

Statistical Testing Methods for Data Analysis in Dental Morphology

Pruebas Estadísticas para el Análisis de Datos en Morfología Dental

Pablo Navarro¹; Imaculada Alemán²; Catalina Sandoval^{4,5}; Camila Matamala³ & Gilda Corsini⁶

NAVARRO, P.; ALEMÁN, I.; SANDOVAL, C.; MATAMALA, C. & CORSINI, G. Statistical testing methods for data analysis in dental morphology. *Int. J. Morphol.*, 38(5):1317-1324, 2020.

SUMMARY: Statistical testing methods are essential tools to measure and analyze the data obtained through research in dental morphology. This descriptive study analyzes the different statistical testing methods in order to suggest its appropriate use and manage in the interpretation and analysis of quantitative and/or qualitative data of dentistry investigations. The knowledge obtained through this study could orient students and/or interested professionals to decide, with theoretical basis, which one to use in the development of their research.

KEY WORDS: Data analysis; Statistics; Methods.

INTRODUCTION

In daily practice, professionals know that their work is not based on the use of statistics, mainly because they are accustomed to working with people in a one-on-one setting, as opposed to with a collection of individuals. This circumstance risks limiting these professionals from developing research in their area. It is important to highlight that researchers, having some notion of statistical procedures, can understand and manage data variability which is often produced by "uncertainty" and "uncertainties" which are often health problem consequences. Statistics are important in part because they support academic research at different levels and subsequently generate rigorous knowledge for the benefit of society. Currently, there is a growing interest in statistical knowledge because many authors are convinced that statistics must be understood through inductive and deductive reasoning. Statistical education is not based solely on formulas and figures, but also interpretation of data so that decisions can be applied correctly.

The nature of the present study is qualitative descriptive, wherein variable relationships with respecti-

ve statistical procedures were analyzed and described based on narrative context. Results are generally relevant and applicative, since the research informed about statistical procedures proposed in dental morphology, to support the professional improvement of researchers in biostatistics.

Statistics is a branch of science that deals with the collection, organisation, analysis of data and drawing of inferences from the samples to the whole population. This requires a proper design of the study, an appropriate selection of the study simple and choice of a suitable statistical test. An adequate knowledge of statistics is necessary for proper designing of an epidemiological study or a clinical trial. Improper statistical methods may result in erroneous conclusions which may lead to unethical practice (Ali & Bhaskar, 2016).

Statistical procedure corresponds to these steps:

- a) Data collection.
- b) Statistical data analysis.
- c) Presentation.

¹ Center for Research in Odontological Sciences (CICO). Department of Adult Dentistry, Dentistry Faculty. Universidad de La Frontera. Temuco, Chile.

² Laboratory of Antropology, Departament of Legal Medicina, Toxicology and Physical Antropology, School of Medicine, University of Granada, Granada, España.

³ Universidad Autónoma de Chile, Chile.

⁴ Department of Orthodontics, Dentistry Faculty, Universidad de La Frontera, Temuco, Chile.

⁵ BioMAT'X Laboratory, Dentistry Faculty. Universidad de los Andes, Santiago, Chile.

⁶ Center for Research and Innovation in Clinical Dentistry (CIDIC). Department of Adult Dentistry, Dentistry Faculty. Universidad de La Frontera. Temuco, Chile.

Quantitative or qualitative information described in research is gathered during data collection. There are various ways to collect data, but it is important to emphasize that this can be performed by simple observation or through other procedures, which are often complex depending on the proposed design.

Statistical data analysis comprises two main statistical methods: descriptive and inferential analysis. The first, more widely known, consists of describing a data set (collect data, present data in tabular or graphical form, summarize measurements, etc.) to interpret the behavior of variables. Inferential analysis consists of applying certain statistical techniques, to generalize or infer results in the analysis population.

