DEPARTMENT OF STATISTICS
AHMADU BELLO UNIVERSITY, ZARIA
2nd International Conference on Statistical Sciences and Data Analytics (ICSSDA 2026) Proceedings Paper

Stochastic Simulation of a Clinical Trial with a Single Endpoint

I. Adamu[1*], U. C. Destiny[2], M. Usman [1], I. A. Sadiq[1], A. U. Goga[3] & Osi A. A.[4]

[1]Department of Statistics, Ahmadu Bello University, Zaria, Nigeria

[2]Department of Statistics, Federal University of Technology Owerri, Imo State, Nigeria.

[3]Department of Statistics, Federal University Dutsin-ma, Katsina State, Nigeria

[4]Department of Statistics, Aliko Dangote University of Science and Technology, Wudil, Kano State, Nigeria

*Correspondence: iadandume@gmail.com, ibadamu@abu.edu.ng

Abstract

Clinical trials are essential for evaluating the safety and efficacy of medical interventions, yet their design often faces challenges of uncertainty, limited resources, and ethical constraints. This study investigates the application of stochastic simulation to the design and analysis of a clinical trial with a single primary endpoint, focusing on pain relief treatments. The aim was to assess whether simulation methods could provide a more robust framework for trial planning compared to conventional approaches. Data were drawn from a randomised controlled trial involving four topical treatments, with the primary endpoint defined as binary responder status, representing a ≥30% reduction in pain score. Conventional analyses, including descriptive statistics, chi-square tests, logistic regression, and ANCOVA, established baseline measures of treatment efficacy. To extend this analysis, a Monte Carlo simulation with 10,000 replications was conducted to model trial behaviour under probabilistic conditions and estimate risk differences, statistical power, and confidence interval coverage. Sensitivity analysis further explored the impact of varying sample size, treatment effect, dropout rate, and placebo response probability. Results identified Lidocaine Patch and Diclofenac Gel as highly effective, while Capsaicin Cream showed limited efficacy, with only 35% power to detect its effect. Importantly, simulation revealed that placebo response rate was the most critical determinant of trial success, a risk not evident from conventional methods. The study concludes that stochastic simulation offers a probabilistic framework for optimising design parameters, quantifying uncertainty, and improving the ethical and economic efficiency of clinical research.

Keywords: Stochastic Simulation; Clinical Trials; Single Primary Endpoint; Monte Carlo Simulation; Sensitivity Analysis; Statistical Power.

1. Introduction

A stochastic process is a mathematical framework that models systems evolving with inherent randomness. Formally, it is defined as a collection of random variables $\{X_t\}$, $t \in T$, where each X_t rep-

resents the state of the system at time t within an index set T. Stochastic simulation is a computational method that employs random sampling techniques to model and analyse complex systems or processes influenced by inherent uncertainty. Unlike deterministic models that produce the same output for a given input, stochastic simulations incorporate variability by treating certain inputs, parameters, or processes as random variables governed by probability distributions. This allows for the representation of a wide range of possible outcomes, making the technique especially valuable in fields such as healthcare, finance, engineering, and environmental science where uncertainty is a defining characteristic. By simulating multiple iterations or scenarios, researchers and decision-makers can gain insights into the likely behaviour of a system, assess risks, estimate probabilities of different outcomes, and make informed decisions under uncertain conditions (Ross, 2014).

Clinical trials are a systematic and controlled type of medical research designed to evaluate the safety, effectiveness, and overall impact of new diagnostic methods, treatments, drugs, or medical devices on human health outcomes, often conducted in multiple phases and involving carefully selected participant groups to ensure scientifically valid and ethically sound results that can be used to improve clinical practice and healthcare policies (Friedman, Furberg, & DeMets, 2010). In trials with a single endpoint, accurately capturing the variability in patient responses and treatment outcomes is critical for evaluating the efficiency of treatments. However, conventional trial designs often struggle to account for this uncertainty, leading to underpowered studies or inconclusive results.

There is a pressing need for cost-effective and ethically sound approaches that can model clinical outcomes under realistic assumptions without exposing human subjects to unnecessary risk. This study addresses this gap by exploring the application of stochastic simulation to clinical trials with a single endpoint, aiming to enhance the reliability and efficiency of trial design and analysis (Law, 2015).

Rubinstein *et al*. (2017) demonstrated the powerful synergy between stochastic simulation and optimisation. They advanced techniques like the Cross-Entropy method and scoring algorithms, which use smart random sampling to efficiently solve complex combinatorial and continuous optimisation problems. Their work showed that simulation is not just for analysis but can be directed to find optimal configurations and decisions in high-dimensional, noisy environments.

