Bucindolol for the Maintenance of Sinus Rhythm in a Genotype-Defined HF Population

GENETIC-AF Trial Investigators; Piccini, Jonathan P.; Abraham, William T.; Dufton, Christopher; Carroll, Ian A.; Healey, Jeff S.; van Veldhuisen, Dirk J.; Sauer, William H.; Anand, Inder S.; White, Michel

Published in:
JACC. Heart failure

DOI:
10.1016/j.jchf.2019.04.004

Document Version
Final author's version (accepted by publisher, after peer review)

Publication date:
2019

GENETIC-AF: Bucindolol for the Maintenance of Sinus Rhythm in a Genotype-Defined Heart Failure Population

Jonathan P. Piccini, MD, MHS, FACC, William T. Abraham, MD, FACC, Christopher Dufton, PhD, Ian A. Carroll, PhD, Jeff S. Healey, MD, FHRS, Dirk J. van Veldhuisen, MD, PhD, FACC, William H. Sauer, MD, FHRS, FACC, Inder S. Anand, MD, PhD, FACC, Michel White, MD, Stephen B. Wilton, MD, MSc, Ryan Aleong, MD, FHRS, FACC, Michiel Rienstra, MD, PhD, Steven K. Krueger, MD, Felix Ayala-Paredes, MD, Yaariv Khaykin, MD, Bela Merkely, MD, PhD, FESC, FACC, Vladimir Miloradović, MD, Jerzy K. Wranicz, MD, PhD, Leonard Ilkhanoff, MD, MS, FHRS, FACC, Paul D. Ziegler, MS, Gordon Davis, MSPH, Laura L. Emery, MSPH, Debra Marshall, MD, FACC, David P. Kao, MD, Michael R. Bristow, MD, PhD, FACC, Stuart J. Connolly, MD, , on behalf of the Genotype-Directed Comparative Effectiveness Trial of Bucindolol and Toprol-XL for Prevention of Atrial Fibrillation/Atrial Flutter in Patients with Heart Failure Trial Investigators

PII: S2213-1779(19)30253-7

DOI: https://doi.org/10.1016/j.jchf.2019.04.004

Reference: JCHF 1081

To appear in: *JACC: Heart Failure*

Received Date: 15 February 2019

Revised Date: 17 April 2019

Accepted Date: 17 April 2019

Please cite this article as: Piccini JP, Abraham WT, Dufton C, Carroll IA, Healey JS, van Veldhuisen DJ, Sauer WH, Anand IS, White M, Wilton SB, Aleong R, Rienstra M, Krueger SK, Ayala-Paredes F, Khaykin Y, Merkely B, Miloradović V, Wranicz JK, Ilkhanoff L, Ziegler PD, Davis G, Emery LL, Marshall D, Kao DP, Bristow MR, Connolly SJ, on behalf of the Genotype-Directed Comparative Effectiveness Trial of Bucindolol and Toprol-XL for Prevention of Atrial Fibrillation/Atrial Flutter in Patients with Heart Failure Trial Investigators, GENETIC-AF: Bucindolol for the Maintenance of Sinus Rhythm in a Genotype-Defined Heart Failure Population, *JACC: Heart Failure* (2019), doi: https://doi.org/10.1016/j.jchf.2019.04.004.

GENETIC-AF: Bucindolol for the Maintenance of Sinus Rhythm in a Genotype-Defined Heart Failure Population

