

WATCHMAN® PROTECT AF Study
Rev. 6

Protocol Synopsis

Title

WATCHMAN® Left Atrial Appendage System for Embolic PROTECTion in Patients with Atrial Fibrillation (PROTECT AF)

Sponsor

Atritech/Boston Scientific

IDE

G020312

Protocol Number

ST1021 Rev 06

Device

WATCHMAN Left Atrial Appendage System
WATCHMAN Device: 21mm, 24mm, 27mm, 30mm, 33mm
ACCESS SYSTEM: Double curve and Single curve Transseptal Access Sheath

Study Design

Multicenter, prospective randomized study stratified by center, with 2 patients randomized to the WATCHMAN group for each control patient, comparing the WATCHMAN device to long-term Coumadin® therapy, demonstrating the treatment arm is non-inferior to the control arm.

Primary Endpoint

Primary Effectiveness: All stroke (including ischemic and hemorrhagic), cardiovascular death (limited to any cardiovascular and unexplained), and systemic embolism.

Primary Safety: Life-threatening events as determined by the Clinical Events Committee which would include events such as device embolization requiring retrieval and bleeding events e.g., pericardial effusion requiring drainage, cranial bleeding events due to any source, gastrointestinal bleeds requiring transfusion, and any bleeding related to the device or procedure that necessitates an operation.

Secondary Endpoint

- TIA [defined as an acute focal neurological event lasting at least 5 minutes and up to 24 hours that is MR imaging negative]. All TIAs will be adjudicated by the CEC.
- Other individual complication rates including, but not limited to MI and death.

For the WATCHMAN group:

- Technical Success defined as successful delivery and release into the LAA including successful recapture and retrieval if necessary;
- Procedure Success defined as technical success and no serious adverse events related to the treatment or procedure within the hospital stay;
- 30-day Major Complication Rate defined as death, stroke, MI or any other

serious adverse events related to the treatment or procedure within the first 30 days or through hospital discharge (whichever is longer);

- Individual complication rates including, but not limited to hematomas, and pseudoaneurysms.

For the Coumadin Control group:

- Non-therapeutic INR >3.0 INR, <2.0 INR and stopped therapy¹
- Excessive anticoagulation INR>4.0
- Bleeding complications: hematuria, rectal bleeding, epistaxis, bleeding from varicose veins, oral bleeding, prolonged bleeding from a laceration, or bruising-hematoma, hemothorax, red eye, thrombosis.

Follow-up Schedule

All randomized study patients will complete follow-up assessments at post-randomization intervals of 45 days, 6 months, 9 months, 12 months and semi-annually thereafter for the duration of the study.

Number of Patients

The study will enroll sufficient patients to accumulate a minimum of 600 patient followup years and a maximum of 1500 patients.

Number of Centers

Up to 60 centers across the United States will participate in the study.

Design Manuscript

Fountain RB, Holmes DR, Chandrasekaran K, et al. The PROTECT AF (WATCHMAN Left Atrial Appendage System for Embolic PROTECTION in Patients with Atrial Fibrillation) trial. *Am Heart J.* 2006 May;151(5):956-61

¹ACC/AHA/ESC Practice Guidelines, Fustest et al., JACC Vol. 48:No. 4 2006, August 15, 2006, e149-246

Statistical Analysis Plan

PROTECT AF: IDE#G020312

Table of Contents

Introduction.....	3
Study and Analysis Overview.....	3
General Statistical Methods.....	3
Efficacy Analyses.....	4
Primary Efficacy Endpoint Analyses.....	4
Criterion for Non-inferiority.....	5
Criterion for Futility (Not Non-inferiority).....	5
Criterion for Superiority.....	5
Sequential Analysis Plan.....	5
Analysis of Two Year Success Rates.....	6
Protection against Possibly Non-Constant Hazard.....	7
Additional Efficacy Analyses.....	7
Safety Analyses.....	7
Primary Safety Endpoint.....	7
Components of the primary safety endpoint.....	8
Unstratified model.....	8
Severity of Adverse Events.....	8
Mortality.....	8
Device and Procedure Related events.....	8
Early vs. Late Implants.....	9
Additional Supportive Analyses.....	9
Sensitivity Analysis of Analysis Populations.....	9
Assessment of Treatment Effect Consistency.....	10
Comparison of Device Types.....	10
Analysis of Protocol Violations.....	10
Patient Accountability.....	10

Introduction

This document describes the analysis that will be performed for the PROTECT AF Trial, IDE #G020312. This document expands on and clarifies analyses that were previously described in the protocol and the prior statistical analysis plan, most recent version IDE Supplement #S029, April 17, 2006. Moreover, this includes analyses that have been discussed previously in multiple communications between Atritech and the FDA. For added readability and clarity, supportive information from previous versions of the analysis plan that was used to design the trial is no longer incorporated in this document. This version supersedes all other descriptions of planned statistical analyses.

