ISSN: 3048-8451 (ONLINE)

INTERNATIONAL RESEARCH JOURNAL OF SCIENTIFIC STUDIES

DECEMBER 2025
VOLUME 2, ISSUE 10

https://www.irjss.com

irjss.editor@gmail.com

Published by Sciencia Publishers

International Research Journal of Scientific Studies
e-ISSN: 3048-8451

RESEARCH ARTICLE DECEMBER 2025, VOLUME 2, ISSUE 10

Analytical Validation of NovoNumeric: An Indigenous, High-Performance Statistical Software for Indian Medical Research

Mohammed Jaseem Ibrahim K.[1]

[1] NovoNumeric Software, Thiruvananthapuram, Kerala, India

Corresponding author email: **drjaseemibrahim@gmail.com**

Cite this article: Mohammed Jaseem Ibrahim K. (2025). Analytical Validation of NovoNumeric: An Indigenous, High-Performance Statistical Software for Indian Medical Research. International Research Journal of Scientific Studies, 2(10), 1–7. **https://doi.org/10.64383/irjss.202501227**

Abstract: *Statistical analysis is a fundamental pillar of medical research, yet many powerful software options come with a steep learning curve or prohibitive costs. NovoNumeric is a novel statistical software for macOS designed to address this by balancing accessibility with high performance. This study provides the first formal analytical validation of NovoNumeric by comparing its analytical output against established gold-standard packages. A comparative validation study was performed using an anonymized clinical dataset (n=174) from a prospective observational study on diabetic foot infections, conducted at Sree Uthradom Thirunal Academy of Medical Sciences, Thiruvananthapuram (Ethical Approval No. 22/IEC/SUTAMS/2023). Parallel analyses—including descriptive statistics, independent t-tests, chi-square tests, Pearson's correlation, and multiple linear regression—were conducted using NovoNumeric (v4.0), R (v4.3.1), and SPSS (v28). The primary outcome was the equivalency of numerical results. For every statistical test performed, NovoNumeric produced identical outputs to R and SPSS, with maximum absolute differences < 1 x 10^{-8}. These findings confirm that NovoNumeric is a valid, reliable, and high-performance tool for core statistical analyses in medical research, demonstrating its suitability as a cost-effective solution for Indian academic and clinical research institutions.*

Keywords: *Data Analysis, Health Technology, Medical Research Software, Statistical Software, Analytical Validation.*

1. Introduction

Statistical analysis is a fundamental pillar of evidence-based medicine, yet a significant gap persists between the analytical needs of medical researchers and the tools available to them. While code-based environments like R [1] and Python offer immense power, they present a steep learning curve for clinicians and health professionals who are not specialist statisticians. Conversely, traditional GUI-based packages [2, 3] can be prohibitively expensive. In Low- and Middle-Income Countries (LMICs), the high licensing costs of proprietary software—often exceeding hundreds of dollars annually per user—pose a significant barrier to independent research and institutional adoption.

This proficiency and accessibility gap creates a bottleneck, potentially slowing research or leading to the misapplication of statistical methods. NovoNumeric (v4.0), a novel statistical software developed natively for macOS, was designed to address this specific challenge. Its core design philosophy is to provide a high-performance analytical engine, leveraging Apple's Accelerate and LAPACK frameworks [4], within an intuitive, user-friendly, and affordable interface. A previous technical paper described the software's architecture and features [5]; however, it has not yet undergone formal, independent validation using real-world clinical data.

The objective of this study was to provide the first formal analytical validation of NovoNumeric. It aimed to determine if NovoNumeric produces statistically identical results to gold-standard, internationally accepted software (R and SPSS) when analyzing a complex, real-world clinical dataset from a tertiary care hospital.

2. Methods

2.1 Study Design and Data Source

A comparative validation study was performed. The data source was a pre-existing, anonymized clinical dataset (n=174) from a prospective observational study conducted in the Department of Pharmacology, Sree Uthradom Thirunal Academy of Medical Sciences (SUTAMS), Thiruvananthapuram. The original dissertation study, titled "Antimicrobial therapy, glycaemic management and the outcomes of diabetic foot infections at a tertiary care hospital in Kerala," was approved by the Institutional Ethics Committee (No. 22/IEC/SUTAMS/2023, dated 23.03.2023). This manuscript represents a secondary analysis of this

approved data. The dataset included a mix of continuous (e.g., HbA1c, age, duration of diabetes) and categorical (e.g., gender, organism type, amputation status) variables. The dataset contained no missing values for the selected variables; thus, complete case analysis was utilized.

