



STATISTICAL PRACTICES IN GENERAL SURGERY RANDOMIZED CONTROL TRIAL STUDY: AN EVALUATION OF PUBLISHED RESEARCH ARTICLES

Dharmendra Kumar Dubey^{1*} and Nancy Parul²

^{1*} Assistant Professor, Department of Community Medicine, Baba Kinaram Autonomous State Medical College Chandauli, Government of Uttar Pradesh, **Email id:** dubey.dharm@gmail.com, **ORCID:** <https://orcid.org/0000-0002-7256-6738>.

² Assistant Professor, Department of General Surgery, Baba Kinaram Autonomous State Medical College Chandauli, Government of Uttar Pradesh, **Email id:** parul.n11@gmail.com, **ORCID:** <https://orcid.org/0009-0003-4273-5347>.

Abstract: This review article evaluates the use and appropriateness of statistical methods in general surgery research articles indexed in PubMed over the past five years (2020–2025). A total of 5,388 original articles were screened, among which 15 randomized controlled trials (RCTs) containing the keyword “General Surgery” in the title were identified. These RCTs featured 23 statistical applications primarily focused on continuous or ordinal outcomes, with the Pearson chi-square test and Student’s t-test being the most commonly used methods. Notably, 26% of the reviewed articles showed deviations from CONSORT guidelines or omitted the “General Surgery” keyword, limiting discoverability and methodological clarity. The most frequent statistical error observed was inadequate reporting of p-values. Furthermore, only 33.3% of the studies clearly specified the statistical software used, with SPSS and STATA being the most frequently mentioned programs. An encouraging trend toward the increased and more appropriate use of statistical methods in general surgery literature, indicating growing awareness among researchers. However, the persistence of reporting issues and inconsistent software disclosure highlights the need for improved adherence to reporting

Key Words: Statistics, General surgery, Randomized control trial

Introduction: Effective critical appraisal of the scientific literature is an essential skill for all surgeons to remain up to date in evidence based surgery. Therefore, a working knowledge of statistical techniques is a prerequisite for critical appraisal. The emergence of evidence-based medicine (EBM) has significantly influenced clinical practice and medical research over the past few decades. Elstein outlines the philosophical and historical development of EBM, emphasizing its role in improving clinical decision-making through systematic evidence appraisal [1]. However, randomized controlled trials (RCTs) represent the highest level of clinical evidence, but their value depends critically on sound statistical design, conduct, and transparency in reporting [2]. Concerns about the quality of medical research are not new.

Altman criticized the pervasiveness of poor research design and statistical misuse, calling it a “scandal” that compromises the reliability of scientific conclusions [3]. Supporting this view, Windish et al. found that internal medicine residents often struggle to interpret statistical results in the literature, which raises concerns about the dissemination and application of evidence in clinical

