Data analysis

The statistical analyses are performed using the software R version 4.0.0 [R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/]. Statistical analysis will be conducted intra- and inter-groups. The main outcome variable was the change of mean PD values at 3 months, while CAL, REC, BOP and the frequency of detection of the five keystone bacteria were regarded as secondary outcomes. A minimal required sample size of 16 patients per group was established to achieve 80% power for detecting a significant mean difference of 1 mm in the reduction of PD between groups, assuming a common standard deviation of 0.8 mm and given significance level α=0.05. The Pitman asymptotic relative efficiency correction was applied in the sample size computation to account for the use of nonparametric comparison tests. To account for possible dropouts, it was decided to enroll at least 18 patients per group.

For each of the quantitative variables PD, REC, CAL, a patient mean value was computed per timepoint, which was further used in the statistical analyses. For quantitative data, intergroup comparisons were made using Kruskal-Wallis tests with post hoc Mann-Whitney tests, and differences within each group between baseline and later timepoints (3, 6, 9 and 12 months) were analyzed using Wilcoxon signed-rank tests. Chi square tests or Fisher's exact tests, as appropriate, were used for comparisons between groups in the case of qualitative data. $p < 0.05$ values were accepted for statistical significance.

Regarding the microbiological status, changes in the score of detection of major keystone bacteria were assessed. The collected PCR data were analyzed as total counts of samples of the bacterial species in the categories <1000, 10^3, 10^4, 10^5, 10^6 and 10^7 (Rusu et al., International Journal of Dental Hygiene, 2014). Non-parametric tests were used for intra-group testing of the microbiological results (Wilcoxon test and Friedman test for two or more different moments of time analyzed) and inter-group testing will be performed using Man-Whitney test and Kruskal-Wallis for two or more groups.