Microtensile test in dental research. Controversial aspects in statistical analysis (experimental unit and premature failures)

Teste de microtração na pesquisa odontológica. Aspectos controversos na análise estatística (unidade experimental e falhas prematuras)

Ivan BALDUCCI¹, Clovis PAGANI¹, Daphne Câmara BARCELLOS¹, Mayra CARDOSO¹
1 – Institute of Science and Technology – UNESP – Univ Estadual Paulista – School of Dentistry – São José dos Campos – SP – Brazil.

ABSTRACT

Objectives: The aim of this study was to present the controversial aspects in performing inferential statistical analysis with regard to selection of the experimental unit and procedures adopted in the case of pre-test failures. Methods: Eighty-seven dissertations and theses in the field of dentistry were evaluated, for which microtensile testing was made available at three public Universities in the State of São Paulo were performed, namely: School of Dentistry, USP; School of Dentistry, Unesp (Araraquara and São José dos Campos) and the Piracicaba School of Dentistry (Unicamp). Results: Regarding the experimental unit, 34 (39.08%) theses and/or dissertations used the “stick” and 53 (60.91%) the “tooth/block” type. With regard to pre-test failures, 9 (10.34%) theses and/or dissertations attributed the value “0”, 6 (6.89%) disregarded the sticks, 2 (2.29%) attributed the minimum value of 2 MPa, 1 (1.14%) the minimum value of 4 MPa, 1 (1.14%) attributed half the minimum value of the group and 68 (78.16%) did not mention the occurrence of pre-test failures. Conclusion: At present the experimental unit is considered to be the tooth (or ceramic block) and not the stick. As it relates to pre-test failures, the following are considered: attribution of a minimum bond strength value under experimental condition or in the group considered, an expressed number and percentage of cases of failures under each experimental condition; failures included in the statistical analysis and detailed explanation of how the authors performed with regard experimental designs (experimental unit and premature failures).

KEYWORDS

Microtensile; Teeth; Sticks; Premature failure; Statistical analysis.

RESUMO

Objetivo: O objetivo dessa pesquisa é apresentar aspectos controversos na condução da análise estatística inferencial quanto à seleção da unidade experimental e quanto ao procedimento adotado diante das falhas pré-teste. Método: Foram avaliadas 87 dissertações e teses na área da odontologia que realizaram ensaio de microtração disponíveis em três universidades públicas do estado de São Paulo: Faculdade de Odontologia da USP (São Paulo, Ribeirão Preto e Bauru), Faculdade de Odontologia da Unesp (Araraquara e São José dos Campos) e Faculdade de Odontologia da Unicamp (Piracicaba e Campinas). Resultados: Quanto à unidade experimental, 34 (39,08%) teses e/ou dissertações utilizaram o “palito” e 53 (60,91%) o “dente/bloco”. Quanto às falhas pré-teste, 9 (10,34%) teses e/ou dissertações atribuíram valor “0”, 6 (6,89%) desconsideraram os palitos, 2 (2,29%) atribuíram o valor mínimo de 2 MPa, 1 (1,14%) atribuíram o valor mínimo de 4 MPa, 1 (1,14%) atribuíram a metade do valor mínimo do grupo e 68 (78,16%) não citaram a ocorrência de falhas pré-testes. Conclusão: Quanto à unidade experimental, atualmente considera-se como unidade experimental o dente (ou bloco cerâmico) e não o palito. Quanto à ocorrência de falhas pré-teste, considera-se: atribuição de um valor mínimo de resistência adesiva na condição experimental ou grupo considerado; expressar o número e percentual dos casos de falhas em cada condição experimental; quais tipos de falha foi incluído na análise estatística e informar exatamente como os autores procederam em relação ao delineamento experimental (unidade experimental e falhas pré-teste).

PALAVRAS-CHAVE

Microtração; Dentes; Palitos; Falha pré-teste; Análise estatística.
INTRODUCTION

Experimental designs for microtensile tests in the field of Dentistry present great differences in approach in terms of statistical analysis. These differences give rise to debates and controversies, especially with regard to the validity of the conclusions presented by the authors.

The test unit or experimental unit is a fraction of the experimental material (individuals, plants, cells, animals) that can be randomly distributed among the groups being studied. In 1994, Sano et al. [1] proposed a methodology for microtensile tests in the field of dentistry, considering the “stick” unit; that is to say, the specimen to be submitted to traction, as the experimental unit. The stick unit was obtained from slices of the experimental material, the tooth. From then on researches that used the microtensile methodology followed the approach used by these researchers.

