Minimising statistical errors in the research domain: Time to work harder and dig deeper!

Vithal K. Dhulkhed, Thrivikrama P. Tantry[1], Madhuri S. Kurdi[2]

Department of Anaesthesia, Krishna Institute of Medical Sciences, Karad, Maharashtra, [1]Department of Anaesthesiology, AJ Institute of Medical Sciences and Research Centre, Kuntikana, Mangaluru, Karnataka, [2]Department of Anaesthesiology, Karnataka Institute of Medical Sciences (KIMS), Hubli, Karnataka, India

Address for correspondence: Dr. Vithal K. Dhulkhed,
Department of Anaesthesia, Krishna Institute of Medical Sciences, Karad, Maharashtra, India.
E-mail drvithalk@hotmail.com

For the advancement of science, it is essential to identify and correct errors that occur in the process of investigation. It is not surprising to know that a significant number of articles even in top-ranking journals continue to have errors in the study design, methodology and statistics.[1] It is interesting to note that in its early years, the original articles submitted for publication to the Indian Journal of Anaesthesia (IJA) lacked sufficient details of data analysis except for a mention of the "mean" and "*P* value <0.05". However, over the years, there has been a significant improvement in their quality.[2-12] The research process has undergone continual evolution with the proper application of sound research ethics, study methodology, ways of interpretation of observations and presentation.

Statistics is the cockpit of research. Understanding its design, layout and components along with an ability to analyse and understand the output, do have a great influence on maintaining the quality and strength of the research. At this juncture, it is time to ask ourselves some questions. Is the right approach being followed in the design and analysis of clinical trials? Is the research question clearly framed? Are the methodologies being appropriately designed with robust statistical techniques to answer the research question? How much is the magnitude of statistical errors in the present times? Have efforts been made to elevate our level of literacy in research methodology and statistical techniques to minimise errors during the research process and for the presentation of the results? The answers to such questions will be highly variable as is the diversity of articles being submitted to the IJA. Even then, the editorial board is committed to encouraging robust methods of research to publish high-quality articles.

STATISTICAL MISTAKES COMMONLY ENCOUNTERED DURING PEER REVIEW

Statistical errors can lead to misleading conclusions which could be unsafe to use in clinical practice. Errors can occur at various levels in a study and such errors are frequently encountered during the peer review process.

ERRORS IN STUDY DESIGN AND METHODOLOGY

The analysis of data is governed by the study design. During peer review, improper or inadequate study designs that do not resonate with the aims of the study, incorrectly formulated/weak research hypothesis, unclear primary/secondary objectives, unexplained measurement variables, imprecise inclusion and exclusion criteria to define the population characteristics from which the sample is derived and lack of randomisation and blinding are often observed.

SAMPLING AND SAMPLE SIZE ERRORS

Sampling techniques and methods are often not mentioned and sampling errors occur when the researcher does not select a sample that represents the entire population of interest.

A majority of times, the sample size of a study is based on the previous literature, clinical knowledge, consultation with experts in the field or a pilot study.[13] Scientifically speaking, it should be calculated for primary as well as important secondary objectives. Many studies do not mention sufficient details of the sample size calculation.[14] The lack of determination of the sample size at the start of the study is a common observation and many authors cook up some false information and incorrect formulae regarding the sample size estimation after being questioned about it during the review process.

ERRORS IN ANALYSIS

The analytical test sometimes does not match the type of data and wrong/weak tests are applied for statistical inference. It is commonly observed that parameters like mean and standard deviation (SD) are calculated for pain scores which are non-parametric data and *t*-test is erroneously applied instead of non-parametric statistical tests. Often, the investigators do not specify whether the test is paired or unpaired (in such tests as the *t*-test, Wilcoxon's test) and whether it is one or two-tailed. Some authors in sheer ignorance mention standard error wrongly in place of SD. The application of independent *t*-test meant for two-group comparison in a multiple group study, its application in place of analysis of variance (ANOVA) or unpaired *t*-test for paired data is incorrect and is frequently observed. Moreover, the power of the statistical test and confidence intervals are not described adequately.[15] The authors quite often do not define '*priori*' sample size for the study which will give adequate power and sufficiently reduce type II error. Only doing *post hoc* analysis for power on the suggestion of the peer reviewer in such a situation is not advisable; nevertheless, in case the authors do it, the explanation has to be provided as to how, why and for what outcome measures and defined effect sizes the *post hoc* analysis has been done. The most common error is not checking the parametricity of the data before choosing the test to be applied and/or application of a parametric test in cases where the data are skewed or for variable data with non-normal distribution.[16] The failure to follow this first step of data-distribution-analysis may lead to serious errors. The application of well-known tests like the Shapiro–Wilk test and Kolmogorov–Smirnov test at this juncture can help to reduce these errors. The application of statistical significance tests for baseline characteristics like gender, height, weight, etc., of the study groups is commonly observed and this practice is not advisable.[17] In case they have a confounding influence on the study findings, this can be adjusted by including these variables in a multivariate analysis. The advent of modern computers with their fast computational power and the development of statistical software packages has made it possible to progress in multivariate data analysis. The failure to use multivariate techniques to adjust for the confounding factors is a frequent observation. Longitudinal or repeatedly measured data at consecutive time points are frequently erroneously analysed with the wrong assumption that each level of measurement is independent of each other and the *t*-test is applied for two levels. Appropriate multivariate techniques are available such as the Generalised Linear Model for repeated measurements which produce robust results. However, the authors need to be aware of the fact that multivariable analysis is different from multivariate analysis and the two terms should not be interchangeably used. In multivariate analysis, multiple outcomes are analysed, whereas, in multivariable analysis, there is a single outcome variable and multiple independent predictor variables.

