Appendix 1

Outcomes definition

Stroke

New focal neurological symptoms lasting at least 24 hours with or without CT scan confirmation.

Thromboembolic event

Local clinical signs of persistent tor transient ischemia (acute loss of blood flow in a peripheral artery) supported by objective evidence of embolism.

Major bleed

Perioperative major bleed is classified according to VARC criteria as:

A) Major Overt bleeding is either associated with a drop in the hemoglobin of ≥3.0 g/dL or requiring transfusion of ≥3 U of whole blood or packed RBCs AND does not meet the criteria of life-threatening or extensive bleeding.
B) Extensive Overt source of bleeding with a drop in hemoglobin of ≥4 g/dL or whole blood or packed RBC transfusion ≥4 U within any 24-h period, or bleeding with a drop in hemoglobin of ≥6 g/dL or whole blood or packed RBC transfusion ≥4 U (BARC type 3b) within 30 days of the procedure.
C) Life-threatening Bleeding in a critical organ, such as intracranial, intraspinal, intraocular, or pericardial, necessitating surgery or intervention, or intramuscular with compartment syndrome OR bleeding causing hypovolemic shock or hypotension (systolic blood pressure <90 mmHg lasting >30 min and not responding to volume resuscitation) or requiring significant doses of vasopressors or surgery.
D) Fatal Bleeding adjudicated as being a proximate cause of death. Severe bleeding adjudicated as being a major contributing cause of a subsequent fatal complication, such as MI or cardiac arrest, is also considered fatal bleeding.

During FU, major bleed is defined as type 3a, 3b, 3c, 5a, and 5b as per BARC criteria.

Hospitalization with heart failure

Admission to the inpatient unit or ward in the hospital for 24 hours, including an emergency department stay. Hospitalizations planned for pre-existing conditions are excluded unless the baseline condition worsens. Symptoms, signs and laboratory evidence of worsening heart failure must be reported.

Sample size determination

The study is a randomized non-inferiority trial for occlusion with half-dose oral anticoagulant (HDOA) *vs.* occlusion with full-dose oral anticoagulant (FDOA).

- Primary endpoint (efficacy): composite endpoint of ischemic stroke, thromboembolic events, cardiovascular death.
- Primary endpoint (safety): composite endpoint of perioperative death, major bleeding.
- Secondary endpoint: incomplete appendage occlusion, ischemic stroke, thromboembolic events, cardiovascular death, overall death.

For two group models (i.e., HDOA as treatment and FDOA as control group with no covariates), we denote the parameter for the treatment group by μ_t and the parameter for the control group by μ_c. The default null and alternative hypotheses are given by

$$H_0: \mu_t - \mu_c \geq \delta$$

and

$$H_1: \mu_t - \mu_c < \delta,$$

where δ is a prespecified constant.

Let Θ_0 and Θ_1 denote the parameter spaces corresponding to $H_0$and H_1. Let $y^{(n)}$ denote the simulated current data associated with a sample size of n and let $\theta=(\mu_t,\mu_c,\tau_c)$ denote the model parameters. Let $\pi^{(s)}(\theta)$ denote the sampling prior and let $\pi^{(f)}(\theta)$ denote the fitting prior. The sampling prior is used to generate the hypothetical data while the fitting prior is used to fit the model after the data is generated. Let $\pi_0^{(s)}(\theta)$ denote a sampling prior that only puts mass in the null region, i.e., $\theta \subset \Theta_0$. Let $\pi_1^{(s)}(\theta)$ denote a sampling prior

https://dx.doi.org/10.21037/acs-2023-afm-18

that only puts mass in the alternative region, i.e., $\theta \subset \Theta_1$. To determine the Bayesian sample size, we estimate the quantity

$$\beta_{sj}^{(n)} = E_s[I\{P(\mu_t-\mu_c<\delta \mid y^{(n)},\pi^{(f)})\geq\gamma\}]$$

where j=0 or 1, corresponding to the expectation taken with respect to $\pi_0^{(s)}(\theta)$ or $\pi_1^{(s)}(\theta)$. The constant $\gamma>0$ is a prespecified posterior probability threshold for rejecting the null hypothesis (e.g., 0.975). The probability is computed with respect to the posterior distribution given the simulated data $y^{(n)}$, and the fitting prior $\pi^{(f)}(\theta)$, and the expectation is taken with respect to the marginal distribution of $y^{(n)}$ defined based on the sampling prior $\pi^{(s)}(\theta)$. Then $\beta^{(n)}_{s0}$ corresponding to $\pi^{(s)}(\theta)=\pi_0^{(s)}(\theta)$ is the Bayesian type I error rate, while $\beta^{(n)}_{s1}$corresponding to $\pi^{(s)}(\theta)=\pi_1^{(s)}(\theta)$ is the Bayesian power.