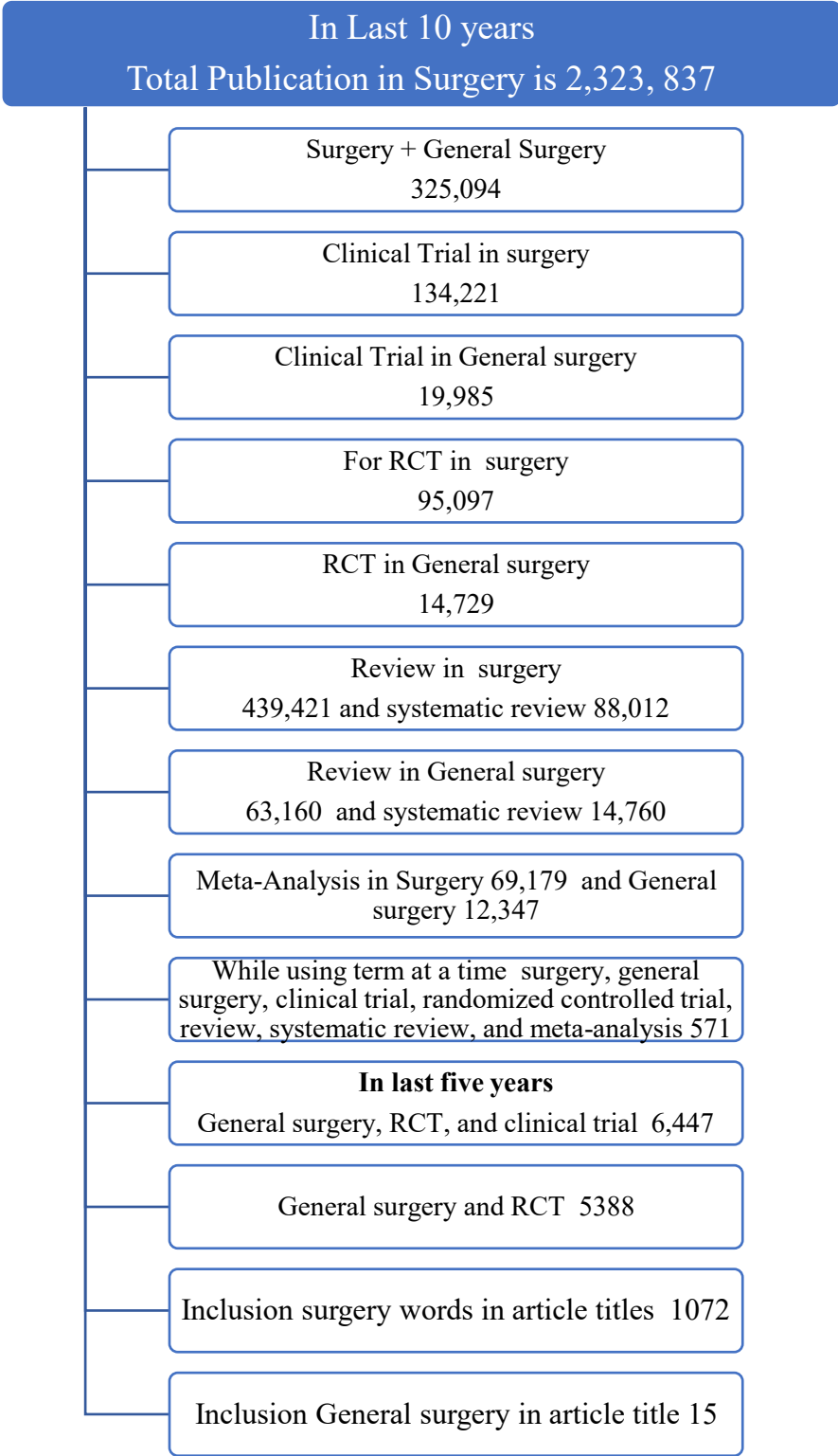
settings [4]. Moreover, Emerson and Colditz highlighted that although statistical methods are widely used in top-tier journals like The New England Journal of Medicine, their application often lacks consistency or depth [5]. Han and Jung reported evolving patterns in the use of statistical methods in plastic surgery journals, noting a preference for basic descriptive and inferential statistics over more complex models [6]. Parmar et al., in their systematic evaluation of addiction science journals, revealed that many journals fail to adequately endorse unbiased and transparent reporting practices, as outlined in their author instructions [7]. Given the persistent gaps in statistical understanding, reporting, and methodological rigor in surgical research, this study aimed to identify the most commonly used statistical techniques in general surgery articles published in PubMed (2020–2025) and assess their appropriateness, emphasizing the need for transparency and improved research standards.

Methods: In surgery, globally, between the past 10 years and up to May 30, 2025, a total of 2,323, 837 articles have been published. Among these, 325,094 articles include the keyword “surgery” in combination with “general surgery.” When filtered for clinical trials, there are 134,221 articles under surgery and 19,985 specifically under general surgery. For randomized controlled trials (RCTs), the numbers are 95,097 for surgery and 14,729 for general surgery. Articles categorized as reviews number 439,421 for surgery and 63,160 for general surgery. Similarly, systematic reviews account for 88,012 articles in surgery and 14,760 in general surgery, while meta-analyses appear in 69,179 surgical articles and 12,347 under general surgery.

When all the keywords—surgery, general surgery, clinical trial, randomized controlled trial, review, systematic review, and meta-analysis—are used together, only 571 articles were identified. In the last five years alone, articles using the keywords general surgery, RCT, and clinical trial together total 6,447. Therefore, if we specify key words general surgery and RCT total 5388 published research article reported.

Inclusion and Exclusion Criteria: This research represents the Randomized control trial in general surgery. Data extracted from PubMed website, using the key word “General surgery and Randomized control trial. Total published article is 5388. After identified list of total 5388 of publications only those articles were considered for statistical evaluation which article title included surgery words. That is 1072. Final statistical evaluation done only those article titles are having words “General Surgery” that is 15. Above details were shown in **flow diagram**.

