

Biostatistics Methods Used in Celiac Disease Population Studies

Mihaela Iancu¹, Mariana Tulinca^{2,*}, Andrei Achimas Cadariu¹

¹Department of Medical Informatics and Biostatistics, Iuliu Hatieganu University of Medicine and Pharmacy Cluj-Napoca, Romania

²Department of Cellular and Molecular Biology, University of Medicine and Pharmacy Targu Mures, Romania

*Corresponding author: tilincamari@yahoo.com

Received December 02, 2014; Revised December 08, 2014; Accepted December 12, 2014

Abstract Introduction. Research on celiac disease is based on a descriptive and inferential variety of techniques that allow summarizing data and extrapolating the target population. The aim of this study was to identify descriptive and inferential statistical methods used in *International Journal of Celiac Disease*. Material and method. The sample of this study was defined based on all original articles published in the journal mentioned above in 2013-2014 period. We evaluated the descriptive statistics, the number and type of inferential statistical methods used. Results. Of the eight articles analyzed, one study was a randomized clinical trial, the remaining studies were cross-sectional studies. The articles have used both quantitative and qualitative data. The indicator of central tendency on data was mean ($n = 5$, 62.5 %) or the median with interquartile range ($n = 1$, 12.5 %) and the dispersion of data was estimated only by the standard deviation in 5 studies (62.5 %). Quantitative data were used in raw form, except for one study using Z scores. Qualitative data were summarized using absolute and relative frequencies ($n = 7$, 87.5 %). The statistical software was mentioned in all studies. All articles except one used the inferential statistical methods or statistical tests. Most articles (87.5 %) used common tests as Student's t- test (37.5), analysis of variance/covariance (25 %), Kruskal -Wallis test (12.5 %), χ^2 test (50 %), and Fisher's exact test (12.5 %). A specified significance level was noted in 62.5 % of the articles and the significance level was estimated in 6 studies (75 %). Conclusion. Results showed that a proper understanding of standard statistical methods leads to a proper interpretation of clinical research hypothesis. If necessary, the advanced statistical analysis should be performed by a multidisciplinary collaboration with specialists in biostatistics or epidemiology.

Keywords: *statistical methods, celiac disease, journals*

Cite This Article: Mihaela Iancu, Mariana Tulinca, and Andrei Achimas Cadariu, "Biostatistics Methods Used in Celiac Disease Population Studies." *International Journal of Celiac Disease*, vol. 2, no. 4 (2014): 134-136. doi: 10.12691/ijcd-2-4-10.

1. Introduction

The existence of statistical programs with user-friendly interface and easy to use but also needs of scientific journals to demonstrate clinical hypothesis using statistical methods led to an increase in the types of statistical techniques used. Therefore, knowledge of biostatistics methods corresponding to a particular study design has become essential to a critical appreciation of the medical literature.

Description and assessing of statistical methods was performed in various medical fields [1,2,3,4] but not in that of celiac disease. The purpose of this study was to identify methods of analysis used descriptive statistics inferential journal articles *International Journal of Celiac Disease*.

2. Material and Methods

The sample of this study was defined based on all original articles published in the journal noted above. We

evaluated a total of 8 articles published in *International Journal of Celiac Disease*, period 2013-2014 in relation to descriptive statistics, the number and type of inferential statistical methods used. Statistical analysis was described by results presented as absolute frequencies or percentage.

3. Results

Most studies were analytical studies, except one which was limited to descriptive analysis [5]. Statistical programs used were SPSS, R, Matlab, MedCalc or Graphpad. Of the 8 studies, 7 were observational studies, cross-sectional and one study was randomized clinical trial.

The articles have used quantitative and qualitative data. The indicator of central tendency on data was mean ($n = 5$, 62.5 %) or median with range associated interquartile range ($n = 1$, 12.5 %). The data dispersion was calculated by standard deviation in 5 studies (62.5 %). Quantitative data were presented in raw form, except for one study using Z scores. Qualitative data were summarized using absolute and relative frequencies ($n = 7$, 87.5 %).