Jonathan P. Piccini, MD, MHS, FACC[1]; William T. Abraham, MD, FACC[2]; Christopher Dufton, PhD[3]; Ian A. Carroll, PhD[3]; Jeff S. Healey, MD, FHRS[4]; Dirk J. van Veldhuisen, MD, PhD, FACC[5]; William H. Sauer, MD, FHRS, FACC[6]; Inder S. Anand, MD, PhD, FACC[7]; Michel White, MD[8]; Stephen B. Wilton, MD, MSc[9]; Ryan Aleong, MD, FHRS, FACC[6]; Michiel Rienstra, MD, PhD[5]; Steven K. Krueger MD[10]; Felix Ayala-Paredes, MD[11]; Yaariv Khaykin, MD[12]; Bela Merkely, MD, PhD, FESC, FACC[13]; Vladimir Miloradović, MD[14]; Jerzy K. Wranicz, MD, PhD[15]; Leonard Ilkhanoff, MD, MS, FHRS, FACC[16]; Paul D. Ziegler, MS[17]; Gordon Davis MSPH[3]; Laura L. Emery, MSPH[3]; Debra Marshall, MD, FACC[3]; David P. Kao, MD[6]; Michael R. Bristow, MD, PhD, FACC[3,6]; Stuart J. Connolly, MD[4] on behalf of the Genotype-Directed Comparative Effectiveness Trial of Bucindolol and Toprol-XL for Prevention of Atrial Fibrillation/Atrial Flutter in Patients with Heart Failure Trial Investigators.

[1]Duke Clinical Research Institute and Duke University Medical Center; [2]Ohio State University Medical Center; [3]ARCA biopharma, Inc.; [4]Population Health Research Institute, McMaster University, [5]University of Groningen, University Medical Center Groningen, The Netherlands; [6]University of Colorado; [7]US Department of Veterans Affairs; [8]Montreal Heart Institute; [9]Libin Cardiovascular Institute of Alberta, University of Calgary; [10]Bryan Heart Institute; [11]Centre Hospitalier Universitaire de Sherbrooke; [12]Southlake Regional Health Centre; [13]Heart and Vascular Center of the Semmelweis University, Budapest; [14]University of Kragujevac and Clinical Center Kragujevac, Serbia; [15]Medical University of Lodz, Poland; [16]Inova Heart and Vascular Institute; [17]Medtronic, PLC;

Running Title: Bucindolol in AF-HFrEF
Word Count: 6303

Funding: ARCA biopharma

Corresponding Author:
Jonathan P. Piccini, MD, MHS, FACC
Duke University
Duke Clinical Research Institute
PO Box 17969
Durham, NC 27710
United States
919-564-9666
919-668-7057 (fax)
jonathan.piccini@duke.edu

Disclosure Information
GENETIC-AF was sponsored by ARCA biopharma. JPP receives research funding from ARCA, Boston Scientific, Gilead, Janssen Pharmaceuticals, Spectranetics, and St Jude Medical and serves as consultant to Allergan, Amgen, GlaxoSmithKline, Johnson & Johnson, Medtronic, and Spectranetics. SBW reports research support from Medtronic, Abbott, and Boston Scientific and

serves as consultant to ARCA. WTA receives consulting fees from ARCA. PDZ is an employee of Medtronic. MRB is an officer and director of ARCA. CD, DAM, IAC, LLE and GWD are employees of ARCA.

Abstract

Objective: To compare the effectiveness of bucindolol and metoprolol succinate for the maintenance of sinus rhythm in a genetically defined heart failure (HF) population with atrial fibrillation (AF).

Background: Bucindolol is a beta-blocker whose unique pharmacologic properties provide greater benefit in HF patients with reduced ejection fraction (HFrEF) who have the $beta_1$-adrenergic receptor (ADRB1) Arg389Arg genotype.

Methods: 267 HFrEF patients with a left ventricular ejection fraction (LVEF) < 0.50, symptomatic AF, and the ADRB1 Arg389Arg genotype were randomized 1:1 to bucindolol or metoprolol and up-titrated to target doses. The primary endpoint of AF/atrial flutter (AFL) or all-cause mortality (ACM) was evaluated by electrocardiogram (ECG) during a 24-week period.