1. Primary endpoint (composite endpoint of ischemic stroke, thromboembolic events, cardiovascular death)

The basic model targets composite rates (a binary outcome) for treatment and control groups with no covariates. Patients with atrial fibrillation and CHA2DS2-VASc ≥3 with left auricle closed in thoracoscopy are randomized on the first day half-dose postoperative anticoagulant *vs.* continue full dose anticoagulant, annual telephone FU for various endpoints.

We consider the non-inferiority design application of Chen *et al.* (46).

Historical information can be borrowed from previously conducted OA and occlusion trials. Data are taken from Della Rocca *et al.* (37) and Cepas-Guillen *et al.* (47). *Table 1* summarizes the historical data.

Let $y_t^{(nt)}=(y_{t1},\ldots,y_{tnt})$ and $y_{c(nc)}=(y_{c1},\ldots,y_{cnc})$ denote the responses from the current trial for HDOA and the FDOA, respectively. The total sample size is $n=n_t+n_c$.

We assume the i-th observation from the test group y_{ti} follows Bern(μ_t), and the i-th observation from the control group y_{ci} follows Bern(μ_c).

A Bayesian sample size determination (SSD) approach incorporates historical data using the power prior with fixed $\alpha_0$and the normalized power for $\alpha_0$modeled as random.

The hypotheses for non-inferiority testing are

H0: $\mu_t-\mu_c\geq\delta$

and

H1: $\mu_t-\mu_c<\delta$,

where δ is a prespecified non-inferiority margin. We set δ=1%.

We choose beta(10^{-4},10^{-4}) for the initial prior for μ_c, which performs similarly to the uniform improper initial prior for $\log(\mu_{c1}-\mu_c)$ used in Chen *et al.* (46) in terms of operating characteristics.

Power is computed under the assumption that $\mu_t=\mu_c$ and type I error rate is calculated under the assumption that $\mu_t=\mu_c+\delta$.

For sampling priors, a point mass prior at μ_c=1% is used for $\pi^{(s)}(\mu_c)$ where 1% is the pooled proportion for the historical control datasets, and a point mass prior at $\mu_t=\mu_c$ is used for $\pi_{(s)}(\mu_t)$.

We use N=10,000, n_t/n_c=1, and γ=0.95 for all computations.

1.1 Power prior with fixed α0

When α0 is fixed, the historical matrix is fixed, each row represents a historical dataset, and the three columns represent the sum of responses, sample size and α0, respectively, of the historical control data. The FDA 2010 Guidance recommends α0=0.05 but this needs to be explored further. In a sensitivity analysis we evaluated a range of α0 values, from 0 to 0.4 by 0.05. Note that α0=0 coincides with non-informative prior and α0=1 with full borrowing.

We consider n_t values ranging from 50 to 100 to achieve the desired power of 0.8.

Since point mass sampling priors are used for μ_t and μ_c, samp.prior.mu.t and samp.prior.mu.c are both scalars.

For Bernoulli outcomes, beta initial priors are used for μ_t and μ_c, with hyperparameters specified by prior.mu.t.shape1, prior.mu.t.shape2, prior.mu.c.shape1 and prior.mu.c.shape2.

We can see that a sample size (test group) greater than 50 is required to achieve a power of at least 0.8 when α0>0.10 (*Figure S1*).

We then compute the type I error rate for these sample sizes.

Since the type I error rate is computed under the assumption that $\mu_t=\mu_c+\delta$, we use a point mass at μ_c=1% for the sampling prior for μ_c, and a point mass at μ_t=1%+1% for the sampling prior for μ_t (*Figure S2*).

2. Secondary endpoint: incomplete appendage occlusion, ischemic stroke, thromboembolic events, cardiovascular death, overall death

We will conduct a Bayesian analysis for the secondary endpoint if non-inferiority between devices is established for the primary endpoint. A Bayesian approach to compare proportions of incomplete appendage occlusion, ischemic

https://dx.doi.org/10.21037/acs-2023-afm-18

stroke, thromboembolic events, cardiovascular death, overall death between the test group and the control group is adopted. The observed data consists of the sample sizes (nt=100 and nc=100) and the number of migraine episodes (stand x_c) in the test and control groups, respectively.

The parameters section defines the unknown probabilities p_t and p_c for the test and control groups. These probabilities are assumed to follow a beta distribution with hyperparameters (0.001, 0.001), representing non-informative diffuse priors.

The model section specifies the likelihood of the observed data given the parameters. The binomial distribution models the number of successes out of the corresponding sample sizes for the test and control groups. The difference in proportions is calculated as the difference between p_t and p_c.