Of the 8 articles, 7 studies used the inferential statistical methods, with an initial significance level ($n = 5$, 62.5 %) and the estimated significance level was specified in 6 studies (75 %). Name of statistical tests was noted in 6 studies (75 %), the common tests used were the Student's

t- test, Chi -square and / or Fisher test, analysis of variance / co - variance (ANOVA / ANCOVA) and Kruskal-Wallis tests. Distribution of statistical techniques has been presented in Table 1.

Table 1. Description of inferential statistical methods

Statistical method	Number of articles	Relative frequency (%)
Quantitative data		
Pearson product-moment	1	12.5
Qualitative data		
Chi-square test (χ^2)	4	50
Fisher's test	1	12.5
Mixed data		
<i>Parametric tests</i>		
Student's t-test	3	37.5
ANOVA/ANCOVA	2	25
Post-test analysis: multiple comparison (Dunnett test)	1	12.5
<i>Nonparametric tests</i>		
Kruskal–Wallis	1	12.5
<i>Regression techniques</i>		
Classification and Regression Tree (CART)	1	12.5
Confidence interval of parameters	1	12.5
Verification of distribution of normality (Kolmogorov- Smirnov test)	1	12.5

4. Discussion

Analytical studies have evaluated either by Chi-square test or Fisher's exact test (applied if at least 20 % of the theoretical frequencies were less than 5) the differences of frequencies in histopathological changes in the upper gastrointestinal tract in children with lymphocytic colitis in groups aged 3-11 years and 12-18 years [6], differences frequencies related to demographic characteristics (such gender, ethnicity) and clinical characteristics (gastrointestinal symptoms, vomiting, abdominal pain, decreased appetite, nausea, diarrhea, constipation, family history, laboratory Investigations: tissue transglutaminase IgA antibody levels, endomysial antibodies, HLA DQ2 and DQ8 genotypes, nutrient deficiency) in patients with CD and control group consisting of children aged less on than 16 years who had undergone endoscopic SBB for investigation of possible CD [7] or frequency differences related to compliance / adherence (gluten-free diet) with respect to gender (M / F), age group (9-11 years / 12-14years) and disease duration (0-6 years / 7-14 years) [8].

The odds ratio (OR) was one of the statistics used to estimate the magnitude of association of two qualitative variables and the 95 % confidence interval of OR were used to estimate the chance of an event for the studied populations [4].

Of the seven observational studies, one study was a diagnostic study which evaluated the performance of a diagnostic combined test IgA / IgG DGP / tTG for young children. indicators Se, Sp, PPV, NPV performance of a diagnostic test combined IgA / IgG DGP / tTG CB screening for young children. Comparison between the combined test and IgA EMA test respectively IgA tTG test was performed by Se, Sp, PPV, NPV indicators and Chi-square test [9].

In the comparative studies we found also the evaluation of differences in values of quantitative characteristics by parametric and nonparametric statistical tests. Assuming that the condition of normality was verified on data, it was possible to compare means of a quantitative characteristic for two independent populations using a parametric test, such as student-t test or ANOVA test for multiple populations. Thus, the t-test was used to evaluate the average of demographic characteristics (height, weight expressed in Z scores) in CD patients compared to controls [3] or the clinical variables such hs-CRP, IL-6 and TNF-alpha in the female and male patients for each of the groups of patients with CD-defined by Marsh classification (Marsh I, II and IIIA, B and C) [10]. ANOVA was used to assess mean serum levels of hs-CRP, IL-6 and TNF-a for all patients with CD [10] or assessing differences in scores itch, skin lesions and Areas of lesions after 7 days of treatment for on 20 DH (dermatitis herpetiformis) in clinical remission Patients who Received enzyme therapy or placebo. If a test result was significantly ($p < 0.05$) ANOVA test was accompanied by a post-test analysis to determine between which of the two populations they are differences. The Dunnett test, with a level of significance of 0.05, was used in the case when compared groups were homogeneous with respect studied characteristic [11].

