Response to Reviewer 1 Comments

Major remarks:
Point 1: Section 3.1 line 259-260: Do the Authors mean that they included in the same analysis data from T1 e T2 all together? If so, this analysis was not totally correct. This way to compute correlation coefficient is not theoretically appropriate in a repeated-measure study, because it ignores the correlation of the two measurements within the same subjects. The appropriate analysis should account for both within-subject and across-subject variability (for example using mixed effect models).

Response 1: Thank you for this remark. We changed our analysis and calculated a mixed model. The correlation estimates with CI and p- values are now given under the figure. We colored the two timepoints within the figure and removed the fitted line.
(Regarding T1 and T2: the included timepoints are actually T0 and T2. The timepoint T1 was an additional appointment during the diet phase (two weeks after start of the diet). At T1 we did not hand out the MEDAS and FFQ to the participants, but only the 24 hours diet recall as shown in Figure 1. To clarify this, we additionally fitted the design of Figure 1 and added information in Figure 2. Furthermore, we added information regarding one missing MEDAS value where appropriate.)
We changed the statistical methods section according to inclusion of the mixed model and added a citation regarding the used model.

Point 2: Figure 5: concordance analysis would be more appropriate for the purpose.

Response 2: we changed the analysis. We calculated the concordance using Kendall´s τ and added the values into figure 5. Statistically significant values are marked with an asterisk. Accordingly, we changed the text (Lines 306 – 311).

Point 3: Table 2: concordance correlation coefficient would be more appropriate, instead of Sperman correlation. In general, to evaluate the concordance between two tools, concordance analysis should be performed. The analysis used by the Authors may overestimate the degree of agreement between the two tools.

Answer: We added the values of the Kendall analysis, but we would like to keep the Spearman values because there are some studies using Spearman to compare such assessment methods – for example Haftenberger et al., 2010. These studies are addressed within the discussion. Keeping these values might provide a better comparability to those studies.

Minor remarks:

Point 1: In the abstract Authors say "The MEDAS score was significantly negatively correlated with periodontal inflammation" but p-values or C.I. were not reported.

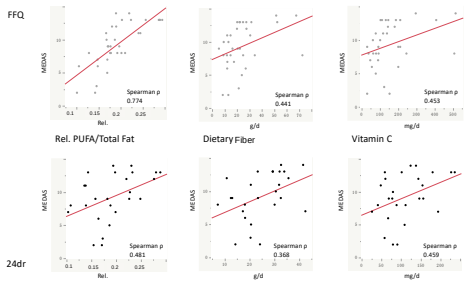
Answer: we added the p- values.

Point 2: In section 2.8: please specify that it is Wilcoxon signed-rank test, in order to distinguishing from the Wilcoxon rank-sum test.

Answer: We used the Wilcoxon signed-rank test within the analysis of figure 5, aiming to test for significant differences between FFQ and 24dr. Now that we changed to the Kendall analysis, we removed the Wilcoxon p-values and marked the significant results of Kendal with an asterisk (see answer to major remark 2). Regarding the analysis of Table 2: this only includes inter-group-comparisons, so the Wilcoxon signed rank test is no longer used. We specified it in section 2.8 to Wilcoxon rank-sum test.

Point 3: Figure 3 and Figure 4: where are the 95% C.I.?

Answer: This might be a technical problem of the doc-file – (presumably a windows / Mac – compatibility problem?). Within the submitted PDF-Version, the CI should be visible, I can also see the CI in my doc-file.
This is a screenshot of my PDF with the CI.



Point 4: Section 3.1 line 260: correlations are statistically significant?

Answer: We added an asterisk to mark statistical significance within figure 4 and added an explanation within the text of section 3.1. The new figure 4 now includes a correction regarding the 24dr - Spearman analysis value.

Point 5: Caption of Table 2: Matching results of intragroup or intergroup?

Answer: "Intergroup" is meant. Within this column we marked matching results of the two intergroup comparison – p – values regarding statistical significance.