Results: The hazard ratio (HR) for the primary endpoint was 1.01 (95% CI: 0.71, 1.42) but trends for bucindolol benefit were observed in several subgroups. Precision therapeutic phenotyping revealed that a differential response to bucindolol was associated with: 1) the interval of time from the initial diagnosis of HF and AF to randomization, and; 2) the onset of AF relative to initial HF diagnosis. In a cohort whose first HF and AF diagnoses were < 12 years prior to randomization, in which AF onset did not precede HF by more than 2 years (N=196) the HR was 0.54 (95% CI: 0.33, 0.87; p=0.011).

Conclusion: Pharmacogenetic-guided bucindolol therapy did not reduce the recurrence of AF/AFL/ACM compared to metoprolol in HFrEF patients, but populations were identified that merit further investigation in future Phase 3 trials.

Key Words: atrial fibrillation; bucindolol; heart failure; beta-blocker; pharmacogenetics; precision medicine

List of Abbreviations

ADRB1 = $beta_1$-adrenergic receptor gene
AF = atrial fibrillation
AFL = atrial flutter
Arg = arginine
DTRI = diagnosis to randomization index
DxT = Time fror initial diagnosis to randomization
HF = heart failure
HFlrEF = HF with lower-range ejection fraction (LVEF < 0.40)
HFmrEF = HF with mid-range ejection fraction (0.40 ≤ LVEF < 0.50)
HFrEF = HF with reduced ejection fraction (LVEF < 0.50)
ICM = insertable cardiac monitor

Introduction

Atrial fibrillation (AF) is a common and serious medical problem associated with significant morbidity and mortality, especially in patients with heart failure (HF) (1). Development of AF is associated with increased risk of adverse cardiovascular outcomes, and when AF occurs in patients with HF these adverse effects are accentuated (2,3). AF and HF often co-exist and have common risk factors, as well as overlapping pathophysiologies (3). Therefore, there is a strong rationale to minimize the occurrence of AF in patients with HF. Antiarrhythmic drugs can reduce AF burden but have many side effects including proarrhythmia, with many agents being contraindicated in HF patients (1). Although catheter ablation shows promise for preventing recurrent AF in HF patients with reduced ejection fraction (HFrEF) (4,5), it may not be suitable or practical for many patients. Thus, there is an unmet need for safe and effective drugs to reduce AF in patients with HF. Beta-blockers are first-line therapy for HFrEF due to their benefits in reducing morbidity and mortality and are widely used in HF patients with AF to control ventricular response rate. In addition, beta-blockers have modest AF prevention effects in HFrEF patients (6).

Bucindolol is a non-selective beta-blocker with mild vasodilator properties and two unique antiadrenergic properties; a moderate sympatholytic effect (7) and inverse agonism for the ADRB1 Arg389 major allele gene product (8), a property which promotes inactivation of constitutively active $beta_1$-adrenergic receptors. The treatment effects of bucindolol appear to be enhanced in patients homozygous for ADRB1 Arg389 (ADRB1 Arg389Arg) (8,9). In advanced HFrEF patients with this genotype, a 74% reduction in the development of AF was observed for patients in sinus rhythm at baseline who received bucindolol compared to placebo (10). Metoprolol and carvedilol do not appear to confer similar clinical benefits in patients with an

ADRB1 Arg389Arg genotype (11,12). Therefore, the GENETIC-AF trial (i.e., **Gen**otype-Directed Comparative **E**ffectiveness **Tri**al of Bu**c**indolol and Toprol-XL for the Prevention of Symptomatic **A**trial **F**ibrillation/Atrial Flutter in Patients with Heart Failure) was designed to evaluate the efficacy of a pharmacogenetically-guided rhythm control intervention with bucindolol compared to metoprolol for the prevention of AF/AFL in an ADRB1 Arg389Arg HFrEF population at risk of AF/AFL recurrence.

Methods

Study Design

GENETIC-AF was a multicenter, randomized, double-blind, comparative efficacy trial in a genotype-defined population with HFrEF, defined as a left ventricular ejection fraction (LVEF) < 0.50 and AF (**Online Supplement**). The trial had an adaptive design allowing for seamless transition from Phase 2B to Phase 3 based on review of interim data. The rationale and design of the trial have been previously reported (13).

