Paper AS08

Is There Really Any Benefit to Stratified Randomisation Compared to Simple Randomisation with Covariate Adjustment?

Pavankumar Bhagat, Boehringer Ingelheim Pharma GmbH & Co. KG, Ingelheim, Germany
Dr. Jack Wilkinson, University of Manchester, Manchester, UK

ABSTRACT
Randomized controlled trials (RCTs) are the gold standard for evaluating interventions in clinical research. Stratified randomization is a common technique used to balance prognostic factors across treatment groups, aiming to increase the precision of treatment effect estimates. However, its effectiveness and efficiency compared to simple randomization with covariate adjustment remain a subject of debate.
A simulation study as part of master's thesis at university of Manchester was performed to compare the performance of stratified randomization and simple randomization with covariate adjustment under various conditions, including different sample sizes, association levels between covariates and outcomes, stratification levels, and treatment effects. By employing a linear regression model, key performance metrics such as power, type I error rate, empirical standard error, bias, and coverage probability were assessed.
These findings suggest that stratified randomization has advantages, only in smaller sample sizes while adjusting for highly prognostic covariates. It showed slightly increased statistical power and reduced standard errors. However, the benefits were more pronounced in specific scenarios, highlighting the need for careful consideration of study design and analysis.

INTRODUCTION
In the realm of clinical trials, the utilization of randomized controlled trials (RCTs) stands as the gold standard for evaluating the effects of interventions [1], and it is a key part of the modern clinical trial design. It helps to make sure that people in different treatment groups are similar, so the results are fair and not biased. The most basic way to do this is called simple randomisation**,** which is simple, quick and easy to perform where patients are allocated to treatment or control based on the single randomisation list prepared by the investigator [2] or simply by computer, by preserving allocation concealment. However, it may result in unequal group sizes in the small trials and leads to random imbalances in the prognostic characteristics.

Simple randomisation have some challenges like the potential for imbalances in treatment groups and prognostic characteristics, such imbalance is very common in small clinical trials, where sicker patients could happen to be in the control arm and other patients in the treatment arm whereas in stratified randomisation, patients are defined by a group or strata based on their prognostic characteristics(like age or disease severity etc.), and then according to the strata, a separate randomisation list is generated in blocks. Here, stratified randomisation balances the groups based on prognostic characteristics, hence it is used to avoid the random imbalances and improve balance in RCT's [3]

BACKGROUND

Many researchers have studied how randomisation affects clinical trial results. Simple randomisation is easy to use and works well in large trials. But in smaller trials, it can lead to uneven groups especially if prognostic patient characteristics (like age or disease severity etc.) are not balanced between treatment and control arms.

In the other hand stratified randomisation overcomes the challenge of imbalance in treatment groups and prognostic characteristics [4]. Stratified randomisation is commonly used to balance covariates, as there is a belief that this may lead to increased power, and possibly a reduction in bias. However, it is not clear that balancing covariates using stratified randomisation leads to any of these benefits compared to simple randomisation with adjustment for the prognostic characteristic in the analysis.

There is no proven evidence from statistical research that indicates any benefit of power or confidence interval when using stratified randomisation while adjusting the covariates at the analysis stage over simple randomisation with co-variate adjustment without any limitations. If we set random imbalance aside, there might be no other benefit between these two methods.

It is beneficial to be aware of this fact as if there is no proven benefit of stratified randomisation.
with adjustments over simple randomisation with adjustments, it would be more recommendable to use simple

randomisation for large trials due to the complexity of stratified randomisation, which is difficult to organize and maintain. Moreover, not adjusting the covariates is even more disadvantageous. This leads to an important question: **Is stratified randomisation always needed, or is simple randomisation good enough in many cases?**

A simulation study with continuous outcome was performed to compare both methods by adjusting the covariates to answer the above research question, and this simulation study follows the ADEMP structure [5].

This paper builds on these ideas by using simulations to compare both methods. We test many different trial setups to see how stratified and simple randomisation perform under different conditions. Our goal is to give clear, practical advice for trial designers.