Flow diagram



Results: Table 1: This table provides insight into how well general surgery RCTs adhere to best statistical and ethical practices. It evaluates various domains, including reporting guidelines, statistical methodology, data presentation, and ethical considerations. The CONSORT (Consolidated Standards of Reporting Trials) guidelines referenced Only 26.7% of the studies. Which provides a standardized framework to improve the transparency and quality of reporting in RCTs. Every study (100%) reported a statistical analysis. This is a positive indicator showing that statistical methods are

universally employed to interpret the data. All articles reported sample size calculations (100%) or data. Some of the studies clearly not reported inclusion and exclusion criteria. Many of the articles were not mentioned the statistical software (33.3%) used for analysis. Informed Consent of the articles reported and it compliance with ethical guidelines. Ethical Considerations is essential information that should be universally reported. Transparency regarding conflicts of interest is crucial for understanding potential biases. Clinical Trial Registration helps prevent selective outcome reporting and ensures accountability.

Most studies reported mean values, and a majority included standard deviation (SD) to convey data spread. Few studies reported median and interquartile range (IQR), which are more appropriate for skewed data. Only a small fraction of studies reported risk ratio and risk reduction. A formal hypothesis guides study design and statistical testing. Its absence may indicate poor planning or weak scientific rationale. The p-value was reported in nearly all studies. Confidence intervals and p-values should ideally be included in all clinical research. Most studies used tables to present data, and footnotes and visual representation of data is critical for reader comprehension, and figure legends are essential for interpretation.

Chi-Square Test was the most commonly reported test, used for assessing categorical variables. Student's t-Test used for comparing means between two groups with normally distributed data. Many other statistical tests, such as paired t-tests, Mann-Whitney U, Wilcoxon signed-rank, Kruskal-Wallis, and repeated measures ANOVA, were each reported by the studies. Regression Models, use of models like Cox proportional hazards, Poisson regression, and multivariate regression was rare. Subgroup Analyses & Effect Size, very few studies conducted subgroup analyses or reported effect sizes. Numbers Needed to Treat (NNT) tells clinicians how many patients need to receive a treatment to prevent one adverse outcome and is highly relevant in practice. Intention-to-Treat Analysis suggests potential biases. ITT maintains randomization benefits and reflects real-world adherence. Visual Analog Scale (VAS) & Likert Scale, these are common tools for measuring subjective outcomes like pain and satisfaction. Other Scales, such as Intensive Care Oral Care Frequency Assessment Scale (ICOCFAS), Braden, Berg Balance, and NRS were used rarely.

The Publication Types was Randomized Controlled Trials, as expected, all studies were RCTs, which is consistent with the study's inclusion criteria. The RCT includes Multicentre Study, Comparative Study, Research Support and Comment.

Table 1. Statistical methods applied in the articles published in PubMed specifically RCT and General surgery (n =15).

Domains Evaluated	Domains Evaluated
Reporting guidelines	Student-t Test
CONSORT	Paired t-test
Statistical reporting	Mann Whitney-U Test
Methodology	Wilcoxon signed rank test
Statistical analysis	Repeated measures ANOVA
Sample size	Kruskal-Wallis tests
Inclusion criteria	One-way analysis of variance, and post-hoc test
Exclusion criteria	Kolmogorov-Smirnov test
Statistical software	Subgroup analysis
Descriptive statistics	Cox proportional hazards models
Mean	Hazard Ratio
SD	Poisson's regression
Median	Multivariate regression analysis
IQR	Fisher exact test
Risk Ratio	Effect size
Risk reduction	Numbers needed to treat (NNT)
Inferential statistics	Intention-to-treat
Hypothesis	Used Scale

p-Value	Intensive Care Oral Care Frequency Assessment Scale (ICOCFAS)
95% CI	Visual Analog Scale (VAS)
Presentation of the results	Braden Scale score
Tables	Likert Scale
Table footnotes	(Berg Balance Scale
Figures	Numerical rating scale (NRS)
Figure legends	Key Words: General Surgery
Ethical requirements	Publication types
Informed consent	Randomized Controlled Trial
Ethical considerations	Multicentre Study
Conflict of interest	Comment
Clinical trial registration	Research Support
Statistical test	Comparative Study
Chi-square test	

Table 2, Summary of Randomized Controlled Trials in General Surgery Indexed in PubMed (2020–2025).

