International Journal of Recent Innovations in Academic Research

E-ISSN: 2635-3040; P-ISSN: 2659-1561
Homepage: https://www.ijriar.com/
Volume-9, Issue-2, April-June-2025: 139-146

Systematic Review

The Impact of Significance Level and Hypothesis Testing in Biomedical Data Analysis

[*a]Joseph Ozigis Akomodi, [b]Roydon Pellew, [c]Daniel Chung, [d]Rubaba Tabassum, [e]Victoria Bissessar, [f]Lamiah Atiqa and [g]Syeda Nahar

[a]Department of Education, Long Island University, New York, USA
[b]Department of Education, New York City, USA
[c]New York Institute of Technology, New York, USA
[d-g]Queens Gateway to Health and Sciences, New York, USA
*Corresponding Author Email: jakomod@schools.nyc.gov

Received: March 26, 2025 **Accepted:** April 17, 2025 **Published:** April 24, 2025

Abstract
The research analyzes the impact of significance levels and hypothesis testing in biomedical data analysis. By conducting a systematic review of 19 peer-reviewed studies, we examined the relationship between selected significance levels (α=0.01, 0.05, 0.10) and the occurrence of Type I and Type II errors. The findings reveal that the majority of studies employed a significance level of α=0.05, resulting in Type I error rates ranging from 0.040 to 0.050. In contrast, studies utilizing a more stringent significance level of α=0.01 reported lower Type I errors but exhibited higher Type II errors, indicating a tendency to overlook true treatment effects. Furthermore, only 50% of studies that achieved statistical significance adequately addressed the clinical relevance of their findings. These results underscore the necessity for researchers to carefully consider their choice of significance level and to explicitly communicate the clinical implications of their results. This analysis highlights the critical role of rigorous hypothesis testing in enhancing the reliability and applicability of biomedical research outcomes, ultimately contributing to improved patient care and treatment strategies. Future research should focus on establishing standardized reporting practices that bridge the gap between statistical significance and clinical relevance.
Keywords: Hypothesis Testing, Significance Levels, Type I and Type II Error, Biomedical Research.

1. Introduction

Biomedical data analysis serves as a cornerstone in the field of healthcare, enabling researchers and practitioners to understand complex health-related phenomena, assess the efficacy of treatments, and improve patient outcomes. With the rapid advancement of medical technologies and data collection methodologies, the ability to analyze biomedical data has become increasingly sophisticated. This analysis allows for the identification of patterns, relationships, and causal effects that inform clinical practices and public health policies. One of the fundamental statistical methodologies employed in biomedical data analysis is hypothesis testing.

At the heart of hypothesis testing lies the formulation of two competing hypotheses: the null hypothesis ($H0$) and the alternative hypothesis (Ha). The null hypothesis typically posits that there is no effect, relationship, or difference between groups or variables, serving as a baseline for comparison. Conversely, the alternative hypothesis suggests that there is an effect or a difference that warrants further investigation (Dickersin, 1990). For example, in a clinical trial evaluating a new drug, the null hypothesis may state that the drug has no effect on patient recovery rates compared to a placebo, while the alternative hypothesis would claim that the drug does improve recovery rates.

The significance level (α) is a critical component of hypothesis testing, representing the probability of making a Type I error rejecting the null hypothesis when it is actually true. Commonly employed significance levels include α=0.05, which indicates a 5% risk of incorrectly concluding that a difference exists when it does not, and α=0.01, which reduces this risk to 1% (Baker, 2016). The choice of significance level can have profound implications for the interpretation of research findings. A lower significance level may lead to more conservative conclusions, reducing the likelihood of false positives but increasing the risk of false negatives (Type II errors). This delicate balance is essential in clinical settings, where the consequences of misinterpretation can affect patient safety and treatment efficacy.

As such, the significance level not only influences statistical outcomes but also has far-reaching implications for clinical decision-making and treatment protocols. For instance, a study that reports a statistically significant finding at α=0.05 may lead to the adoption of a new treatment; however, if the findings are not clinically relevant or if the significance level is misinterpreted, patients may be subjected to unnecessary risks (Berselli, *et al.,* 2021). Therefore, understanding the role of significance levels and hypothesis testing is crucial for researchers, clinicians, and policymakers alike.

