STATISTICAL ANALYSIS PLAN

Patent foramen ovale with the association with perioperative ischemic stroke

Primary Investigator: Matthias Eikermann, MD PhD

Statisticians: Timothy T. Houle, PhD, Sara Burns, MS

Original Date: Jan 31, 2017

Updated: Mar 2, 2017

Note: Revision after consultation with external collaborators. Costs of hospital stay added as secondary outcome and falsification analysis added.

Updated: Jun 23, 2017

Note: Revision in response to peer review (NEJM). Additional confounders added to the model: coronary artery disease, congestive heart failure, pulmonary edema, pulmonary hypertension, cardiomyopathy, congenital heart disease, valvular heart disease, hypercoagulable state, deep vein thrombosis, pulmonary embolism, and systemic embolic phenomenon.

Updated: Nov 15, 2017

- Note: Revision in response to peer review (JAMA). Several post hoc subgroup analyses were added for the primary outcome: patients who underwent transthoracic vs transesophageal echocardiography, and patients who underwent agitated saline testing. Additional falsification testing in the subgroup of patients who had history of echocardiogram was included. A mixed effects logistic regression model to account for the impact of variation from different healthcare facilities was added as post hoc analyses. Exploratory analyses to adjust of different surgical services and duration of data availability were added. Finally, the association between PFO and post-discharge stroke were added in response to reviewer concerns.

Primary Objective

- To investigate whether or not an association exists between PFO and perioperative ischemic stroke.
    - H_0: p(stroke|PFO+) = p(stroke|PFO-)
    - H_1: p(stroke|PFO+) ≠ p(stroke|PFO-)

The primary hypothesis is that the probability (p) of ischemic stroke for those with patent foramen ovale (PFO+) is *greater than* the probability of stroke for those without patent foramen ovale (PFO-). Although the hypothesis is directional, the statistical hypothesis testing will be two-tailed.

Secondary Objectives

- To investigate whether or not an association exists between PFO and:
    - 30-day readmission
    - 30-day mortality
    - costs of hospital stay
    - stroke subtype by Oxford Community Stroke Project
    - stroke-related neurologic deficit (NIHSS)

Statistical Methodology

Data Sources

Data will be obtained from the MetaVision Anesthesia Information Management System (AIMS) (iMDsoft, Dedham, MA), the Research Patient Data Registry (RPDR), and Enterprise Performance Systems Inc (EPSi) (Allscripts Healthcare) at Massachusetts General Hospital. The AIMS prospectively collects intraoperative data including physiological parameters such as blood pressure, ventilator and respiratory indices, administered drug doses, and fluid volumes. This is matched to the patients' demographic data and pre-/post-procedural condition using RPDR, a centralized clinical data warehouse that compiles health records and billing data from various Partners hospital systems specifically for research purposes. Information on hospital length of stay and costs will be collected through EPSi. Work relative value units (work RVU), a marker of operation procedural complexity, will be recorded.

Outcomes

The primary outcome is perioperative ischemic stroke within 30 days of surgery. The secondary outcomes are 30 day hospital readmission, 30 day mortality, and costs of hospital stay. Hospital readmission is defined as an in-patient readmission to the same hospital.

Statistical Methods

All analyses will be performed with pre-specified endpoints and statistical methods. Associations between PFO status and primary and secondary outcomes will be examined using both unadjusted (crude) and adjusted methods. Unadjusted analyses will be conducted using chi-square tests for categorical variables and Student's t-test or Wilcoxon rank-sum tests for continuous variables.

The primary analysis will utilize multivariable logistic regression modeling to evaluate the relationship between having a PFO and perioperative ischemic stroke while controlling for confounding variables selected a priori based on data in the published literature and biological plausibility. This model will be conducted using forced variable entry and evaluated using calibration tests (Hosmer-Lemeshow) and discrimination indices (AUC). Covariates included in the model will include baseline patient characteristics such as age, sex, body mass index, ASA physical status classification, and Charlson comorbidity index; coexisting conditions such as history of cigarette smoking, hypertension, diabetes mellitus, dyslipidemia, coronary artery disease, myocardial infarction, congestive heart failure, pulmonary edema, pulmonary hypertension, cardiomyopathy, congenital heart disease, atrial fibrillation, valvular heart disease, chronic obstructive pulmonary disease, migraine, chronic kidney disease, hypercoagulable state, deep vein thrombosis, pulmonary embolism, and systemic embolic phenomenon; prescription within 28 days before surgery of beta-blockers, statins, anti-platelet agents, and anticoagulants. Factors relating to surgery, including emergency surgery status, inpatient surgery, high risk surgical service, duration of surgery, intraoperative hypotensive minutes, intraoperative dose of vasopressors, intraoperative fluid volumes, requirement for packed red blood cells transfusion, and work relative value units – a marker of procedural complexity, will also be adjusted by forcing them into the model.

