

An Introduction to Statistics: Choosing the Correct Statistical Test

Priya Ranganathan¹

ABSTRACT

The choice of statistical test used for analysis of data from a research study is crucial in interpreting the results of the study. This article gives an overview of the various factors that determine the selection of a statistical test and lists some statistical tests used in common practice.

Keywords: Biostatistics, Research, Statistics as topic.

Indian Journal of Critical Care Medicine (2021): 10.5005/jp-journals-10071-23815

In a previous article in this series, we looked at different types of data and ways to summarise them.¹ At the end of the research study, statistical analyses are performed to test the hypothesis and either prove or disprove it. The choice of statistical test needs to be carefully performed since the use of incorrect tests could lead to misleading conclusions. Some key questions help us to decide the type of statistical test to be used for analysis of study data.²

WHAT IS THE RESEARCH HYPOTHESIS?

Sometimes, a study may just describe the characteristics of the sample, e.g., a prevalence study. Here, the statistical analysis involves only *descriptive statistics*. For example, Sridharan et al. aimed to analyze the clinical profile, species distribution, and susceptibility pattern of patients with invasive candidiasis.³ They used descriptive statistics to express the characteristics of their study sample, including mean (and standard deviation) for normally distributed data, median (with interquartile range) for skewed data, and percentages for categorical data.

Studies may be conducted to test a hypothesis and derive inferences from the sample results to the population. This is known as *inferential statistics*. The goal of inferential statistics may be to assess differences between groups (comparison), establish an association between two variables (correlation), predict one variable from another (regression), or look for agreement between measurements (agreement). Studies may also look at time to a particular event, analyzed using survival analysis.

ARE THE COMPARISONS MATCHED (PAIRED) OR UNMATCHED (UNPAIRED)?

Observations made on the same individual (before–after or comparing two sides of the body) are usually *matched or paired*. Comparisons made between individuals are usually *unpaired or unmatched*. Data are considered paired if the values in one set of data are likely to be influenced by the other set (as can happen in before and after readings from the same individual). Examples of paired data include serial measurements of procalcitonin in critically ill patients or comparison of pain relief during

Department of Anaesthesiology, Critical Care and Pain, Tata Memorial Centre, Homi Bhabha National Institute, Mumbai, Maharashtra, India

Corresponding Author: Priya Ranganathan, Department of Anaesthesiology, Critical Care and Pain, Tata Memorial Centre, Homi Bhabha National Institute, Mumbai, Maharashtra, India, Phone: +91-22-24177000, e-mail: drpriyaranganathan@gmail.com

How to cite this article: Ranganathan P. An Introduction to Statistics: Choosing the Correct Statistical Test. *Indian J Crit Care Med* 2021; 25(Suppl 2):S184–S186.

Source of support: Nil

Conflict of interest: None

sequential administration of different analgesics in a patient with osteoarthritis.

WHAT ARE THE TYPE OF DATA BEING MEASURED?

The test chosen to analyze data will depend on whether the data are categorical (and whether nominal or ordinal) or numerical (and whether skewed or normally distributed). Tests used to analyze normally distributed data are known as *parametric tests* and have a *nonparametric* counterpart that is used for data, which is distribution-free.⁴ Parametric tests assume that the sample data are normally distributed and have the same characteristics as the population; nonparametric tests make no such assumptions. Parametric tests are more powerful and have a greater ability to pick up differences between groups (where they exist); in contrast, nonparametric tests are less efficient at identifying significant differences. Time-to-event data requires a special type of analysis, known as survival analysis.

HOW MANY MEASUREMENTS ARE BEING COMPARED?

The choice of the test differs depending on whether two or more than two measurements are being compared. This includes more than two groups (unmatched data) or more than two measurements in a group (matched data).