In 1999, Pashley et al. (1999) [2] explained that the methodology developed by Sano et al. [1] had the advantage of obtaining a large sample from few teeth. However, according to Camargo et al. [3], this methodology promotes a statistical flaw, because the data were treated in an independent manner (independent variable) when in truth they are dependent. An independent variable is a variable manipulated and modified by the researcher, whereas a dependent variable is a variable that the researcher wishes to measure or record. Furthermore, according to Camargo et al. [3], this statistical flaw generates consequences, such as inadequate standard errors and undue increase in the statistical power by enlarging sample size.

Moreover, in 2007, Roulet & Van Meerbeek [4] also observed the problematic fact that various samples coming from the same tooth are not independent samples. In general, several teeth (from 2 to 3) have one of their tooth surfaces flattened, and the treatments are applied on these surfaces. After this, the teeth are cut into various sticks (from 4 to 16), which are submitted to tensile testing until rupture, in a universal test machine.

In the dental literature, another statistical flaw or difficulty in analysis of the data obtained has also been observed: the occurrence of pre-test failures. According to Roulet & Van Meerbeek [4], during sample preparation, it is not rare for many sticks to fail before they can be submitted to tensile testing. According to Mine et al. (2009) [5], the occurrence of pre-test failures can be explained by traumatic preparation technique, which generates excessive stress during the preparation of sticks; or, materials with low bond strength values in the microtensile test, which present higher rates of adhesive failures, and, consequently, present high rates of pre-test failures [4, 6].

How does one proceed in cases of these pre-test failures? Does one attribute the value zero or a predetermined minimum value to them? The consequence of a high number of premature failures, whether or not the value zero is attributed, is probable deviation in the supposition of normality of the data obtained, and consequently, other distributions must be considered.

Therefore, there are two questions that have been the focus of our considerations: How should the experimental unit in a microtensile test be considered, by tooth/block or stick? How should one proceed in cases of premature failures?

It is necessary to conduct a study in this area in order to: observe the best approach to a statistical analysis, present safe guidance in planning and conducting the research and how to analyze its data, encourage the researcher to assimilate the basic concept of an experimental design. Therefore, the aim of this qualitative-descriptive exploratory research was to evaluate master’s dissertations and doctoral theses from public universities in the State of São Paulo with regard to the different aspects of performing
statistical analysis in microtensile tests. The hypothesis tested was: There is no difference between experimental designs (experimental unit and premature failures) for microtensile tests from public universities in the State of São Paulo.

MATERIAL AND METHODS

A total of 87 master’s dissertations and doctoral theses available at three public universities in the State of São Paulo were selected, namely: FO-USP (São Paulo, Bauru and Ribeirão Preto available at http://www.teses.usp.br/), FO-Unicamp (Campinas and Piracicaba available at http://cutter.unicamp.br/) and FO-Unesp (Araraquara and São José dos Campos available at http://unesp.br/cgb/conteudo.php?conteudo=562). The search was conducted with the following criteria: All years, dissertation and/or thesis in the field of dentistry, microtensile testing performed in the methodology.

RESULTS

With regard to origin of the institution, FO-USP presented 11 (12.64%) dissertations and/or theses [7-17]; FO-Unesp presented 30 (34.49%) dissertations and/or theses[18-47]; and FO-Unicamp presented 46 (52.87%) defended in the field of dentistry, which performed microtensile testing in the methodology [48-94].

Experimental Unit

Table 1 shows the number of dissertations and/or theses that used the “stick” or “tooth/block” experimental units for statistical analysis, in the period from 1999 to 2011.

It may be observed that 34 (39.08%) dissertations and/or theses used the “stick” as experimental units [15,17,18,20,21,23-33,35-37,39-43,50,54,61,62,64,67-69,84,85] and 53 (60.91%) dissertations and/or theses used the “tooth/block” as experimental units [4-14,16,19,22,44-49,51-53,55-60,63,65,66,70-83,86-91]. All dissertations and/or theses (11) from FO-USP used the “tooth/block” as experimental units; 10 dissertations and/or theses from FO-Unicamp used the “sticks” and 36 dissertations and/or theses used the “tooth/block” as experimental units; 24 dissertations and/or theses from FO-Unesp used the “sticks” and 6 dissertations and/or theses used the “tooth/block” as experimental units.