METHODOLOGY

To compare stratified and simple randomisation, we used a simulation-based approach. Simulations allow us to test many different trial setups and see how each method performs under controlled conditions. We followed the ADEMP framework, which helps organize simulation studies clearly. ADEMP stands for: Aims, Data Generating Mechanism, Estimands, Methods, and Population [5].

Aims

The goal of this study was to:

•Evaluate the power benefit while using stratified randomisation plus co-variate adjustments versus simple randomisation plus co-variate adjustments in the analysis stage, which is the primary focus of this simulation study/research.

• Evaluate the benefit in confidence interval for stratified randomisation with co-variate adjustment when compared to simple randomisation with co-variate adjustment.

Data Generating Mechanism.

In this section, we outline how data was simulated including the values of the key inputs. Here, we generated several independent simulated datasets, each representing different sample sizes or phases of clinical trials, and these generated datasets were independent of each other allowing us to assess the performance of different methods under varying conditions. Like

- Under different sample sizes or different phases of the clinical trials.
- With different association levels of the covariates with the outcome, i.e., Weak, Moderate, and strong correlation between baseline covariates used for stratification and the outcome.
- with different stratification proportions, equal proportions, or unequal proportions.
- with different effects of treatment i.e., if there is a true treatment effect or if there is no treatment effect.

Estimands

We focused on estimating the average treatment effect the mean difference in outcomes between the two groups. This was calculated using a linear regression model that included the treatment indicator and covariates.

Methods

We used R programming language to run the simulations and various R packages like Blockrand, dplyr, tidyr and ggplot2 were used in this simulation study. Each scenario was repeated 100,000 times to avoid monte carlo errors and get reliable results. For each simulated trial, we:

1. Generated covariate data for all participants.
2. Assigned treatments using either simple or stratified randomisation.
3. Simulated outcomes based on treatment and covariates.
4. Estimated the treatment effect using linear regression (Analysis of Covariance).

Performance Measures

Since we aim to compare the two different randomisation methods with covariate adjustments (Simple v/s Stratified) for the continuous outcome, in this simulation study we focused on numerical quantities that are used to assess the performance of the method like power, empirical standard error, type-1 error rate, bias and coverage.

Simulation Scenarios and Assumptions

To evaluate the performance of randomisation methods under controlled conditions, we designed our simulation study based on the following scenarios and assumptions:

- Study was simulated as a two-arm parallel design trial.
- Treatment variables were generated from a uniform distribution across both simple and stratified randomisation methods.

- Baseline covariates were simulated as normally distributed random variables with different means. In stratified randomisation, these covariates were used to define strata with two levels (B = 0 and B = 1), allowing for both equal and unequal stratum sizes.
- For simple randomisation, treatment was assigned using a uniform random variable. Whereas in stratified randomisation, block randomisation was applied within each stratum, with block sizes determined by the number of treatment levels (two: control [0] and treatment [1]) and sample sizes (n = 50, 100, 250).
- The true treatment effect was set to $\theta = 0.6$ for power estimation and $\theta = 0$ for assessing type I error rates by maintaining a significance level of 0.05, and simulation assumed complete data availability, with no missing values and no imputation procedures applied.
- To explore the impact of stratification, four levels of correlation between the baseline covariate and the outcome were considered:
    - Weak: log (1.2).
    - Moderate: log (2).
    - Strong: log (9).
- Three different distributions of patients across the two strata were evaluated to assess their influence on performance metrics:
    - Equal: 50% / 50%.
    - Unequal: 70% / 30%.
    - Unequal: 80% / 20%.
- . Outcome variables follow a standard normal distribution with a fixed standard deviation of 0.5.

These scenarios and assumptions provided a structured framework for evaluating the robustness and efficiency of randomisation strategies, ensuring consistency in analysis and interpretation across varying experimental conditions.

