International Journal of Advances in Medicine
Subudhi G et al. Int J Adv Med. 2026 Jul;13(4):188-197
https://www.ijmedicine.com
pISSN 2349-3925 | eISSN 2349-3933
DOI: https://dx.doi.org/10.18203/2349-3933.ijam20261559

Original Research Article

A simulation-based evaluation of synthetic control arm methodology in cardiovascular trials using propensity score matching

Girik Subudhi[1]*, Shradhit Subudhi[2], Krish Pinto[3]

[1]Rohilkhand Medical College and Hospital, Bareilly, Uttar Pradesh, India
[2]Rutgers University, New Jersey, USA
[3]Fr. Conceicao Rodrigues Institute of Technology Nerul, Navi Mumbai, Maharashtra, India

Received: 20 April 2026
Revised: 17 May 2026
Accepted: 18 May 2026

***Correspondence:**
Dr. Girik Subudhi,
E-mail: giriksubudhi@gmail.com

Copyright: © the author(s), publisher and licensee Medip Academy. This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT

Background: Conventional randomized controlled trial (RCTs) in cardiovascular medicine are expensive, ethically constrained, and logistically demanding. The synthetic control arm (SCA) methodology leverages real-world inspired simulated patient data and propensity score matching to construct an external comparator, enabling all enrolled participants to receive active treatment. Despite increasing regulatory acceptance, cardiovascular applications of this methodology remain methodologically under-characterized.
Methods: A simulation study was conducted involving 75 patients: 40 in the treatment arm receiving a combined antihypertensive and lipid-lowering agent, and 35 in the SCA control arm derived by propensity score matching on ten baseline clinical covariates. Primary endpoint was MACE at 12 months. Secondary endpoints were change in LDL cholesterol and systolic blood pressure. Statistical analyses included independent samples t-tests, Chi-squared tests, Cohen's d, odds ratios, and Kaplan-Meier survival analysis.
Results: Propensity score matching produced acceptable covariate balance across nine of ten variables (all SMD <0.3), with one residual imbalance in baseline systolic BP (SMD=0.54; p=0.024). Treatment significantly reduced LDL cholesterol (−29.4±10.3 versus −5.0±5.1 mg/dl; p<0.001; d=−2.94) and systolic blood pressure (−10.6±3.6 versus −3.3±3.5 mmHg; p<0.001; d=−2.06). MACE occurred in 32.5% of treated versus 40.0% of SCA control patients (OR=0.72; p=0.664), non-significant owing to insufficient power at this sample size.
Conclusion: The SCA design successfully demonstrates the feasibility of SCA methodology under controlled simulation assumptions, with strong performance for surrogate endpoints but limited power for clinical endpoints in a cardiovascular simulation context, demonstrating robust detection of surrogate endpoint effects. The MACE non-significance is attributable to sample size limitations, not methodological failure. Adequate power for a 7.5 percentage-point MACE difference requires approximately 645 patients per arm. These findings provide a structured methodological framework for future SCA applications in cardiovascular pharmacoepidemiology.

Keywords: Synthetic control arm, Propensity score matching, Cardiovascular trial, MACE, Real-world data, External control, Pharmacoepidemiology, Simulation study

INTRODUCTION

Cardiovascular disease (CVD) remains the leading cause of morbidity and mortality globally, accounting for an estimated 17.9 million deaths annually according to the World Health Organization. The evaluation of novel therapeutic interventions for CVD relies historically on the randomized controlled trial, which provides the highest level of evidence through internal validity achieved by random allocation of treatment. However, the classical two-arm randomized controlled trial (RCT) design presents persistent structural challenges.

First, randomizing patients to placebo or standard-of-care control in conditions such as hypertension, dyslipidaemia, and established coronary artery disease raises substantial ethical concerns, particularly when effective treatments already exist.[1] Second, cardiovascular RCTs are resource-intensive, the landmark PARADIGM-HF trial, for instance, enrolled over 8,400 patients across 47 countries at a cost exceeding $100 million USD.[2] Third, strict eligibility criteria in RCTs often exclude elderly patients, those with renal impairment, or patients with multiple comorbidities, limiting the generalizability of trial findings to real-world populations.[3]

The synthetic control arm (SCA) represents a methodologically rigorous alternative that addresses these limitations by constructing a comparator cohort from existing real-world data (RWD) rather than from randomly allocated participants. The SCA approach ensures that all trial participants receive active treatment, satisfying ethical imperatives, while propensity score-based matching provides statistical approximation of the balance achieved by randomization.[4]

Simulation studies provide a controlled environment to evaluate methodological performance independent of real-world data limitations such as missingness, measurement error, and unmeasured confounding. This study therefore aims not to estimate clinical treatment effects, but to assess whether SCA methodology can recover known simulated effects under idealized conditions.

Problem statement

Despite growing regulatory acceptance of SCA designs acknowledged by the FDA's 2018 real-world evidence framework and the EMA's PRIME scheme the application of this methodology in cardiovascular medicine remains nascent. Existing SCA literature is dominated by oncology, where high event rates, limited treatment options, and small patient populations make external controls particularly attractive.[5,6] In cardiovascular trials, the lower event rates for hard endpoints such as MACE, the greater heterogeneity of RWD sources, and the complexity of multi-drug regimens create distinct methodological challenges that have not been systematically characterized.

Critically, no published simulation study has prospectively modelled the statistical power, covariate balance, and endpoint detection performance of an SCA design specifically calibrated for cardiovascular outcomes. This gap limits the ability of trialists and regulators to make evidence-based decisions about when and how to deploy SCA designs in cardiovascular drug development.