Four separate Markov chains will be run with a total number of iterations set to N=10,000 and 1,000 iterations used for warm-up or burn-in.

Instead of controlling the Type I error rate, Bayesian analysis allows us to assess the posterior probabilities of hypotheses and make decisions based on those probabilities. *Figure S3* shows the posterior distribution of the difference corresponding to the scenario in which the observed migraine rates $x_t/n_t=x_c/n_c$ are both set at 15% according to the historical information.

The Bayesian model is specified using the Stan modeling language through R software v4.3.0 (R Core Team. 2023. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing).

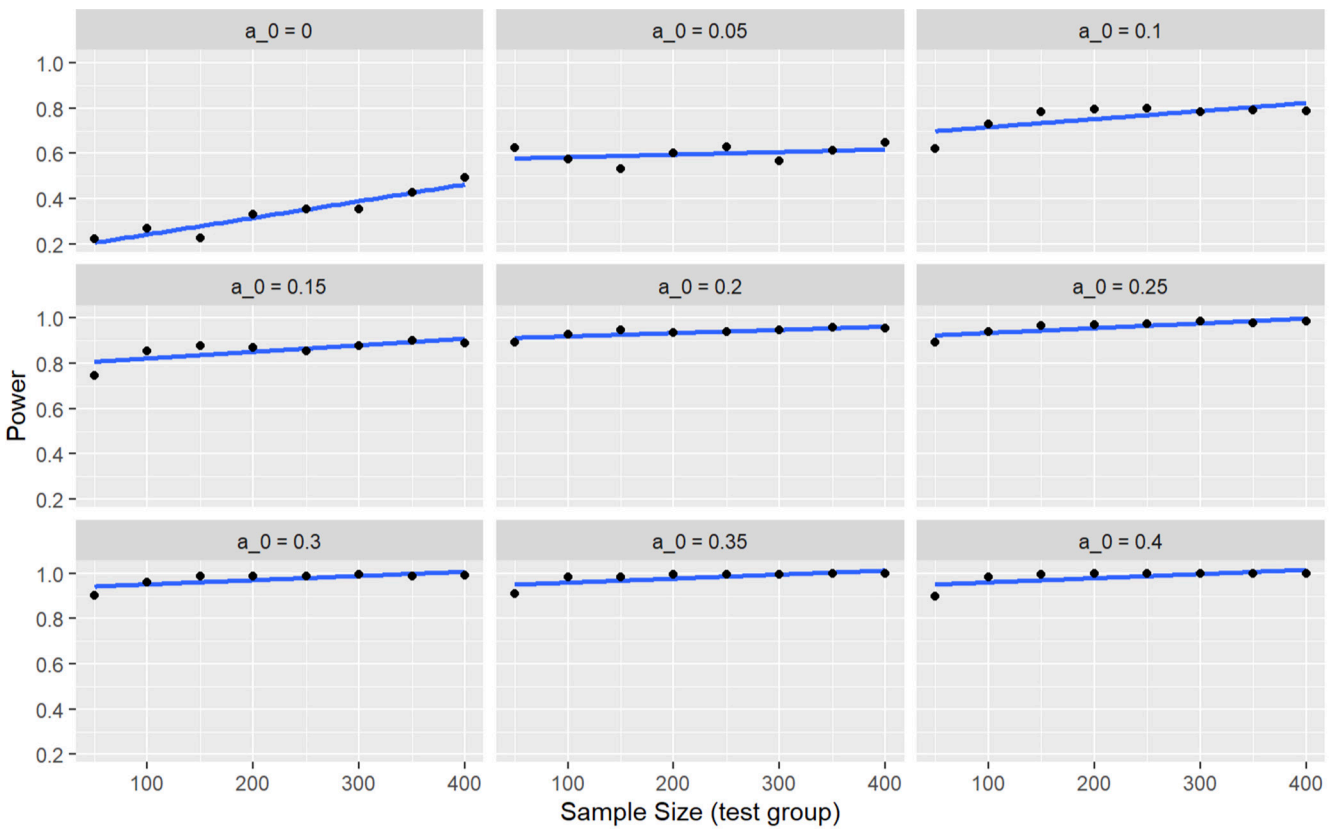

Figure S1 Power curve for sample sizes ranging from 50 to 400.