S.No.	Title	First Author	Journal	Publication Year
1	Safety and Efficacy of Tranexamic Acid in General Surgery	Park LJ, et al.	JAMA Surgery	2025
2	Effects of cold spray on thirst, frequency of oral care, and pain of general surgery intensive care unit patients	Gungor S, et al.	Scientific Reports	2024
3	Effect of a prophylactic dressing for sacral pressure injuries in non-critically ill patients after general surgery: A randomized controlled trial	Yeo H, et al.	Worldviews on Evidence-Based Nursing	2023
4	Telemedicine versus face-to-face follow up in general surgery: a randomized controlled trial	Fink T, et al.	ANZ Journal of Surgery	2022
5	The investigation of effect on foot plantar massage on functional recovery in older adults with general surgery, randomized clinical trial	Saltan A, et al.	Aging Clinical and Experimental Research	2024
6	Impact of pharmacist interventions on drug-related problems in general surgery patients: a randomised controlled trial	AbuRuz S, et al.	European Journal of Hospital Pharmacy	2021
7	Video consent significantly improves patient knowledge of general surgery procedures	Bremer K, et al.	Surgical Endoscopy	2024
8	Feasibility of Prospectively Comparing Opioid Analgesia With Opioid-Free Analgesia After Outpatient General Surgery: A Pilot Randomized Clinical Trial	Do U, et al.	JAMA Network Open	2022
9	A Clinical Trial of a Video Intervention Targeting Opioid Disposal After General Surgery: A Feasibility Study	Lewis J, et al.	Journal of Surgical Research	2021

10	The Impact of Simulation Training on Operative Performance in General Surgery: Lessons Learned from a Prospective Randomized Trial	Naples R, et al.	Journal of Surgical Research	2022
11	Disinfection with single or double usage of new antiseptic olanexidine gluconate in general surgery: a randomized study	Yamamoto M, et al.	Langenbeck's Archives of Surgery	2020
12	Real-Time Pain Control Education After Outpatient General Surgery: A Randomized Controlled Trial	Lee WG, et al.	Journal of Surgical Research	2025
13	Comparing the effects of rosemary aromatherapy and music therapy on anxiety levels in patients undergoing general surgery: A randomized controlled clinical trial	Sayadi Mank-Halati M, et al.	Explore: The Journal of Science and Healing	2024
14	Economic impact of outpatient follow-up using telemedicine vs in-person visits for patients in general surgery: A secondary analysis of a randomized clinical trial	Ferret G, et al.	Cirugía Española (English Edition)	2024
15	Use of Low-Cost Task Trainer for Emergency Department Thoracotomy Training in General Surgery Residency Program	Misra A, et al.	Journal of Surgical Education	2024

Table2, presents 15 randomized controlled trials (RCTs) published between 2020 and 2025, all of which explicitly include the term "General Surgery" in their article titles. The studies were published in a range of peer-reviewed journals, including high-impact titles such as JAMA Surgery, Scientific Reports, and the Journal of Surgical Research, indicating a broad dissemination of general surgery research across both surgical and interdisciplinary platforms.

Discussion: In this review, we analyzed articles published in PubMed between 2020 and 2025, focusing on the use and types of statistical methods and software packages. The results revealed an increasing trend in the application of statistical techniques and the moderate use of statistical software. When focusing on general surgery, Balasubramanian et al. demonstrated that the reporting standards for RCTs are often suboptimal, particularly in the documentation of methodology and statistical analysis [8]. The misinterpretation of statistical significance particularly reliance on p-values has received substantial attention. Wasserstein and Lazar, on behalf of the American Statistical Association, issued a statement discouraging overdependence on p-values as sole indicators of significance, urging researchers to consider effect sizes and confidence intervals for more meaningful interpretations [9]. Kyriacou further discusses how the p-value has evolved over time and its limitations in modern research contexts [10]. This observation is echoed in surgical literature more broadly, where issues like underreporting of power analysis, lack of intention-to-treat analysis, and unclear primary outcome measures persist [11,12].

Notably, 10 out of the 15 articles reviewed did not specify which statistical software was used. This lack of disclosure is concerning, as different statistical programs can produce varying results due to differences in default algorithms and computational methods. Clear reporting of the software used is essential for transparency, reproducibility, and accurate interpretation of findings. Williams et al. examined statistical techniques commonly used in general surgery literature and stressed the importance of statistical literacy among surgeons to improve research quality and interpretation [12]. Statistical methods should be clearly described in the Methods section, including the software used and significance levels set. Some articles presented results without specifying methods, making

interpretation difficult. Despite this, they were classified as using statistics. Journal guidelines emphasize detailing both the type and purpose of statistical analyses to ensure transparency, reproducibility, and proper scientific interpretation. Studies with small sample sizes, such as early-stage surgical research or animal experiments, may not require complex statistical tests. While advanced analyses aren't always necessary, clear data summarization using appropriate descriptive statistics is valuable. When needed, nonparametric methods can offer suitable alternatives for small-sample statistical analysis. Similarly, Robinson et al. systematically reviewed surgical RCTs over more than a decade and found that while the volume of RCTs has increased, methodological quality and reporting standards remain inconsistent [13].