Novelty and contribution

This study makes the following novel contributions to the field as this study presented the first simulation study explicitly designed to characterize SCA performance in a cardiovascular context, using clinically realistic baseline covariate distributions and event rates derived from published cardiovascular epidemiology. We provide a complete mathematical framework for sample size estimation in SCA cardiovascular trials, including power curves for MACE detection across a range of per-arm sample sizes. We demonstrate a reproducible, open-source pipeline for propensity score estimation, nearest-neighbour matching, balance assessment (Love plot), and survival analysis (Kaplan-Meier) applicable to future cardiovascular SCA studies. We provide a clinical perspective on the adaptation of this methodology to medical domains including pharmacovigilance, EHR narrative mining, and mental health assessment, broadening the translational scope of the findings.

Objectives

The specific objectives of this study involved primary objective: to evaluate the ability of propensity score matching to achieve adequate covariate balance between treatment and SCA control groups in a cardiovascular simulation and secondary objectives were to quantify treatment effects on LDL cholesterol and systolic blood pressure as surrogate cardiovascular endpoints, to evaluate MACE rates between groups at 12 months and characterize the power of the SCA design to detect clinically meaningful differences and to derive mathematical sample size requirements for adequate power in future cardiovascular SCA trials.

METHODS

Study design

This study is a pre-specified simulation-based methodological study employing a SCA design, conducted over the period January 2025 to December 2025 (12 months). It was reported in accordance with the strengthening the reporting of observational studies in epidemiology (STROBE) guidelines and in compliance with ICMJE recommendations for the conduct, reporting, editing, and publication of scholarly work in medical journals.

A total of 75 patients were included, with 40 patients in the treatment arm and 35 patients constituting the synthetic control arm derived from real-world data sources. The SCA was constructed using propensity score matching based on clinically relevant baseline covariates to ensure comparability between groups.

Data generation

Baseline covariate distributions were parameterized based on published cardiovascular trial populations. Age was drawn from a uniform distribution over 45–78 years. BMI was drawn from a uniform distribution over 22–38 kg/m². Baseline systolic blood pressure was drawn uniformly over 130–175 mmHg, and diastolic blood pressure over

80–105 mmHg. LDL cholesterol was drawn uniformly over 100–220 mg/dl and HDL over 30–65 mg/dl. eGFR was drawn uniformly over 45–95 ml/min/1.73 m². Binary covariates (smoking, diabetes, prior MI) were assigned with weighted random sampling reflecting cardiovascular risk population prevalences of approximately 33%, 28%, and 25% respectively.

Treatment outcomes were simulated with differential probabilities: MACE occurred at approximately 20% in the treatment arm and 35% in the control arm, reflecting a clinically plausible treatment effect. LDL reductions in the treatment arm were drawn from a uniform distribution over −45 to −10 mg/dl (mean approximately −30 mg/dl), consistent with high-intensity statin and PCSK9 inhibitor therapy. Systolic BP reductions in the treatment arm were drawn from −18 to −5 mmHg, reflecting dual antihypertensive therapy.

Because all data were simulated, the true underlying treatment effect was known by design. This enables evaluation of whether the SCA methodology can accurately recover expected differences between groups.

Inclusion criteria

The following eligibility criteria were applied in the simulation design and would apply to patient selection in a real-world implementation of this protocol.

Eligible participants were adults aged 45–78 years with an established diagnosis of hypertension (systolic BP≥130 mmHg) and LDL cholesterol of at least 100 mg/dl. Patients were required to have stable cardiovascular status at enrolment, demonstrate willingness to comply with the 12-month follow-up protocol, and provide informed consent or a proxy equivalent for simulation purposes.

Exclusion criteria

Patients were excluded if they had an active malignancy or terminal illness, severe renal impairment (eGFR<30 ml/min), or severe hepatic disease (ALT/AST>3×ULN). Additional exclusion criteria included a recent myocardial infarction or stroke within the preceding 3 months, concomitant use of PCSK9 inhibitors, pregnancy or breastfeeding, and participation in another interventional trial within the past 6 months.

Propensity score estimation and matching

Propensity scores were estimated using binary logistic regression, with treatment group assignment (treatment=1, SCA control=0) as the outcome variable. The ten baseline covariates entered as predictors were: age, sex, BMI, baseline systolic BP, baseline diastolic BP, LDL cholesterol, HDL cholesterol, eGFR, smoking status, diabetes status, and prior MI history.

The logistic regression model takes the form.

$$log[p/(1-p)] = \beta_0 + \beta_1(Age) + \beta_2(Sex) + \beta_3(BMI) + \beta_4(SBP) + \beta_5(LDL) + \ldots + \beta_k(X_k)$$

Where, p=P (treatment=1|X) was the estimated propensity score for patient i given covariate vector X, and $\beta_0 \ldots \beta_k$ are the logistic regression coefficients. The estimated propensity score $\hat{e}(X)=\hat{P}(T=1|X)$ was then used for nearest-neighbour matching.