If the data did not have a Gaussian distribution, the assessing of differences related to the values of a quantitative variable for multiple populations was performed using the Kruskal-Wallis non-parametric test [1].

One single study developed a predictive model for CD prediction. Classification and Regression using the method tree (CART) was performed to identify predictors of CD based on metabolic profile. This analysis revealed that the most important predictor of CD was lactate metabolite (with a lower level than 0.1026), so it was

revealed a reduction in blood serum level of milk compared to healthy controls [12].

5. Conclusions

Research findings should be viewed in the context of the following limits: in general were used small samples of patients, estimates of indicators were achieved only in the form of point estimates without associated confidence intervals, and if multiple comparisons were performed by t test, the authors did not use a Bonferroni correction to minimize risk of type I errors.

Future studies should examine the consistency and stability of current results but also creating the predictive models based on multivariate analysis techniques such as binary logistic regression or cluster analysis, structural equation modelling, advanced methods that allow identification, ranking of the clinical CD predictors in children population and probability estimation of CD occurrence in this population.

References

- [1] A. Petrie. Statistics in orthopaedic papers. *J Bone Joint Surg* 2006; 88: 1121-1136.
- [2] Rao VP. Statistical methods for survival analysis in oesophageal cancer. *European Journal of Cardio-thoracic Surgery* 2012; 42(4): 759.
- [3] Fisher RA. Statistical methods in Genetics. *International Journal of Epidemiology* 2010; 39(2): 329-335.
- [4] Yin Y, Zhang, H, Qu, P. Statistical methods for clinical trials in oncology. *Chinese Journal of New Drugs* 2014; 23(8): 889-895.
- [5] Jarmila Kabátová., Epidemiology of Celiac Disease in Slovakia: Life Conditions of Celiac Disease Patients in Slovakia. *International Journal of Celiac Disease* 2014; 2(2): 38-39.
- [6] Grzybowska-Chlebowczyk U, Kajor M, Więcek S, Kowol J, Chlebowczyk W, Woś H. Celiac Disease in Course of Lymphocytic Colitis in Children. *International Journal of Celiac Disease* 2014; 2(3): 100-104.
- [7] Wendy Miller, Andrew S Day. A Retrospective Application of the European Society for Paediatric Gastroenterology Hepatology and Nutrition Guidelines for Diagnosis of Coeliac Disease in New Zealand Children. *International Journal of Celiac Disease* 2014; 2(4): 126-130.
- [8] Kvetoslava Rimarova. Differences in Quality of Life and Compliance with Diet in Children with Celiac Disease in Slovak Republic. *International Journal of Celiac Disease* 2014; 2(4): 131-133.
- [9] Belei O, Marginean O. The Usefulness of IgA/IgG DGP/tTG Screen Assay for Celiac Disease Detection among Symptomatic and at Risk Young Children. *International Journal of Celiac Disease* 2013; 1(1): 23-26.
- [10] Abangah Ghobad, Asadollahi Ruhangiz, Rahmani Asghar, Asadollahi Khairollah, Teimouri hadi. Plasma Inflammatory Cytokines as Predictor Markers for Severity of Histological Variations of Intestine among Celiac Patients. *International Journal of Celiac Disease* 2014; 2(2): 54-57.
- [11] Agnieszka Zebrowska, Hugh J. Cornell, Finlay A. Macrae, Anna Sysa-Jedrzejowska, Elzbieta Waszczykowska, Teodor Stelmasiak. The Effect of Enzyme Therapy on Skin Symptoms and Immune Responses in Patients with Dermatitis Herpetiformis. *International Journal of Celiac Disease* 2014; 2(2): 58-63.
- [12] Fariba Fathi, Fatemeh Ektefa, Kaveh Sohrabzadeh, Afsaneh Arefi-Oskouie, Mohsen Tafazzoli, Kamran Rostami, Mohammad-Reza Zali, Mohammad Rostami-Nejad. A Metabonomics Study on Celiac Disease by CART. *International Journal of Celiac Disease* 2014; 2(2): 44-46.