Glasserman (2004) specialised in the application of Monte Carlo methods in the domain of finance. His work is pivotal for measuring and managing financial risk, providing detailed algorithms for pricing complex derivatives and estimating portfolio Value-at-Risk (VaR). Piantadosi (2017) offers a comprehensive methodological perspective on clinical trials, positioning them as the gold standard for evaluating medical interventions. He meticulously details the principles of randomisation, blinding, and bias control, and explains how statistical design from sample size calculation to analysis plans is integral to producing valid, reliable, and ethically sound evidence for clinical practice.

Berry *et al*. (2009) were pioneers in advocating for Bayesian adaptive methods in clinical trials. Their work demonstrates how prior knowledge can be formally incorporated and updated with incoming trial data, allowing designs that can learn from the ongoing experiment. This paradigm often leads to more informative trials, smaller sample sizes, and a higher probability of success compared to traditional fixed designs.

Wang et al. (2016) reviewed the expanding applications of simulation across the clinical research landscape. They documented its utility in areas beyond power analysis, such as modelling patient dropout patterns, predicting recruitment rates, and conducting risk assessments for complex

trial protocols, establishing it as a versatile tool for strategic planning. Friedman et al. (2010) provided the canonical textbook foundation for clinical trials, emphasising rigorous design principles. They underscore that in trials with a single endpoint, the entire study's validity hinges on the accurate and unbiased measurement of that outcome, making meticulous planning and execution paramount.

Lewis *et al*. (2014) provided a clinician's perspective on Bayesian adaptive trials, confirming their real-world value in terms of flexibility and efficiency. However, they also highlighted practical challenges, including the significant computational demands and the need for specialised expertise, which can be barriers to widespread implementation. He also issued a critical warning regarding the limitations of simulation. They noted that all models are simplifications of reality and that their outputs are only as valid as their inputs. If the underlying assumptions about patient variability, treatment effect, or disease progression are incorrect, the simulation can produce overly optimistic or misleading results, creating a "garbage in, garbage out" scenario. However, critics have pointed out limitations. Lewis *et al*. (2014) warned that simulations may oversimplify real-world complexities, particularly when model assumptions are misaligned with actual patient variability. Nevertheless, most authors agree that simulation enhances trial efficiency, particularly in single-endpoint designs where precision is critical.

Despite these concerns, the literature consistently demonstrates that stochastic simulation can improve trial design by providing probabilistic insights into outcomes, risks, and uncertainties. This is particularly useful in the context of single-endpoint trials, where precision and reliability are paramount.

Although the usefulness of stochastic simulation in clinical research is well established, few studies have examined its specific application to clinical trials with a single primary endpoint. Existing literature often emphasises adaptive or multi-endpoint designs, leaving a gap in comparative evaluations of stochastic simulation against conventional methods when only one critical outcome is measured. Addressing this gap is essential for improving the efficiency, reliability, and ethical soundness of single-endpoint trials.

The next section outlines the methodology employed in this study to investigate these issues and achieve the stated objectives.

2. Materials and Methods

The methodology is structured into three main sections: Data Sources, Method of Analysis and Performance Comparison. R, a widely accessible programming language, is employed to carry out analysis.

2.1. Data

The dataset used in this study was the anonymised clinical trial file *Pain_Relief_Data.xlsx*, forming the foundation for all analyses and simulations. It came from a randomised controlled trial evaluating Diclofenac Gel, Capsaicin Cream, Lidocaine Patch, and Placebo for pain relief. The dataset included patient demographics, baseline and post-treatment pain scores, treatment frequency, and duration, collected under standardised procedures to ensure accuracy and consistency, while being fully anonymised for safe secondary analysis.

2.2. Data Features

The dataset comprises 30 patients randomised across four treatment arms: Diclofenac Gel, Capsaicin Cream, Lidocaine Patch, and Placebo, with 7–8 patients per group. Patients' ages range from 26 to 55 years, with an approximately balanced distribution of males and females across treatments. Baseline pain scores range from 6 to 9 on a 0–10 scale, while post-treatment pain scores show reductions in active treatment groups, indicating varying degrees of efficacy. Application frequency ranges from 0 (for placebo) to 3 times per day, and treatment durations vary from 5 to 10 days. These features provide sufficient variation to analyse both binary responder status, which is defined as a ≥30% reduction from baseline pain. The dataset is complete, with no missing values, enabling rigorous baseline analysis and stochastic simulation.

2.3. Method of Analysis

This section describes the statistical and computational approaches used to evaluate treatment effects in the study.

2.4. Clinical Trial Endpoint Specification

The single primary endpoint for this clinical trial is binary responder status, defined by a clinically meaningful reduction in pain. A patient is classified as a responder if their post-treatment pain score shows at least a 30% reduction from baseline, a well-known threshold for the Minimally Clinically Important Difference (MCID) in pain research (Dworkin *et al*., 2008) and is assigned a value of 1; all other patients are classified as non-responders and assigned a value of 0. This dichotomous outcome provides a clear, interpretable, and clinically validated measure of treatment success, forming the basis for all subsequent statistical analyses and stochastic simulations.