Research is at times mistaken for gathering information, documenting facts, and rummaging for information (Leedy & Ormrod, 1974). Research is the process of collecting, analyzing, and interpreting data in order to understand a phenomenon (Leedy & Ormrod). The research process is systematic in that, defining the objective, managing the data, and communicating the findings occur within established frameworks and in accordance with existing guidelines. The frameworks and guidelines provide researchers with an indication of what to include in the research, how to perform the research, and what types of inferences are probable based on the data collected.

Research originates with at least one question about one phenomenon of interest. For example, what competencies might inhibit or enhance the accession of women into senior leadership positions? Or what leadership factors might influence the retention choices of registered nurses? Research questions, such as the two preceding questions, help researchers to focus thoughts, manage efforts, and choose the appropriate approach, or perspective from which to make sense of each phenomenon of interest (Williams, 2007).

The three common approaches to conducting research are quantitative, qualitative, and mixed methods. The researcher anticipates the type of data needed to respond to the research question. For instance, is numerical, textual, or both numerical and textual data needed? Based on this assessment, the researcher selects one of the three approaches to conduct research. Researchers typically select the quantitative approach to respond to research questions requiring numerical data, the qualitative approach for research questions requiring textual data, and the mixed methods approach for research questions requiring both numerical and textual data (Leedy & Ormrod).

MATERIAL AND METHOD

This study corresponds to a cross-sectional descriptive-observational design. We reviewed and collected information through documentary analysis, addressing aspects in statistical test development and applications in dentistry. We used unpublished data corresponding to studies carried out between 2014 and 2017 by researchers at the Center for Research in Odontological Sciences (CICO).

RESULTS

Descriptive Statistics. This procedure describes and indicates preliminary data behavior. Data are presented in tables and graphs, by which central tendency measures, also called summary measures, are calculated for post analysis.

Within this type of analysis, it is essential to consider the following procedures: data presentation, frequency distribution tables and summary measures (mean, median, mode, quartiles, range, variance, standard deviation and variation coefficients).

In this context, variables also play an important role, defined as those characteristics that can be taken by different observations, and are not necessarily numerical. Some examples of non-numerical variables are sex, treatment type, gingival recession, etc. Variables can also be classified as qualitative, corresponding to the attributes or qualities of study subjects, and quantitative, representing numerical values of study subject characteristics. Qualitative variables can be classified as nominal, or those whose category does not have a pre-established order (sex, marital status, vital status, etc.) On the other hand, qualitative variables can be classified in ordinals, with pre-established orders (educational level, patient disease, etc.). Quantitative variables can be classified as continuous or discrete. Continuous quantitative variables are those that take infinite values within a range of numbers (weight, height, etc.), whereas discrete quantitative variables take a finite number of values within a range of numbers (number of cavities, number of patients treated, etc.).

Descriptive measures. The following are descriptive measures, to be used depending on variable type and study objectives:

- Qualitative variables: Rate, ratio and proportion.
- Quantitative variables: Central tendency (averages, modes, medians) and variability (range, variance, standard deviation, variation coefficients) measures.

Statistical inference. Quantitative research in health sciences involves the collection of data from different sources. This data includes demographic variables (e.g. age, gender), biological variables (e.g. weight, blood pressure), risk factors for disease (e.g. smoking status, obesity), outcome variables (e.g. survival data, length of hospital stay), etc. The purpose of statistical analysis is to process this ‘raw’ data into an organized form so that it provides the required information in summary (descriptive statistics). The inferential statistical analysis makes generalizations about the population based on a ‘representative sample’ taken from the study population; it also compares the results between different subgroups of the data sampled to determine any difference or association between the predictor (independent) and outcome (dependent) variables based on the objectives/hypotheses of the study. The purpose of inferential statistics is to determine with a degree of confidence whether the observed differences are statistically significant or may be due to chance alone using the P-value/ confidence interval (Omair, 2014).