Secondary models will be conducted using distributions and linking functions appropriate to the outcomes under study. For example, the association between PFO and ischemic stroke subtype, using the Oxfordshire Community Stroke Project (OCSP) classification, will be analyzed with a multinomial logistic regression model. The association between PFO and costs of hospital stay will be analyzed with a negative binomial regression model and log link.

Data management and statistical analyses will be performed in Stata software, version 13 (StataCorp LP) and R Studio software, version 3.2.5 (R Foundation for Statistical Computing). Where appropriate, statistical significance will be interpreted using a two-tailed P value of less than 0.05.

Planned Sensitivity Analyses

To address potential bias in the diagnosis of PFO that may not be considered in the multivariable confounder model and to further consider the average treatment effect for those treated (ATT), a logistic regression model that estimates the likelihood of diagnosed PFO will be constructed based on coexisting medical conditions that may subject patients to echocardiography studies. The association of PFO with perioperative ischemic stroke will then be tested in a 5:1 propensity-score-matched cohort. Participants will be matched using the 'Matchit' package in R Studio, using a 1:5 matching ratio with nearest neighbor and sampling without replacement with a caliper set to 0.20. A direct comparison in

the incidence between those with stroke in both PFO exposure groups will be examined using a non-conditional (i.e., no blocking using pairings) logistic regression model.

The primary analysis will be repeated in a subgroup including only patients with a history of a documented echocardiogram in the same institution, in order to further control for unmeasured differences that biased the referral for evaluation by echocardiogram.

A probability score for the baseline risk of perioperative ischemic stroke independent of PFO diagnosis will be created based on comorbid conditions and surgical factors that are traditional risk factors for stroke, and the PFO-stroke association will be re-examined for heterogeneity across this baseline risk. Effect modification on the association between PFO and perioperative ischemic stroke by patient's baseline stroke risk will be tested by estimating a model with this risk score, PFO status, and introducing an interaction term (risk x PFO status) into the multivariable regression model.

Falsification testing will be performed with three postoperative outcomes selected based on a common contributing etiology of non-thrombotic tissue ischemia, but unlikely to be causally related to the presence or absence of PFO. The association between PFO status and each of these outcomes will be evaluated using the same multivariable model as the primary analysis. It is expected that no meaningful association between PFO and these outcomes will be observed.

The complete case method will be used in the primary statistical analysis. This approach assumes that the missing data are ignorable conditional on the covariates in the model. To further examine the potential impact of excluding cases with missing data, the primary analysis will be repeated with the entire cohort using the technique of multiple imputations by chained equations – 5 iterations with 5 imputations per iteration will be performed.

Planned Exploratory Analyses

Effect modification on the association between PFO and perioperative ischemic stroke by the occurrence of postoperative deep vein thrombosis (DVT) will be tested by adding an interaction term (PFO x DVT) to the primary multivariable model.

The relationship between PFO and other perioperative complications of embolic etiology, such as acute limb ischemia and renal artery embolism, will also be explored. Specifically, the association between PFO and:

- Perioperative systemic embolic complications – a composite of acute embolic events in the extremities, kidneys, spleen, splanchnic circulation, and retina within 30 days after surgery,
- Acute limb ischemia,
- Renal artery embolism, and
- Acute intestinal vascular insufficiency will be examined.

These binary outcomes will all be evaluated using the same logistic regression model as the primary analysis.

Sample Size and Power

The analyses will be conducted based on all available data from the observational period. A statistical power calculation was conducted to define the power available to detect a clinically meaningful effect size. For this calculation, we will define an odds ratio of 2.0 as a clinically meaningful association between PFO and perioperative ischemic stroke. Assuming an observed PFO rate of 1.0%, a one-sided alpha level of 0.025, and an events rate of 0.5% perioperative ischemic stroke within 30 days after surgery, we achieved 94.3% power to detect an odds ratio of 2.0 or greater. Previous studies have examined the PFO-stroke association in smaller sample sizes, but would have had modest levels of statistical power using our event rates.

Post Hoc Sensitivity Analyses

- Subgroup analysis of patients who underwent transthoracic vs transesophageal echocardiography
- Subgroup analysis of patients who underwent agitated saline testing
- Falsification testing in subgroup of patients who had history of echocardiography
- Mixed effects logistic regression to account for the impact of variation from different healthcare facilities

Post Hoc Exploratory Analyses

- Association between PFO and perioperative new onset atrial fibrillation
- Association between PFO and perioperative myocardial infarction
- Association between PFO and post-discharge stroke
- Adjustment for different surgical services
- Adjustment for duration of data availability