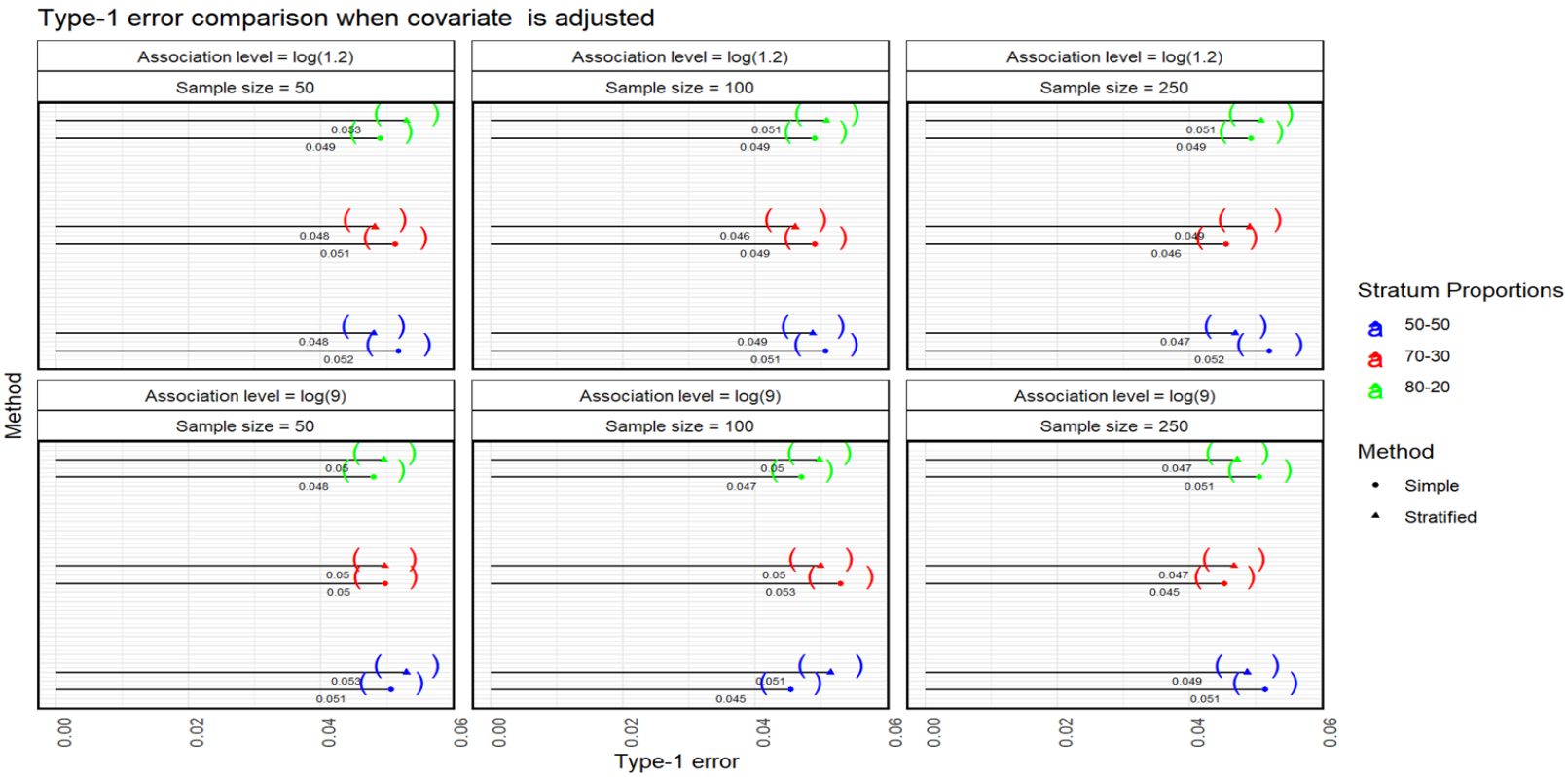
RESULTS

Here, we present the results of our simulations study for a continuous outcome comparing stratified and simple randomisation by adjusting the covariates during analysis, and these results demonstrate the performance measures observed such as type I error rate, power, empirical standard error (SE), coverage, and bias. In the results, for the figures, we only considered the association level as Log (1.2) and Log (9) to simplify the graphs and fit in the paper as there were no differences in the results between Log (1.2) and Log (2).

