**Topic: Total Dissolved Solids (TDS) Removal – Analysis of possible effects of Biofilm treatment.**

**Executive Summary**

Analysis on the provided statistical test output did not support the research hypothesis, namely that the mean TDS in the biofilm-treated water (treatment) is lower than water filtered by the current method (control). The strongest statistical evidence indicated biofilm-treated water was at best statistically similar compared to the existing method, and at worst, less effective at filtration than existing method. Furthermore, the significant variation between the control and treatment groups, especially the high outliers in the treatment sample, produce further concern as to the efficacy of the biofilm filtration as represented in the provided sample.

**Toplines for Wider Presentation**

Two statistical tests were performed on control and treatment (biofilm) samples – a T-test and a Bayesian BEST procedure. The output of the t-test included a null hypothesis significance test (NHST) and an associated confidence interval. The NHST indicated that the calculated mean difference generated from the t-test (0.14925) has a high probability (p value of 0.8823) of being generated from samples that were statistically the same, meaning that there is no indicator from this test that the control group is statistically different than the sample taken from the biofilm procedure. This test also generates a confidence interval of approximately -12 to 14, which can be understood as a proxy for uncertainty around the mean difference estimated from this method.

From the Bayesian procedure, the population mean difference between the control and biofilm groups was estimated at 8.2, with a 95% likelihood that the mean difference between the control and biofilm groups is within the range of 0.0193 to 16.2 TDS. If anything, these results indicate that the most likely mean estimation has the biofilm population at +8 TDS from the control, directly contradicting the research hypothesis. It is advised that the methods by which these samples were obtained be audited, and further research and development on the biofilm treatment be pursued before any discussion of taking this product to market.

**Summary of Statistical Test Output**

Lower and and upper bounds of the (frequentist or t-test) 95% confidence interval of the mean difference:

-12.35187 14.30250

Point estimate of the mean difference (t-test):

t = 0.14925

Outcome of the null hypothesis significance test on the difference of means -

p-value = 0.8823, null hypothesis = there is no difference between samples taken with and without the biofilm treatment



**Lower Bound of 95% HDI:** 0.0193

**Upper Bound of 95% HDI:** 16.2

**% values in posterior distribution of mean differences above and below 0:** 2.8% below, 97.2 above.

**Technical Notes**

The provided materials for this analysis included output from a data frame with 32 observations and 2 variables – control and treatment – where values represented recorded levels of Total Dissolved Solids (TDS) in various water samples. The control group had a mean of 32 TDS, and the treatment 31. The range of the control group was approximately 4 – in contrast, the range of the treatment group was approximately 200.

****

Further graphical examination of the recorded TDS values helps highlight the need for additional scrutiny on the provided samples, and represents a potential confound in the output of this analysis. In the plot of TDS distribution across the control sample (1) and the biofilm treatment (2), the presence of the outlier value near 200 is well outside the expected standard deviations from the mean. For this reason it is recommended that the raw data be re-examined to ensure that this value is A) correctly recorded and B) from the same sample. Only after this value has been re-examined should the conclusions in this analysis be used to help guide the next steps for the client.