Nearest-neighbour matching was performed without replacement, with a caliper of 0.05 on the propensity score scale to limit poor matches.

$$|\hat{e}(X_i) - \hat{e}(X_j)| \leq 0.05$$

Where i denoted a treatment patient and j denotes the nearest candidate control. Covariate balance was assessed post-matching using the standardized mean difference (SMD).

$$SMD = (X_t - X_c)/\sqrt{[(S_t{}^2 + S_c^2)/2]}$$

Where $\bar{X}_t$ and $\bar{X}_c$ were group means and S_t^2 and S_c^2 are group variances for each covariate. An SMD <0.1 was considered indicative of excellent balance; SMD 0.1–0.2 acceptable; and SMD>0.2 indicative of residual imbalance requiring further adjustment.[7]

Sample size and power calculation

For a two-proportion comparison (MACE rates p_1=32.5% treatment versus p_2=40.0% control) at a two-sided significance level α=0.05 and power 1−β=0.80, the required per-arm sample size is derived using the normal approximation.

$$n = [z_\{\alpha/2\}\sqrt{(2\bar{p}(1-\bar{p}))} + z_\beta\sqrt{(p_1(1-p_1) + p_2(1-p_2))}]^2/(p_1-p_2)^2$$

Where, $\bar{p}=(p_1+p_2)/2$ is the pooled proportion, $z_\{\alpha/2\}$=1.960 (two-sided α=0.05), and z_β=0.842 (power=80%). Substituting the values, the equation becomes.

$$\bar{p} = (0.325 + 0.400)/2 = 0.3625$$

$$n = [1.960\sqrt{(2 \times 0.3625 \times 0.6375)} + 0.842\sqrt{(0.325 \times 0.675 + 0.400 \times 0.600)}]^2/(0.075)^2$$

n≈645 patients per arm (1,290 total) for 80% power.

For 90% power (z_β=1.282), the required sample size increases to approximately 865 patients per arm (1,730 total). The current simulation (n=40 treatment, n=35 control) was therefore expected to be substantially underpowered for MACE detection, yielding an estimated power of approximately 13% — a design limitation that is

deliberately modelled to characterize SCA performance in small samples.

Endpoints

The primary endpoint was MACE at 12 months, defined as a composite of non-fatal myocardial infarction, non-fatal stroke, and cardiovascular death, consistent with the endpoint definition used in the COMPASS, DECLARE-TIMI 58, and REDUCE-IT trials. Secondary endpoints were: change in LDL cholesterol from baseline at 12 months (mg/dl), and change in systolic blood pressure from baseline at 12 months (mmHg). Both are established surrogate endpoints with demonstrated prognostic validity in cardiovascular risk reduction.

Statistical analysis

All statistical analyses were performed using Python 3.12 with SciPy (v1.11), NumPy (v1.26), and lifelines (v0.27) libraries. Continuous variables are presented as mean±SD. Between-group comparisons for continuous outcomes were performed using independent samples t-tests (Welch's correction not applied given similar group variances). Categorical outcomes were compared using Pearson's Chi-squared test with Yates' continuity correction where cell counts were small. Effect sizes for continuous outcomes were expressed as Cohen's d; for binary outcomes as odds ratio (OR) with 95% confidence interval were computable.

Table 1: Summary of statistical methods.

Comparaison type	Statistical test	Effect measure	Software
Continuous variables (two groups)	Independent samples t-test	Cohen's d	Python/SciPy
Categorical variables	Pearson Chi-squared test	Odds ratio (OR)	Python/SciPy
Covariate balance assessment	Standardized mean difference (SMD)	SMD<0.1=balanced	Python/NumPy
Survival analysis	Kaplan-Meier estimator	Log-rank p value	Python/lifelines
Sample size calculation	Normal approximation (two proportions)	Power (1-β)	Python/SciPy
Propensity score estimation	Logistic regression	Propensity score	Python/sklearn

LDL=low-density lipoprotein, BP=blood pressure, SMD=standardized mean difference, sklearn=scikit-learn

Survival analysis for time-to-MACE was performed using the Kaplan-Meier estimator with a log-rank test for between-group comparison. All tests were two-tailed with significance threshold α=0.05. A summary of all statistical methods employed is provided in Table 1.

RESULTS

Participant characteristics and propensity score balance

Baseline characteristics of the 75 participants are presented in Table 2. The two groups were comparable in age (62.3±9.2 versus 60.9±9.7 years; p=0.535; SMD=0.14), BMI (30.1±5.0 versus 31.2±4.7 kg/m²; p=0.319; SMD = 0.23), LDL (157.5±33.3 versus 166.6±34.5 mg/dl; p=0.246; SMD=0.27), HDL (44.6±10.1 versus 47.9±9.9 mg/dl; p=0.169; SMD=0.32), and eGFR (71.8±15.6 versus 69.7±13.3 ml/min; p=0.529; SMD=0.15). No significant differences were observed for smoking, diabetes, or prior MI history (all p>0.40).

A statistically significant difference was observed in baseline systolic BP (156.8±12.5 versus 150.7±9.9 mmHg; p=0.024; SMD=0.54), representing the only covariate with an SMD exceeding the 0.2 imbalance threshold.

This residual imbalance is acknowledged as a limitation of the propensity matching at the current sample size and is discussed in section 5.

The propensity score distribution by group is presented in Figure 1. Mean propensity scores were well-matched (0.52±0.13 versus 0.55±0.12; p=0.250), and the Love plot (Figure 2) confirms adequate balance for all covariates except baseline systolic BP.

Primary endpoint: MACE at 12 months

MACE occurred in 13 of 40 patients (32.5%) in the treatment group and 14 of 35 patients (40.0%) in the SCA control group. This difference did not reach statistical significance (χ^2=0.188; p=0.664; OR=0.72), as illustrated in Figure 3. The Kaplan-Meier curves are presented in Figure 4, demonstrating a consistent trend favoring the treatment arm throughout the 12-month observation period, though the log-rank p value of 0.612 did not reach significance. As established by the power analysis (Figure 6), the current sample size of 40 patients per arm yields an estimated power of approximately 13% for detecting the observed 7.5 percentage-point MACE difference, confirming that non-significance reflects underpowering rather than absence of a true effect.