Type I Error Rate

By simulating the study for the scenario where treatment effect was assumed to be zero to estimate type-1 error rate. both randomisation methods-maintained type I error rates at the nominal 5% level across all investigated scenarios Neither sample size, stratum proportions, nor association levels impacted Type I error control. Confidence intervals consistently contained the 0.05 threshold for both methods, indicating no distinguishable difference in false-positive rate control.

Figure.1



Power

Stratified randomisation showed slightly increased power than simple randomisation, particularly for small sample sizes with unequal stratum proportions (70%-30%) and strong association level (Log9) However, the practical significance of this advantage is limited:

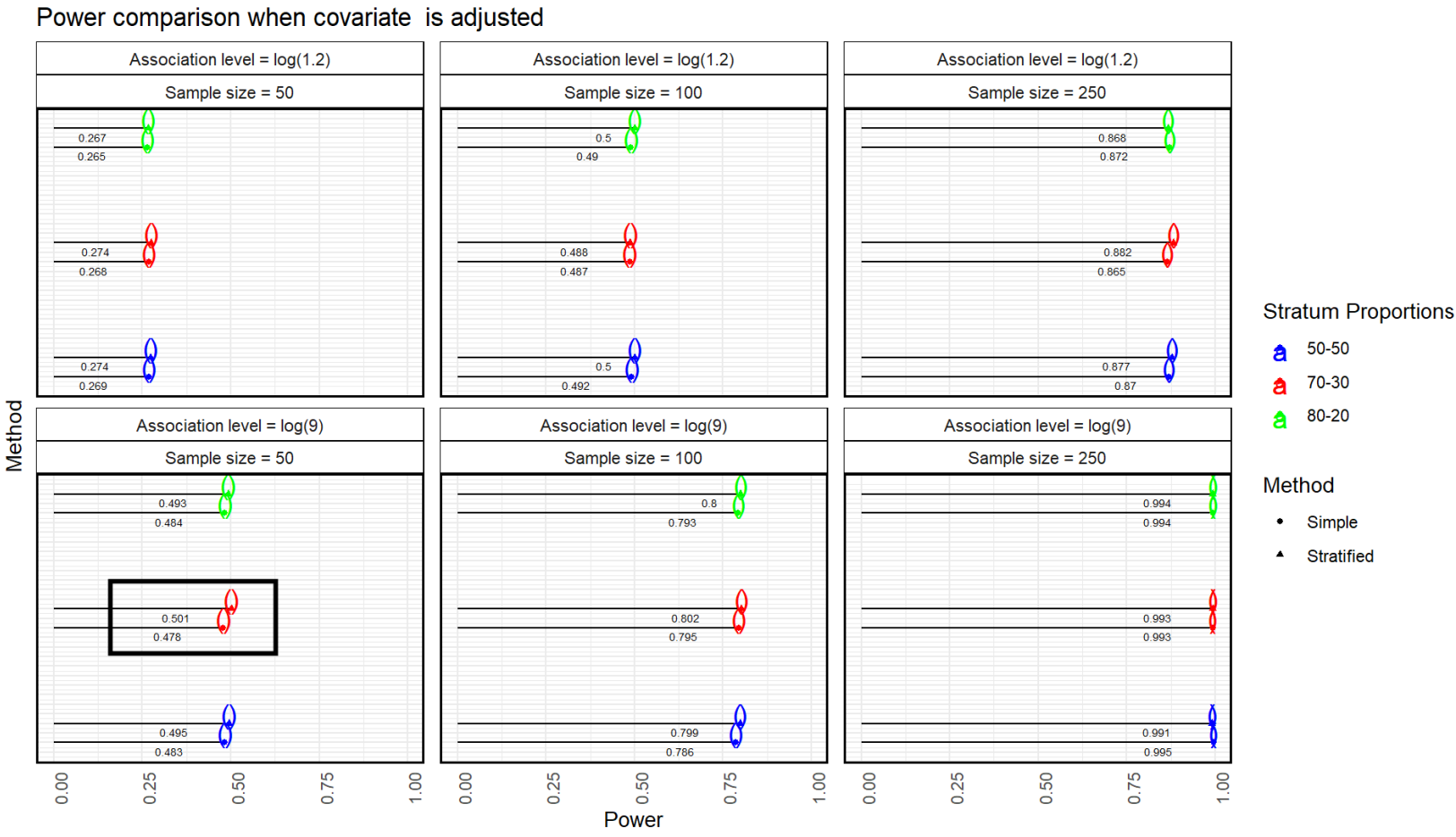
- n=50: Power difference ≈0.01 (1%) in all scenarios except one where association level of outcome with covariate is strong (log 9) and stratum proportion is unequal (70%-30%) where the power difference is ≈0.02 (5%)