Table 1 – Prevalence of dissertations and/or theses which, for statistical analysis, considered the experimental unit for statistical the STICK/BLOCK. Period of evaluation: 1999 to 2011

<table>
<thead>
<tr>
<th>Year</th>
<th>USP-SP</th>
<th>UNESP</th>
<th>UNICAMP</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1999-2001</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2002</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2003</td>
<td>0</td>
<td>2</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>2004</td>
<td>0</td>
<td>7</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>2005</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>2006</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2007</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2008</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>2009</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2010</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>2011</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>TOTAL</td>
<td>11</td>
<td>6</td>
<td>36</td>
<td>53</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>USP-SP</th>
<th>UNESP</th>
<th>UNICAMP</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1999-2001</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2002</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2003</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>2004</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2005</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>12</td>
</tr>
<tr>
<td>2006</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2007</td>
<td>3</td>
<td>2</td>
<td>6</td>
<td>11</td>
</tr>
<tr>
<td>2008</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>2009</td>
<td>1</td>
<td>0</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>2010</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>2011</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>TOTAL</td>
<td>11</td>
<td>6</td>
<td>36</td>
<td>53</td>
</tr>
</tbody>
</table>
It was observed that from 2005, the number of dissertations and/or theses that used the “tooth/block” as experimental unit was higher than the number of dissertations and/or theses that used the “stick” as experimental unit.

Pre-test Failures

With regard to the occurrence of pre-test failures in the dissertations and/or theses, it was observed that 9 or 10.34% dissertations and/or theses (5 from FO-USP; 1 from FO-Unesp; 3 from FO-Unicamp) attributed the value zero [4,7,8,10,11,35, 49,65,68]; 6 or 6.89% dissertations and/or theses (3 from FO-Unesp; 3 from FO-Unicamp) disregarded the sticks [34,36,42,57,84,86]; 2 or 2.29% dissertations and/or theses (2 from FO-Unesp) attributed the minimum value of 2 MPa to pre-test failures [17,21]; 1 (1.14%) dissertation and/or thesis (1 from FO-USP) attributed half the minimum value of MPa to pre-test failures found in the group [5]. The other 68 or 78.16% dissertations and/or theses (5 from FO-USP; 24 from FO-Unesp; 41 from FO-Unicamp) did not mention the occurrence of pre-test failures.

DISCUSSION

Experimental Unit

The hypothesis tested was rejected for experimental unit because there was a difference between experimental design (experimental unit) for microtensile tests from public universities in the State of São Paulo.

According to the results obtained in this study (Table 1), all dissertations and/or theses from the School of Dentistry of São Paulo, Bauru and Ribeirão Preto (USP) considered the tooth as the experimental unit throughout the period of evaluation. Ten dissertations and/or theses from FO-Unicamp (Campinas and Piracicaba) used the “sticks” and 36 dissertations and/or theses from FO-USP (São Paulo) and FO-Unesp (São Paulo) used the “tooth/block” as experimental unit.
theses used the “tooth/block” as experimental units, and that, Piracicaba School of Dentistry (Unicamp) considered the tooth as the experimental unit from 2008. Contrastingly, 24 dissertations and/or theses from the Schools of Dentistry of São Jose dos Campos and Araraquara (Unesp) used the “sticks”, and only 6 dissertations and/or theses used the “tooth/block” as experimental units. The Schools of Dentistry of São Jose dos Campos and Araraquara (Unesp) did not clearly present the approach chosen.

Based on the aforementioned information, it could be said that there is no consensus with regard to the adoption of the “tooth” or “stick” experimental unit. The results shown in Figure 1 demonstrate that until 2004, the number of dissertations and/or theses that used the “stick” as experimental unit was higher than the number of dissertations and/or theses that used the “tooth/block” as experimental unit. Therefore, one notes the adoption of sticks in studies at the end of the 1990s, perhaps based on Shono et al., (1999)[95] and Bouillaguet et al., (2003) [96], who obtained various sticks from a few teeth (3 or 4), treating them as independent variables.

However, Figure 1 demonstrates that from 2005 the number of dissertations and/or theses that used the “tooth/block” as experimental unit was higher than the number of dissertations and/or theses that used the “stick” as experimental unit. The correct choice in methodology was brought into use in the different schools in a similar timeframe, indicating at least a minimum comprehension of its importance by the authors and, may be based on Loguercio et al., (2005)[97], which observed that the variance of the errors (intratooth) is higher than the variance of the random effect (intertooth) variability, and concluded that ‘sticks from the same tooth cannot be considered as an experimental unit, since it does not fulfill all ANOVA requirements’. In 2007, Eckert et al. [98] demonstrated that there was a correlation between beams taken from the same tooth for microtensile testing, therefore, beams obtained from the same tooth are statistically independent and would overstate the statistical significance of study results.