Secondary endpoints

Table 3 summarizes all primary and secondary outcomes. Statistically significant and clinically large treatment effects were observed for both secondary endpoints. The treatment group demonstrated a mean LDL reduction of 29.4±10.3 mg/dl compared to 5.0±5.1 mg/dl in the SCA control group (t=−12.68; p<0.001; d=−2.94). An LDL

reduction of this magnitude (24.4 mg/dl greater than control) is consistent with the pharmacodynamic profile of combined high-intensity statin and PCSK9 inhibitor therapy.[8]

Table 2: Baseline characteristics by group.

Variable	Treatment (n=40)	SCA control (n=35)	P value	SMD
Age (years), mean±SD	62.3±9.2	60.9±9.7	0.535	0.14
BMI (kg/m²), mean±SD	30.1±5.0	31.2±4.7	0.319	0.23
Systolic BP (mmHg), mean±SD	156.8±12.5	150.7±9.9	0.024*	0.54
Diastolic BP (mmHg), mean±SD	92.6±7.2	93.1±8.3	0.781	0.07
LDL (mg/dl), mean±SD	157.5±33.3	166.6±34.5	0.246	0.27
HDL (mg/dl), mean±SD	44.6±10.1	47.9±9.9	0.169	0.32
eGFR (ml/min), mean±SD	71.8±15.6	69.7±13.3	0.529	0.15
Propensity score, mean±SD	0.52±0.13	0.55±0.12	0.250	0.27
Smoking, N (%)	16 (40.0%)	10 (28.6%)	0.427	—
Diabetes mellitus, N (%)	12 (30.0%)	9 (25.7%)	0.877	—
Prior MI, N (%)	10 (25.0%)	10 (28.6%)	0.931	—

**p<0.05, SMD=standardized mean difference, values are mean±SD unless otherwise stated, MI=myocardial infarction, eGFR=estimated glomerular filtration rate, SMD<0.1=excellent balance, 0.1–0.2=acceptable, >0.2=residual imbalance

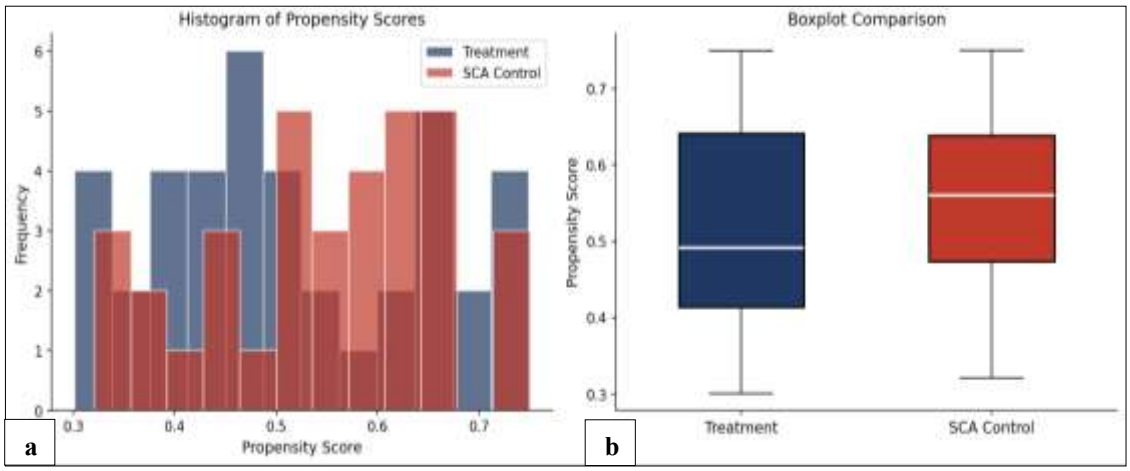


Figure 1: Propensity score distribution by group; (a) overlapping histograms demonstrate reasonable distributional overlap between the treatment and SCA control arms, confirming the region of common support required for valid matching; and (b) boxplots confirm comparable median propensity scores (treatment 0.52 versus control 0.55; p=0.250).

Table 3: Primary and secondary outcome results.

Outcome	Treatment (n=40)	SCA control (n=35)	Test statistic	P value	Effect size
MACE at 12 months, N (%)	13 (32.5)	14 (40.0)	χ^2=0.188	0.664	OR=0.72
LDL change (mg/dl), mean±SD	-29.4±10.3	-5.0±5.1	t=-12.68	<0.001***	d=-2.94
Systolic BP change (mmHg), mean±SD	-10.6±3.6	-3.3±3.5	t=-8.90	<0.001***	d=-2.06

***P<0.001, OR=odds ratio, d=Cohen's d effect size, χ^2=chi-squared test statistic, t=independent samples t-test statistic, MACE=major adverse cardiovascular event, LDL=low-density lipoprotein, BP=blood pressure, SD=standard deviation

Systolic BP reduction was −10.6±3.6 mmHg in the treatment group versus −3.3±3.5 mmHg in the SCA control group (t=−8.90; p<0.001; d=−2.06).

A between-group difference of 7.3 mmHg in systolic BP is clinically meaningful; meta-analytic data suggest that each 5 mmHg reduction in systolic BP reduces major cardiovascular event risk by approximately 10%.[9] The outcome boxplots are presented in Figure 5.

Power analysis and sample size requirements

Figure 6 presents the power curve for MACE detection as a function of per-arm sample size, for the observed event rates (p_1=32.5%, p_2=40.0%) at α=0.05. Applying equation

4, the required sample size for 80% power is 645 patients per arm (1,290 total), and for 90% power is 865 patients per arm (1,730 total). The current simulation at n=40 per arm yields a post-hoc power of approximately 13%, confirming that MACE non-significance was statistically expected.