Patients were randomly assigned to receive bucindolol or metoprolol and were up-titrated to target doses **(Online Table 1)**. Following up-titration, electrical cardioversion (ECV) was performed if needed to establish sinus rhythm prior to the start of follow-up. During the 24-week follow-up period, heart rhythm was monitored by 12-lead electrocardiogram (ECG) every 4 weeks (**Online Figure 1**). A prospectively defined device substudy permitted continuous heart rhythm monitoring to assess AF burden. Substudy participants had a pre-existing Medtronic pacemaker or defibrillator with an atrial lead or were implanted with a Medtronic Reveal LINQ insertable cardiac monitor (ICM) prior to the start of follow-up. After week 24, patients continued to receive blinded study drug and had clinic visits every 12 weeks for assessments of efficacy and safety.

Patients had HFrEF with a LVEF < 0.50 assessed in the past 12 months, symptomatic paroxysmal or persistent AF in the past 180 days and were receiving optimal anticoagulation therapy for stroke prevention. Patients were genotyped at screening and those who were ADRB1 Arg389Arg were eligible for randomization.

Exclusion criteria included New York Heart Association (NYHA) Class IV symptoms, clinically significant fluid overload, permanent AF (ongoing AF event >1 year), antiarrhythmic therapies in past 7 days, prior atrioventricular node ablation, high-grade atrioventricular block, catheter ablation for AF or atrial flutter (AFL) in past 30 days, and prior intolerance or contraindication to beta-blocker therapy. Details of the trial entry criteria have been previously reported (13).

The active comparator, metoprolol succinate (Toprol-XL), is a selective $beta_1$-adrenergic receptor blocker indicated for the treatment of HF. Metoprolol was selected as the active comparator to ensure continuity with previous HF trials and because it has demonstrated effectiveness in preventing AF in HFrEF patients (14,15), but does not appear to confer enhanced benefits in patients with an ADRB1 Arg389Arg genotype (11,12).

Patients were randomized (1:1) to treatment with bucindolol or metoprolol, which was over-encapsulated to maintain blinding. Since bucindolol is administered twice-daily (bid), and metoprolol is given once-daily (qd), a placebo dose was included for the metoprolol arm and all study drugs were administered twice-daily. Randomization was centralized and stratified by HF etiology (ischemic, non-ischemic), LVEF (< 0.35, ≥ 0.35), device type (ICM, pacemaker/defibrillator, no device), and rhythm at randomization (sinus rhythm, AF/AFL), using 16,000 randomly generated numbers and a block size of four. Study drug was titrated weekly to obtain a target dose of 100 mg bid (50 mg bid if < 75 kg) for bucindolol (16) and 200 mg qd for

metoprolol (17). For more details see **Online Table 1**. Patients experiencing AF/AFL during follow-up remained on blinded study drug and could undergo ECV, ablation, or initiate therapy with amiodarone or dofetilide.

ADRB1 Arg389Gly genotype was determined by RT-PCR in DNA extracted from whole blood. Systemic venous plasma norepinephrine was assayed by high-pressure liquid chromatography with electrochemical detection and venous plasma NT-proBNP was measured by electrochemiluminescence immunoassay.

Study design, conduct, and performance were overseen by a 11-member Steering Committee and was monitored by a 3-member Data and Safety Monitoring Committee (DSMB) who also performed the interim efficacy analysis (committee composition in **Online Supplement**). The protocol was approved by the Institutional Review Board/Ethics Committee and all patients provided written informed consent.