Up to what point is this dilemma, stick or tooth/block important? When one considers the “stick” instead of the “tooth” as experimental unit, the problem of pseudoreplication occurs. Pseudoreplication takes place when a single sample is divided into various samples [99], and these samples are treated in an independent manner, when in fact, they are dependent [99]. Thus, the result obtained is not necessarily due to the factor under study (e.g., restorative technique, restorative material) but the result of inherent spatial variation among the studied teeth. Pseudoreplica is the result of the use of inferential statistics to test the effects of treatments of given experiments whose treatments are not replicated, or whose replicas are not statistically independent [100]. That is to say, pseudoreplication is not a problem of sample planning of itself only (or of sampling) but a particular combination of the experimental design (or sampling) and the respective statistical analysis that are inappropriate for testing the hypothesis of interest [101].

Therefore, when considering the stick, the design does meet the basic principle of replicas, thus, invalidates the conclusions obtained in the statistical test. When considering an unreal sample size there is (I) an increase in the statistical value of the test; (II) a reduction in the p-value obtained and (III) incorrect conclusions are established about the force of evidence against the hypothesis of nullity (HO). With pseudoreplication, it is not possible to obtain an estimate of the experimental error [100]; consequently, one cannot obtain a standard error value required for a multiple comparison test of means for any test.

When one considers the tooth/block as the experimental unit, the usual practice is to obtain the mean value of the sticks (specimens) in each tooth/block. The mean value becomes variable in analysis of the statistical test.
There are advantages in the choice of the mean value of the tooth/block as the variable under analysis, especially when it satisfies the condition of homoscedasticity (errors that present constant variance) of the ANOVA model. One of these is that it avoids the problem of pseudoreplication, according to Hurlbert [100]. Another advantage, according to Picquelle and Mier [102], is that the probability of error Type I (error rate to reject a true hypothesis H0) virtually doesn’t change in relation to the nominal value of 5%, even in designs with groups of different sample sizes. That is to say, an unequal number of units (teeth/blocks) under each experimental condition, or unequal number of specimens (sticks) in each unit, or the unequal number of units under an unequal number of specimens.

For Roulet & Van Meerbeek [4], instead of considering the calculation of the mean value of the sticks in each tooth, in the ANOVA model one could consider the factor tooth as a random effect, as De Munck et al. [103]; Farias [32], Kreidler [40] and Gamborgi et al. [104] did in their studies.

However, the factor “tooth/block” which presents a random effect does not represent a factor of interest to the researcher in terms of multiple comparison of means analysis; that is, there is no sense in performing a Tukey test, for example, for a random effect. Knowing the contribution of the random factor that could explain the total variables of the data may be useful at the stage of experimental planning. For instance, if the researcher wishes to test the effect of different treatments on the bond strength in bovine teeth, he/she will face a dilemma: How many teeth and how many measures (sticks) will be needed. What if after conducting a pilot study in six teeth, with four measurements (sticks) per tooth the researcher analyzes the data using one-way ANOVA and evaluate the components of variance with reference to the tooth and the sticks. If there is a high percentage of variance among the teeth, but not among the sticks; then there is a great deal of variance from one tooth to another. In this case, the experiment could be planned to include a larger number of teeth. If there is a high percentage of variance among the STICK, but not among the TEETH; then there is a great deal of variance among the sticks. In this case, the experiment could be planned to include a larger number of sticks per tooth [97, 98].

Considering the mean value of each tooth as the variable under analysis, the same result is obtained when one considers the tooth as a random factor. In practical terms of a dental research, consider the factor tooth as a random effect in the ANOVA model, as recommended by Roulet & Van Meerbeek [4], is equivalent to the approach in relation to calculation of the mean value of the sticks in each tooth. These two approaches (unit means analysis and nested analysis) do not differ in terms of statistical test power [102].

**Premature Failures**

The hypothesis tested was rejected for premature failures because there is difference between experimental design (premature failures) for microtensile tests from public universities in the State of São Paulo.

Most of the dissertations and/or theses did not mention the occurrence of pre-test failures (78.16%), 10.34% attributed the value zero MPa, (percentage?)attributed the minimum value of 2 MPa, 6.89% disregarded the sticks, 1.14% attributed the minimum value of 4 MPa and 1.14% attributed half the minimum value of MPa to pre-test failures found in the group. For Roulet & Meerbeck [4], the most important data from the incidence of premature failures is to report the number and to describe how they were handled’, therefore, the procedure of the majority dissertations and/or theses evaluated in this study (78.16%), which did not mention the occurrence of pre-test failures, should be avoided.