The person who is verifying the results can make use of the same software program used by the author or may choose another suitable software package for analysis with more robust statistical techniques. However, many articles fail to give detailed information about the computer software program used for the analysis. Insufficient knowledge of the mathematical concepts underlying the statistical technique and statistical principles can lead to improper use of these packages.

ERRORS IN PRESENTATIONS AND INTERPRETATIONS OF RESULTS

Disparities in the statistical results of the analysis in the abstracts and main body of the text and between the values in the text and tables/figures are found in many of the articles published in some of the frontline journals.

Some researchers when describing the statistical methods used mention that such and such tests were used in the analysis as appropriate. This is not correct and the tests and the distinct data sets or variables to which the tests are applied for the analysis should be elaborated except in the abstract part. The research

articles commonly use *P* values <0.05, <0.01 while indicating the significance of the test results. Ideally, exact calculated *P* values should be mentioned. A poor *P* value interpretation is another observation. Many researchers are unaware of the difference between clinical significance and statistical significance. Even a small difference in the means of the two groups can lead to a statistically significant result if the sample size is unethically large. But this small value of the difference may not be clinically relevant or significant. Hence, the sample size should be calculated with a clinically useful effect size.[18]

It will be clinically useful if the confidence intervals (CIs) of the outcome measures like the mean and odds ratios are mentioned. The CIs are clinically useful as they provide the range of values around the parameter that can be expected on the application of the intervention in a clinical scenario. Missing data, wrong use of CIs, not conducting sub-group analysis and inadequate graphical/numerical description of the basic data add fuel to the fire of statistical errors.

SOME GLIMPSES INTO THE STATISTICS OF THE ORIGINAL ARTICLES PUBLISHED IN IJA

Ninety per cent of the articles in IJA mention only the basic statistical tests viz. Student's *t*-test, Fisher's Exact test and occasionally correlation statements for analysis.[19] A few articles do utilise the ANOVA (one-way or factorial), Receiver Operating Characteristic (ROC) curve, Bland Altman plots or Interclass Correlation Coefficients.[20-22] The value of additional basic or advanced tests is best decided by a dedicated statistician. The basic tests like the multivariate ANOVA, linear regression or multiple regression, when absolutely essential are often not applied by the researchers. The utilisation of univariate analysis and multiple regression by a few authors, however, is noteworthy.[23-25] What constitutes additional or advanced statistics is a matter of subjective opinion. Multiple regression with at least three predictors, factor analysis, and cluster analysis (cluster observations or k-means clustering), discriminant analysis and a time-series analysis do constitute additional tests. Logistic regression, COX/longitudinal regression, survival methods, mixed effect models/data transformation, bootstrapping, propensity score matching and sensitivity analysis, too, can be considered as advanced methods. These statistical tests which are seldom reported in IJA need to be incorporated into the analytical methodology at the data collection stage.[26,27] Additionally, conclusions based on correlation statistics lag behind those on regression statistics due to the inclusion of a single predictor rather than multiple ones.[28] Nonetheless, a wide variety of statistical tests including the Student's *t*-test, Chi-square test, Fisher's Exact test, ANOVA, Mann–Whitney test, repeated multivariate ANOVA, Wilcoxon's signed-rank test, ROC curve, univariate and forward stepwise multivariate logistic regression analysis have been applied in the studies published in this issue of the IJA.[29-36]

The importance of the graphic representation of the data as against a text format cannot be ignored.[37] Non-reporting of missing data and failure to suggest alternate hypotheses are other frequent avoidable errors by the research authors. Arriving at conclusions entirely divergent from those suggested by statistical analysis too is, unfortunately, more common than what meets the eye.