Statistical Analyses

For the interim analysis, the endpoint of interest was time to first event of AF/AFL or all-cause mortality (ACM) during a 24-week follow-up period. The primary endpoint for the planned Phase 3 study was time to symptomatic AF/AFL or ACM, with symptoms captured by a study-specific questionnaire (**Online Supplement**). A clinical events committee, blinded to treatment assignment, adjudicated the first occurrence of the AF/AFL endpoint, including the association of new or worsening symptoms. Sample size for Phase 3 assumed a 60% event rate in the metoprolol arm, a 25% relative risk reduction with bucindolol, and accrual of 330 primary events in approximately 620 patients for 90% power at alpha=0.01.

The efficacy analysis was conducted according to intention-to-treat with censoring at 24 weeks for patients not experiencing an event. Hazard ratio (HR) and 95% confidence interval

(CI) values were determined by Cox proportional hazards models with adjustment for the four randomization strata, and treatment as a covariate. Testing for superiority was performed using a 2-sided significance level of 0.05. Patients who died prior to start of follow-up and patients who failed to establish sinus rhythm post-ECV were assigned an event on day 1. Patients were censored on day 1 if they were in AF/AFL and the ECV procedure was not performed, or if they withdrew from the study prior to start of follow-up.

Variables identified in the GENETIC-AF Statistical Analysis Plan (SAP, **Online Supplement**) that were potential predictors of the primary endpoint were investigated by precision therapeutic phenotyping. Hypothesis-based (e.g., AF duration, AF type, LVEF, NYHA Class, NT-proBNP, norepinephrine) and hypothesis-free (e.g. HF duration, initial study dose) elements were included in the multivariate methodology, which was applied to both obvious and non-obvious data to identify a therapeutic phenotype appropriate for investigating in Phase 3. To examine the relationship between HF duration and bucindolol effectiveness for reducing HF events, we analyzed data from the BEST trial (16) and pharmacogenetic substudy (8) for the endpoint of time to all-cause mortality or first HF hospitalization (ACM/HFH).

Time to first event of AF/AFL or ACM was assessed in the device substudy following similar methodology for the primary endpoint, with an AF/AFL event prospectively-defined as AF burden $\geq$ 6 hours per day as recorded by continuous monitoring. Six hours of AF burden has previously been shown to be associated with an increased rate of hospitalization for HF (18). Due to the smaller sample size in the substudy, treatment effect estimates were determined based on Cox proportional hazards models with no adjustment for randomization strata.

Normally distributed continuous variables were analyzed by t-tests or ANOVA where appropriate. Neurohormonal changes from baseline and DTRI data were analyzed by the

Wilcoxon signed rank test, and between group differences by the Wilcoxon rank sum test. Categorical variable differences were assessed by Chi square or Fisher's exact test.

An interim analysis examined data from the initial Phase 2B population. If the DSMB determined that the data were consistent with pre-trial assumptions, the trial was to seamlessly proceed to Phase 3 (see **Online Supplement** for SAP). To aid in signal detection, Bayesian predictive probability of success estimates (19,20) were generated and compared to prespecified thresholds for each potential outcome (i.e., Phase 3 transition, Phase 2B completion, or futility). Based on the interim analysis the DSMB recommended completion of Phase 2B, and the data from this population are presented below.

Results

Population and Baseline Characteristics

The trial was conducted in 92 centers in 6 countries (Canada, Hungary, The Netherlands, Poland, Serbia, and the United States) between April 2014 and December 2017. A total of 760 patients were screened (**Figure 1**); 362 (48%) failed screening due to genotype, 73 (9.6%) did not meet other eligibility criteria, and 58 (7.6%) failed due to other reasons (e.g., withdrawal of consent, lost to follow-up). The remaining 267 patients were randomized to study drug and up-titrated to target doses. Compliance was >90% in both groups, with a higher proportion of patients attaining target dose for bucindolol compared to metoprolol (84% and 72%, respectively; $p = 0.035$).