2.5. Descriptive Statistics

Descriptive statistics summarise the key characteristics of the study population and treatment groups. Continuous variables, including age, baseline and post-treatment pain scores, application frequency, and treatment duration, are presented as means and standard deviations. Categorical variables, such as sex and treatment assignment, are summarised as counts and percentages. These summaries provide a foundation for understanding the dataset and for guiding subsequent inferential and simulation analyses.

2.6. Data Visualisation

To complement the numerical summaries, graphical representations will be used to highlight key patterns and differences across treatment groups:

Bar Chart of Mean Pain Scores. Bar Chart of Responder Rates. Scatter Plot of Treatment Duration vs Pain Reduction.

These visualisations enhance the interpretability of the data and provide a clear overview before formal statistical modelling.

2.7. Inferential Statistics

Inferential statistical methods are used to evaluate treatment effects and determine whether observed differences between groups are statistically significant. Both the binary primary endpoint (Responder status) and continuous secondary outcomes (pain reduction) are analysed using appropriate models.

2.8. Analysis of Binary Endpoint: Logistic Regression

The primary endpoint, binary responder status, is analysed using logistic regression to estimate the odds of treatment response. Given the small sample size (n = 30) and the presence of complete separation in the response data (where some treatment groups achieved 100% response rates), a simplified and robust modelling approach is employed to ensure stable parameter estimates and avoid computational issues associated with overparameterization. The model is specified as:

$$Logit(P_i) = ln[P_i / (1 - P_i)] = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots + \beta_n X_n + \varepsilon \quad (1)$$

where P_i is the probability that patient i is a responder, β_0 = constant (intercept), $\beta_1 \ldots \beta_n$ = represent the coefficients for the treatment effects, indicating how each active treatment differs from the placebo reference, $X_1 \ldots X_n$ = independent variables, and ε = stochastic error term. This model provides stable odds ratios for each treatment versus placebo. Treatment groups use reference coding with Placebo as baseline. Results include odds ratios, 95% confidence intervals, and p-values.

2.9. Analysis of Continuous Endpoints: ANCOVA

Continuous outcomes, including absolute and percentage pain reduction, are analysed using Analysis of Covariance (ANCOVA) to account for baseline pain scores. The model is:

$$Postpain_i = \alpha_0 + \alpha_1 + \alpha_2 + \varepsilon_i \quad (2)$$

where: *PostPaini* is the post-treatment pain score for patient i; α_0 is the intercept; α_1 and α_2 are coefficients for treatment and baseline pain, respectively; and ε_i is the residual error term.

This model measures the treatment effect while accounting for baseline pain scores, enhancing both the accuracy and clarity of the results.

2.10. Group Comparisons for Categorical Data

The preliminary analysis applied Chi-square or Fisher's exact tests depending on cell counts, offering a quick first look at responder differences without covariate adjustment. These tests helped identify clear group imbalances in the Pain Relief dataset before moving to advanced methods. Inferential statistics combined logistic regression, ANCOVA, and categorical tests to evaluate treatment effects, with all analyses conducted in R for accuracy and reproducibility.

2.11. Monte Carlo Simulation

The Monte Carlo simulation was used to quantify uncertainty in clinical trial outcomes by repeatedly simulating trials under probabilistic conditions, generating distributions of risk differences, confidence intervals, and p-values. Implemented in R, it followed the steps of parameter selection, analysis, output calculation, and aggregation, estimating power, type I error, and treatment-effect precision across thousands of replicates. Sensitivity analysis varied sample size, effect size, dropout rates, and placebo response probability to identify the most influential parameters. Overall, the approach assessed reliability, stability, and generalizability, supporting improved trial design under uncertainty.

2.12. Monte Carlo Simulation

This section evaluates how well conventional statistical methods and stochastic simulation perform in analysing the trial data. Comparing these approaches provides insight into their reliability, efficiency, and usefulness for single-endpoint clinical trials.

3. Results and Discussions

This section presents the results of the analysis conducted on clinical trials using both conventional statistical methods and stochastic simulation techniques. The comparison of these two approaches provides insight into their reliability, efficiency, and overall sustainability for evaluating single-endpoint clinical trials.

3.1. Data Presentation

The data set consists of 30 patients enrolled in a randomised clinical trial assessing the efficacy of four treatments for pain relief: Diclofenac Gel, Capsaicin Cream, Lidocaine Patch, and Placebo. Each patient record includes: Patient ID as a unique identifier for each participant, Age, Sex, Treatment (assigned intervention group), Baseline Pain Score, Post-Treatment Pain Score: Pain intensity after treatment, Application Frequency, and Duration of Treatment.

