Statistical Analysis

The main outcome variable was the reduction of PPD at 3 and 6 months after intervention, while BOP, CAL, REC, FMPS, FMBS, pocket closure and detection scores of the five selected bacterial species changes (*A. actinomycetemcomitans, P. gingivalis, T. forsythia, P. intermedia* and *T. denticola*) were regarded as secondary outcomes. It was established that 19 patients per group are needed to detect a significant mean difference of 1 mm in PPD reduction between groups, assuming a common standard deviation of 1 mm, 80% power, and a significance level of 0.05. The Pitman asymptotic relative efficiency correction was used in the sample size calculations to account for the use of nonparametric tests. Considering an anticipated drop-out rate of ~10%, it was decided to enroll at least 21 patients in each group. For each of the quantitative variables PPD, REC, CAL, a patient mean value was computed per timepoint which was further used in the statistical analyses. Differences between groups for variables measured on a continuous or ordinal scale were analyzed using Kruskal–Wallis tests, with post–hoc Mann–Whitney pairwise tests, as necessary. Proportions were compared by chi square tests. Assessment of intragroup differences between successive time points for quantitative variables was performed using Friedman tests, with subsequent Wilcoxon signed-rank tests for pairwise comparisons. The Bonferroni correction was used to account for multiple comparisons. *p* values < 0.05 were accepted for statistical significance. The statistical analyses were performed using the software R version 4.1.2 [*citare: R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/*].Changes in the detection frequency scores of the main keystone bacteria were evaluated in terms of the microbiological status. Results were noted and categorized into one of four groups: 0 = nondetectable, 1 = detectable $< 10^4$ (10^3 for *A.a.*), 2 = 10^4 -10^5 (10^3 -10^4 for *A.a.*), 3 = 10^5 -10^6 (10^4 -10^5 for *A.a.*), and 4 $\geq 10^7$ (10^6 for *A.a.*) [39]. Using the Wilcoxon signed rank test, intragroup comparisons of detection scores of pathogen species between the baseline and 6-month reevaluation timepoints were made. For intergroup comparisons of the detection scores at each timepoint, the Kruskal-Wallis test was applied.