Study and Analysis Overview

This is a randomized, prospective, multi-center study to determine the safety and effectiveness of the WATCHMAN[®] Left Atrial Appendage System (WATCHMAN[®]) device. The Control group consists of patients receiving warfarin drug therapy. The primary endpoint is a composite effectiveness criterion consisting of the occurrence of all stroke, systemic embolism, and cardiovascular death. The statistical objective is to determine if the WATCHMAN[®] device is non-inferior to the Control group with respect to the event rate for the composite endpoint. An additional analysis may be conducted to assess whether the WATCHMAN[®] rate is superior to the control rate.¹

General Statistical Methods

Statistical analyses will be analyzed in an “intent-to-treat” approach unless otherwise noted, with each patient analyzed as being part of their group regardless of the actual treatment received. All patients not having an event or lost to follow-up will be censored at the time of the last documented follow-up visit or last known status. Patient years will be calculated from the date of randomization to the appropriate event or censoring date (for patients without an event) for each patient and aggregated over analysis groups. Event rates will be calculated as the number of events per 100 patient years of follow-up. In addition to the set of intent-to-treat analyses, another set of analyses will be performed to provide additional information on the safety and efficacy of the WATCHMAN[®] device.

Descriptive statistics will be generated for the data collected at baseline, during the procedure and at follow-up. For continuous variables, the mean, standard deviation, median, range and 95% credible intervals will be reported. Credible intervals (95%) for the difference between means will be used to compare groups. For proportions, 95% credible intervals will be reported. Credible intervals (95%) for the ratio between the treatments will be used to compare the groups.

Any data from the study not described in detail in this analysis plan, including but not limited to core lab assessments, quality-of-life measures, and neurological assessments, will be analyzed using these methods.

¹ IDE Supplement #S024, November 10, 2005

Efficacy Analyses

Primary Efficacy Endpoint Analyses

A Bayesian model is proposed for evaluation of the statistical objective. The primary endpoint analyses will be performed for all randomized patients, following the principal of intent-to-treat. Sequential evaluation of the statistical objective allows for early stopping for futility or non-inferiority if the study data give clear indications for the decision. The first evaluation is planned after 600 patient years of follow-up. Subsequent evaluations are to be made after each additional 150 patient years up to a maximum of 1500 patient years of follow-up. No evaluation will take place until 300 subjects have reached one year of follow-up and 100 subjects have reached two years of follow-up. Patients are allocated randomly in a 2:1 Treatment: Control ratio.

Event rate distributions for the WATCHMAN[®] group and Control group are computed using a Bayesian model. The CHADS₂ stratification is used to control for patient risk of future stroke. We combine CHADS₂ = 5 and 6 because of the small number of subjects in each group, for notational purposes we label this group as CHADS₂ = 5.² For notation, we let N_{ij} be the number of patient-years of follow-up for the i^{th} group and the j^{th} CHADS₂ stratum, where $i = 1$ for the WATCHMAN[®] group in the current trial and $i = 2$ for the Control group in the current trial. Let X_{ij} be the corresponding number of observed events. We assume the X_{ij} follow the Poisson distribution with expectations $N_{ij}\lambda_{ij}$. For each CHADS₂ stratum j for each treatment group, we form a Bayesian posterior distribution based on the Poisson-Gamma conjugate prior relationship. A gamma(0.001, 0.001) prior distribution for each "lambda" is assumed. Specifically, the posterior distributions for λ_{1j} and λ_{2j} are taken to be Gamma($X_{1j}+0.001$, $N_{1j}+0.001$)/ $N_{1j}+0.001$ and Gamma($X_{2j}+0.001$, $N_{2j}+0.001$) / $N_{2j}+0.001$, respectively.