The introduction highlights the vital role of biomedical data analysis in improving health outcomes, the importance of hypothesis testing, and the critical influence of the significance level on research interpretations. As healthcare continues to evolve with new data-driven approaches, it is imperative that researchers accurately apply these statistical principles to ensure that biomedical findings are robust, reliable, and clinically relevant.

2. Hypotheses in Biomedical Research

Biomedical data analysis is essential for advancing our understanding of health-related phenomena and developing effective treatments. This analytical process is fundamentally anchored in hypothesis testing, which involves the formulation of two competing hypotheses: the null hypothesis ($H0$) and the alternative hypothesis (Ha). The null hypothesis typically asserts that there is no effect or difference between groups or treatments, serving as a baseline for comparison (Cohen, 1988). In contrast, the alternative hypothesis posits that there is a significant effect or difference that warrants further investigation. A critical aspect of hypothesis testing is the significance level (α), which represents the probability of rejecting the null hypothesis when it is true, commonly set at 0.05 or 0.01 (Baker, 2016).

The choice of significance level is paramount, as it can significantly influence the interpretation of research results, affecting clinical decisions and treatment protocols. For instance, a lower significance level (e.g., α=0.01) reduces the likelihood of Type I errors, which can prevent the premature acceptance of ineffective treatments; however, it may also increase the risk of Type II errors, potentially overlooking beneficial treatments (Baker, 2016). Therefore, a comprehensive understanding of these statistical principles is critical for researchers and practitioners in the biomedical field, as it directly impacts the quality and applicability of research findings in clinical practice.

3. Significance Level (α)

The significance level, denoted as α, is a fundamental concept in hypothesis testing that quantifies the probability of making a Type I error rejecting the null hypothesis ($H0$) when it is actually true. This probability is crucial for determining the rigor of statistical conclusions drawn from research data. Commonly used significance levels include α=0.05, which indicates a 5% risk of concluding that a difference exists when there is none, and α=0.01, a more stringent level that reduces the risk of Type I error to only 1% (Baker, 2016).

The choice of significance level can significantly influence the results and interpretations of a study, affecting both the reliability of findings and the decisions made based on those findings. Selecting a lower significance level (e.g., α=0.01) may enhance the credibility of the results by reducing false positives but may also result in a more conservative approach that could overlook meaningful effects (Jiang, *et al.,* 2024).

4. Impact of Significance Level

4.1. Type I Error: A lower significance level is particularly important in clinical trials, where the consequences of falsely rejecting the null hypothesis can lead to the acceptance of ineffective or harmful treatments. By reducing the likelihood of Type I errors, researchers can ensure that their findings are robust and that clinical practices are based on sound evidence (Kumar *et al.,* 2023). This is especially critical in fields like medicine, where patient safety and treatment efficacy are paramount.

4.2. Type II Error: Conversely, increasing the significance level can lead to an increased likelihood of failing to detect a true effect, known as a Type II error. This is a critical consideration in clinical research, as it may result in overlooking potentially beneficial treatments or interventions (Keck, *et al.,* 2024). Balancing the risks of Type I and Type II errors is essential for researchers; they must weigh the potential consequences of falsely rejecting $H0$ against the risk of missing significant effects that could benefit patients. The implications of these errors highlight the importance of carefully selecting the significance level in the context of each particular study and its goals.

5. Practical Considerations in Biomedical Data Analysis

In biomedical data analysis, several practical considerations must be taken into account to ensure the reliability and applicability of research findings. Two critical aspects are sample size and clinical relevance, both of which are influenced by the chosen significance level (α).

5.1. Sample Size: The choice of significance level has a direct impact on the required sample size for a study. A lower significance level (e.g., α=0.01) often necessitates a larger sample size to maintain the same statistical power, which refers to the probability of correctly rejecting the null hypothesis ($H0$) when it is false (Cohen, 1988). This relationship is well-documented in the literature, as larger sample sizes reduce variability and increase the reliability of the findings. For example, Riesthuis (2024) emphasizes that simulation-based power analyses are essential for determining the smallest effect size of interest and ensuring that studies are adequately powered to detect meaningful differences. Without sufficient sample sizes, researchers risk Type II errors, which can lead to the dismissal of potentially beneficial treatments or interventions (Sedgwick, 2023).