Inferential Statistics are used to deduce or obtain generalizations regarding population parameters, accounting for the information obtained from samples. A parameter is a numerical characteristic of one or more populations (population mean, variance, proportion, etc.). In addition, it should be noted that there is a marked division in statistical techniques.

Parametric and non parametric statistics

Parametric Statistics. We can freely say that most people who use statistics are more familiar with parametric than nonparametric techniques. Parametric tests are based on the assumption that the data follow a normal or “bell-shaped” distribution. Parametric methods are often those for which we know that the population is approximately normal, or we can approximate using a normal distribution after we invoke the Central Limit Theorem. There are two parameters for a normal distribution: the mean and the standard deviation. Parametric tests are usually appropriate when examining either interval data or ratio data. Altman states that “parametric methods require the observations within each group to have an approximately Normal distribution if the data do not satisfy these conditions a nonparametric method should be used”.

According to the Central Limit Theorem, when the sample size is larger than 30, normality is not a main condition for a standard t (Student) or z hypothesis test: even though the individual values within a sample might follow an unknown, non-normal distribution, the simple means (as long as the sample sizes are at least 30 will follow a normal distribution (Stojanovic *et al.*, 2018).

Significance tests. Null hypotheses "H0" (no significant differences between samples), and alternative hypotheses "H1" (significant differences between samples) are always considered in statistical significance tests. More information is obtained when null hypotheses can be rejected, meaning that statistics vary among different samples, with a probability higher than 95 %. If the null hypothesis cannot be rejected ($p > 0.05$), information is lost because it cannot be said that samples are the same or different because probability is less than 95 %.

Student t test. In most biomedical research, investigators hypothesize about the relationships of various factors, collect data to test those relationships, and try to draw conclusions about those relationships from the data collected. In many cases, investigators test relationships by comparing the average level of a factor between 2 groups or between 1 group and a standard reference.

In some research projects, the study design includes only a single sample, and the goal may be to determine whether the outcome measure for the population from which the sample was drawn has same mean as some standard population. Determining an appropriate standard for comparison for these designs is often an issue. Nonetheless, when well-established standards exist, investigators may wish to use these standards for maximal comparability. In this situation, we might perform a 1-sample (not 1-sided) t test (Davis & Mukamal, 2006).

Example: (unpublished data)

Is lingual cortical thickness in Class I skeletal patients different from 0? (Table I).

It can be concluded with a significance of 5 % that lingual cortical thickness is different from 0 ($p \leq 0.001$).

Table I. Lingual cortical thickness in Class I skeletal patients – First example.

	N	Mean	Standard deviation
Lingual cortical	16	2.8063	0.91248
T test for a sample			
	T	gl	p
Lingual cortical	12.302	15	0.000

Example: (unpublished data)

Is lingual cortical thickness in skeletal Class I patients different from 2.5 mm? (Table II).

Table II. Lingual cortical thickness in Class I skeletal patients – Second example.

	N	Mean	Standard deviation
Lingual cortical	16	2.8063	0.91248

T test for a sample			
	T	gl	p
Lingual cortical	1.342	15	0.199

It can be concluded with a significance of 5 % that lingual cortical thickness is approximately 2.5 mm ($p = 0.199$).

Student's t-test for independent samples. The independent t-test, also called the two sample t-test or student's t-test, is an inferential statistical test that determines whether there is a statistically significant difference between the means in two unrelated samples on the same continuous, dependent variable. The test also asks whether a difference between two samples averages is unlikely to have occurred because of random chance in sample selection. A difference is more likely to be meaningful and “real” if (i) the difference between the averages is large, (ii) the sample size is large, and (iii) responses are consistently close to the average values and not widely spread out (the standard deviation is low). Note that before performing any Independent t-Test the following assumptions must be satisfied:

- i. Independence: Observations within each sample must be independent (they don't influence each other)
- ii. Normal Distribution: The scores in each population must be normally distributed
- iii. Homogeneity of Variance: The two populations must have equal variances (the degree to which the distributions are spread out is approximately equal) (Usman, 2016).