Posterior distributions for the event rates for WATCHMAN[®] and the Control group in the current trial are formed from the posterior distributions of the λ_{1j} and λ_{2j} as follows: Let W_j be the observed proportion of follow-up time in each stratum, summed across the WATCHMAN[®] and Control groups, i.e.,

$$W_j = \frac{N_{1j} + N_{2j}}{\sum_j (N_{1j} + N_{2j})}$$

The posterior distribution for the WATCHMAN[®] event rate is formed from the posterior distributions of the λ_{1j} as

$$\sum_j W_j \lambda_{1j} ,$$

^{2 2} Response to questions from 4-6-2004 and 4-29-2004 submitted in Pre-IDE submission on 5-29-2004

and the posterior distribution for the Control group in the current study is formed from the posterior distributions of the λ_{2j} as

$$\sum_j W_j \lambda_{2j} .$$

Comparisons between the WATCHMAN[®] and Control group are made by computing posterior probabilities from these two distributions. For instance, taking the criterion for non-inferiority to be that the event rate for the WATCHMAN[®] group should be statistically no worse than 2 times the event rate for the Control group, the posterior probability of that event is directly calculable from these two posterior distributions.

The first primary efficacy endpoint event for each patient will be considered a terminal event in the Bayesian model. Subsequent events in the same patient are tracked for reporting purposes, but are censored in the endpoint analysis.³

Criterion for Non-inferiority

The criterion for establishing non-inferiority at an evaluation time is defined as a result where the posterior probability that the event rate for the WATCHMAN[®] group is less than two times the event rate for the Control group is at least 0.975. In addition, to demonstrate non-inferiority, the posterior probability that the event rate for the WATCHMAN[®] group is less than the event rate for the control group must be at least 0.05.

Criterion for Futility (Not Non-inferiority)

The criterion for establishing futility (not non-inferiority) at an evaluation time is defined as a result where the posterior probability that the event rate for the WATCHMAN[®] group is greater than or equal to the event rate for the Control group is 0.95 or greater. In addition, to demonstrate non-inferiority, the posterior probability that the event rate for the WATCHMAN[®] group is less than the event rate for the control group must be at least 0.05.

Criterion for Superiority

The criterion for establishing superiority at an evaluation time is defined as a result where the posterior probability that the event rate for the WATCHMAN[®] group is less than the event rate for the Control group is at least 0.95. The superiority test will only be performed if the trial has been stopped for non-inferiority; there will not be an attempt to test for superiority if the non-inferiority test fails.

Sequential Analysis Plan

As indicated above, a Sequential Analysis Plan is proposed with initial evaluation after 600 patient years of follow-up, with subsequent evaluations after each additional 150 patient years, up to a maximum of 1500 patient years of follow-up. At each evaluation, posterior distributions for the event rates for the WATCHMAN[®] group and the Control group will be calculated. The Criterion for Futility will be applied first and if the posterior probability that the event rate for the WATCHMAN[®] group is greater than or equal to the event rate for the Control group is 0.95 or greater, the WATCHMAN[®] device will be declared “Not Non-inferior.” If the Criterion for Futility is not met, the Criterion for Non-inferiority will

³ IDE Supplement #S012, December 17, 2004

be applied and if the posterior probability that the event rate for the WATCHMAN[®] group is less than two times the event rate for the Control group is at least 0.975, the WATCHMAN[®] device will be declared “Non-inferior.”

If neither “Not Non-inferior” nor “Non-inferior” are declared, the decision for the evaluation time will be “Undecided,” and additional 150 patient years of follow-up will be collected before the next evaluation time, up to a limit of 1500 patient years of follow-up. If after the maximum of 1500 patient years of follow-up the WATCHMAN[®] is not established as “Non-inferior”, the device will be considered “Not Non-inferior.”

For documentation purposes, the results of any interim analyses for the primary efficacy endpoint will be recorded, stored, and submitted at the time of the final application.

Analysis of Two Year Success Rates⁴

Patient success rates after two years of follow-up will be compared between randomized groups via Kaplan-Meier estimates at the two-year time point. Additional analyses will incorporate covariate information to increase the precision of these estimates; this may include weighted Kaplan-Meier estimates or proportional hazards regression models. We will look at the following covariates for statistical relevance: gender, age, and CHADS₂ score, AF category, left atria size and left ventricular ejection fraction. Atritech will also consider other variables of interest identified during the study conduct. In addition to comparison of device and control groups by the methods given above, the feasibility study (PILOT) of 50 patients will provide information on the two-year failure rate for the device. This information will be provided to supplement the device information from the pivotal study.

Protection against Possibly Non-Constant Hazard

The potential for non-constant hazard function across time exists. To allow for the possibility of non-constant hazard, a test for equality of event rates across time will be made separately for each treatment group using six month intervals. We will present an analysis using a Bayesian piecewise constant hazard rates model. CHADS₂ scores will be incorporated in a Cox proportional hazards modeling approach.

