Appendix A: Supplementary Methods

Appendix A1. Analysis Details

This analysis includes all trials that were identified for a systematic review of studies looking at recurrent stroke with patent foramen ovale (PFO). The SCOPE PI (Kent) was part of the team that performed this systematic review, which was updated in August 2019 for guideline development by the American Academy of Neurology (AAN). And subsequently updated to September 2021 for this article. Based on this systematic search performed of Medline and Embase, these studies represent the *totality of available randomized evidence on the use of percutaneous implanted devices for PFO closure versus medical therapy in patients with PFO-associated cerebral ischemic events*. Complete information about the search strategy and systematic review can be found in the original guidance.[1] **Appendix B1** shows a PRISMA flowchart of all studies identified.

All RCTs identified in the systematic review provided individual patient-level study data. Data entered into the central SCOPE database were a limited dataset (LDS), with all high-level patient identifiers removed. All data were collected under the aegis and supervision of the SCOPE Steering Committee, and integrated and stored at the Predictive Analytics and Comparative Effectiveness (PACE) Center at Tufts Medical Center, Boston, MA. The data were harmonized and analyzed by two statisticians at the PACE Center, Tufts Medical Center to ensure they accurately matched the values reported by the trials. **Appendix A5** describes variables that were harmonized, including ASA and shunt size. There were no issues identified in checking IPD.

The PI of this study (Kent) developed an initial list of variables based on variables used in a prior 3-trial individual patient meta-analysis[2] and variables that make up the Risk of Paradoxical Embolism (RoPE) Score[3,4]. The list was further expanded and refined at an investigator meeting in February 2020. **Appendix Table 1** displays the variables collected.

Appendix B4 provides the patient-level characteristics for each study, and note where data was missing.

All analyses were conducted using SAS (Version 9.4) and R (Version 4.0.2).

Examination of proportional hazards assumption

Proportional hazards assumptions were assessed using graphical and statistical test-based methods. Visual assessment of the log-log survival curve for each treatment group in each trial was used to detect violations of proportionality. Time-dependent covariates — interactions between the predictors and log(time) — were included to assess proportionality for each predictor. Additionally, tests of proportional hazards assumption was based on scaled Schoenfeld residuals for each predictor and overall (global test).[5] No visual or statistical violation of proportional hazards was observed.

Handling of missing data

Missing values for covariates were imputed using fully conditional specification methods (predictive mean matching for continuous variables and discriminant function method for all dichotomous variables) to generate 10 complete data sets.[6] The imputation model for each variable with missing values included all pre-specified covariates and the outcome. Analyses were conducted in each of the 10 compete data sets separately and pooled using Rubin's Rules.

Random effects Cox proportional hazards regression

Study-specific random effects were modeled using SAS PROC PHREG procedure using the RANDOM statement to fit a shared frailty model for clustered data.[7] The log-normal distribution of shared frailty was used and the common variance parameter (covariance estimate = 0.13; asymptotic standard error = 0.12) was estimated using residual maximum likelihood.

Assessment of linear assumption

The functional form of continuous variables (age and RoPE Score) was assessed for linearity using higher order polynomial terms (i.e., quadratic). These higher order terms were tested for statistical significance and model fit was assessed by differences in likelihood ratio compared to models with a linear relationship. We found no evidence of statistically significant non-linear associations with the treatment effect.

Appendix Table 1. Variables of Interest.

Category	Variable
Clinical Variables	Age (at time of stroke)
	Sex
	Coronary artery disease
	Diabetes
	Hypertension
	Hyperlipidemia
	Prior spells: number, date(s), event(s)
	Smoking status: current
	Body Mass Index
	Index event: stroke or TIA
	Index event: date
	Medication at index event: statin, antiplatelet, anticoagulant, CP/HRT
Echocardiographic Variables	Mobility of septum: normal, hypermobile
	PFO size: large, not large
	Shunt at rest: yes, no
Neuroradiology Variables	Index stroke seen: yes, no

	Location: superficial, deep
	Size: large, small/not seen
	Multiple: yes, no (not seen = single)
	Prior stroke: yes, no
Treatment Variables	Warfarin (anticoagulant, Coumadin)
	Antiplatelets
Follow-Up Variables	Date of last follow-up
	Duration of follow-up
	Recurrent stroke
	Recurrent TIA
	Date of recurrent event
	Death
	Date of death
	Cause of death
	PFO closure (treatment)
	Atrial Fibrillation, all and after 45 days (safety)
	Major Bleeding (safety)
	Procedural complication (safety)
Cohort Designation and Randomization	Intent-to-treat group (closure vs. medical therapy)
	Per-protocol group (closure vs. medical therapy vs. excluded)
	As-treated group (closure vs. medical therapy vs. excluded)

CP, contraceptive pill; HRT, hormone replacement therapy; PFO, patent foramen ovale; TIA indicates transient ischemic attack.