Conclusion:

The present review highlights both progress and persistent gaps in the application and reporting of statistical methods in randomized controlled trials (RCTs) within general surgery literature from 2020 to 2025. While it is encouraging that all reviewed studies employed statistical analyses, and many adhered to basic reporting norms such as sample size estimation and p-value reporting, there remains considerable room for improvement in methodological rigor and transparency.

Despite widespread adoption of evidence-based medicine in surgical practice, these findings echo longstanding concerns about statistical literacy and reporting quality. Given that RCTs represent the highest standard of clinical evidence, their reliability is heavily contingent on appropriate study design, statistical analysis, and transparent reporting. Without these components, conclusions drawn from surgical trials risk being misleading or irreproducible. Furthermore, omitting "General Surgery" as a keyword limits article discoverability, which may hinder knowledge translation. Future research should emphasize adherence to CONSORT guidelines, full disclosure of analytic tools, and improved statistical training among researchers to strengthen the evidence base and ensure the validity of clinical recommendations.

Key highlights:

Although this study focused on evaluating publications related to surgery and randomized controlled trials (RCTs), it was noted that many of the authors were not affiliated with general surgery departments. There was also a lack of clarity regarding the use of parametric versus non-parametric tests. It is recommended that, when physiotherapists or professionals from other disciplines author such articles particularly those involving clinical scales at least one co-author should be from a surgery department to ensure clinical relevance. Additionally, the statistical software used should be explicitly stated, and "General Surgery" should be included as a keyword to improve the article's visibility and categorization.

Disclosure: The authors have nothing to disclose.

Conflict of interest: There is no conflict of interest.

Funding source: There is no funding available for this task.

Authors' Contribution: All authors contributed equally.

Reference

1. Elstein A. On the origins and development of evidence-based medicine and medical decision making. *Inflamm Res*. 2004;53(Suppl 2):S184–S189.
2. Cook JA. The challenges faced in the design, conduct and analysis of surgical randomised controlled trials. *Trials*. 2009;10:9.
3. Altman DG. The scandal of poor medical research. *BMJ*. 1994;308(6924):283.
4. Windish DM, Huot SJ, Green ML. Medicine residents' understanding of the biostatistics and results in the medical literature. *JAMA*. 2007;298(9):1010–1012.

5. Emerson JD, Colditz GA. Use of statistical analysis in the New England Journal of Medicine. *N Engl J Med*. 1983;309(12):709–713.
6. Han K, Jung I. Trends in statistical methods in articles published in Archives of Plastic Surgery between 2012 and 2017. *Arch Plast Surg*. 2018;45(3):207–213. doi:10.5999/aps.2018.00010
7. Parmar A, Dubey DK, Balhara YPS, Mishra AK. Do addiction science journals endorse unbiased reporting of research? A systematic evaluation of instructions for authors. *Subst Use Misuse*. 2019. doi:10.1080/10826084.2019.1610444
8. Balasubramanian SP, Wiener M, Alshameeri Z, Tiruvoipati R, Elbourne D, Reed MW. Standards of reporting of randomized controlled trials in general surgery: can we do better? *Ann Surg*. 2006;244(5):663–667. doi:10.1097/01.sla.0000217640.11224.05
9. Wasserstein RL, Lazar NA. The ASA’s statement on p-values: context, process, and purpose. *Am Stat*. 2016;70:129–133.
10. Kyriacou DN. The enduring evolution of the p value. *JAMA*. 2016;315:1113–1115.
11. Solomon MJ, Laxamana A, Devore L, McLeod RS. Randomized controlled trials in surgery. *Surgery*. 1994;115(6):707–712.
12. Williams PJM, Murphy P, Van Koughnett JAM, et al. Statistical techniques in general surgery literature: what do we need to know? *J Am Coll Surg*. 2018;227(4):450–454.e1. doi:10.1016/j.jamcollsurg.2018.07.656
13. Robinson NB, Fremes S, Hameed I, et al. Characteristics of randomized clinical trials in surgery from 2008 to 2020: a systematic review. *JAMA Netw Open*. 2021;4(6):e2114494. doi:10.1001/jamanetworkopen.2021.14494