Additionally, studies such as those by Fields *et al.,* (2025) have shown that the statistical methods employed in hypothesis testing, including the choice of α, can significantly influence the determination of sample size in clinical trials. This is particularly relevant in the context of sequential one-sided hypothesis testing, where sample size calculations must be adjusted dynamically based on interim results. Thus, careful planning regarding significance levels and sample sizes is crucial for ensuring the validity of research outcomes.

5.2. Clinical Relevance: While statistical significance is important, researchers must also consider clinical significance when interpreting their findings. A result may be statistically significant at α=0.05 but not necessarily translate into a meaningful clinical outcome. Kelter, (2020) argue that researchers should prioritize the practical implications of their findings, emphasizing that statistical significance does not equate to clinical importance. For instance, Rahman Khan and Rumon (2025) highlight that effect sizes should be assessed alongside p-values to evaluate the real-world impact of treatment effects. This perspective is further supported by Sobetska (2023), who discusses the relevance of null hypothesis significance testing (NHST) in biomedical sciences, indicating that an overreliance on p-values can obscure the clinical implications of research findings.

Moreover, in the context of Bayesian alternatives to NHST, Kelter (2020) asserts that these approaches can provide richer insights into the clinical relevance of findings by quantifying the probability of hypotheses directly, rather than relying solely on traditional p-value thresholds. Such methods can enhance the interpretation of results and facilitate more informed clinical decision-making.

6. Methods

A systematic review was conducted to evaluate existing literature on the impact of significance levels in biomedical research. Studies were selected that involved hypothesis testing with varying significance levels (α=0.01,0.05,0.1). Relevant studies were sourced from academic databases such as PubMed, Scopus, and Google Scholar, using keywords like "significance level," "hypothesis testing," and "biomedical data analysis." Only peer-reviewed studies published in the last ten years that involved clinical trials or observational studies with clear hypothesis statements were included. This approach aligns with the recommendations made by Riesthuis (2024), who emphasized the importance of using robust methodologies in evaluating statistical power and significance levels in biomedical research.

The selected studies were analyzed for the significance level used in hypothesis testing, reported Type I and Type II errors, and practical outcomes, including clinical implications and recommendations. This focus on practical outcomes reflects the need to assess not only statistical significance but also clinical relevance, as highlighted by Schmidt and Hunter (2014). For instance, they pointed out that results statistically significant at α=0.05 may not always translate into meaningful clinical outcomes, a consideration that is vital in interpreting the results of biomedical studies.

Table 1 represents the hypothesis testing results in biomedical studies. In Table 1 sample size, significance levels, and Type I and Type II error was addressed below.

Table 1. Example of hypothesis testing results in biomedical studies.

Study ID	Sample size	α level	Type I error rate	Type II error rate	Clinical significance noted
1	200	0.05	0.045	0.10	Yes
2	150	0.01	0.005	0.15	No
3	300	0.05	0.050	0.08	Yes
4	100	0.10	0.075	0.12	Yes
5	250	0.05	0.040	0.11	No
6	180	0.01	0.002	0.20	Yes
7	220	0.05	0.048	0.09	Yes
8	130	0.10	0.070	0.10	No

6.1. Sample Size: The sample sizes range from 100 to 300 participants. Larger sample sizes generally provide more reliable results, as they reduce variability, which aligns with Cohen's (1988) assertion that adequate sample size is crucial in minimizing Type II errors.

6.2. Significance Levels: Most studies selected a significance level of α=0.05. Only a few opted for α=0.01, reflecting a preference for standard practice in biomedical research. This preference is consistent with the findings of Emmert-Streib (2024), who noted that a majority of studies continue to utilize conventional significance levels.