- n>50: Differences negligible; both methods achieve nearly same power.

for large sample sizes, association levels and stratum sizes had no impact on the results, in all scenarios results were virtually identical.

Key Finding: Stratified randomisation offers statistically detectable but clinically negligible power advantages only in smaller studies (n=50) where association level is strong and stratum sizes are unequal, whereas for sample size>50, both methods perform equivalently in all scenarios.

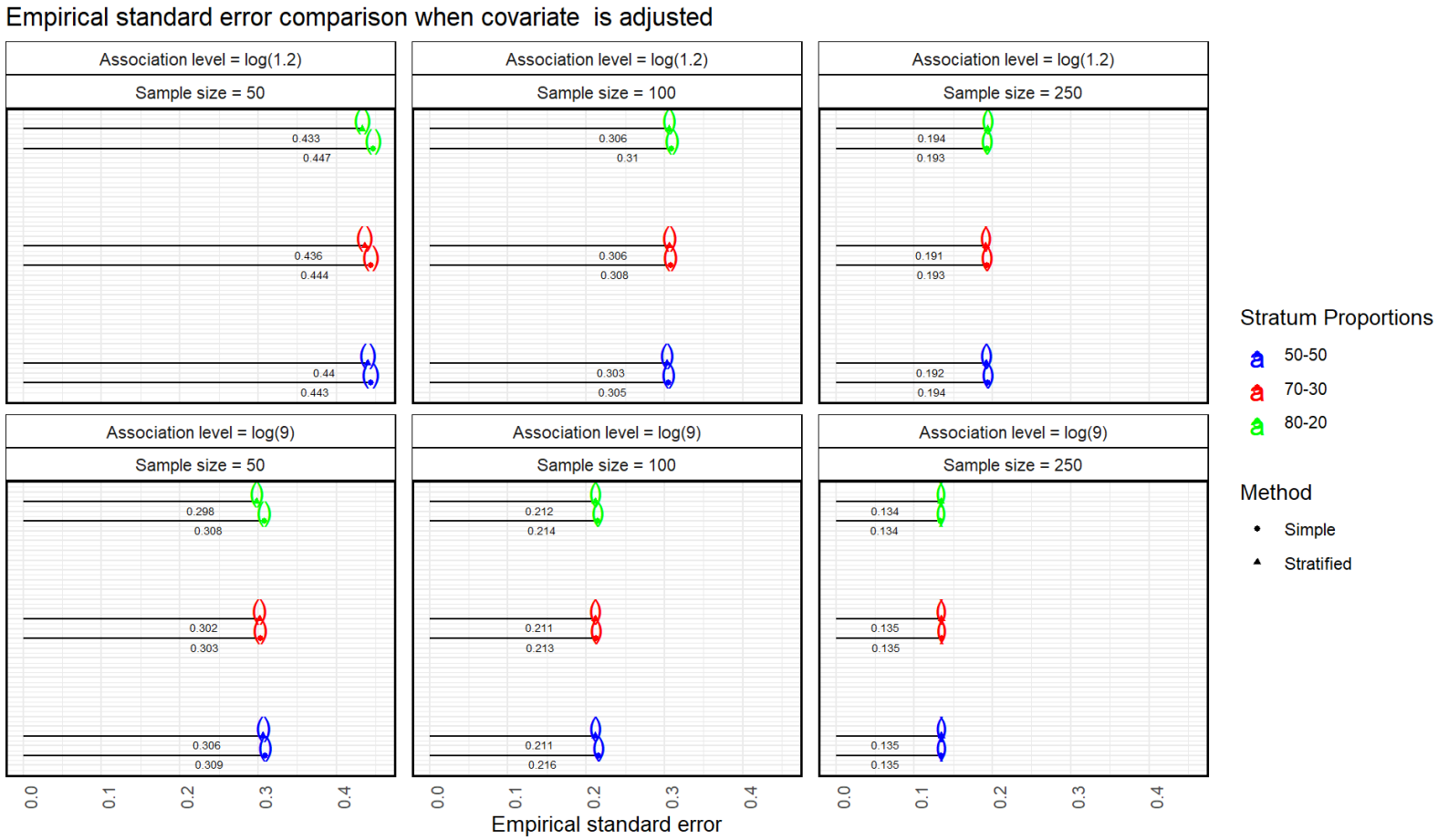
Figure.2



Empirical Standard Error

Empirical standard error decreased certainly with increasing sample size for both methods, with overlapping confidence intervals across all scenarios Treatment effect presence (θ=0 vs. θ=0.6) did not influence empirical standard error comparisons. Stratum proportions and association levels showed minimal impact on SE estimates.

