Statistics: a nuisance, a tool, or a must?

Dear Readers,

"Do not believe statistics you did not modify yourself!" is a common saying. This shows how misunderstood statistics is in the population. Unfortunately, this is also true for some part of the scientific community. The medical/dental language and also the pathways of thought are much different from those of statisticians. Therefore, the two groups often do not understand each other, and for physicians and dentists, statistics becomes a nuisance – they know it is necessary, but they neither like nor try to understand it. Certainly, this behavior is wrong.

Statistics is a tool which needs to be mastered by the scientists, because with this tool we are able to estimate the probability that chance alone could produce the results we have obtained, which may hence not be due to the experimental conditions. Furthermore, statistics, if properly done, may allow us to draw general conclusions about our data. Therefore, in every scientific paper, statistics are a must.

However, statistics themselves are full of traps and pitfalls, which lead to their faulty application, very often causing problems with manuscripts; just to name a few:

- General conclusions can only be drawn if the sample is truly random.
- Analysis of variance (ANOVA) requires normal distribution of data.
- If there are significant interactions with 2- or 3-way ANOVA, it is not appropriate to draw main effect conclusions.
- Unless you deliberately use tests for repeated data (eg, before/after on the same specimens), data must be independent.
- Authors mix parametric statistics (mean ± SD) with nonparametric statistics (box-plot including median, 25th and 75th percentile).
- Faulty handling of missing data.

In adhesive dentistry, the use of the microtensile bond strength test (μTBS) has become one of the frequently used test methods. However, due to the setup of this test, several problems may occur. Besides typical test-parameter problems, such as the kind of jig employed and the way the specimen is fixed to the jig, the trimmed or non-trimmed specimen design, the specimen size, etc, basic statistical rules may be violated as well. This is the issue dealt with in this editorial.

It is the duty of the *Journal of Adhesive Dentistry's* editorial office to discuss such problems and to try to provide our authors with some viable solutions.

Description of potential statistical problems

For μTBS, typically a few teeth (2 to 3) are ground flat on one surface (most often the occlusal one), after which, based on the experimental setup, the materials are applied onto this surface. Then, the teeth are cut into several sticks (4 to 16), which are then subjected to tensile stress up to fracture in a universal testing machine. During the sample preparation, it is not infrequent that many sticks break before they can be subjected to the tensile stress, and are often referred to as "pre-test failures".

Problem 1: Since several sticks come from the same tooth, specimens are not independent; however, a weak correlation of μTBS has been found between beams.[1] It has been shown that the variation within the tooth, probably due to differences in the bonding quality between superficial and deep dentin, and between central and peripheral dentin, is larger than the intertooth variation.[2] Nevertheless, assuming independence of beams overstates the statistical significance levels for comparisons between materials.[1]

Problem 2: How should the pre-test failures (PTFs) be handled? This is often done in two ways:

- Assign them the value of zero. Though commonly not done, a predetermined minimal value is preferred, since such PTFs do not fall apart spontaneously, but are due to stress imposed on the bond during the actual specimen processing. A certain bond strength is thus present, and could perhaps be calculated using regression statistics.

The consequence of this is that it is very likely that the normal distribution of data is lost, especially with a high number of PTFs, with the further consequence that ANOVA should not be used.

- Ignore them. Simply compute the values of the samples that were able to be tested. If there are too many pre-test failures, it is obvious that the values reported are too high. It should also be noted that PTFs are typically associated with relatively low bond strength data measured for those specimens that did not fail prior to testing.

To date, it has been the policy of JAD to require the authors to report the number of PTFs and to describe how they were handled, because much information could be derived from the incidence of PTFs.

Possible solutions

One obvious solution to problem 1 would be to use only one stick per tooth, if possible always from the same area, otherwise picked with a random process. However, besides the fact that extracted teeth are very precious and this approach takes away a major advantage of µTBS (vs conventional shear bond strength testing), it seems quite impractical.

Another way to handle the problem would be to use every tooth as its own control. This means the tooth surface for bonding is divided into 2 or, for practical reasons, a maximum of 4 sectors, each receiving a different treatment. Then, one stick from each sector is obtained. This can be handled with the appropriate statistics.

Another way would be to calculate the average µTBS for the sticks originating from each tooth and to use these averages for the statistical analysis. This would guarantee independence of data, and at the same time reduce the variability.

The best way to handle problem 1 would be to know which stick comes from which tooth, and then use linear mixed models (ANOVA with random effects) to analyze the data. Another way would be to apply survival analysis like the Weibull model or Cox proportional hazard using the force that is required for bond failure.[1] This approach is also recommended by the ISO Technical Specification on testing adhesion to tooth structure.[3]

A solution to problem 2 would be a failure analysis of the pre-test failures. If there is an adhesive failure, the lowest measured value could be assigned. If it is a cohesive failure in dentin or enamel (remote from the interface in one of the adherents), the specimen may be discarded, because we can exclude that failure of adhesion was the reason. However, in certain adhesive systems that present high bond strength values, cohesive failure may occur within the adhesive layer; thus, the real bond strength might be even higher than the measured value.

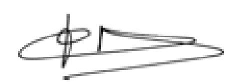
However, as mentioned above, materials that have a high proportion of pre-test failures also commonly have a low µTBS. Otherwise, materials with a high µTBS do not have any pre-test failures. Therefore, a “gold-standard” three-step adhesive, for which high µTBSs have repeatedly and consistently been reported in literature, should always be enclosed in the study setup as control.

In any case, the researchers should check for normal distribution (eg, using scatter diagrams, box-plots or tests such as Kolomogorov-Smirnov) and run the statistics accordingly thereafter.

Finally, it is important that the authors exactly report how they dealt with the statistical problems.

We are convinced that this recommended statistical approach would help to eliminate confusion, reporting of questionable results and conclusions, and thus increase the quality of the *Journal of Adhesive Dentistry*.

J.-F. Roulet,
Editor-in-Chief



B. Van Meerbeek,
Editor-in-Chief

REFERENCES

1. Eckert GJ, Platt JA. A statistical evaluation of microtensile bond strength methodology for dental adhesives. Dent Mater 2007;23:385-391.
2. Loguercio AD, Barroso LP, Grande RH, Reis A. Comparison of intra- and inter-tooth resin-dentin bond strength variability. J Adhes Dent 2005;7:151-158.
3. ISO. Dental materials - Testing of adhesion to tooth structure. Technical Specification 2003:No. 11405.

ACKNOWLEDGMENTS

We thank Dr. S. Heintze, Prof. M. Özcan, Prof. V. Rousson, Dr. J. De Munck, and Drs. A. Poitevin for their constructive discussions and their advice in the above treated matter.