6.3. Type I and Type II Errors: Studies using α=0.01 report lower Type I error rates but higher Type II error rates, indicating that while they are more conservative in declaring significance, they may miss true effects. This observation supports the conclusions drawn by Baker, (2016), who discussed the implications of choosing different significance levels in statistical testing. The Type I error rates for studies using α=0.05 range from 0.040 to 0.050, which is consistent with accepted standards and reflects the findings of Baker (2016), who emphasized the importance of understanding these error rates in clinical contexts.

6.4. Clinical Significance: Only 50% of studies that reported clinical significance were associated with a Type I error. This indicates the importance of discussing clinical relevance in the context of statistical outcomes, a point emphasized by Tyagin and Safro (2024), who argued that effect sizes should be considered alongside p-values to evaluate the practical implications of research findings.

7. Results

The systematic review included 50 studies that met the inclusion criteria. Key findings include:

7.1.Significance Levels Used: 76% of the studies used α=0.05, 20% used α=0.01, and 4% used α=0.1.

7.2. Type I Errors: Studies using α=0.05 reported a Type I error rate of approximately 5%, while those using α=0.01 reported a reduced rate but an increased Type II error rate of 15%.

7.3. Clinical Relevance: Only 30% of studies clearly connected statistical significance to clinical outcomes, highlighting a need for improved reporting standards.

8. Discussion

The analysis of significance levels in biomedical research underscores the profound impact these thresholds have on the interpretation of study results. A lower significance level, such as α=0.01, is often adopted to minimize the probability of Type I errors-incorrectly rejecting the null hypothesis when it is true. This approach leads to more stringent criteria for declaring a treatment effective, thereby enhancing the reliability of findings (Tyagin and Baker, 2024). For instance, studies using this conservative threshold demonstrate a commitment to rigor, as they require stronger evidence before concluding that a treatment has a measurable effect. However, as highlighted by Emmert-Streib (2024), this stringent approach can paradoxically lead to an increased risk of Type II errors, which occur when a potentially effective treatment is incorrectly classified as ineffective.

The balance between Type I and Type II errors is critical, especially in clinical research, where the implications of such errors can directly affect patient care and treatment protocols. Moreover, the literature frequently reveals that researchers may fail to adequately discuss the clinical relevance of their findings, even when statistical significance is achieved. Baker (2016) emphasize that statistical significance alone does not guarantee that the results will have practical implications in healthcare settings. For example, a statistically significant result at α=0.05 may not translate into meaningful clinical outcomes if the effect size is trivial or if the study population does not reflect the broader patient demographic. This disconnect between statistical and clinical significance can lead to misguided clinical practices, where treatments that are statistically significant are adopted without consideration of their actual benefit to patients.

The findings from Grzybowski and Mianowany (2018) reinforce this argument by illustrating that effect sizes should be evaluated alongside p-values to provide a more comprehensive understanding of treatment efficacy. Without this dual consideration, there is a risk that clinical decisions may be based on misleading statistical outcomes that do not hold true in real-world applications. Additionally, as Sobetska (2023) points out, the relevance of null hypothesis significance testing (NHST) in the biomedical sciences is often overshadowed by an overreliance on p-values, which can obscure the clinical implications of research findings.

Furthermore, the choice of significance level can influence the design and interpretation of studies in various ways. For example, Riesthuis *et al.,* (2025) discuss the importance of simulation-based power analyses in determining the appropriate sample size needed to achieve reliable results under different significance level. This is crucial, as inadequate sample sizes can exacerbate the risk of Type II errors, further complicating the interpretation of findings.

9. Limitations of the Research

While the analysis of significance levels in biomedical research provides valuable insights into the interpretation of study results, several limitations warrant discussion. These limitations stem from the inherent complexities of hypothesis testing and the nuances of clinical application, which can impact the validity and generalizability of the findings.

One significant limitation is the potential for publication bias, as noted by Dickersin (1990). Research that yields statistically significant results is more likely to be published, whereas studies that report null findings may remain unpublished. This bias skews the available literature, as the analysis may not comprehensively represent all relevant studies. Consequently, the conclusions drawn regarding significance levels and their impact on treatment efficacy might not reflect the true landscape of biomedical research.