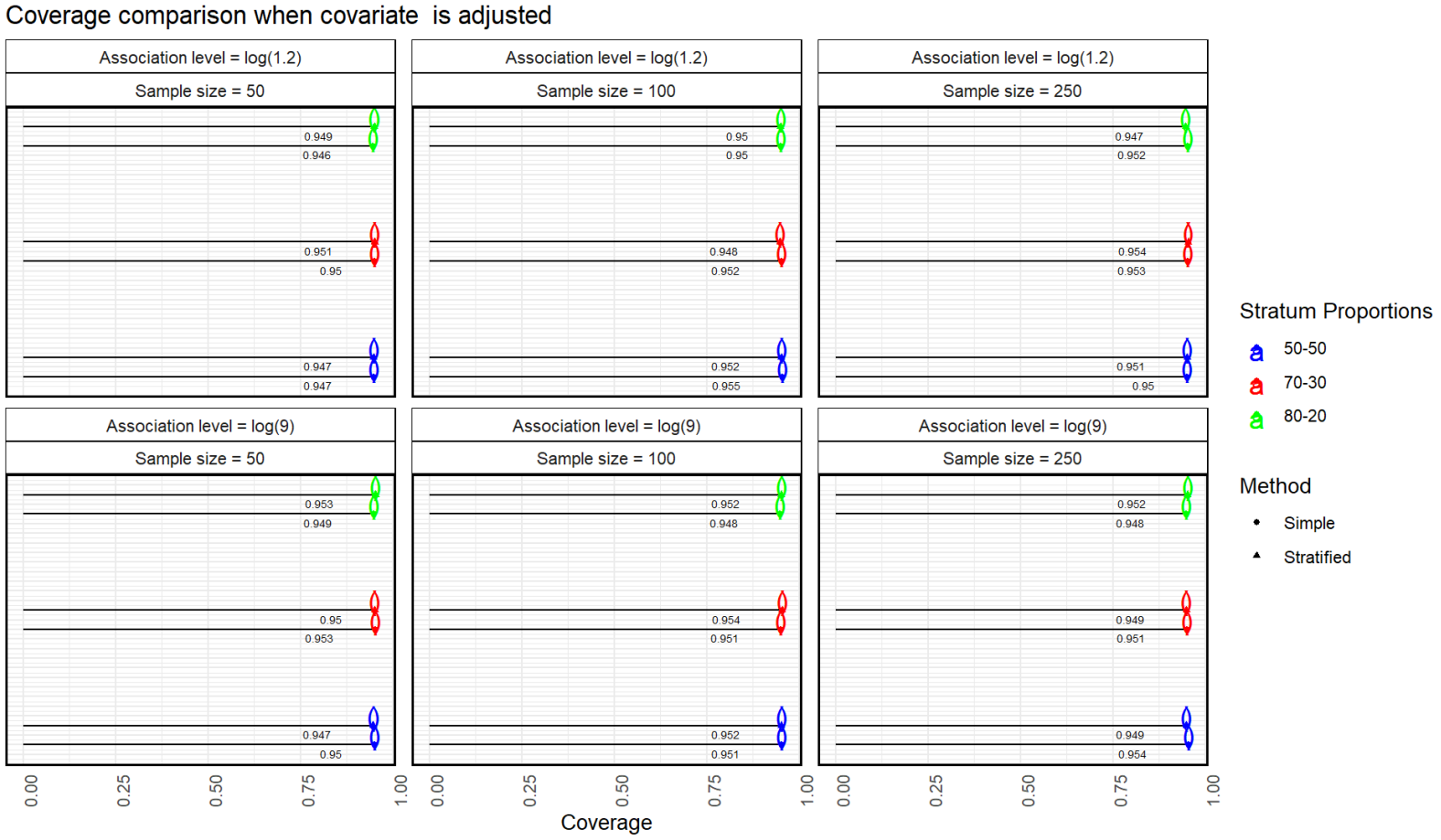
Figure.3



Coverage

Both methods achieved coverage rates approximating the nominal 95% level across all conditions Coverage performance was consistent regardless of treatment effect, sample size, stratum proportions, or association levels. Confidence intervals for both methods demonstrated adequate capture of true parameter values. Both randomisation methods provide reliable interval estimation with proper coverage properties; method selection does not impact inferential validity.

