prob>Chi-Square | Significantly different/NOT |
| 3ms | BMLD | 125 | 35 | 86.8 | 3038.5 | 2.71 | 4 | 0.60829 | Not significantly different |
| 250 | 35 | 88.6 | 3101 |
| 500 | 35 | 84.61 | 2961.5 |
| 750 | 35 | 99.4 | 3479 |
| 1000 | 35 | 80.57 | 2820 |
| 18ms |  | 125 | 35 | 100.8 | 3528.5 | 8.87 | 4 | 0.04445 | Significantly different |
| 250 | 35 | 100.6 | 3521 |
| 500 | 35 | 84.01 | 2940.5 |
| 750 | 35 | 83.69 | 2929 |
| 1000 | 35 | 70.89 | 2481 |
| 48ms | BMLD | 125 | 35 | 113.7 | 3982 | 13.25 | 4 | 0.01011 | Significantly different |
| 250 | 35 | 89.72 | 3140.5 |
| 500 | 35 | 84.57 | 2960 |
| 750 | 35 | 77.21 | 2702.5 |
| 1000 | 35 | 74.71 | 2615 |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | **Rank** | | | **Test Statistic** | | | |
|  |  | **N** | **Mean Rank** | **Sum Rank** | **Chi-Square** | **DF** | **Prob>Chi-Square** |  |
| **125 Hz** | BMLD | 35 | 19.91 | 697 | 62.3 | 2 | 2.93E-14 | Significantly different |
| SoNo | 35 | 71.7 | 2509.5 |
| SpiNo | 35 | 67.38 | 2358.5 |
| **250 Hz** | BMLD | 35 | 21.25 | 744 | 58.4 | 2 | 2.02E-13 | Significantly different |
| SoNo | 35 | 73.2 | 2562 |
| SpiNo | 35 | 64.54 | 2259 |
| **500 Hz** | BMLD | 35 | 19.82 | 694 | 63.2 | 2 | 1.88E-14 | Significantly different |
| SoNo | 35 | 73.07 | 2557.5 |
| SpiNo | 35 | 66.1 | 2313.5 |
| **750 Hz** | BMLD | 35 | 20.71 | 725 | 60.4 | 2 | 7.62E-14 | Significantly different |
| SoNo | 35 | 73.45 | 2571 |
| SpiNo | 35 | 64.82 | 2269 |
| **1000 Hz** | BMLD | 35 | 21.7 | 759.5 | 55.6 | 2 | 8.11E-13 | Significantly different |
| SoNo | 35 | 70.34 | 2462 |
| SpiNo | 35 | 66.95 | 2343.5 |