Example: (unpublished data)

Is there a statistically significant difference between Class I and Class III lower cortical thickness? (Table III).

Table III. Cortical thickness between Class I and Class III.

Class	N	Mean	Standard deviation	p
Class I	108	3.0231	0.90807	0.931
Class II	174	3.0339	1.07894	

It can be concluded with a significance of 5 % that there is no significant difference between Class I and Class III lower cortical thickness ($p = 0.931$).

Student t-Test for related samples. Dependent sample t-test, sometimes called the paired sample t-test is used when the observations in the two populations of interest are collected in pairs. Two samples are dependent (or consist of matched pairs) if the members of one sample can be used to determine the members of the other sample. Words such as dependent, repeated, before and after, matched pairs, paired and so on are hints for dependent samples. According to University of Arizona Military Reach (2009), dependent Samples t-test is used to compare two groups of scores and their means in which the participants in one group are somehow meaningfully related to the participants in the other group. One common example of such a relation is in a pretest post-test research design. Participants at the pre-test are the same participants at the post-test and the scores between pre- and post-test are meaningfully related. What it means is that the scores between pre- and post-test are dependent on each other (Gerald, 2018).

Example: (unpublished data) (Table IV). When applying the t-test for related samples, only significant differences were observed between right and left canines in men ($p = 0.008^*$). If related according to arch sides, no significant differences were observed in women.

Table IV. Comparison of the FMMF between the different teeth related to the side of the arch.

Sex	Tooth	Measure (Newton)		Significance p
		Right	Left	
Men	Molar	705.40	690.72	0.533
	Premolar	522.96	510.24	0.598
	Canine	347.72	297.08	0.008*
Women	Molar	475.64	457.40	0.342
	Premolar	434.00	429.28	0.810
	Canine	242.28	221.76	0.071

One-way Analysis of Variance. Analysis of variance (ANOVA) is one of the most frequently used statistical methods in medical research. The need for ANOVA arises from the error of alpha level inflation, which increases Type I error probability (false positive) and is caused by multiple comparisons. ANOVA uses the statistic F, which is the ratio of between and within group variances. The main interest of analysis is focused on the differences of group means; however, ANOVA focuses on the difference of variances. The illustrated figures would serve as a suitable guide to understand how ANOVA determines the mean difference problems by using between and within group variance differences. The differences in the means of two groups that

are mutually independent and satisfy both the normality and equal variance assumptions can be obtained by comparing them using a Student's t-test. However, we may have to determine whether differences exist in the means of 3 or more groups. Most readers are already aware of the fact that the most common analytical method for this is the one-way analysis of variance (ANOVA) (Kim, 2017).

Example: (unpublished data)

One-way ANOVA was applied for a factor to see if significant differences existed between measurements (condyles and distances) according to age groups, for which it could be observed that there were no significant differences: 1) 11 to 28 years; 2) 29 to 45 years; 3) 46 to 62 years; 4) 63 to 79 years. (Table V).

Table V. One-way ANOVA according age groups and condyle.

Relation	p
Age v/s Right condyle	0.740
Age v/s Left condyle	0.443

Non parametric statistics

When data fails to fulfil the assumptions of a parametric test, researchers opt for non-parametric tests given the fact that they are less restrictive. The test can also be used for small sample size of <30. Therefore, non-parametric tests are used when the variables are non-metric. Although non-parametric tests do not require restrictive statistical assumptions, it is recommendable that the sample obtains a criteria that is; computing the same aggregates for instance, mode, median, range. Furthermore, non-parametric tests also assume that the variables are measured on a nominal or ordinal scale. The tests can further be categorised depending on whether one or two samples are involved (Kataike *et al.*, 2017).