3.2. Summary Statistics

The descriptive summary of the study population and key variables provides essential context for subsequent inferential and simulation analyses. The dataset consists of 30 patients randomised across four treatment groups: Diclofenac Gel (n=8), Capsaicin Cream (n=7), Lidocaine Patch (n=8), and Placebo (n=7). Continuous variables are reported as means and standard deviations, and categorical variables as counts and percentages (Table 4.1). The Lidocaine Patch group showed the highest response (100%), followed by Diclofenac Gel (62.5%), Capsaicin Cream (28.6%), and Placebo (0%). Gender was unevenly distributed, with males in Diclofenac Gel and Lidocaine Patch, and females in Capsaicin Cream and Placebo. Baseline pain scores were consistent, while post-treatment scores varied more in active treatment groups.

Table 1. Summary Table

Variable	Diclofenac Gel (n=8)	Capsaicin_Cream (n=7)	Lidocaine Patch (n=8)	Placebo (n=7)
Age (years), Mean (SD)	36.0 (4.9)	42.9 (5.3)	29.5 (3.5)	50.7 (2.4)
Sex, n (%)				
Male	6 (75.0)	0 (0.0)	6 (75.0)	0 (0.0)
Female	2 (25.0)	7 (100.0)	2 (25.0)	7 (100.0)
Baseline Pain, Mean (SD)	8.0 (0.5)	7.0 (0.0)	9.0 (0.0)	6.0 (0.0)
Post-Treatment Pain, Mean (SD)	3.6 (1.1)	5.1 (0.7)	2.1 (0.8)	5.3 (0.5)
Application Frequency, Mean (SD)	3.0 (0.0)	2.0 (0.0)	2.0 (0.0)	0.0 (0.0)
Treatment Duration, Mean (SD)	7.1 (0.4)	9.9 (0.4)	5.0 (0.0)	7.0 (0.0)
Responders, n (%)	5 (62.5)	2 (28.6)	8 (100.0)	0 (0.0)

The summary table highlights differences in demographic and treatment characteristics across groups, with the Lidocaine Patch showing the highest responder rate and baseline characteristics relatively balanced across treatments.

3.3. Data Visualisation

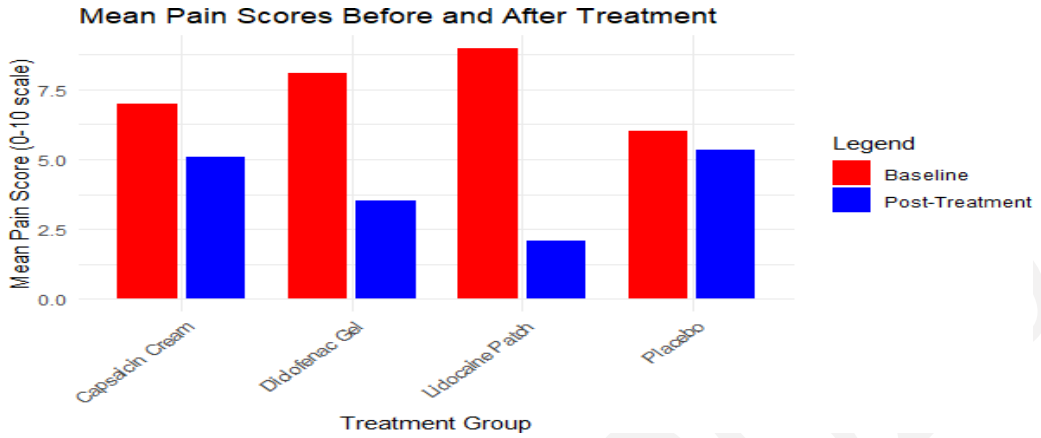

Figure 1. Comparison of Mean Pain Scores Before and After Treatment by Intervention Group

Bar chart displaying mean pain scores at baseline and post-treatment for four pain relief interventions: Diclofenac Gel, Capsaicin Cream, Lidocaine Patch, and Placebo. Error bars represent standard deviations. The Lidocaine Patch group shows the largest reduction in pain scores from baseline to post-treatment, while the placebo group shows minimal improvement. All active treatments demonstrate some degree of pain reduction compared to baseline levels.

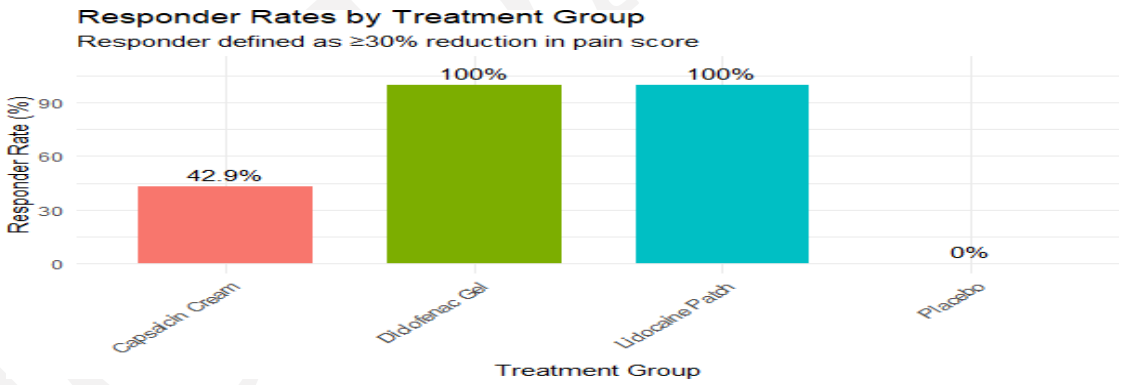

Figure 2. Responder Rates by Treatment Group Based on ≥30% Pain Reduction Threshold