Baseline characteristics were well-balanced between treatment groups (**Table 1)**. Mean LVEF was 0.36±0.10, 72% had NYHA II or III symptoms at baseline, 51% had persistent AF, and plasma NT-proBNP were elevated at baseline (median = 801 pg/ml; inter quartile range

(IQR): 384, 1420). ECV was required in 46% of patients to establish sinus rhythm prior to follow-up start. About half (48%) of all patients had implanted monitoring devices, which included ICMs inserted for the trial (16%) and pre-existing pacemakers or defibrillators (32%).

Efficacy Outcomes

A total of 143 events were observed for the efficacy endpoint, including 121 AF/AFL events, 19 ECV failures, and 3 deaths. Nearly all AF/AFL events were adjudicated as symptomatic by a blinded clinical events committee (114/121; 94%). Event rates were similar for the bucindolol and metoprolol groups (54% and 53%, respectively), with a HR of 1.01 (95% CI: 0.71, 1.42) for the covariate-adjusted Cox proportional hazards model (**Figure 2**). In a prespecified analysis (**Online Supplement,** Statistical Analysis Plan and Phase 2B Amendment) of regional subgroups (**Table 2, Online Figure 3**), a trend for bucindolol benefit compared to metoprolol was observed in the U.S. subgroup (HR = 0.70; 95% CI: 0.41, 1.19), which was not seen in Canada (HR = 1.52; 95% CI: 0.68, 3.43) or in Europe (HR = 1.01; 95% CI: 0.48, 0.48, 2.14).

Device Substudy

The device substudy included 69 patients from the U.S. (N=42), Canada (N=21), and Europe (n=6) who underwent continuous atrial rhythm monitoring. Cardiac monitors were inserted in 43 patients for the trial, whereas, 26 patients had pre-existing pacemakers or implantable cardioverter defibrillators (ICDs). The baseline characteristics of the substudy were well-balanced between the two groups and were generally similar to the overall population (**Table 1**); however, the substudy had a higher proportion of males (93% vs. 82%), persistent AF (64% vs. 51%), and AF at the time of randomization (65% vs. 51%), compared to the overall population.

An analysis of time to first event of AF/AFL or ACM was conducted in the device substudy following similar methodology for the primary endpoint. As shown in the **Figure 3**, a trend for

bucindolol benefit compared to metoprolol was observed by device-based detection (HR = 0.75; 95% CI: 0.43, 1.32). Similar results were observed when the substudy population was assessed by intermittent, clinic-based 12-lead ECGs (HR = 0.69; 95% CI: 0.38, 1.23); however, the device-detected endpoint generally occurred earlier than the ECG-based endpoint (median = 6.5 days; $p < 0.0001$). For detection of subsequent ECG-determined AF, AF burden ≥ 6 hours had a sensitivity of 100%, a specificity of 87% and an accuracy of 96%.

Patient Characteristics and Treatment Response by Region

The differences in treatment response observed in the U.S. and non-U.S. cohorts prompted examination of baseline characteristics by region (**Online Table 2**). In general, the non-U.S. cohort had less severe HF compared to the U.S. cohort, as demonstrated by significantly higher LVEF (0.39 vs. 0.33), systolic blood pressure (126 v. 120 mmHg), and NYHA class I symptoms (39% vs. 17%), as well as significantly lower plasma NT-proBNP (1135 vs. 1380 pg/mL) and NYHA class III symptoms (5% vs. 26%). Notably, patients in the non-U.S. cohort had a more recent diagnosis of HF (**Table 2, Online Table 2**), with a mean time from HF diagnosis to randomization that was less than half of that in the U.S. group (2.0 vs. 4.5 years); whereas, mean time from AF diagnosis to randomization was similar between the two groups (3.8 vs. 3.4 years).