Pearson's Chi Squared (C2) test. The Chi-square statistic is a non-parametric (distribution free) tool designed to analyze group differences when the dependent variable is measured at a nominal level. Like all non-parametric statistics, the Chi-square is robust with respect to the distribution of the data. Specifically, it does not require equality of variances among the study groups or homoscedasticity in the data. It permits evaluation of both dichotomous independent variables, and of multiple group studies. Unlike many other non-parametric and some parametric statistics, the calculations needed to compute the Chi-square provide considerable information about how each of the groups performed in the study. This richness of detail

allows the researcher to understand the results and thus to derive more detailed information from this statistic than from many others. The Chi-square is a significance statistic and should be followed with a strength statistic. The Cramer's V is the most common strength test used to test the data when a significant Chi-square result has been obtained. Advantages of the Chi-square include its robustness with respect to distribution of the data, its ease of computation, the detailed information that can be derived from the test, its use in studies for which parametric assumptions cannot be met, and its flexibility in handling data from both two group and multiple group studies. Limitations include its sample size requirements, difficulty of interpretation when there are large numbers of categories (20 or more) in the independent or dependent variables, and tendency of the Cramer's V to produce relative low correlation measures, even for highly significant results (McHugh, 2013).

Example: (unpublished data). The Pearson Chi-Square test was performed to determine the relationship between sex (male, female) and facial cone (triangular, quadrilateral). No association was indicated ($p = 0.240$) (Table VI).

Table VI. Contingency table sex vs facial cone.

		Facial cone		
		Triangular	Quadrilateral	
Sex	Men	Count % of the total	19 27.1 %	6 8.6 %
	Women	Count % of the total	28 40.0 %	17 24.3 %
Total		Count % of the total	47 67.1 %	23 32.9 %

Mann-Whitney U test (for two independent samples).

Mann-Whitney U test is a non-parametric statistical technique. It is used to analyze differences between the medians of two data sets. It can be used in place of a t-test for independent samples in cases where the values within the sample do not follow the normal or t-distribution but also when the distribution of values is unknown. In order for the Mann-Whitney U test to be applied, values need to be measurable on an ordinary scale and comparable in size. The fact that all values are compared makes it distinct from the t-test, which compares the sample means. The Mann-Whitney U is also used to test the null hypothesis, subject to both samples coming from the same basic set or having the same median value (Milenovic, 2011).

Example: (unpublished data). We aimed to determine if there were significant differences in distances between mesiobuccal canal 1 and mesiobuccal canal 2 according to sex (male, female). After performing the normality test, we observed that samples did not behave normally (Table VII).

Table VII. Normality tests.

	Sex	Kolmogorov-Smirnov*			Shapiro-Wilk		
		Statistical	gl	Sig.	Statistical	gl	Sig.
MB1-MB2	Women	0.261	14	0.011	0.784	14	0.003
	Men	0.282	11	0.015	0.807	11	0.012

We performed the Mann-Whitney U test for independent samples, indicating that the distances between mesiobuccal canal 1 and mesiobuccal canal 2, do not vary significantly between males and females ($p = 0.979$) (Table VIII).

Table VIII. Hypothesis test summary.

	Null hypothesis	Test	Sig.	Decision
1	The distribution of MB1-MB2 is the same among sex categories.	Mann-Whitney U test of independent samples	0.979 ¹	Retain the null hypothesis

The asymptotic significances are shown. The level of significance is .05
¹The exact significance for this test is shown

Kruskal-Wallis test (for three or more independent samples). The Kruskal-Wallis test is a nonparametric technique with which to analyze the variance. In other words, it analyzes whether there is a difference in the median values of three or more independent samples. The Kruskal-Wallis test is similar to the Mann-Whitney test in that it ranks the original data values. That is, it collects all data instances from the samples and ranks them in increasing order. If two scores are equal, it uses the average of the two ranks to be given (Nahm, 2016).

Example: (unpublished data).