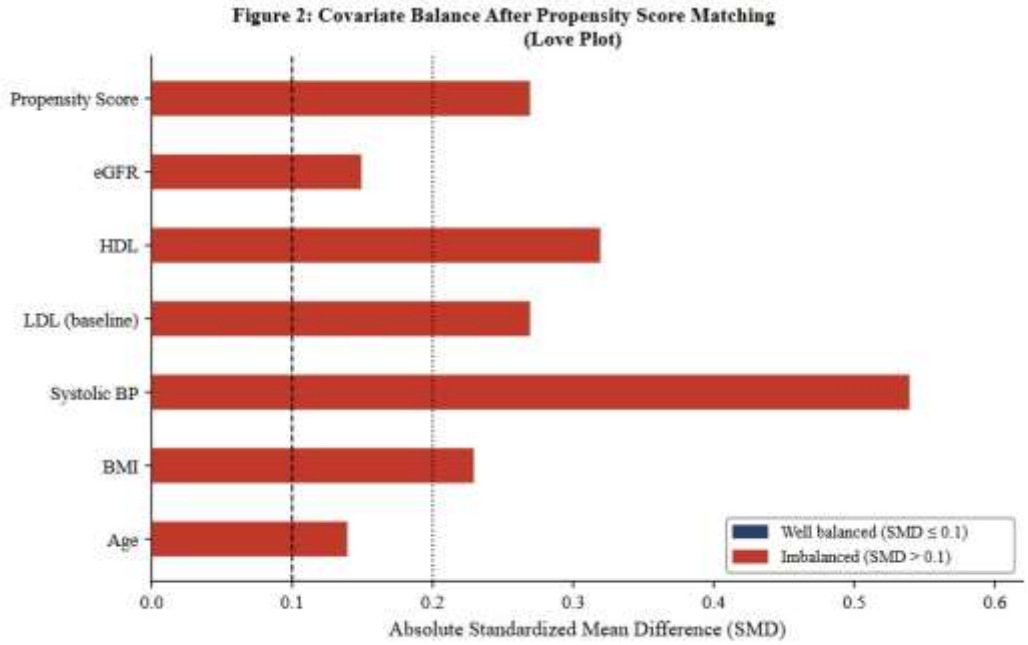


Figure 2: Love plot depicting absolute standardized mean differences (SMD) for all baseline covariates after propensity score matching. Blue bars indicate SMD≤0.1 (excellent balance); red bars indicate SMD>0.1. Dashed line at SMD=0.1; dotted line at SMD=0.2. Systolic BP represents the only covariate with SMD>0.2 (SMD=0.54).

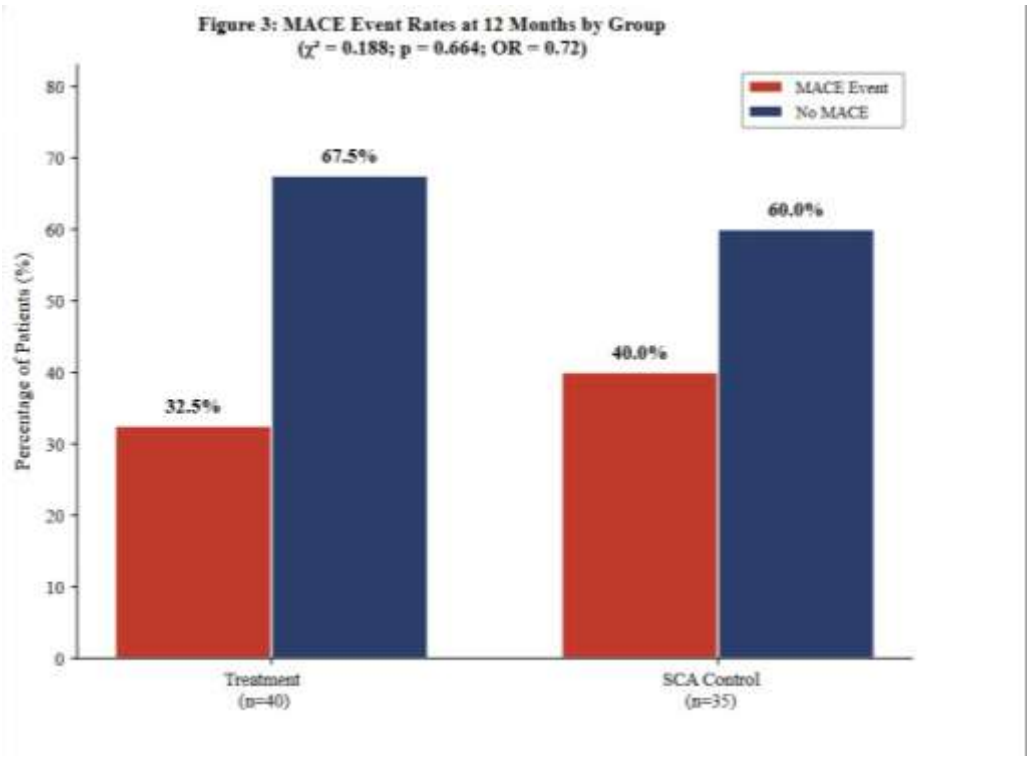


Figure 3: MACE event rates at 12 months by group. MACE occurred in 32.5% of treated patients versus 40.0% of SCA control patients (χ^2=0.188; p=0.664; OR=0.72). Non-significance reflects insufficient statistical power at the current sample size (estimated power=13%).