https://dx.doi.org/10.21037/acs-2023-afm-18

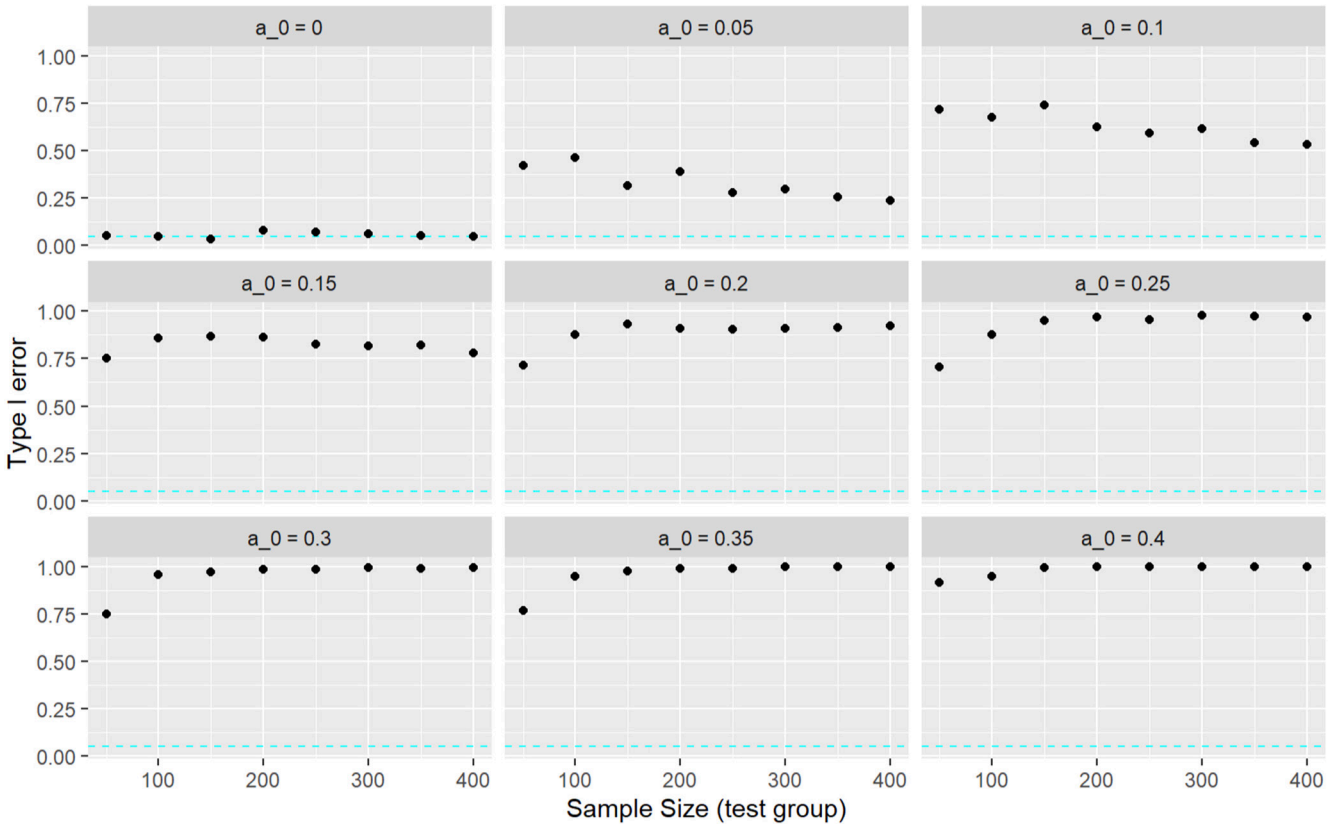

Figure S2 Type I error curve for sample sizes ranging from 50 to 400.

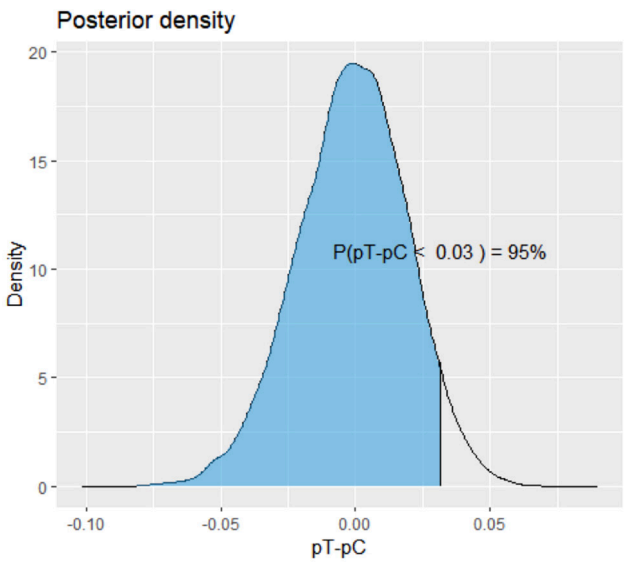

Figure S3 Curve showing the posterior distribution of the difference for observed rates $x_t/n_t=x_c/n_c$ set at 15%.

https://dx.doi.org/10.21037/acs-2023-afm-18