We aimed to determine if significant differences exist between nasal index (known magnitude measuring nose width for classification according to size; a quotient or relation, which classifies certain subraces or members, to gain knowledge) and skeletal class (class I, class II and class III). It was observed that samples did not behave normally by performing the normality test, for which the Kruskal-Wallis test was applied, indicating significant differences between nasal index and skeletal classes ($p \leq 0.001$) (Table IX).

Table IX. Contrast statistics^{a,b}

	Nasal Index
Chi-squared	15.910
gl	2
Asymptotic significance	0.000

a. Kruskal-Wallis test
 b. Grouping variable: Skeletal Class

Wilcoxon signed-rank test. Wilcoxon signed rank test is a rank based alternative to the parametric t test that assumes only that the distribution of differences within pairs be symmetric without requiring normality, the Wilcoxon signed rank test is used to test the null hypothesis that the median of a distribution is equal to some value and can be used in place of a one sample t-test, a paired t-test or for ordered categorical data where a numerical scale is inappropriate but where it is possible to rank the observations. To use the Wilcoxon signed rank sum test, we first find the difference between the observation and the hypothesized median in the one sample problem or the difference between the paired observations in the paired sample problems. We then take the absolute values of these differences and rank them, either from the smallest to the largest, or from the largest to the smallest, always taking note of the ranks of the absolute values with positive differences and those with negative differences. The requirement that the populations from which the samples are drawn are continuous, makes it possible to state at least theoretically, that the probability of obtaining zero differences or tied absolute values of the differences is zero (Cochran, 1950).

Example: (unpublished data). We aimed to determine significant differences between hyperplastic and non-hyperplastic condyle width measurements, for which the Shapiro-Wilk normality test was performed ($p = 0.027$). Samples did not meet the normality test, therefore the nonparametric Wilcoxon test for related samples was applied ($p = 0.552$), indicating no differences between measurements (Table X).

Cochran's Q test. In statistics, in the analysis of two-way randomized block designs where the response variable can take only two possible outcomes (coded as 0 and 1), Cochran's Q test is a non-parametric statistical test to verify whether k treatments have identical effects. It is named after William Gemmill Cochran. Cochran's Q test should not be confused with Cochran's C test, which is a variance outlier test. Put in simple technical terms, Cochran's Q test requires that there only be a binary response (e.g. success/failure or 1/0) and that there be more than 2 groups of the same size. The test assesses whether the proportion of successes is the same between groups. Often it is used to assess if different observers of the same phenomenon have consistent results (interobserver variability) (Shoukri, 2004).

Table X. Hypothesis test summary.

	Null hypothesis	Test	Sig.	Decision
1	The median of the differences between med 1 C HCS and med 1 CNHCS equals 0.	Wilcoxon test of the ranges with sign of related samples	0.552	Retain the null hypothesis

The asymptotic significances are shown. The level of significance is .05

Example : (unpublished data). We compared three treatments for Bruxism (T1, T2 and T3). Fourteen patients were randomized and effectiveness (result one) and non-effectiveness (zero result) were determined for each treatment (Table XI).

Table XI. Frequencies and contrast statistics.

Frequencies		
	0	1
Treatment 1	4	10
Treatment 2	3	11
Treatment 3	11	3

Contrast statistics	
N	14
Q of Cochran	14.250 ^a
g1	2
Asymptotic significance	0.001

a. 0 is treated like a success

Example 2: (unpublished data). In our case, it has a significance of 0.001; at less than 0.05 the null hypothesis is rejected. Answers in the three treatments are different.

Friedman’s test. The Friedman test provides another method for testing to detect a shift in location of a set of ?? populations. Like other non-parametric tests, it does not require any assumptions about populations. Consider a situation where a complete randomized design was used to compare the reaction times of subjects under the influence of one or two drugs. When the effect of the drug is short-lined and when the drug effect varies greatly from person to person it may be beneficial to employ a randomized block design. Using the subjects as blocks, we would hope to enhance the variability among subjects and thereby increase the amount of information in the experiment (Friday *et al.*, 2019).

