Potentially Large yet Uncertain Benefits
A Meta-analysis of Patent Foramen Ovale Closure Trials

Georgios D. Kitsios, MD, MS, PhD; David E. Thaler, MD, PhD; David M. Kent, MD, MS

Until recently, the evidence base for patent foramen ovale (PFO) closure with a device to prevent future cerebrovascular events in patients with cryptogenic stroke and presumed paradoxical embolism consisted of a large body of >50 observational studies that cumulatively showed substantial and unequivocal benefits with this intervention. On the basis of this observational data and the compelling pathophysiologic rationale, PFO closure to prevent stroke recurrence had become routine in many centers, and off-label closure of PFOs hampered enrollment in ongoing trials.

However, within the past year, the first true clinical experiments on PFO closure were published, and 3 consecutive randomized clinical trials (RCTs) investigating 2 different devices (STARFlex in CLOSURE-I and Amplatzer PFO Occluder in RESPECT and PC trial) failed to show evidence of statistically significant benefit in their primary outcome. Yet, despite the consistently null overall outcome of these trials, on closer inspection, their results seem well calibrated to sustain the controversy; in particular, the results of the RESPECT and PC trial, published simultaneously, seem highly suggestive of benefit for mechanical closure over medical therapy. Thus, we aimed to explore more thoroughly the results from the available randomized evidence via meta-analysis, which can sometimes overcome the uncertainty in individual trials through their pooling.

We considered the outcome of recurrent stroke as our primary outcome (as opposed to also including transient ischemic attacks), given its more standardized definition and its long-term clinical import. We additionally explored composite outcomes in sensitivity analyses, and apart from the intention-to-treat analysis, we also synthesized the per-protocol results of the trials to appreciate the effects of crossovers and dropouts. For the statistical synthesis, we considered the generally accepted random effects model, which accounts for possible between-trial treatment-effect heterogeneity. In sensitivity analysis excluding the CLOSURE-I trial, we also explored a fixed effects model for the 2 RCTs using the same device, appropriate only under a strong assumption of a single true effect of closure using the Amplatzer device.

First, we examined the absolute event rates of recurrent stroke, incidence rates with their 95% exact Poisson confidence intervals (95% CIs) in device and medical groups separately (Figure 1). By random effects meta-analysis, the summary incidence rate of stroke was relatively low in both groups; in the device groups, it was 0.76 (95% CI, 0.30–1.96) per 100 person-years versus 1.30 (95% CI, 0.94–1.81) per 100 person-years in medical groups. Statistically significant heterogeneity of incidence rates in the device groups (P for heterogeneity=78%; P=0.01) may reflect between-trial differences in study design. Important differences may have included variation in the stringency of patient selection, medical therapy received on either arm, the presence of true, device-specific effects of closure, or differences in outcome ascertainment methods (which were, for example, much more assiduous in RESPECT than in the PC trial). In fact, a concerning signal for differential referral for outcome adjudication in the device group was noted in the PC trial, which may, in part, contribute to the exceedingly low incidence rate in this trial.

Next, we conducted a series of meta-analyses for the treatment effects estimated as hazard ratios. In our primary results for the outcome of stroke, closure showed a nonsignificant 45% reduction in risk with all 3 RCTs combined (summary hazard ratio, 0.55 [95% CI, 0.26–1.18]; Figure 2). The results for the outcome of stroke were also nonsignificant when the 2 Amplatzer device trials were combined without the CLOSURE-I trial included, or when the per-protocol results were synthesized (Table). Yet the absence of significance is not universally the case across all analyses. Importantly, for the composite primary outcome for which each RCT was powered (nonidentically defined across studies), a borderline statistically significant effect was found for the intention-to-treat analysis only (summary hazard ratio, 0.67 [95% CI, 0.44–1.00]). The treatment effect was larger still and more statistically significant when the 2 Amplatzer device trials (RESPECT and PC trial) were synthesized with a fixed effects model (summary hazard ratio, 0.41 [95% CI, 0.19–0.88]), although a random effects model is probably more appropriate given differences in the design and conduct of these trials. No significant effects were found for the composite of stroke or transient ischemic attack.