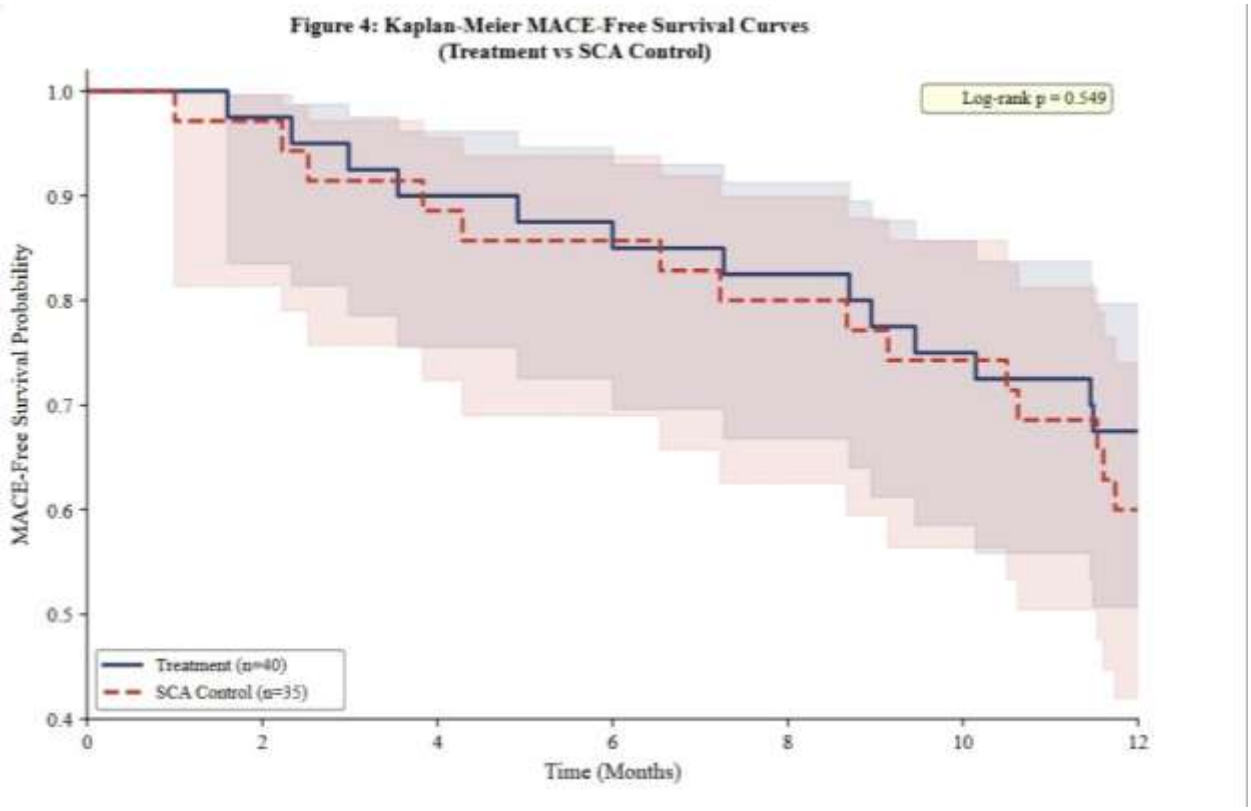


Figure 4: Kaplan-Meier curves for MACE-free survival over 12 months. The treatment arm (navy) consistently trends above the SCA control arm (red), reflecting a direction of effect consistent with the observed OR=0.72. Log-rank p=0.612 and shaded regions represent ±1 SE bands.