How does one proceed in cases of these pre-test failures? According to Zander-Grande
et al. [105], the inclusion of premature failures resulted in statistically significant difference compared to those without the inclusion of premature failures, therefore, the practice of disregarding the premature failures, excluding them from statistical analysis does not portray reality and underestimates the real bond strength [106]. For Mine et al.[5], the act of attributing the value as 0 MPa, penalizes the adhesive too severely, whilst attributed the minimum value (as 2 or 4 MPa) is somewhat arbitrary, because according to Reis et al. [107], specimens that de-bonded before being actually tested could present an estimated “bond strength” value that would be somewhere in the range between zero and the minimum bond strength value that was measured in that specific study'.

The goal of these approaches is to offer statistical analysis values that meet the presuppositions of a parametric approach: normality and uniformity of the residual values in the analysis of variance model, which is the prevalent test in dental research involving microtensile tests.

With regard to the dissertations and/or tests evaluated, it was found that it was not usual practice to present the statistical analysis as justification when disregarding sticks with pre-test failure or attributing value zero. Salameh [108] eliminated all the pre-test failures from statistical analysis, with the purpose of obtaining normal distribution of the data. Therefore, Salameh [108] emphasizes the supposition of normality, which is now known to be not as important as the condition of homoscedasticity, as the ANOVA models are robust.

Roulet & Meerbeck [4] explain that if premature failure was an adhesive failure, the lowest measured value could be assigned, corroborated by some authors [107,109,110] who have attributed a predetermined minimum value, that is, the lowest microtensile value obtained within the condition or experimental group. Also, Mine et al. [5] explain that the most important is to reduce the incidence of premature failures, supporting the slices with instance gypsum or alginate during preparation of sticks, and avoiding excessive stress on it.

Therefore, the following strategies are the contemporary trend in the case of premature failures: (I) attribution of a minimum bond strength value, within the values of the sticks actually tested under the experimental condition or the group considered; (II) reporting of the number and percentage of cases of failure under each experimental condition; (III) indicating the type of failure; that is to say, if the cases considered referred to adhesive failures only, or if they also referred to cohesive and/or mixed failures and (IV) comparison of the analysis excluding the sticks with the analysis including the sticks. Finally, without knowing that “each case is a case”, it is possible to express the reality of the phenomenon that occurs with valid information for a statistical analysis that must always be performed using good sense.

Final Consideration

As regards the experimental unit, 34 (39.08%) theses and/or dissertations used the “stick” and 53 (60.91%) the “tooth/block” type. With regard to pre-test failures, 9 (10.34%) theses and/or dissertations attributed the value “0”, 6 (6.89%) disregarded the sticks, 2 (2.29%) attributed the minimum value of 2 MPa, 1 (1.14%) the minimum value of 4 MPa, 1 (1.14%) attributed half the minimum value of the group and 68 (78.16%) did not mention the occurrence of pre-test failures.

In the experimental designs in microtensile tests, contemporary practice in dentistry is to correctly consider the tooth (or ceramic, resin block, etc.) an experimental unit and not the stick. Premature failures should be attributed the lowest microtensile values obtained under the condition or experimental group in order to meet the presuppositions of a parametric approach. In addition, it is important to report the number and percentage of cases of failure
under each experimental condition, which failures were included in the statistical analysis, and give detailed report on how the authors performed experimental designs (experimental unit and premature failures).

REFERENCES


58. Carvalho FG. Chemical, microbiological and mechanical analysis of adhesive system containing antibacterial monomer and fluoride.
bonded to dentin. Piracicaba. [Thesis]. Faculdade de Odontologia de Piracicaba, Unicamp; 2009.


88. Schneider R. Influence of surface treatment and conditioner in the microtensile bond strength between commercially pure titanium.
Microtensile test in dental research. Controversial aspects in statistical analysis (experimental unit and premature failures)

Balducci I et al.

100. Huribert SH. Pseudoreplication and the design of ecological field experiments. Ecological Monographs 1984; 54(2) 187-211.

Ivan Balducci (Corresponding address)
Av. Eng. Francisco José Longo, nº 777,
Jardim São Dimas, São José dos Campos – SP
12245-000
email: ivan@ict.unesp.br

Date submitted: 2013-02-27
Accept Submission: 2013-06-24