The tables of time-series data are more pleasing and more comprehensive in graphic format and often eliminate the need for lengthy explanations.[38] A common mistake is the absence of error bars representing SD or interquartile ranges in the graphs irrespective of whether they are bar charts or line graphs. An error bar can be shown only on one side of the line graph to keep it simple. Many authors fail to mention the number of subjects (sample) for each time point on the *x*-axis. The insertion of an asterisk (*) for data comparisons having a statistically significant *P* value in the graph itself makes it self-explanatory.[39]

While all reputed journals including the IJA welcome systematic reviews and meta-analyses registered in the PROSPERO, very few have actually been published. Is it the lack of enthusiasm or the complexity of the statistical techniques involved that discourage the authors from venturing into such a pursuit? Nevertheless, presently, the researchers have the luxury of access to free software to help with the statistical analysis.[40] The common errors observed in a meta-analysis are non-recognition of the heterogeneity involved in the included studies, failure to evaluate the same through meta-regressions, ignoring publication bias and not performing sensitivity analysis to derive meaningful conclusions. Several Cochrane articles and guidelines are available which can guide the researchers in conducting a meta-analysis. A special issue of IJA dedicated to meta-analysis should become a reality in the coming days.

High impact factor journals like Anesthesia and Analgesia and Anesthesiology publish articles where the conclusions are made through aggressive and advanced statistical methods. Nonetheless, the use of 't' tests, categorical tests and ANOVA in the original articles published in these journals is showing a declining trend in favour of more important regression or multivariate analytical techniques.[41]

HOW TO IMPROVE THE QUALITY OF STATISTICS IN ORIGINAL RESEARCH ARTICLES?

It is important that the study data should be verifiable and the analysis and results should be reproducible. For instance, it may be contested on suspicion that suppression or misreporting of trial data in a study might have led to false conclusions regarding the safety and efficacy of a drug. It is desirable to provide raw data or individual participant data so that it can help to reanalyse and authenticate the results.[42] The involvement of a statistician in the early phase of designing a study can help in strengthening the study methodology and producing reliable results.The study investigators and authors should follow the principles of essential statistical methods and description of results as per the revised guidelines of the International Committee of Medical Journal Editors.[43] The inclusion of the statistical experts in the editorial team and publishing statistics-related topics more frequently in our journal (such as 'statistical reviews') will certainly help in improving the quality of our research. A national panel of expert biostatisticians can be formed so that the researchers and reviewers can contact them for statistical aspects of research studies.

Statistics is the invisible messenger of the quality of research. Just as one tries to understand the language of poetry, one should learn to understand the language of statistics and use it appropriately with minimal errors to draw distinct, meaningful and clear conclusions.

Financial support and sponsorship
Nil.

Conflicts of interest
There are no conflicts of interest.

REFERENCES

1. Nieuwenhuis S, Forstmann B, Wagenmakers EJ. Erroneous analyses of interactions in neuroscience: A problem of significance. Nat Neurosci 2011;14:1105-7.

2. Sinha C, Kumar A, Kumar A, Kumari P, Singh JK, Jha CK. Deep versus superficial erector spinae block for modified radical mastectomy: A randomised controlled pilot study. Indian J Anaesth 2021;65:97-101.

3. Singariya G, Choudhary K, Kamal M, Bihani P, Pahuja H, Saini P. Comparison of analgesic efficacy of intrathecal 1% 2-chloroprocaine with or without fentanyl in elective caesarean section: A prospective, double-blind, randomised study. Indian J Anaesth 2021;65:102-7.

4. Jadon A, Sinha N, Chakraborty S, Singh B, Agrawal A. Pericapsular nerve group (PENG) block: Afeasibility study of landmark based technique. Indian J Anaesth 2020;64:710-3.

5. Abduallah MA, Ahmed SA, Abdelghany MS. The effect of postoperative ultrasound-guided transmuscular quadratus lumborum block on postoperative analgesia after hip arthroplasty in elderly patients: A randomised controlled double-blind study. Indian J Anaesth 2020;64:887-93.

6. Magoon R, Kaushal B, Chauhan S, Bhoi D, Bisoi AK, Khan MA. A randomised controlled comparison of serratus anterior plane, pectoral nerves and intercostal nerve block for post-thoracotomy analgesia in adult cardiac surgery. Indian J Anaesth 2020;64:1018-24.