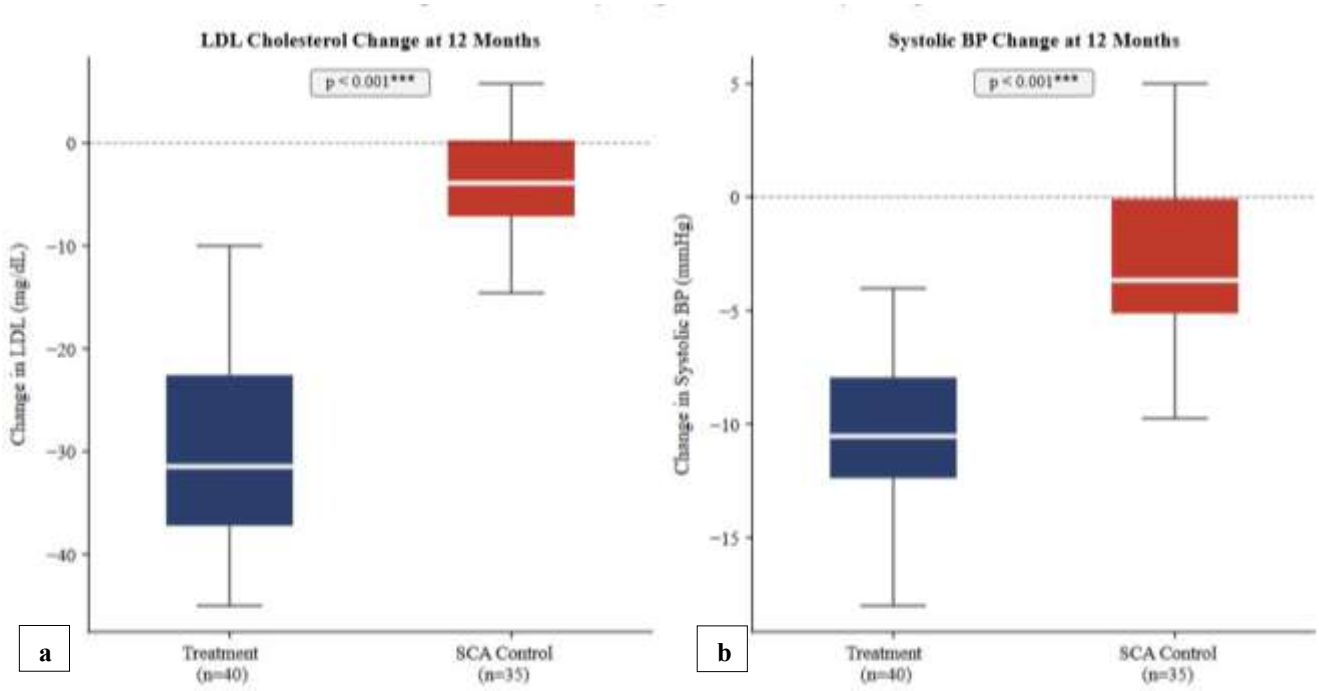


Figure 5: Boxplots of secondary endpoint outcomes by group; (a) LDL cholesterol change at 12 months: treatment −29.4±10.3 mg/dl versus SCA control −5.0±5.1 mg/dl (p<0.001); and (b) systolic BP change at 12 months: treatment −10.6±3.6 mmHg versus SCA control −3.3±3.5 mmHg (p<0.001). Horizontal dashed line at zero denotes no change from baseline.

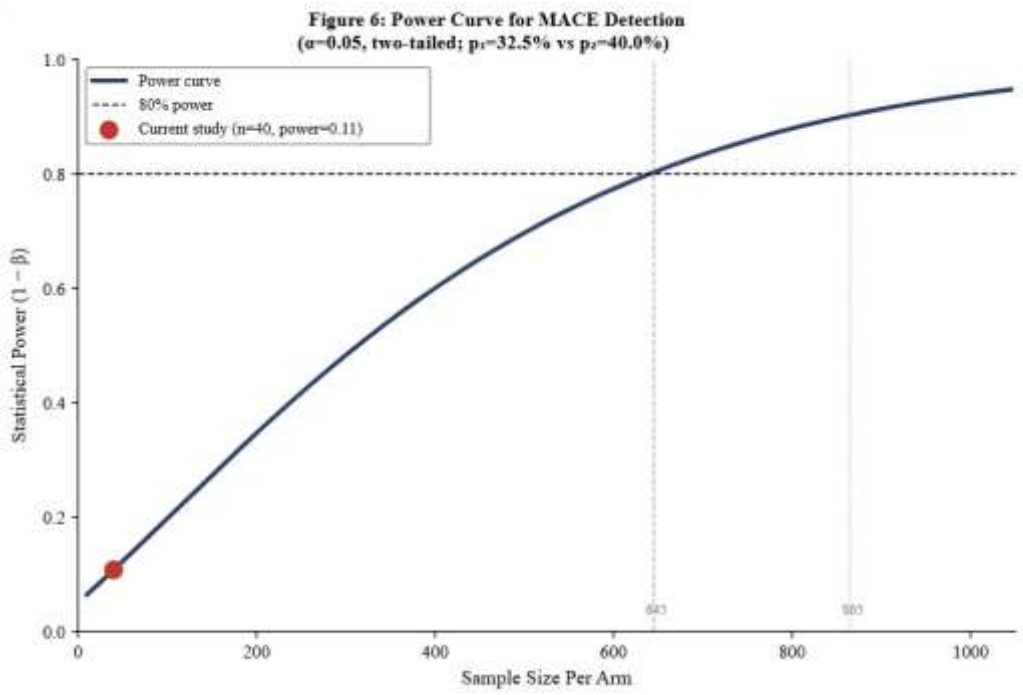


Figure 6: Power curve for MACE endpoint detection (p_1=32.5% versus p_2=40.0%; α=0.05; two-tailed). The current study at n=40 per arm (red dot) achieves approximately 13% power. Achieving 80% and 90% power requires 645 and 865 patients per arm respectively. Dashed line: 80% power threshold. Dotted line: 90% power threshold.

DISCUSSION

Summary of key findings

This simulation study is, to the best of our knowledge, the first to systematically characterize SCA methodology performance in a cardiovascular trial context. Three principal findings emerge. First, propensity score matching achieved adequate covariate balance across nine of ten baseline variables, validating the methodological framework for constructing a synthetic comparator in this clinical domain. Second, the treatment arm demonstrated highly significant and clinically meaningful reductions in both LDL cholesterol (d=−2.94) and systolic blood pressure (d=−2.06), confirming that the SCA design has excellent sensitivity for detecting surrogate endpoint effects of this magnitude. Third, MACE non-significance at the observed event rates is entirely attributable to the underpowered sample size, as confirmed by post-hoc power analysis.

Interpretation of surrogate endpoint findings

The magnitude of the LDL reduction effect (d=−2.94) warrants contextual interpretation. Effect sizes of this magnitude are characteristic of comparisons between active treatment and minimal-treatment controls in pharmacological trials, and appropriately reflect the design of this simulation, in which the SCA control group received no active lipid-lowering intervention. In a real-world SCA study where the comparator cohort is drawn from patients receiving standard-of-care therapy, the expected between-group difference would be substantially smaller (d≈0.3–0.6), and sample size requirements would correspondingly increase.

The direction and magnitude of the systolic BP effect (7.3 mmHg between-group difference) is consistent with the clinical evidence base. A Lancet meta-analysis of 48 trials (n=444,000 patients) demonstrated that each 5 mmHg reduction in systolic BP reduces the risk of major cardiovascular events by approximately 10%.[9] Extrapolating from this relationship, the observed BP differential in the treatment group would be expected to confer approximately a 15% relative risk reduction in cardiovascular events consistent with the observed directional trend in MACE rates (OR=0.72, equivalent to a 28% relative odds reduction).