2.2 Comparative Analysis Protocol

A parallel analysis protocol was established. The same dataset was loaded into three separate statistical packages: NovoNumeric (v4.0), R (v4.3.1) via RStudio [6,7], and SPSS (v28.0) [2]. The same set of statistical tests was performed in all three programs, targeting common analyses in medical research. The tests included:

- Descriptive Statistics: Mean, Standard Deviation (SD), and N for continuous variables.
- Frequency Analysis: Counts and percentages for categorical variables.
- Bivariate Analysis (Categorical): Chi-square test (e.g., Organism Type vs. Amputation Status).
- Bivariate Analysis (Continuous): Pearson's correlation (e.g., Age vs. HbA1c).
- Group Comparison: Independent samples t-test (e.g., HbA1c between amputation vs. no amputation groups).
- Multivariate Modelling: Multiple linear regression to model Systolic Blood Pressure (dependent variable) using Age, HbA1c, and Duration of Diabetes as independent predictors.

2.3 Independent Verification and Reproducibility

To ensure objective verification, the analysis followed a strict validation protocol. Outputs were generated independently in each software environment. For R, script-based execution was used to ensure reproducibility (script available in Supplementary Material S1). For NovoNumeric and SPSS, detailed command logs and output exports were archived.

Numerical outputs (coefficients, p-values, standard errors, and confidence intervals) were compared across platforms. Floating-point equality was assessed with a maximum absolute difference tolerance of 1×10^{-8}, adhering to IEEE 754 standards for double-precision floating-point arithmetic. No random seeding was required as bootstrapping methods were not employed in this validation set; all algorithms were deterministic.

2.4 Outcome Measures

The primary outcome was the analytical equivalency of results between NovoNumeric and the gold-standard

packages. All statistical outputs were compared to at least eight decimal places. Statistical significance was set at $p<0.05$.

3. Results

The validation was successful across all tested parameters. NovoNumeric (v4.0) produced numerical outputs that were identical to those generated by both R (v4.3.1) and SPSS (v28.0) for all analyses performed. The maximum absolute difference between NovoNumeric outputs and the reference standards was 1×10^{-8}.

For descriptive statistics, the mean HbA1c was identical (9.74 ± 1.52 %), as was the mean age (59.6 ± 10.1 years) across all three platforms. In frequency analysis, all counts and percentages matched precisely. For the chi-square test comparing organism type (Gram-positive vs. Gram-negative) against amputation status (Yes/No), all three packages produced an identical chi-square value of 0.418 and a p-value of 0.518.

The independent t-test comparing mean HbA1c between the amputation group (n=48) and the no-amputation group (n=126) yielded an identical t-statistic of 2.195 and a p-value of 0.030 in all three programs. Similarly, Pearson's correlation between age and HbA1c was identical ($r = 0.171$, $p = 0.024$).

The most rigorous test, multiple linear regression, also showed perfect concordance. As shown in **Table 1**, the coefficients, standard errors, and p-values for all predictors (Age, HbA1c, Duration of Diabetes) and the intercept were identical across NovoNumeric, R, and SPSS.

Table 1. Comparison of Multiple Linear Regression Outputs for Modelling Systolic BP (n=174)

Variable	Package	Coefficient (B)	Std. Error	p-value
(Intercept)	NovoNumeric 4.0	97.415	4.882	<0.001
	R 4.3.1	97.415	4.882	<0.001
	SPSS 28.0	97.415	4.882	<0.001
HbA1c (%)	NovoNumeric 4.0	1.833	0.501	<0.001
	R 4.3.1	1.833	0.501	<0.001
	SPSS 28.0	1.833	0.501	<0.001
Age (years)	NovoNumeric 4.0	0.179	0.052	0.001
	R 4.3.1	0.179	0.052	0.001
	SPSS 28.0	0.179	0.052	0.001
Duration (years)	NovoNumeric 4.0	0.081	0.070	0.248
	R 4.3.1	0.081	0.070	0.248
	SPSS 28.0	0.081	0.070	0.248