Figure.4

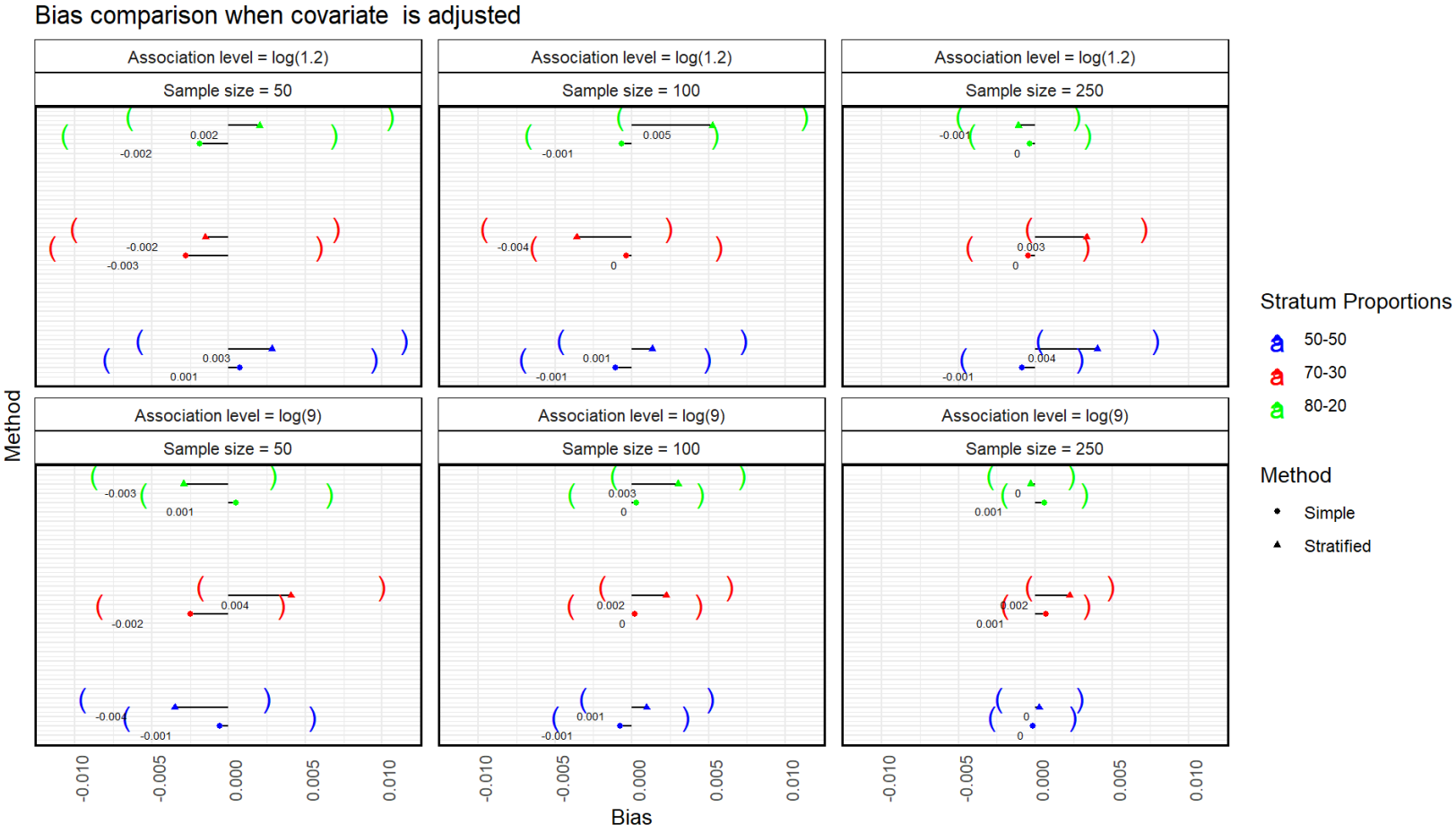


Bias

Both methods demonstrated low bias across all scenarios, with estimates centered near zero. Bias magnitude varied slightly by sample size and stratum proportions but showed no consistent pattern favoring either method. Treatment effect presence did not systematically alter bias patterns.

Key Finding: Neither method demonstrates superiority in bias reduction; both provide unbiased treatment effect estimates across tested conditions.

Figure.5



CONCLUSION

This simulation study compared stratified and simple randomisation both with covariate adjustment for a continuous outcome across a range of clinical trial scenarios. The results showed that stratified randomisation slightly increases statistical power in small trials only when the covariates are highly prognostic, whereas in larger trials, both methods performed similarly, with minimal bias and consistent Type I error rates.
Our simulations used ideal conditions like complete data, with only two stratum, and fixed treatment effects like 0 and 0.6. Real trials may face challenges like missing data, protocol deviations, or unexpected changes. Also, we only tested this for continuous outcome.

Further research is needed to explore this for different outcomes like binary, survival and future studies could explore more complex trial designs with more strata factors.

Both methods are statistically valid and can be used confidently in clinical research. The choice between them should depend on the trial's size, complexity, and goals. By choosing the right randomisation strategy, researchers can improve trial quality without adding unnecessary steps. This