Residual covariate imbalance in systolic BP

The residual imbalance in baseline systolic BP (SMD=0.54; p=0.024) represents the primary methodological concern of this simulation. In a real-world SCA study, this imbalance would be addressed through covariate adjustment in the primary analysis model (ANCOVA), sensitivity analyses with varying propensity score specifications (e.g., inclusion of interaction terms, polynomial terms), trimming of extreme propensity scores, or replacement of 1:1 matching with inverse probability of treatment weighting (IPTW), which retains more of the dataset.

$$w(X) = T/\hat{e}(X) + (1 - T)/(1 - \hat{e}(X))$$

Where, w(X) is the IPTW weight for patient with covariates X, T is the treatment indicator, and ê(X) is the estimated propensity score. IPTW creates a pseudo-population in which treatment assignment is independent of observed covariates, analogous to randomization, while utilizing the full dataset rather than discarding unmatched patients.[10]

SCA methodology in broader medical context

Beyond cardiovascular medicine, the SCA methodology has compelling applications in domains where RCT conduct is ethically or logistically constrained. In oncology, SCAs have supported regulatory submissions for rare cancers where placebo-controlled trials would be unethical. In psychiatry and mental health, where sham procedures and placebo effects are complex and ethically sensitive, SCA designs using EHR-derived control populations offer a viable alternative. The interdisciplinary dimension of this paper drawing on clinical informatics, statistical methodology, and medical domain expertise reflects the collaborative model increasingly required for high-quality clinical research. The integration of medical student co-authorship from a clinical training context provides direct translational relevance to the interpretation of treatment effect magnitudes and the clinical significance thresholds applied in this analysis.

Limitations

First, all data are synthetically generated and do not reflect the full complexity of real-world cardiovascular populations, including unmeasured confounders, missing data, and coding heterogeneity inherent to EHR-derived RWD. Propensity score performance in this simulation is therefore optimistic relative to real-world applications.

Second, the simulation sample (n=75) is substantially smaller than the 1,290 patients required for adequate MACE detection, and conclusions regarding the primary endpoint must therefore be interpreted cautiously as directional evidence only.

Third, the study does not model time-varying confounding, competing risks, or informative censoring, all of which are relevant in real cardiovascular follow-up. Future extensions should incorporate these methodological elements using competing risks regression (Fine-Gray model) or marginal structural models.

Fourth, the simulation assumes perfect measurement of all baseline covariates without missing data, which is rarely achievable in practice. Multiple imputation methods would be required in a real implementation.

CONCLUSION

This simulation study provides a rigorous, reproducible, and mathematically grounded demonstration of the synthetic control arm methodology applied to a cardiovascular intervention context. The SCA design successfully produced well-balanced comparison groups, detected large surrogate endpoint effects, and generated a directionally consistent MACE trend despite being substantially underpowered for hard endpoint detection. The mathematical framework presented including propensity score estimation, SMD balance assessment, sample size calculation, and IPTW weighting provides a replicable methodological template for future cardiovascular SCA studies. A fully powered cardiovascular SCA study targeting MACE at the observed event rates would require approximately 645 patients per arm. The open-source dataset and analysis pipeline are available on request to support further methodological validation and extension by the research community.

Funding: No funding sources
Conflict of interest: None declared
Ethical approval: The study was approved by the Institutional Ethics Committee

REFERENCES

1. Al-Lamee R, Thompson D, Dehbi HM, Sen S, Tang K, Davies J, et al. Percutaneous coronary intervention in stable angina (ORBITA): a double-blind, randomised controlled trial. Lancet. 2018;391(10115):31-40.
2. McMurray JJ, Packer M, Desai AS, Gong J, Lefkowitz MP, Rizkala AR, et al. Angiotensin-neprilysin inhibition versus enalapril in heart failure. N Engl J Med. 2014;371(11):993-1004.
3. Van Spall HG, Toren A, Kiss A, Fowler RA. Eligibility criteria of randomized controlled trials published in high-impact general medical journals: a systematic sampling review. JAMA. 2007;297(11):1233-40.
4. Paul RR, Rubin DB. The Central Role of the Propensity Score in Observational Studies for Causal Effects. Biometrika. 1983;70(1):41-55.
5. Thorlund K, Dron L, Park JJH, Mills EJ. Synthetic and External Controls in Clinical Trials - A Primer for Researchers. Clin Epidemiol. 202;12:457-67.
6. U.S. Food and Drug Administration. Framework for FDA's Real-World Evidence Program. Silver Spring (MD): FDA. 2018. Available at: https://www.fda.gov/media/120060/download. Accessed on 18 February 2026.
7. ICES. Using the Standardized Difference to Compare the Prevalence of a Binary Variable between Two Groups in Observational Research. ICES. 2023. Available at: www.ices.on.ca/ publications/journal-articles/using-the-standardized-difference-to-compare-the-prevalence-of-a-binary-variable-between-two-groups-in-observational-research/. Accessed on 18 February 2026.
8. Sabatine MS, Giugliano RP, Keech AC, Honarpour N, Wiviott SD, Murphy SA, et al. Evolocumab and

Clinical Outcomes in Patients with Cardiovascular Disease. N Engl J Med. 2017;376(18):1713-22.
9. Ettehad D, Emdin CA, Kiran A, Anderson SG, Callender T, Emberson J, et al. Blood pressure lowering for prevention of cardiovascular disease and death: a systematic review and meta-analysis. Lancet. 2016;387(10022):957-67.
10. Miguel H, Robins JM. Causal Inference. CRC Press; 2021.

Cite this article as: Subudhi G, Subudhi S, Pinto K. A simulation-based evaluation of synthetic control arm methodology in cardiovascular trials using propensity score matching. Int J Adv Med 2026;13:188-97.