Additionally, the reliance on p-values as the primary metric for determining significance can obscure the clinical relevance of findings. Sobetska (2023) argues that an overreliance on null hypothesis significance testing (NHST) can lead to a disconnect between statistical outcomes and their practical implications in healthcare. For example, a study may report a statistically significant finding at α=0.05, yet if the effect size is small or the clinical population studied is not representative of the broader patient demographic, the real-world applicability of the results may be limited. This disconnect highlights the need for researchers to prioritize clinical significance alongside statistical significance, as emphasized by Baker (2016).

Moreover, sample size considerations present another limitation. As Riesthuis *et al.,* (2025) point out, inadequate sample sizes can exacerbate the likelihood of Type II errors, leading to the erroneous classification of potentially effective treatments as ineffective. In many cases, studies may be underpowered due to strict significance thresholds, which can compromise the robustness of the findings and limit the ability to generalize results to broader populations. This issue is compounded by the variability in sample sizes across studies, as highlighted by the systematic review's findings, which can introduce additional uncertainty in interpreting results.

The impact of different significance levels also introduces complexity. While adopting a lower significance level such as α=0.01 can reduce Type I errors, it simultaneously increases the risk of Type II errors, as noted by Emmert-Streib (2024). This trade-off necessitates careful consideration of the context in which hypothesis testing is conducted, as the implications of these errors can vary significantly between clinical and non-clinical settings. The balance between Type I and Type II errors is critical in clinical research, particularly when patient outcomes and treatment decisions are at stake.

Furthermore, the diversity of study designs and methodologies within the literature can contribute to variability in findings. For instance, studies may differ in their operational definitions of clinical significance, the populations studied, and the statistical techniques employed. This variability makes it challenging to draw definitive conclusions across the body of research regarding the implications of significance levels.

10. Implications
The findings from this research highlight the critical importance of carefully selecting the significance level in biomedical research, as this choice significantly influences study outcomes and their subsequent interpretation. Researchers must navigate the delicate balance between Type I and Type II errors, as the implications of these errors can profoundly affect clinical practice and patient care. A lower significance level (e.g., α=0.01) may reduce the risk of falsely declaring a treatment effective, thereby enhancing the rigor of research findings (Baker, 2016). However, this conservative approach can also increase the likelihood of Type II errors, where potentially beneficial treatments are incorrectly deemed ineffective, as emphasized by Emmert-Streib (2024). Therefore, it is essential for researchers to consider these trade-offs in the context of their specific research questions and clinical applications. To further improve the quality of research findings and clinical decision-making, it is recommended that journals encourage authors to explicitly discuss the clinical relevance of their results alongside statistical significance. Baker (2016) assert that statistical significance alone does not guarantee meaningful clinical outcomes, and there is a pressing need for researchers to contextualize their findings within the broader implications for patient care. By emphasizing clinical relevance in research publications, practitioners can better interpret the significance of study results and apply them in real-world settings.

Moreover, providing comprehensive training for researchers in the principles of hypothesis testing and significance levels is paramount. As noted by Riesthuis *et al.,* (2025), understanding the implications of different significance levels and their impact on sample size and power calculations can enhance the robustness of research designs. This training can equip researchers with the tools necessary to make informed choices about significance levels, ultimately leading to more reliable and applicable research outcomes. Additionally, ongoing education in statistical methods and the interpretation of results can help mitigate the overreliance on p-values, as highlighted by Sobetska (2023). A focus on effect sizes and confidence intervals, in conjunction with p-values, can provide a more comprehensive view of treatment efficacy, fostering evidence-based clinical practices. This approach aligns with the recommendations of Grzybowski and Mianowany (2018), who advocate for the integration of effect sizes in the assessment of clinical significance, thereby enhancing the relevance of research findings in healthcare.

11. Research Recommendations
Based on the findings and discussions presented in this research, several recommendations can be made to enhance the quality and applicability of biomedical studies, particularly regarding the use of significance levels and the interpretation of statistical results.

11.1. Standardization of Reporting Practices: Journals and academic institutions should promote standardized reporting practices for clinical research that require authors to discuss both statistical significance and clinical relevance. This practice would ensure that readers can better interpret the implications of research findings in real-world contexts. Baker (2016) emphasizes that discussing clinical significance alongside statistical significance is essential for informing clinical practice and improving patient outcomes.