TESTS FOR COMPARISON

Table 1 lists the tests commonly used for comparing unpaired data, depending on the number of groups and type of data. As an example, Megahed and colleagues evaluated the role of early bronchoscopy in mechanically ventilated patients with aspiration pneumonitis.⁵ Patients were randomized to receive either early bronchoscopy or conventional treatment. Between groups, comparisons were made using the unpaired *t*-test for normally distributed continuous variables, the Mann–Whitney *U*-test for non-normal continuous variables, and the chi-square test for categorical variables. Chowhan et al. compared the efficacy of left ventricular outflow tract velocity time integral (LVOTVTI) and carotid artery velocity time integral (CAVTI) as predictors of fluid responsiveness in patients with sepsis and septic shock.⁶ Patients were divided into three groups— sepsis, septic shock, and controls. Since there were three groups, comparisons of numerical variables were done using analysis of variance (for normally distributed data) or Kruskal–Wallis test (for skewed data).

A common error is to use multiple unpaired *t*-tests for comparing more than two groups; i.e., for a study with three treatment groups A, B, and C, it would be incorrect to run unpaired *t*-tests for group A vs B, B vs C, and C vs A. The correct technique of analysis is to run ANOVA and use post hoc tests (if ANOVA yields a significant result) to determine which group is different from the others.

Table 2 lists the tests commonly used for comparing paired data, depending on the number of groups and type of data. As discussed above, it would be incorrect to use multiple paired *t*-tests to compare more than two measurements within a group. In the study by Chowhan, each parameter (LVOTVTI and CAVTI) was measured in the supine position and following passive leg raise. These represented paired readings from the same individual and comparison of prereading and postreading was performed using the paired *t*-test.⁶ Verma et al. evaluated the role of physiotherapy on oxygen requirements and physiological parameters in patients with COVID-19.⁷ Each patient had pretreatment and post-treatment data for heart rate and oxygen supplementation recorded on day 1 and day 14. Since data did not follow a normal distribution, they used Wilcoxon’s matched pair test to compare the prevalues and postvalues of heart rate (numerical variable). McNemar’s test was

Table 1: Tests for comparison of unpaired data

Type of data	Two groups	More than two groups
Nominal	Chi-square test or Fisher’s exact test	
Ordinal or skewed	Mann–Whitney <i>U</i> -test (Wilcoxon rank sum test)	Kruskal–Wallis test*
Normally distributed	Unpaired <i>t</i> -test	Analysis of variance (ANOVA)*

*To be followed by post hoc testing

Table 2: Tests for comparison of paired data

Type of data	Two groups	More than two groups
Nominal	McNemar’s test	Cochran’s Q
Ordinal or skewed	Wilcoxon signed rank test	Friedman test*
Normally distributed	Paired <i>t</i> -test	Repeated measures ANOVA*

*To be followed by post hoc testing

used to compare the presupplemental and postsupplemental oxygen status expressed as dichotomous data in terms of yes/no. In the study by Megahed, patients had various parameters such as sepsis-related organ failure assessment score, lung injury score, and clinical pulmonary infection score (CPIS) measured at baseline, on day 3 and day 7.⁵ Within groups, comparisons were made using repeated measures ANOVA for normally distributed data and Friedman’s test for skewed data.

TESTS FOR ASSOCIATION BETWEEN VARIABLES

Table 3 lists the tests used to determine the association between variables. Correlation determines the strength of the relationship between two variables; regression allows the prediction of one variable from another. Tyagi examined the correlation between ETCO₂ and PaCO₂ in patients with chronic obstructive pulmonary disease with acute exacerbation, who were mechanically ventilated.⁸ Since these were normally distributed variables, the linear correlation between ETCO₂ and PaCO₂ was determined by Pearson’s correlation coefficient. Parajuli et al. compared the acute physiology and chronic health evaluation II (APACHE II) and acute physiology and chronic health evaluation IV (APACHE IV) scores to predict intensive care unit mortality, both of which were ordinal data. Correlation between APACHE II and APACHE IV score was tested using Spearman’s coefficient.⁹ A study by Roshan et al. identified risk factors for the development of aspiration pneumonia following rapid sequence intubation.¹⁰ Since the outcome was categorical binary data (aspiration pneumonia— yes/no), they performed a bivariate analysis to derive unadjusted odds ratios, followed by a multivariable logistic regression analysis to calculate adjusted odds ratios for risk factors associated with aspiration pneumonia.