Example : (unpublished data). We used Friedman's test to compare differences in pain levels for post-operated patients who underwent orthognathic surgery, for which follow-up measurements were taken at 3, 5 and 7 days post operation; pain levels were measured according to the following scale: none, low, medium, strong and very strong (Table XII).

Table XII. Ranges and contrast statistics.

Ranges	
	Average range
3rd day	2.04
5th day	2.18
7th day	1.79

Contrast statistics	
N	14
Chi-Squared	1.476 ^a
g1	2
Asymptotic significance	0.478

a. Friedman test

According to results, with a significance of 0.478, there is no statistical evidence to reject the null hypothesis, therefore postoperative pain levels are not different on the third, fifth and seventh days.

DISCUSSION

All health care professionals and medical researchers face the challenge of keeping abreast of a body of knowledge that is expanding at an astonishing rate. The current views on the causes, mechanisms, and treatment methods of diseases are advancing too rapidly for any physician or researcher to achieve personal experience with all the new findings. This has led to a growing reliance on the published literature to learn about new discoveries that can ultimately influence diagnostic evaluations, therapeutic decisions and public health guidelines.

An important function of any medical research journal is the effective dissemination of new findings to its target audience. To be an effective consumer, a journal reader should be familiar with the methodological aspects, especially when the techniques, such as statistical procedures, are invoked to clarify findings or summarize raw data. Statistical methods play an important role in medical publications. This is reflected in the high proportion of articles that are essentially statistical in character. Most papers published in medical journals contain some element of statistical methods, analysis and interpretation (Horton

& Switzer, 2005). Statistical review has also become an important and integral part of the editorial process (Greenwood & Freeman, 2015).

Because of an increasing dependence on the medical literature, it is essential to include statistical education in medical, dental and health care (undergraduate and postgraduate) training as part of the essential topics to support understanding of new research findings. Additionally, clinicians and graduated readers of medical journals should know the frequency with which various statistical concepts are reported in journals that are important to their sub-fields. This helps readers to identify the major statistical skills needed to critically evaluate their literature. Those responsible for training future practitioners and researchers to invest their resources most efficiently should ask the following questions: How often are various statistical techniques reported in the journals of a specific sub-field? Which statistical methods are mentioned most often in their journals compared to more visible journals? Do readers of clinical versus basic science journals need different statistical expertise? Has the use of statistical techniques changed over time or are there new methods that are currently applied more often? (Nieminen *et al.*, 2017).

Therefore, there is a specific protocol or procedure for processing and presenting study results, included in certain designs of appropriate statistical procedures.

Statistic methods have experienced significant growth in recent years with applications in research design strategies and statistic analysis methods, which are more sensitive to different designs. Guides and/or manuals have been developed to guide new statistical procedures and present various methodological issues specific to those designs. Applications of more specific statistical techniques and new software to complement various research designs are increasing.

This work aims to consolidate statistical knowledge and provide an updated approach to statistical testing, which is important and necessary for data management in morphology and dentistry. On the other hand, we present adequate statistical testing options for different design types, providing the possibility of developing research ideas, in the shortest possible time, as well as a data analysis by means of more exact techniques essential to obtain results of a required quality that comply with adequate standards. In addition, our testing methods facilitate the presentation of results in journals, thereby securing a place for investigations which contribute to Dental research.

"After all, high statistics are the numerical expression of common sense" (Karl Pearson)

NAVARRO, P.; ALEMÁN, I.; SANDOVAL, C.; MATAMALA, C. & CORSINI, G. Pruebas estadísticas para el análisis de datos en morfología dental. *Int. J. Morphol.*, 38(5):1317-1324, 2020.