Note: All results are identical. Model Summary: $R^2 = 0.291$, *$F(3, 170) = 23.26$, $p < 0.001$ (Values identical across packages). 95% CIs were also identical: (Intercept) [87.781 – 107.049]; HbA1c [0.844 – 2.822]; Age [0.077 – 0.281]; Duration of Diabetes [-0.057 – 0.219].*

4. Discussion

This study provides the first formal analytical validation for NovoNumeric. The key finding is that NovoNumeric's analytical engine is robust, accurate, and reliable. Its outputs are numerically identical to established, globally trusted statistical software, adhering to standard double-precision accuracy.

The importance of this finding is twofold. First, it provides confidence to researchers that they can use this tool for their own studies, including for publication in peer-reviewed journals. The software correctly implements standard statistical algorithms as described in foundational texts [8,9]. Second, it demonstrates the viability of a low-cost health technology innovation. NovoNumeric provides an alternative that is both financially accessible and, as this study shows, statistically sound.

Furthermore, NovoNumeric is optimized for Apple Silicon (M-series) chips. While formal benchmarking was not the primary objective of this study, the software leveraged the native Accelerate framework to deliver immediate execution for all tested datasets, comparable to the highly optimized performance of R in terminal environments.

4.1 Limitations

It is important to acknowledge the scope of this validation. This study validated six core statistical functions: descriptive statistics, frequencies, chi-square, t-tests, correlation, and multiple linear regression. Advanced analytical methods such as logistic regression, survival analysis (Kaplan-Meier, Cox Proportional Hazards), and non-parametric equivalents (e.g., Mann-Whitney U) were not included in this phase. Consequently, these findings represent a foundational validation of the software's core engine, and future studies are required to validate these expanded modules.

5. Software Availability

- ✓ Software Homepage: www.novonumeric.com
- ✓ System Requirements: macOS 12.0 (Monterey) or later; optimized for Apple Silicon.
- ✓ License: Commercial

6. Conclusion

NovoNumeric (v4.0) is an analytically validated and reliable statistical software. Its analytical outputs for core statistical functions are identical to those of R and SPSS. This validation supports its adoption by the medical research community as a cost-effective,

user-friendly, and accurate tool for advancing clinical and academic research.

Conflict of Interest

The author is the developer of NovoNumeric software. This validation study was designed to be an objective comparison against external gold-standard software to ensure transparency.

References:

[1] R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing; 2023. Available from: https://www.R-project.org/

[2] IBM Corp. IBM SPSS Statistics for Windows, Version 28.0. IBM Corp; 2021.

[3] StataCorp. Stata Statistical Software: Release 17. StataCorp LLC; 2021.

[4] Apple Inc. Accelerate. Apple Developer Documentation. 2023. Available from: https://developer.apple.com/documentation/accelerate

[5] Ibrahim MKJ. NovoNumeric: An Accessible, High-Performance Statistical Software for Advancing Medical Research on macOS. International Research Journal of Scientific Studies. 2025; 2(8): 1-9. https://doi.org/10.64383/irjss.20250924

[6] Posit team. RStudio: Integrated Development Environment for R. Posit; 2023. Available from: http://www.posit.co

[7] Wickham H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag; 2016.

[8] Motulsky H. Intuitive biostatistics: a nonmathematical guide to statistical thinking. 4th ed. Oxford University Press; 2018.

[9] Field A. Discovering statistics using IBM SPSS statistics. 5th ed. Sage publications; 2018.

[10] International Committee of Medical Journal Editors (ICMJE). Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals. 2023. Available from: https://www.icmje.org

Author(s) Profile:

1. Dr. Mohammed Jaseem Ibrahim K. is a medical, with a keen interest in research and a passion for technology, his work focuses on bridging the gap between complex statistical methods and accessible user interfaces. He is the creator of NovoNumeric, a native macOS application designed to make advanced statistical analysis intuitive for researchers and students. His interests include computational statistics, native software development,

and improving analytical workflows in the scientific community.

Article History:

Submitted: 2025-11-09
Reviewed: 2025-11-24
Accepted: 2025-12-08