7. Singh S, Kumar G, Akhileshwar. Ultrasound-guided erector spinae plane block for postoperative analgesia in modified radical mastectomy: A randomised control study. Indian J Anaesth 2019;63:200-4.

8. Bakshi SG, Gawri A, Divatia JV. McGrath MAC video laryngoscope versus direct laryngoscopy for the placement of double-lumen tubes: A randomised control trial. Indian J Anaesth 2019;63:456-61.

9. Koundal V, Rana S, Thakur R, Chauhan V, Ekke S, Kumar M. The usefulness of Point of care ultrasound (POCUS) in preanaesthetic airway assessment. Indian J Anaesth 2019;63:1022-8.

10. Verma K, Malawat A, Jethava D, Jethava DD. Comparison of transversus abdominis plane block and quadratus lumborum block for post-caesarean section analgesia: A randomised clinical trial. Indian J Anaesth 2019;63:820-6.

11. Bajwa SJ, Bajwa SK, Kaur J, Singh G, Arora V, Gupta S, *et al.* Dexmedetomidine and clonidine in epidural anaesthesia: A comparative evaluation. Indian J Anaesth 2011;55:116-21.

12. Swami SS, Keniya VM, Ladi SD, Rao R. Comparison of dexmedetomidine and clonidine (α2 agonist drugs) as an adjuvant to local anaesthesia in supraclavicular brachial plexus block: A randomised double-blind prospective study. Indian J Anaesth 2012;56:243-9.

13. Dhulkhed VK, Dhorigol MG, Mane R, Gogate V, Dhulkhed P. Basic statistical concepts for sample size estimation. Indian J Anaesth 2008;52:788-93.

14. Glantz SA. Biostatistics: How to detect, correct and prevent errors in the medical literature. Circulation 1980;61:1-7.

15. Strasak AM, Zaman Q, Pfeiffer KP, Göbel G, Ulmer H. Statistical errors in medical research--A review of common pitfalls. Swiss Med Wkly 2007;137:44-9.

16. Lee S, Kang H. Statistical and methodological considerations for reporting RCTs in medical literature. Korean J Anesthesiol 2015;68:106-15.

17. Senn S. Testing for baseline balance in clinical trials. Stat Med 1994;13:1715-26.

18. Page P. Beyond statistical significance: Clinical interpretation of rehabilitation researchliterature. Int J Sports Phys Ther 2014;9:726-36.

19. Ayyanagouda B, Ajay BC, Joshi C, Hulakund SY, Ganeshnavar A, Archana E. Role of ultrasonographic inferior venacaval assessment in averting spinal anaesthesia-induced hypotension for hernia and hydrocele surgeries-A prospective randomised controlled study. Indian J Anaesth 2020;64:849-54.

20. Doctor JR, Chandan P, Shetty N, Gala K, Ranganathan P. Ultrasound-guided assessment of gastric residual volume in patients receiving three types of clear fluids: A randomised blinded study. Indian J Anaesth 2021;65:289-94.

21. Panjiar P, Bhat KM, Yousuf I, Kochhar A, Ralli T. Study comparing different airway assessment tests in predicting difficult laryngoscopy: A prospective study in geriatric patients. Indian J Anaesth 2021;65:309-15.

22. Singh N, Ponde V, Jagannathan B, Rao PB, Dixit A, Agarwal G. Development and validation of a questionnaire to study practices and diversities in plexus and peripheral nerve blocks. Indian J Anaesth 2021;65:197-201.

23. Jain A, Singariya G, Kamal M, Kumar M, Jain A, Solanki RK. COVID-19 pandemic: Psychological impact on anaesthesiologists. Indian J Anaesth 2020;64:774-83.

24. Kohli M, Garg N, Sindwani G, Tempe D, Pamecha V, Pasupuleti SS. Effect of positive cumulative fluid balance on postoperative complications after living donor liver transplantation: A retrospective analysis. Indian J Anaesth 2021;65:383-9.

25. Balakrishnan K, Srinivasaraghavan N, Venketeswaran MV, Ramasamy T, Seshadri RA, Raj EH. Perioperative factors predicting delayed enteral resumption and hospital length of stay in cytoreductive surgery with hyperthermic intraperitoneal chemotherapy: Retrospective cohort analysis from a single centre in India. Indian J Anaesth 2020;64:1025-31.

26. Mishra RK, Pandia MP, Kumar S, Singh GP, Kalaivani M. The effect of anaesthetic exposure in presurgical period on delayed cerebral ischaemia and neurological outcome in patients with aneurysmal subarachnoid haemorrhage undergoing clipping of aneurysm: A retrospective analysis. Indian J Anaesth 2020;64:495-500.