Additional Efficacy Analyses

For each component of the primary efficacy endpoint, the number of events, number of patients experiencing each event, the event rate per 100 patient years of follow-up, and the corresponding 95% credible interval for the rate event rate will be displayed by treatment group. These event rates will be analyzed using the same Bayesian methods as the composite endpoint.

Safety Analyses⁵

Primary Safety Endpoint

The primary safety endpoint is the occurrence of major bleed and device embolization requiring retrieval, specifically defined as follows:

Primary Safety: Life-threatening events as determined by the Clinical Events Committee which would include events such as device embolization requiring retrieval and bleeding events such as pericardial effusion requiring drainage, cranial bleeding events due to any source, gastrointestinal bleeds requiring transfusion, and any bleeding related to the device or procedure that necessitates an operation.⁶

The Bayesian model will be used to analyze this endpoint in the same manner as the primary composite efficacy endpoint given above, including possible non-constancy of hazard, though no formal criterion for non-inferiority or futility is defined. To compare the rates of safety events for the device and control arms, the posterior probability that the device arm has a lower rate of safety events than the control arm will be computed.⁷ In addition, a two-sided 95% credible interval will be calculated for the ratio of the control rate of safety events to the device rate of safety events. Consistent with the analysis of the primary efficacy endpoint, the first primary safety endpoint event for each patient will be considered a terminal event in the Bayesian model. Subsequent events in the same patient are tracked for reporting purposes, but are censored in the endpoint analysis.

⁵ IDE Supplement #S029, April 17, 2006

⁶ Pre-IDE Submission, May 29, 2004

⁷ IDE Supplement #S014, March 4, 2005

In addition to the above analysis, the following analyses will be performed for the primary safety endpoint:

Components of the primary safety endpoint

For each component of the primary safety endpoint (including device embolization requiring retrieval and bleeding events such as pericardial effusion requiring drainage, cranial bleeding events due to any source, gastrointestinal bleeds requiring transfusion, and any bleeding related to the device or procedure that necessitates an operation) the number of events, number of patients experiencing each event, the event rate per 100 patient years of follow-up, and the corresponding 95% credible interval for the rate event rate will be displayed by treatment group. These event rates will be analyzed using the same Bayesian methods as the composite endpoint.

Unstratified model

The primary safety endpoint analysis will be repeated, without the CHADS₂ stratification. This supportive analysis will be provided as a sensitivity analysis and to provide an overall assessment of safety.

Severity of Adverse Events

Adverse event rates will be assessed for treatment versus control subjects according to varying severity of events. The primary safety analysis includes only life threatening adverse events. As defined by the CEC, any events deemed by the Clinical Event Committee to be “non-events”⁸ will be listed in a separate table presented by treatment group, but will not be counted as adverse events in additional analyses. One supportive safety analysis will be performed that includes all adverse events. All adverse events will be summarized by treatment group in table format displaying the number of each adverse event, the number of subjects experiencing that adverse event, and the corresponding rate per 100 patient years of follow-up. Additionally, safety events not classified as primary safety endpoints will be summarized by treatment group in table format. This table will also include a row for the overall rate of such adverse events.

Mortality

Mortality rates will be calculated for each group, stratified by CHADS and overall. The time to death will be calculated for each subject and those who have not expired will be censored at the date of last known vital status. Time to death will be compared across treatment groups in Cox proportional hazards regression models, including an unstratified model as well as a model stratified for baseline CHADS.

Device and Procedure Related events⁹

Because the control group has no device implanted, there is expected to be a difference in device and procedure related adverse events between treatment groups. Therefore, an evaluation of non-device and non-procedure related adverse events are of interest.

All adverse events determined to be device or procedure related will be summarized in tabular format, listing each adverse event according to relatedness (device or procedure). All non-procedure related and non-device related adverse events will be summarized in table format according to type of adverse event

⁸ CEC meeting minutes, October 27, 2005 and August 8, 2007

⁹ IDE Supplement #S027, January 27, 2006

(e.g. diagnosis or key term) and treatment group. For each type of event, the number of events, number of subjects with the event, and the rate per 100 patient years of follow-up will be presented for each treatment group, along with the difference in rates and corresponding 95% credible intervals.