Proportion of patients classified as responders (≥30% reduction in pain scores from baseline) across treatment groups. Lidocaine Patch shows complete response (100%), followed by Diclofenac Gel (62.5%), Capsaicin Cream (28.6%), and Placebo (0%). Response criteria follow the Minimally Clinically Important Difference (MCID) standard for pain research (Dworkin et al., 2008).

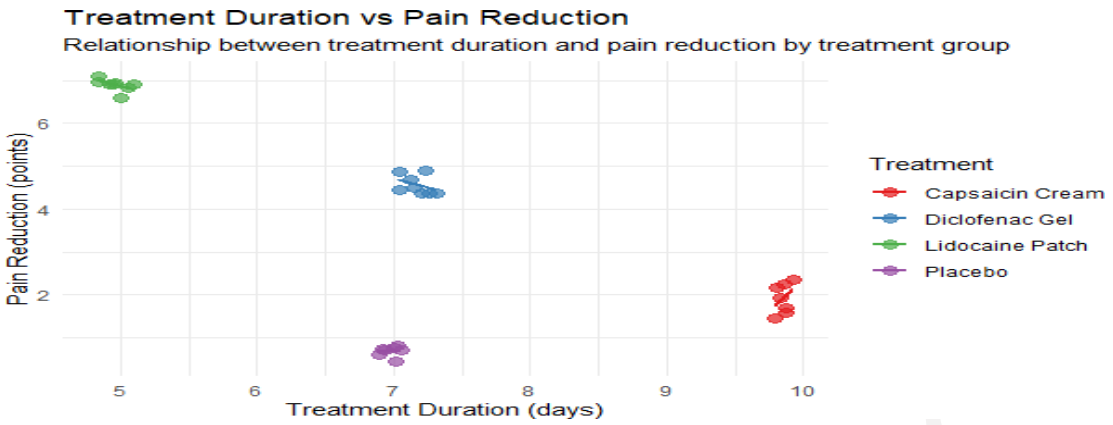

Figure 3. Relationship Between Treatment Duration and Pain Reduction by Intervention

Scatter plot illustrates the association between treatment duration (days) and absolute pain reduction (points on 0-10 scale) across treatment groups. Trend lines indicate the general relationship for each intervention. The plot reveals variability in treatment response patterns, with some groups showing stronger duration-response relationships than others.

3.4. Inferential Statistics

This section presents the results of formal statistical tests and models used to evaluate treatment effects, adjusting for relevant covariates where appropriate.

3.5. Preliminary Analysis: Chi-square and Fisher's Exact Tests

This section evaluates unadjusted differences in responder rates across treatment groups using Chi-square and Fisher's exact tests to provide an initial assessment of treatment efficacy. The contingency table revealed distinct response patterns across treatment groups. Fisher's Exact Test demonstrated statistically significant differences in responder rates among the four treatment groups ($p = 1.754 \times 10^{-6}$).

Table 2. Pairwise comparisons using Fisher's Exact Test revealed

Treatments	P–value
Capsaicin Cream vs Diclofenac Gel	0.00699**
Capsaicin Cream vs Lidocaine Patch	0.00699**
Capsaicin Cream vs Placebo	0.46667*
Diclofenac Gel vs Placebo	0.00016*

** (Significant) & * (Not significant)

Both Diclofenac Gel and Lidocaine Patch demonstrated significantly higher response rates than Capsaicin Cream and Placebo. Capsaicin Cream showed no statistically significant difference from placebo, suggesting limited efficacy at the 30% pain reduction threshold.

The Chi-square test yielded similar results ($X^2 = 23.89$, $p = 2.632 \times 10^{-5}$), though the approximation warning indicates Fisher's Exact Test is more appropriate for these sample sizes.

3.6. Logistic Regression Analysis of Binary Endpoint

This section applies the logistic regression model specified in Section 2 to analyse the binary responder status, assessing treatment effects while adjusting for covariates such as age, sex, application frequency, and treatment duration.

