Comments on the letter from Kjell Pettersson and Reiner Mengel by the editorial team of *EJOI*

We first address some general considerations about the letter from Pettersson and Mengel on the shortcomings of the article by Albouy et al (2008). In general, animal models appear to be indispensable to advancing our knowledge about disease and the optimal management of disease. However, the outcome of those studies should be considered with caution because the results can vary significantly, and animal studies do not reliably predict human outcomes. A major limitation of the Albouy et al study is the difficulty in drawing a valid conclusion from an animal study consisting of only six dogs, a very small sample size. A proper power calculation is mandatory for estimating the appropriate sample size for a study. In addition, this experimental model of peri-implantitis generated with ligatures could differ considerably from the human situation.

For evaluating whether different implant types are more or less susceptible to peri-implantitis, a more scientifically sound approach would be to compare them in a randomised clinical trial with an appropriate sample size and sufficient follow-up. At the protocol level, how the data will be processed for statistics must be clear and specific. The Pettersson and Mengel letter emphasised two interesting issues in the Albouy paper that illustrate how a change in the analysis can alter outcomes. Two implants were lost and were given the maximal possible value for bone loss, which is conceptually correct although not usually done in dental implant studies. However, this choice of assigning or not assigning the maximum value for bone loss likely changed the results of the Albouy paper (i.e. if the maximum value for bone loss is not assigned, there might not be statistically significant differences in bone level changes between different implant types). If there was an *a priori* decision to evaluate the data in this manner, then there is no problem; however, it could be considered misleading if this decision was taken *a posteriori*, after the results were known. The second example is the decision about which baseline to use for evaluating spontaneous disease progression. With each of the possible baselines (before or after induction of peri-implantitis), the results could change again (i.e. using the time prior to peri-implantitis induction as the baseline results in no statistically significant differences in bone level changes between implant types). Again, the results depend on a decision the authors made about which baselines to use. While it is sensible to use bone level after induction of peri-implant bone loss as the baseline, we note that this parameter differed among different implant types. Therefore, in the presence of very weak data (from only six dogs), the results depend considerably on which parameters and methods were chosen for the statistical analysis rather than on the actual data.

Other shortcomings, as highlighted by Pettersson and Mengel, are as follows:

- The description of the statistical analysis in the original paper is not sufficiently clear. The authors should take into account the clustering of the implants in dogs. The authors should also present confidence intervals of the difference between implant types.
- This vagueness of the description leads to two technical aspects. We do not know whether the authors used the appropriate Student Newman–Keuls analysis or if they checked the assumptions of the statistical model.
- The authors of the original paper randomised the position of implants to take into account spurious variables. They also consider with the randomisation the different positions of the implants. It is clear that with only six dogs, there is an imbalance among the groups.

In conclusion, no reliable inferences are possible from such a small study. Small animal studies should not be used as surrogates for large clinical trials because their results can be misleading. If the question cannot be answered by a clinical trial, then large animal studies should be conducted; however, when possible, as in this case, conducting a proper clinical trial is the better choice.

Marco Esposito, Michele Nieri and Jerome Lindeboom