TESTS FOR AGREEMENT BETWEEN MEASUREMENTS

Table 4 outlines the tests used for assessing agreement between measurements. Gunalan evaluated concordance between the National Healthcare Safety Network surveillance criteria and CPIS

Table 3: Tests for assessing the association between variables

Type of data	Test
Correlation	
Both variables normally distributed	Pearson’s correlation coefficient
One or both variables ordinal or skewed	Spearman’s or Kendall’s correlation coefficient
Nominal data	Chi-square test; odds ratio or relative risk (for binary outcomes)
Regression	
Continuous outcome	Linear regression analysis
Categorical outcome (binary)	Logistic regression analysis

Table 4: Tests for assessing agreement between measurements

Type of data	Test
Categorical data	Cohen’s kappa
Numerical data	Intraclass correlation coefficient (numerical) and Bland–Altman plot (graphical display)

for the diagnosis of ventilator-associated pneumonia.¹¹ Since both the scores are examples of ordinal data, Kappa statistics were calculated to assess the concordance between the two methods. In the previously quoted study by Tyagi, the agreement between ETCO₂ and PaCO₂ (both numerical variables) was represented using the Bland–Altman method.⁸

TESTS FOR TIME-TO-EVENT DATA (SURVIVAL ANALYSIS)

Time-to-event data represent a unique type of data where some participants have not experienced the outcome of interest at the time of analysis. Such participants are considered to be “censored” but are allowed to contribute to the analysis for the period of their follow-up. A detailed discussion on the analysis of time-to-event data is beyond the scope of this article. For analyzing time-to-event data, we use survival analysis (with the Kaplan–Meier method) and compare groups using the log-rank test. The risk of experiencing the event is expressed as a hazard ratio. Cox proportional hazards regression model is used to identify risk factors that are significantly associated with the event.

Hasanzadeh evaluated the impact of zinc supplementation on the development of ventilator-associated pneumonia (VAP) in adult mechanically ventilated trauma patients.¹² Survival analysis (Kaplan–Meier technique) was used to calculate the median time to development of VAP after ICU admission. The Cox proportional hazards regression model was used to calculate hazard ratios to identify factors significantly associated with the development of VAP.

SUMMARY

The choice of statistical test used to analyze research data depends on the study hypothesis, the type of data, the number of measurements, and whether the data are paired or unpaired. Reviews of articles published in medical specialties such as family medicine, cytopathology, and pain have found several errors related to the use of descriptive and inferential statistics.^{12–15} The statistical technique needs to be carefully chosen and specified in the protocol prior to commencement of the study, to ensure that the conclusions of the study are valid. This article has outlined the principles for selecting a statistical test, along with a list of tests used commonly. Researchers should seek help from statisticians while writing the research study protocol, to formulate the plan for statistical analysis.