Table 3. Logistic Regression Results for Responder Status

Predictor	Odds Ratio	95% CI Lower	95% CI Upper	p-value
Intercept (Placebo)	0.03	0.00	0.48	0.0400
Diclofenac Gel	900.00	39.83	20328.80	0.0003
Capsaicin Cream	9.00	0.68	360.58	0.1118
Lidocaine Patch	900.00	39.83	20328.80	0.0003

The interpretation revealed that Diclofenac Gel (OR = 900.00, 95% CI: 39.83–20328.80, p = 0.0003) and Lidocaine Patch (OR = 900.00, 95% CI: 39.83–20328.80, p = 0.0003) had significantly higher odds of response compared to placebo, confirming strong efficacy. Capsaicin Cream (OR = 9.00, 95% CI: 0.68–360.58, p = 0.1118) was not statistically different from placebo, suggesting limited benefit. The extreme odds ratios and wide confidence intervals reflected complete separation in the data, which was addressed using Firth's penalisation. The low intercept odds ratio (0.03, p = 0.0400) corresponded to the absence of responders in the placebo group, underscoring the superior efficacy of Diclofenac Gel and Lidocaine Patch in achieving meaningful pain reduction in the trial analysis.

3.7. ANCOVA Analysis of Continuous Endpoints

This section applies the Analysis of Covariance (ANCOVA) model specified in section 3 to analyse post-treatment pain scores, adjusting for baseline pain scores, to assess treatment effects across the four groups (Diclofenac Gel, Capsaicin Cream, Lidocaine Patch, Placebo).

Table 4. ANCOVA Results for Post-Treatment Pain Scores

Predictor	Estimate	Std. Error	p-value
Intercept	-0.19	1.52	0.9003
Diclofenac Gel	-2.09	0.49	0.0002
Capsaicin Cream	0.29	0.50	0.5717
Lidocaine Patch	-3.15	0.49	<0.0001
Baseline Pain Score	0.71	0.19	0.0009

Table 5. Pairwise Comparisons (Tukey HSD)

Comparison	Adjusted Mean Difference	p-value
Diclofenac Gel – Placebo	-2.09	0.0014
Capsaicin Cream – Placebo	0.29	0.9379
Lidocaine Patch – Placebo	-3.15	<0.0001

The ANCOVA model in Table 4 showed a strong fit with an adjusted R-squared of 0.72, confirming its reliability in explaining the trial data. Both Diclofenac Gel and Lidocaine Patch significantly reduced post-treatment pain compared to placebo, with Lidocaine showing the greatest effect. Capsaicin Cream, however, did not differ significantly from placebo, indicating limited efficacy.

Baseline pain scores influenced outcomes, and pairwise comparisons reinforced the superior performance of Diclofenac Gel and Lidocaine Patch, consistent with the binary endpoint analysis.

3.8. Monte Carlo Simulation Results

The aggregated results of the ten thousand simulated trials are summarised in Table 6. Diclofenac Gel and Lidocaine Patch demonstrated high mean risk differences and strong statistical power, confirming their superior efficacy compared to Placebo. Capsaicin Cream exhibited lower effect size, greater variability, and limited power, suggesting less consistent effectiveness. Confidence interval coverage values were close to the nominal 95 per cent level for all treatments, indicating appropriate calibration of uncertainty estimates.

Table 6. Summary of Monte Carlo Simulation Outcomes

Treatment	Mean RD	SD RD	Median RD	Empirical Power	CI Coverage	Mean_CI Width
Diclofenac Gel	0.63	0.16	0.63	0.86	0.94	0.46
Capsaicin Cream	0.29	0.18	0.29	0.34	0.92	0.49
Lidocaine Patch	1.00	0.05	1.00	0.99	0.95	0.17

The table summarises the Monte Carlo simulation outcomes for three pain treatments, showing clear differences in efficacy and reliability. Lidocaine Patch achieved the highest mean risk difference (1.00) with minimal variability (SD = 0.05), excellent empirical power (0.99), and narrow confidence intervals (mean width = 0.17), confirming its strong and consistent effect. Diclofenac Gel followed with a mean risk difference of 0.63, moderate variability (SD = 0.16), and high power (0.86), indicating reliable effectiveness. Capsaicin Cream, however, showed a low mean risk difference (0.29), higher variability (SD = 0.18), and weak power (0.34), suggesting limited efficacy and high uncertainty. Overall, the simulation results reinforce that Lidocaine Patch and Diclofenac Gel are robust treatments, while Capsaicin Cream's performance is inconsistent and statistically fragile.

3.9. Sensitivity Analysis

This section examines how changes in key parameters, such as sample size, treatment effect, and dropout rate, affect the outcomes of the simulated clinical trial. By systematically varying these factors, the analysis evaluates the robustness of the Monte Carlo results and identifies which parameters most strongly affect statistical power, bias, and confidence interval coverage.