Early vs. Late Implants

Adverse event rates (for overall and primary safety events) will be summarized stratified according to early vs. late enrollments. For these analyses, “early” treatment patients will be defined as the first three patients receiving a WATCHMAN® device at a site. This will include a combination of both roll-in and randomized patients as the roll-in component of the study was started mid-trial and combining these groups will provide additional information on the safety of the device. As appropriate, the proportion of patients experiencing events and event rates per 100 patient years will be presented, along with 95% credible intervals.

To assess whether or not a training effect or differential adverse event rates among early and late implants exist, regression models will be fit that include an indicator for early versus late implants. Additional regression models for the risk of experiencing adverse events will be fit and potential predictors may include site enrollment number (patient 1, 2, 3, etc.) and enrollment date.

Additional Supportive Analyses

Sensitivity Analysis of Analysis Populations

For non-inferiority trials, an intent-to-treat analysis can result in biased estimates¹⁰. Additionally, since the design of WATCHMAN® device is intended to “prevent the embolization of thrombi that may form in the LAA” and thereby “prevent the occurrence of ischemic stroke and systemic thromboembolism”¹¹, the interpretation of an analysis of patients receiving the WATCHMAN® device that includes data prior to the occlusion of the LAA is problematic.

Therefore, to provide further support to the primary analyses and to better understand the safety and efficacy of the WATCHMAN® device, additional analyses will be performed for different cohorts of patients and different analysis parameters. These analyses will be secondary in nature and will be used as an additional assessment of safety and efficacy of the device compared to warfarin. These analyses will be performed for the primary safety and efficacy endpoints, as well as for key secondary endpoints. At a minimum, a per-protocol analysis will be performed to help assess the treatments efficacy in a way that most closely reflects the scientific model underlying the protocol.¹²

An example that would cause interpretation problems with an intent-to-treat analysis would be a patient who is randomized to the WATCHMAN® arm who is not successfully implanted. Given that this patient could not possibly receive any benefit of the device, including them with patients randomized to the WATCHMAN® arm who do receive a device would bias measures of the treatment effect. In order to address this particular case, one additional analysis will only include implanted patients as “treated”.

¹⁰ Guidance for Industry: E9 Statistical Principles for Clinical Trials; US Department of Health and Human Services, Food and Drug Administration, CBER, CDER: Section 5B (5.2) page 31, September 1998

¹¹ PROTECT AF Study Protocol ST1021, Rev 06, Section 1.2, page 6

¹² Guidance for Industry: E9 Statistical Principles for Clinical Trials; US Department of Health and Human Services, Food and Drug Administration, CBER, CDER: Section 5B (5.2.2-3) page 30-31, September 1998

Another example of a problem with the interpretation of an intent-to-treat analysis as defined in the protocol is the time lag between implant of the WATCHMAN® device and the cessation of warfarin therapy following occlusion of the left atrial appendage. Prior to this occlusion, patients randomized to the WATCHMAN® arm continue on warfarin therapy and therefore do not yet derive benefit of the device in terms of prevention of stroke. One analysis that will be performed to address this issue will only analyze treated patients as those implanted with WATCHMAN® device who discontinue warfarin therapy; events following discontinuation will be analyzed

Assessment of Treatment Effect Consistency

Consistency of the primary endpoints between centers will be evaluated and, if differences are detected, they will be modeled using a hierarchical model or covariate analysis. Statistical analyses for the effects of covariates will also be performed to look at the effect baseline and procedural variables have on the study endpoints. For rates in patient years, Poisson regression models will be used. For binary response variables, a logistic regression model will be used. For time to failure analysis, Cox regression models will be used. Some specific covariates that are of interest when looking at the primary endpoint include CHADS2, age, gender, AF category, left atria size and left ventricular ejection fraction.

Comparison of Device Types

Comparisons of events across treatment groups will be performed, using regression models adjusted for device type (i.e. ‘short’ vs. ‘other’) and also with regression models employing an interaction term for device type and treatment group. Events included in these analyses will include primary safety and efficacy endpoints. Additionally, non-primary safety events will be analyzed. Event rates along with 90% credible intervals will be presented by treatment group and stratified by device type¹³. Kaplan-Meier plots will be used to illustrate the time-to-first event among WATCHMAN patients, stratified by device type.

Analysis of Protocol Violations

Tables will display the number of protocol violations by type, by treatment group and comments will be made as to the affect of violations on the primary safety and efficacy endpoints.

Patient Accountability

The numbers and percent of patients’ enrolled, implanted, withdrawn, or lost-to follow-up will be provided. Additionally, Kaplan-Meier plots of time to withdrawal/lost-to-follow-up stratified by treatment group will be provided.