study provides practical guidance to help make that decision with confidence. In conclusion, this study provides valuable insights into the performance of simple and stratified randomisation methods with covariate adjustments in controlling type I error rates, power, bias, coverage, and empirical standard error. However, there is still much to explore. Future research should aim to address the limitations of the current study and continue to explore ways to improve the design and analysis of clinical trials.

REFERENCES

1. B. C. Kahan and T. P. Morris, "Reporting and analysis of trials using stratified randomisation in leading medical journals: Review and reanalysis," BMJ (Clinical research ed.), vol. 345, no. e5840, 2012

2. A. Feinstein and J. Landis, "The role of prognostic stratification in preventing the bias permitted by random allocation of treatment," Journal of Chronic Diseases, vol. 29, no. 4, pp. 277–284,1976

3. P. Johansson and M. Schultzberg, "Rerandomization: A complement or substitute for stratification in randomized experiments?" Journal of Chronic Diseases, vol. 218

4. S. Green and D. Byar, "The effect of stratified randomization on size and power of statistical tests in clinical trials," Journal of Chronic Diseases, vol. 31, no. 6-7, pp. 445–454, 1978

5. C. M. Morris TP White IR, "Using simulation studies to evaluate statistical methods," Statistics in Medicine, vol. 38, no. 11, pp. 2074–2102, 2019

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Pavankumar Bhagat
Email: Pavankumar.bhagat@boehringer-ingelheim.com

Boehringer Ingelheim Pharma GmbH & Co. KG
Binger Str. 173 | 55216 Ingelheim am Rhein
Website: www.boehringer-ingelheim.com