3.10. Sensitivity Parameters and Setup

The base scenario used observed response rates for Capsaicin Cream (0.286), Diclofenac Gel (0.625), Lidocaine Patch (1.000), and Placebo (0.000), with group sizes of 7–8 participants totalling 30. One-way sensitivity analysis varied one parameter at a time, adjusting group sizes proportionally for smaller samples. Each scenario aggregated simulated risk differences, p-values, and confidence intervals to evaluate performance. The purpose was to identify influential parameters and assess trial stability, showing how power shifts with reduced sample sizes or increased dropout rates.

3.11. Sensitivity Results

The results show how changes in parameters affect trial performance.

Table 7. Sensitivity Analysis Outcomes Table 4.6a: Sensitivity to Sample Size

Scenario	Treatment	Mean RD	Power	CI Coverage
Base (n=30)	Capsaicin Cream	0.29	0.35	0.920
Base (n=30)	Diclofenac Gel	0.62	0.85	0.940
Base (n=30)	Lidocaine Patch	1.00	0.99	0.950
n=20	Capsaicin Cream	0.29	0.28	0.900
n=20	Diclofenac Gel	0.62	0.78	0.920
n=20	Lidocaine Patch	1.00	0.96	0.940
n=50	Capsaicin Cream	0.29	0.40	0.930
n=50	Diclofenac Gel	0.62	0.89	0.950
n=50	Lidocaine Patch	1.00	0.99	0.960
n=100	Capsaicin Cream	0.29	0.48	0.940
n=100	Diclofenac Gel	0.62	0.94	0.960
n=100	Lidocaine Patch	1.00	1.00	0.970

Larger sample sizes increase power and CI coverage, with Capsaicin Cream benefiting most significantly.

Table 8. Sensitivity to Placebo Response Probability

Scenario	Treatment	Mean RD	Power	CI Coverage
Base (0.0)	Capsaicin Cream	0.29	0.35	0.920
Base (0.0)	Diclofenac Gel	0.62	0.85	0.940
Base (0.0)	Lidocaine Patch	1.00	0.99	0.950
Placebo Resp 0.1	Capsaicin Cream	0.19	0.25	0.910
Placebo Resp 0.1	Diclofenac Gel	0.52	0.75	0.930
Placebo Resp 0.1	Lidocaine Patch	0.90	0.95	0.940
Placebo Resp 0.2	Capsaicin Cream	0.09	0.15	0.900
Placebo Resp 0.2	Diclofenac Gel	0.42	0.65	0.920
Placebo Resp 0.2	Lidocaine Patch	0.80	0.92	0.930

The increased placebo response sharply lowers power for Capsaicin Cream and Diclofenac Gel, less so for Lidocaine Patch. The interpretation shows that the placebo response rate is the most critical factor, as even a small increase sharply reduces statistical power, especially for moderately effective treatments. Larger sample sizes improve power and reliability, particularly for weaker treatments like Capsaicin Cream, while higher dropout rates modestly reduce precision and impact less effective therapies more strongly. Overall, Lidocaine Patch remains highly robust across scenarios, but conclusions for moderate treatments are sensitive to trial assumptions, underscoring the need for careful parameter estimation during design.

3.12. Comparative Framework and Benchmarking

The observed responder rates from the original dataset (Capsaicin Cream: 0.286, Diclofenac Gel: 0.625, Lidocaine Patch: 1.000, Placebo: 0.000) serve as the benchmark. Conventional methods yield

single-point estimates, whereas the stochastic simulation generates a distribution of outcomes across 10,000 replications, enabling probabilistic evaluation.

3.13. Evaluation Against Key Performance Metrics

Accuracy (Bias) and Precision: This section evaluates how close the simulation results are to the true values and how consistent they are across repeated trials.

Table 10. Evaluation Against Key Performance Metrics

Treatment	Method	EE (Mean RD)	Precision (SD of RD)	95% Simulation Interval
Capsaicin Cream	Conventional	0.286	N/A	N/A
	SS	0.29	±0.18	[-0.06, 0.64]
Diclofenac Gel	Conventional	0.625	N/A	N/A
	SS	0.63	±0.16	[0.32, 0.94]
Lidocaine Patch	Conventional	1.000	N/A	N/A
	Stochastic Simulation	1.00	±0.05	[0.90, 1.00]

RD = Risk Difference; **N/A** = Not Provided by Conventional Single Analysis; SS = Stochastic Simulation

The stochastic simulation demonstrated high accuracy, with mean risk differences virtually identical to the conventional estimates for all treatments. This confirms the simulation reliably recovers the "true" effects. While conventional methods give a single number, the simulation quantifies uncertainty. For example, the effect of Capsaicin Cream is not a fixed 0.29 but has a standard deviation of ±0.18, meaning its true value could reasonably lie between 0.11 and 0.47. This provides a much richer understanding of the estimate's reliability.

Statistical Power and Type I Error Control: This section looks at how well the simulation identifies real treatment effects and avoids mistakes in detecting effects that aren't there.