ORCID

Priya Ranganathan <https://orcid.org/0000-0003-1004-5264>

REFERENCES

- Ranganathan P, Gogtay NJ. An introduction to statistics - data types, distributions and summarizing data. *Indian J Crit Care Med* 2019;23(Suppl 2):S169–S170. DOI: 10.5005/jp-journals-10071-23198.
- Nayak BK, Hazra A. How to choose the right statistical test? *Indian J Ophthalmol* 2011;59(2):85–86. DOI: 10.4103/0301-4738.77005.
- Sridharan S, Gopalakrishnan R, Nambi PS, Kumar S, Nandini S, Ramasubramanian V. Clinical profile of non-neutropenic patients with invasive candidiasis: a retrospective study in a tertiary care center. *Indian J Crit Care Med* 2021;25(3):267–272. DOI: 10.5005/jp-journals-10071-23748.
- Hopkins S, Dettori JR, Chapman JR. Parametric and nonparametric tests in spine research: why do they matter? *Global Spine J* 2018;8(6):652–654. DOI: 10.1177/2192568218782679.
- Megahed MM, El-Menshaway AM, Ibrahim AM. Use of early bronchoscopy in mechanically ventilated patients with aspiration pneumonitis. *Indian J Crit Care Med* 2021;25(2):146–152. DOI: 10.5005/jp-journals-10071-23718.
- Chowhan G, Kundu R, Maitra S, Arora MK, Batra RK, Subramaniam R, et al. Efficacy of left ventricular outflow tract and carotid artery velocity time integral as predictors of fluid responsiveness in patients with sepsis and septic shock. *Indian J Crit Care Med* 2021;25(3):310–316. DOI: 10.5005/jp-journals-10071-23764.
- Verma CV, Arora RD, Mistry HM, Kubal SV, Kolwankar NS, Patil PC, et al. Changes in mode of oxygen delivery and physiological parameters with physiotherapy in covid-19 patients: a retrospective study. *Indian J Crit Care Med* 2021;25(3):317–321. DOI: 10.5005/jp-journals-10071-23763.
- Tyagi D, Manjunath BG, Jakka S, Chandra S, Chaudhry D. Correlation of PaCO₂ and ETCO₂ in COPD patients with exacerbation on mechanical ventilation. *Indian J Crit Care Med* 2021;25(3):305–309. DOI: 10.5005/jp-journals-10071-23762.
- Parajuli BD, Shrestha GS, Pradhan B, Amatya R. Comparison of acute physiology and chronic health evaluation II and acute physiology and chronic health evaluation IV to predict intensive care unit mortality. *Indian J Crit Care Med* 2015;19(2):87–91. DOI: 10.4103/0972-5229.151016.
- Roshan R, Sudhakar GD, Vijay J, Mamta M, Amirtharaj J, Priya G, et al. Aspiration during rapid sequence induction: prevalence and risk factors. *Indian J Crit Care Med* 2021;25(2):140–145. DOI: 10.5005/jp-journals-10071-23714.
- Gunalan A, Sistla S, Sastry AS, Venkateswaran R. Concordance between National Healthcare Safety Network (NHSN) Surveillance Criteria and Clinical Pulmonary Infection Score (CPIS) Criteria for Diagnosis of Ventilator-associated Pneumonia (VAP). *Indian J Crit Care Med* 2021;25(3):296–298. DOI: 10.5005/jp-journals-10071-23753.
- Hasanzadeh Kiabi F, Alipour A, Darvishi-Khezri H, Aliasgharian A, Emami Zeydi A. Zinc supplementation in adult mechanically ventilated trauma patients is associated with decreased occurrence of ventilator-associated pneumonia: a secondary analysis of a prospective, observational study. *Indian J Crit Care Med* 2017;21(1):34–39. DOI: 10.4103/0972-5229.198324.
- Yim KH, Nahm FS, Han KA, Park SY. Analysis of statistical methods and errors in the articles published in the Korean journal of pain. *Korean J Pain* 2010;23(1):35–41. DOI: 10.3344/kjp.2010.23.1.35.
- Bahar B, Pambuccian SE, Barkan GA, Akdas Y. The use and misuse of statistical methods in cytopathology studies: review of 6 journals. *Lab Med* 2019;50(1):8–15. DOI: 10.1093/labmed/lmy036.
- Nour-Eldein H. Statistical methods and errors in family medicine articles between 2010 and 2014-Suez Canal University, Egypt: a cross-sectional study. *J Fam Med Prim Care* 2016;5(1):24–33. DOI: 10.4103/2249-4863.184619.