To quantify the relationship between the initial development of HF and AF, an index termed the *diagnosis to randomization index* (DTRI) was derived from information provided in case report forms. This index represents the differences between the HF duration (i.e., the time of HF diagnosis to randomization) and the AF duration (i.e., the time of AF diagnosis to randomization), with positive values representing HF onset prior to AF and negative values representing AF onset prior to HF. As shown in **Table 2**, the U.S. and non-U.S. cohorts had significant differences in the relative timing of HF and AF onset as measured by mean DTRI

($p < 0.0005$). The U.S. cohort, on average, had HF for more than a year prior to developing AF; whereas, the non-U.S. cohort had a diagnosis of AF for nearly 2 years prior to developing HF. Interestingly, bucindolol response for the primary endpoint correlated with mean DTRI ($\rho = -0.93$, $p = 0.020$), with poor response seen in populations having long-standing AF prior to the development of HF (i.e., Hungary and Canada) and good response in populations with concurrent or previous onset of HF prior to the development of AF (i.e., U.S., Poland, and Serbia).

Baseline Characteristics Predicting Endpoint Frequency and/or Interaction with Treatment

Cox proportional hazards regression modeling was performed to explore prespecified variables (SAP, Online Supplement) that were potential predictors of the primary endpoint (**Online Table 3**). Three variables violated the Cox model proportionality of hazards assumption. Of these, atrial rhythm at randomization was previously addressed by randomization stratification, as was heart rate, which generally correlates with atrial rhythm. The third variable, prior treatment with class III anti-arrhythmic drugs, was not previously identified and was included as a covariate in all subsequent analyses to account for non-proportional influence on baseline hazard.

On multivariate analysis, ten variables predicted the occurrence of the primary endpoint. In addition to the initial dose of study drug, which was based on beta blocker therapy prior to enrollment, the two-predictor model identified five variables related to the degree or duration of HF (i.e., systolic blood pressure, HF duration, HF etiology, NT-proBNP, and NYHA Class) and four variables related to heart rhythm (i.e., rhythm at randomization, baseline heart rate, AF type, and the number of prior ECVs). The only predictor by treatment interaction variable having a p-value < 0.05 was duration of time from initial AF diagnosis to randomization (i.e., AF DxT).

The time from initial HF diagnosis to randomization (i.e., HF DxT) was a significant predictor for the occurrence of primary endpoint but did not predict treatment or treatment by predictor interactions in Cox modeling of the primary endpoint (**Online Table 3**). However, since AF DxT predicted bucindolol response for the prevention of AF recurrence, we examined data from the placebo-controlled BEST HF trial (16) to determine whether HF DxT had a similar relationship to bucindolol response for the HF endpoint, ACM or first HF hospitalization (HFH). As shown in **Online Figure 3**, an attenuation of treatment response for the BEST ACM/HFH endpoint is observed in cohorts with greater values of HF DxT upper bound (i.e., inclusion of long-standing HF prior to randomization). This strong, negative correlation was observed in both the entire cohort (N = 2708; r = -0.82; 95% CI: -0.92, -0.59) and for the ADRB1 Arg389Arg subgroup (N = 493; r = -0.79; 95% CI: -0.91, -0.54).

Effect of Duration and Relative Onset of AF and HF on Treatment Effect

To further examine the effects of AF and HF *duration* identified in the above analyses, a 3-dimensional plot was constructed with treatment effect (i.e., 1-hazard ratio) for the GENETIC-AF primary endpoint as the dependent variable (z-axis), and HF DxT (x-axis) and AF DxT (y-axis) as independent variables. As shown in the **Central Illustration (A)**, an attenuation of treatment effect was associated with increasing values of both AF and HF DxT. When equivalent DxT values (both HF and AF DxT values had to be < the timepoint duration on the x axis) were used to examine the combined effects of AF and HF duration (**Online Figure 4**), a strong negative correlation was observed (r = -0.94; 95% CI: -0.97, -0.89), with substantial attenuation of treatment effect seen with the inclusion of a small proportion of patients with both AF and HF durations greater than 12-15 years.