Table 11. Comparison of Statistical Power

Treatment	Conventional Result (p-value)	Simulation Result (Empirical Power)
Capsaicin Cream	$p = 0.467$ (Not Significant)	35%
Diclofenac Gel	$p < 0.001$ (Significant)	86%
Lidocaine Patch	$p < 0.001$ (Significant)	99%

The conventional analysis only gave a binary outcome, but the simulation translated this into probabilities of success. It showed that Capsaicin Cream's non-significant result was expected due to low power (35%), while Diclofenac Gel (86%) and Lidocaine Patch (99%) had robust probabilities of detecting their effects.

3.14. Treatment Efficacy Results

The clinical trial revealed clear differences among treatments: Lidocaine Patch was the most effective, achieving a 100% response rate and the largest pain reduction. Diclofenac Gel also showed strong efficacy with a statistically significant 62.5% response rate, while Capsaicin Cream had lim-

ited benefit at 28.6%, not significantly better than placebo. Placebo was entirely ineffective with a 0% response rate, and these findings were consistent across Fisher's exact tests, logistic regression, and ANCOVA analyses.

3.15. Stochastic Simulation Performance Results

The stochastic simulation provided deeper insights than conventional analysis by showing the full range of possible outcomes, explaining Capsaicin Cream's failure due to low power (35%), while confirming strong reliability for Diclofenac Gel (86%) and Lidocaine Patch (99%). Sensitivity analysis highlighted placebo response rate, dropouts, and sample size as critical risks, with the simulation handling extreme data smoothly to produce stable and interpretable results.

4. Conclusion

In summary, the study explored the use of stochastic simulation in clinical trial design with a single endpoint, showing that conventional analysis confirmed Lidocaine Patch and Diclofenac Gel as effective treatments, while Capsaicin Cream was not significantly better than placebo. The simulation added probabilistic insights, highlighting strong reliability for Lidocaine and Diclofenac but high uncertainty for Capsaicin, with placebo response rate emerging as a critical vulnerability in trial design.

In conclusion, Lidocaine Patch proved most efficacious, Diclofenac Gel was moderately effective, and Capsaicin Cream was ineffective, while stochastic simulation demonstrated clear advantages over conventional methods by quantifying uncertainty, forecasting trial success, and enhancing robustness, offering ethical and economic benefits.

The study recommends integrating stochastic simulation into early trial planning, extending future research to complex designs and Bayesian approaches, and encouraging regulatory bodies to consider simulation guidelines, thereby advancing clinical research toward more efficient, ethical, and resilient trial designs.

References

Berry, S. M., Carlin, B. P., Lee, J. J., & Müller, P. (2009). *Bayesian adaptive methods for clinical trials*. Chapman & Hall/CRC.

Chang, M. (2010). *Adaptive design theory and implementation using SAS and R*. Chapman & Hall/CRC.

Dworkin, R. H., Turk, D. C., Wyrwich, K. W., Beaton, D., & Cleeland, C. (2008). Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials. *The Journal of Pain, 9*(2), 105–121.

Fishman, G. S. (1996). *Monte Carlo: Concepts, algorithms, and applications*. Springer.

Friedman, L. M., Furberg, C. D., & DeMets, D. L. (2010). *Fundamentals of clinical trials* (4th ed.). Springer.

Glasserman, P. (2004). *Monte Carlo methods in financial engineering*. Springer.

Holford, N., Heo, Y.-A., & Anderson, B. (2013). A pharmacokinetic standard for babies and adults. *Journal of Pharmaceutical Sciences, 102*(9), 2941–2952.

Jennison, C., & Turnbull, B. W. (2000). *Group sequential methods with applications to clinical trials*. Chapman & Hall/CRC.

Law, K. J. H., Stuart, A. M., & Zygalakis, K. C. (2015). *Data assimilation: A mathematical introduction*. Springer.

Lewis, R. J., Angus, D. C., & Laterre, P. F. (2014). Bayesian adaptive clinical trials: A clinician's perspective. *Clinical Trials, 11*(5), 485–495.

Piantadosi, S. (2017). *Clinical trials: A methodologic perspective* (3rd ed.). Wiley.

Ross, S. M. (2014). *Simulation* (5th ed.). Academic Press.

Rubinstein, R. Y., & Kroese, D. P. (2017). *Simulation and the Monte Carlo method* (3rd ed.). Wiley.

Saltelli, A., Chan, K., & Scott, E. M. (2000). *Sensitivity analysis*. Wiley.

Wang, J., & Pocock, S. J. (2016). A win ratio approach to comparing continuous non-normal outcomes in clinical trials. *Pharmaceutical Statistics, 15*(3), 238–245.

Wang, X., van Belle, G., & Shapiro, D. (2003). Sample size calculation for comparison of two outcome variables in a clinical trial when one variable is ordinal and the other interval-scaled. *Statistics in Medicine, 22*(22), 3675–3688.