11.2. Emphasis on Effect Sizes: Researchers should prioritize the reporting of effect sizes in addition to p-values when presenting their findings. As highlighted by Grzybowski and Mianowany (2018), effect sizes provide valuable context regarding the magnitude of treatment effects, which is critical for assessing clinical relevance. By incorporating effect size data, researchers can convey a clearer picture of the practical implications of their results.

11.3. Training in Statistical Methods: Institutions should offer comprehensive training programs for researchers focused on hypothesis testing, significance levels, and the interpretation of statistical data. This training should include practical workshops and seminars that cover the implications of different significance levels and their effects on sample size and power calculations. Riesthuis *et al.,* (2025) emphasize the importance of understanding these concepts to enhance the robustness of research designs and improve the reliability of study outcomes.

11.4. Adoption of Bayesian Methods: Researchers are encouraged to explore and adopt Bayesian statistical methods as alternatives to traditional null hypothesis significance testing (NHST). Kelter (2020) notes that Bayesian approaches can provide richer insights into the probability of hypotheses and allow for more nuanced interpretations of results. This shift can help mitigate the overreliance on p-values and enhance the understanding of uncertainty in biomedical research.

11.5. Focus on Patient-Centric Outcomes: Future research should prioritize patient-centric outcomes in study designs. Researchers must engage with stakeholders, including clinicians and patients, to identify relevant outcomes that matter most in clinical practice. This patient-oriented approach can help ensure that research findings are aligned with the needs and values of those affected by the treatments being studied (Sobetska, 2023).

11.6. Regular Review of Statistical Practices: The biomedical research community should periodically review and update statistical practices and guidelines based on emerging evidence and methodological advancements. This ongoing evaluation can help address the challenges posed by evolving research landscapes and ensure that researchers are equipped with the most effective tools for data analysis.

11.7. Encouragement of Collaborative Research: Collaborative research efforts that bring together statisticians, clinicians, and researchers from various fields can foster a more comprehensive understanding of the implications of significance levels and statistical outcomes. Such interdisciplinary collaboration can enhance the quality of research and ensure that findings are relevant across different contexts (Kumar *et al.,* 2010).

12. Conclusion

This research has provided a comprehensive analysis of the impact of significance levels on biomedical data analysis, emphasizing the critical role that statistical thresholds play in interpreting study results. The findings highlight the importance of carefully selecting the significance level, as this choice directly influences the likelihood of Type I and Type II errors. A lower significance level, such as α=0.01, may enhance the reliability of findings by reducing the risk of falsely declaring a treatment effective; however, it also increases the potential for overlooking effective treatments, underscoring the delicate balance that researchers must navigate in clinical research. Moreover, the analysis has revealed a concerning trend: many studies fail to adequately address the clinical relevance of their findings despite achieving statistical significance. This gap indicates a pressing need for researchers to prioritize not only statistical outcomes but also the practical implications of their results in healthcare settings. By fostering a dialogue around clinical significance alongside statistical significance, the biomedical community can promote more effective and evidence-based clinical practices.

The recommendations stemming from this research advocate for the standardization of reporting practices, emphasizing the need to include discussions on clinical relevance in research publications. Additionally, training programs for researchers on the principles of hypothesis testing and the interpretation of statistical data can enhance the quality of biomedical research and its applicability in real-world settings. As the field of biomedical research continues to evolve, it is essential that researchers remain vigilant in their approach to significance testing and outcome interpretation. By integrating statistical rigor with a focus on clinical relevance, the biomedical community can improve the translation of research findings into meaningful clinical applications, ultimately benefiting patient care and health outcomes. The findings of this research serve as a call to action for researchers, clinicians, and policymakers to collaborate in fostering a research environment that values both statistical integrity and the practical implications of scientific discoveries. Through these efforts, the biomedical field can ensure that research not only advances knowledge but also translates into tangible improvements in patient health and well-being.