We further explored sources of potential treatment-effect heterogeneity with subgroup meta-analyses based on available data from RCTs for their primary outcomes (Figure 3).
However, no significant tests of interaction were found (all $P$ values $>0.10$) and CIs were wide, indicating substantial uncertainty around available estimates, despite some interesting yet nonsignificant signals, such as the case of substantial shunt on echocardiography. Nonetheless, our analyses show that biological plausibility of paradoxical embolism (eg, younger age or presence of atrial septal aneurysm) does not necessarily translate into clinically detectable effect modification.

Our meta-analyses are limited by the fact that the primary composite outcome definitions (and also the actuarial component events) varied across studies, and thus, interpretation of some of our summary estimates should be viewed only as exploratory. Furthermore, our analyses for potential treatment-effect modifiers are restricted to subgroups analyzed by the individual component trials because this meta-analysis is based on reported results, not individual patient data.

The main message of these analyses is that the uncertainty of the individual trials is not resolved by combining their results through meta-analysis. The wide CIs of the individual trials are transmitted in the main results of the combined analysis and appear broad enough to accommodate the mutually conflicting viewpoints on the benefits of closure. The uncertainty is only underscored by the conflicting results of different meta-analytic approaches.

Despite the absence of formal statistical significance in our main analysis, the magnitude of these summary effect sizes is reminiscent of the large effect sizes seen in observational studies. In particular, meta-analyses of the Amplatzer device trials provided strong summary effects that fell just short of the significance threshold. Larger sample sizes and longer follow-ups could have potentially solidified the benefits of PFO closure. However, the very low numbers of available data points (eg, a single stroke event in the device group of the PC Trial) makes these results extremely vulnerable to even slight shifts of event numbers between groups, either because of any type of bias or because of random error, and the effect estimate might attenuate substantially (or even completely) with more outcomes. Thus, closing the path for further randomized evidence on the basis of this promising signal leaves substantial uncertainty about the best options for our patients.
Survivors of cryptogenic stroke are typically of young age, with the index event in itself being a harbinger of mortality and recurrence potentially catastrophic; whatever risk is conferred by PFO in those with an index event presumably persists for life and accrues over time. Withholding an intervention that can potentially cut this risk in half for the sake of statistical purity can seem unjustified and unethical. However, any risks of an implanted device in the septum of a heart would also be near-permanent and accrue over time, including device-related thrombus formation or increase in the risk of atrial fibrillation. What we owe patients, then, is not just the option to close their hole with an appropriate device, but better information in a reasonable time frame about the risks and benefits of doing so. Unfortunately, in this case, meta-analysis is no substitute for more randomized data.

Sources of Funding
This study was partially funded by grants UL1 RR025752, R01 NS062153, and R21 NS079826, all from the National Institutes of Health.

Disclosures
Drs Kent and Thaler have consulted for WL Gore Associates. Dr Thaler is a consultant to AGA Medical Corporation. The other author reports no conflict.
References


Key Words: foramen ovale, patent  meta-analysis  patent foramen ovale closure  secondary prevention  stroke
Potentially Large yet Uncertain Benefits: A Meta-analysis of Patent Foramen Ovale Closure Trials
Georgios D. Kitsios, David E. Thaler and David M. Kent

Stroke. 2013;44:2640-2643; originally published online July 18, 2013;
doi: 10.1161/STROKEAHA.113.001773
Stroke is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2013 American Heart Association, Inc. All rights reserved.
Print ISSN: 0039-2499. Online ISSN: 1524-4628

The online version of this article, along with updated information and services, is located on the
World Wide Web at:
http://stroke.ahajournals.org/content/44/9/2640