27. Lim WY, Fook-Chong S, Wong P. Comparison of glottic visualisation through Supraglottic airway device (SAD) using bronchoscope in the ramped versus supine 'sniffing air' position: A pilot feasibility study. Indian J Anaesth 2020;64:681-7.

28. Mittal AK, Jaipuria J, Patel A, Bhatnagar V, Chawla R, Singh S. Utility of lung ultrasound for extravascular lung water volume estimation during cytoreductive surgery and hyperthermic intraperitoneal chemotherapy. Indian J Anaesth 2021;65:458-64.

29. Bindu HM, Dogra N, Makkar JK, Bhatia N, Meena S, Gupta R. Limited condylar mobility by ultrasonography predicts difficult direct laryngoscopy in morbidly obese patients: An observational study. Indian J Anaesth 2021;65:612-8.

30. Priya H, Sripriya R, Ravishankar M, Karthikeyan P, Charulatha R. Baska Mask is non-inferior to tracheal tube in preventing airway contamination during controlled ventilation in elective nasal surgeries: A randomised controlled trial. Indian J Anaesth 2021;65:586-92.

31. Goel K, Luthra N, Goyal N, Grewal A, Taneja A. Comparison of norepinephrine and phenylephrine infusions for maintenance of haemodynamics following subarachnoid block in lower segment caeserean section. Indian J Anaesth 2021;65:600-5.

32. Gupta P, Chaudhari SH, Nagar V, Jain D, Bansal A, Dutt A. Prospective analysis of goal-directed fluid therapy vs conventional fluid therapy in perioperative outcome of composite resections of head and neck malignancy with free tissue transfer. Indian J Anaesth 2021;65:606-11.

33. Bafna U, Sharma P, Singhal RK, Gurjar SS, Bhargava SK. Comparison of hypotensive properties of dexmedetomidine versus clonidine for induced hypotension during functional endoscopic sinus surgery: A randomised, double-blind interventional study. Indian J Anaesth 2021;65:579-85.

34. Jadon A, Mohsin K, Sahoo RK, Chakraborty S, Sinha N, Bakshi A. Comparison of supra-inguinal fascia iliaca versus pericapsular nerve block for ease of positioning during spinal anaesthesia: A randomised double-blinded trial. Indian J Anaesth 2021;65:572-8.

35. Ray A, Sen IM, Bhardwaj N, Yaddanapudi S, Mathews P. Videolaryngoscopic versus direct laryngoscopic paraglossal intubation for cleft lip/palate reconstructive surgeries: A randomised controlled trial. Indian J Anaesth 2021;65:593-9.

36. Viola CT, Joselyn AS, Sukumar A, Sahajanandan R. Preoperative anxiety among patients scheduled for elective surgical procedures during the COVID-19 pandemic – A cross-sectional study in a tertiary care teaching hospital in India. Indian J Anaesth 2021;65:619-25.

37. Lee S. Avoiding negative reviewer comments: Common statistical errors in anesthesia journals. Korean J Anesthesiol 2016;69:219-26.

38. Saha M, Saxena KN, Wadhwa B. Comparative study of recovery of airway reflexes and cognitive function following sevoflurane versus desflurane anaesthesia. Indian J Anaesth 2021;65:282-8.

39. Ravishankar M, Mathew DM, Hemanthkumar VR, Srinivasan P. Quantifying influence of epidural analgesia on entropy guided general anaesthesia using sevoflurane – A randomised controlled trial. Indian J Anaesth 2020;64:131-7.

40. Tantry TP, Karanth H, Shetty PK, Kadam D. Self-learning software tools for data analysis in meta-analysis. Korean J Anesthesiol 2021. doi: 10.4097/kja. 21080. Online ahead of print.

41. Staffa SJ, Zurakowski D. Recent trends in utilization of statistical methods in anesthesia research: 2012-2017. Trends Anes Surg 2018;1:1-2.

42. Bajwa Sukhminder Jit Singh. Basics, common errors and essentials of statistical tools and techniques in anesthesiology research. J Anaesthesiol Clin Pharmacol 2015;31:547-53.

43. Moher D, Schulz KF, Altman DG. The CONSORT statement: Revised recommendations for improving the quality of reports of parallel group randomised trials. Lancet 2001;357:1191-94.

Announcement

Old Issues of IJA

Limited copies of old issues of IJA from 2013 are available in IJA office. Members interested can contact Immediate Past Editor In Chief (editorija@yahoo.in/ijadivatia@gmail.com / 98690 77435)