RESUMEN: Los métodos de pruebas estadísticas son herramientas esenciales para medir y analizar los datos obtenidos a través de la investigación en morfología dental. Este estudio descriptivo analiza los diferentes métodos de pruebas estadísticas con el fin de sugerir su uso apropiado y gestionarlo en la interpretación y análisis de datos cuantitativos y / o cualitativos de las investigaciones odontológicas. El conocimiento obtenido a través de este estudio podría orientar a los estudiantes y / o profesionales interesados a decidir, con base teórica, cuál utilizar en el desarrollo de su investigación.

PALABRAS CLAVE: Análisis de los datos; Estadísticas; Métodos.

REFERENCES

- Ali, Z. & Bhaskar, S. B. Basic statistical tools in research and data analysis. *Indian J. Anaesth.*, 60(9):662-9, 2016.
- Cochran, W. G. The comparison of percentages in matched samples. *Biometrika*, 37(3-4):256-66, 1950.
- Davis, R. B. & Mukamal, K. J. Hypothesis testing. *Circulation*, 114:1078-82, 2006.
- Friday, I. A.; Runyi, E. F.; Chukwudike, C. N. & Bright, C. O. Multiple comparison of demand and sale rate of brands of drinks in cross river state using non-parametric test statistics. *Am. J. Math. Stat.*, 9(1):23-9, 2019.
- Gerald, B. A brief review of independent, dependent and one sample t-test. *Int. J. Appl. Math. Theor. Phys.*, 4(2):50-4, 2018.
- Greenwood, D. C. & Freeman, J. V. How to spot a statistical problem: advice for a non-statistical reviewer. *BMC Med.*, 13:270, 2015.
- Horton, N. J. & Switzer, S. S. Statistical methods in the journal. *N. Engl. J. Med.*, 353(18):1977-9, 2005.
- Kataike, J.; Kulaba, J. & Gellynck, X. What junior researchers must know before and after data collection: difference between parametric and nonparametric statistics. *Int. J. Sci. Res.*, 6(6):653-8, 2017.
- Kim, T. K. Understanding one-way ANOVA using conceptual figures. *Korean J. Anaesthesiol.*, 70(1):22-6, 2017.
- Leedy, P. & Ormrod, J. *Practical Research: Planning and Design*. 7th ed. Upper Saddle River, Merrill Prentice Hall and SAGE Publications, 1974.
- McHugh, M. L. The chi-square test of independence. *Biochem. Med. (Zagreb)*, 23(2):143-9, 2013.
- Milenovic, Z. Application of Mann-Whitney U Test in research of professional training of primary school teachers. *Metodicki Obzori*, 11(6):73-9, 2011.
- Nahm, F. S. Nonparametric statistical tests for the continuous data: the basic concept and the practical use. *Korean J. Anesthesiol.*, 69(1):8-14, 2016.
- Nieminen, P.; Virtanen, J. I. & Vähänikkilä, H. An instrument to assess the statistical intensity of medical research papers. *PLoS One*, 12(10):e0186882, 2017.
- Omair, A. Understanding the process of statistical methods for effective data analysis. *J. Health Spec.*, 2(3):100-4, 2014.
- Shoukri, M. *Measures of Interobserver Agreement*. Boca Raton, Chapman & Hall/CRC, 2004.
- Stojanovic, M.; Andjelkovic-Apostolovic, M.; Milosevic, Z. & Ignjatovic, A. Parametric versus nonparametric tests in biomedical research. *Acta Med. Median.*, 57(2):75-80, 2018.
- Usman, M. On Consistency and limitation of independent t-test Kolmogorov Smirnov Test and Mann Whitney U test. *IOSR J. Math.*, 12(4):22-7, 2016.
- Williams, C. Research methods. *JBER*, 5(3), 2007. DOI: <https://www.doi.org/10.19030/jber.v5i3.2532>

Corresponding author:

Pablo Navarro

Center for Research in Odontological Sciences (CICO).

Department of Adult Dentistry

Dentistry Faculty

Universidad de La Frontera

Temuco - CHILE

Received: 21-01-2020

Accepted: 22-04-2020

Email: pablo.navarro@ufrontera.cl