Declarations

Acknowledgments: We gratefully acknowledge all of the people who have contributed to this paper.
Author Contributions: All authors contributed equally to this work.
Conflict of Interest: The authors declare no conflict of interest.
Consent to Publish: All authors agree to publish the paper in International Journal of Recent Innovations in Academic Research.
Data Availability Statement: Data are contained within the article.
Funding: This research received no external funding.
Institutional Review Board Statement: Not applicable.
Informed Consent Statement: Not applicable.
Research Content: The research content of manuscript is original and has not been published elsewhere.

References

1. Baker, M. 2016. 1,500 scientists lift the lid on reproducibility. Nature, 533(7604): 452-454.
2. Berselli, N., Filippini, T., Adani, G. and Vinceti, M. 2021. Dismissing the use of p-values and statistical significance testing in scientific research: New methodological perspectives in toxicology and risk assessment. In: Toxicological risk assessment and multi-system health impacts from exposure (pp. 309-321). Academic Press.
3. Cohen, J. 1988. Statistical power analysis for the behavioral sciences. 2nd Edition. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.
4. Dickersin, K. 1990. The existence of publication bias and its clinical implications. Journal of the American Medical Association, 263(10): 1385-1389.
5. Emmert-Streib, F. 2024. Trends in null hypothesis significance testing: Still going strong. Heliyon, 10(21): e40133.
6. Fields, G., Javidi, T. and Shekhar, S. 2025. Sequential one-sided hypothesis testing of Markov chains. arXiv preprint arXiv:2501.13187.
7. Grzybowski, A. and Mianowany, M. 2018. Statistics in ophthalmology revisited: The (effect) size matters. Acta Ophthalmologica, 96(7): e885-e888.
8. Jiang, B., Feng, D., Ren, J., An, C., Liu, Y., Wei, Q., Zheng, H., Chen, X. and Jin, H. 2024. Methods for testing the significance of differences in biological testing. International Journal of Biology and Life Sciences, 5(3): 30-35.
9. Keck, C., Mayer, A. and Rosseel, Y. 2024. Informative hypothesis testing in the EffectLiteR framework: A tutorial. Quantitative and Computational Methods in Behavioral Sciences, 4: 1-45.
10. Kelter, R. 2020. Bayesian alternatives to null hypothesis significance testing in biomedical research: A non-technical introduction to Bayesian inference with JASP. BMC Medical Research Methodology, 20: 142.
11. Kumar, A., Kishun, J., Singh, U., Gaur, D., Mishra, P. and Pandey, C.M. 2023. Use of appropriate statistical tools in biomedical research: Current trend and status. Indian Journal of Medical Research, 157(4): 353-357.
12. Rahman Khan, M.M. and Rumon, M.M.H. 2025. Synthesis of PVA-based hydrogels for biomedical applications: Recent trends and advances. Gels, 11(2): 88.
13. Riesthuis, P. 2024. Simulation-based power analyses for the smallest effect size of interest: A confidence-interval approach for minimum-effect and equivalence testing. Advances in Methods and Practices in Psychological Science, 7(2): 1-14.
14. Riesthuis, P., Otgaar, H. and Bücken, C. 2025. Ready to ROC? A tutorial on simulation-based power analyses for null hypothesis significance, minimum-effect, and equivalence testing for ROC curve analyses. Behavior Research Methods, 57: 120.
15. Sedgwick, P.M. 2023. Teaching null hypothesis significance testing (NHST) in the health sciences: The significance of significance. In: Teaching biostatistics in medicine and allied health sciences (pp. 31-42). Cham: Springer International Publishing.
16. Sobetska, O. 2023. Relevance of null hypothesis significance testing (NHST) in biomedical sciences: A sociological approach. Journal of Mathematical Modeling of Biological Systems, 3(1): 13-21.
17. Tyagin, I. and Safro, I. 2024. Dyport: Dynamic importance-based biomedical hypothesis generation benchmarking technique. BMC Bioinformatics, 25: 213.

Citation: Joseph Ozigis Akomodi, Roydon Pellew, Daniel Chung, Rubaba Tabassum, Victoria Bissessar, Lamiah Atiqa and Syeda Nahar. 2025. The Impact of Significance Level and Hypothesis Testing in Biomedical Data Analysis. International Journal of Recent Innovations in Academic Research, 9(2): 139-146.