To examine the effects of the *relative onset* of AF and HF on treatment effect, a 3-dimensional plot was constructed with treatment effect as the dependent variable (z-axis), and the absolute value of DTRI lower bound (i.e., years of AF prior to HF) and DTRI upper bound (i.e., years of HF prior to AF) and as independent variables. As shown in **Central Illustration (B)**, there is an attenuation of treatment effect associated with increasing absolute values of DTRI lower and upper bound (i.e., increasing time between the initial presentations of AF and HF). When equivalent absolute values for DTRI lower and upper bounds were used to examine the concept of contemporaneous AF and HF development (**Online Figure 5A**), there was a nearly linear, negative correlation with treatment effect (r = -0.96; 95% CI: -0.98, -0.92).

Prevention of AF Recurrence in the Precision Therapeutic Selected Phenotype

Duration and relative onset of AF and HF are indirectly related characteristics that may have additive and/or overlapping effects. Therefore, we examined their use in combination to identify a precision therapeutic phenotype appropriate for further study. Details of the precision therapeutic phenotype analyses are presented in the **Online Supplement**.

In the example presented below, we selected a population with an AF and HF DxT < 12 years (i.e., DxT12 cohort), as this cutoff retained a high proportion (86%) of the overall population while minimizing attenuation of the observed treatment effect. We then applied a DTRI lower bound of -2 years (i.e., AF not preceding HF by more than 2 years; DxT12/DTRI-2 cohort), as this cutoff retained 85% of the DxT12 cohort. As shown in **Online Figure 6**, restriction of DTRI upper bound (i.e., years of HF prior to AF) was not required when examined in a DxT12 background.

Patient characteristics of the DxT12 and DxT12/DTRI-2 cohorts are shown in **Online Table 4**. Patients excluded by the DxT12 criteria had characteristics consistent with longstanding AF

and HF; whereas the population excluded by the DTRI > -2 years criteria had characteristics consistent with longstanding AF as primary diagnosis and treatment history, with primarily mild left ventricular dysfunction. Of note, patients who had contemporaneous development of both AF and HF (i.e., DTRI values within 2 years of zero) are the majority of those included in the 230 patient DxT12 cohort ("DTRI included"); whereas DTRI patients with values ±2 years are conspicuously absent from the 37 patient cohort excluded by the DxT12 criteria, i.e. those with the first diagnosis of both AF and HF ≥12 years prior to randomization (**Online Figure 5B**). The accumulation of a substantial number (> 10) of patients with DTRI values ±2 years does not occur until the DxT cutoff is restricted to < 6 years (data not shown).

The primary endpoint of time to first event of AF/AFL/ACM for the DxT12/DTRI-2 cohort (N=196) is shown in **Figure 4**. In HFrEF patients (LVEF < 0.50) the HR was 0.54 (95% CI: 0.33, 0.87) by ECG-based detection, with similar results observed by device-based detection (HR = 0.59; 95% CI: 0.30, 1.19; N=49). In HF patients with mid-range ejection fraction (HFmrEF; LVEF ≥ 0.40 and < 0.50) the HR was 0.42 (95% CI: 0.21, 0.86; p = 0.017) and in HF patients with lower-range ejection fraction (HFlrEF; LVEF < 0.40) the HR was 0.69 (95% CI: 0.33, 1.43; p = 0.32). Device-based estimate for HFmrEF and HFlrEF are not presented due to the small sample size. See **Online Table 5** for more details.

Effects on Norepinephrine and NT-proBNP

Plasma norepinephrine at baseline was similar in the bucindolol (682 ± 348 pg/ml, n=128) and metoprolol (664 ± 359 pg/ml, n=134) groups. At 4 weeks, there was a significant decrease from baseline in the bucindolol group (-124 ± 26 pg/ml; p < 0.001) that was not observed in the metoprolol group (-36 ± 32 pg/ml; p = 0.30). The change from baseline at 4 weeks was significantly different between the two groups (p = 0.012).

Plasma NT-proBNP was non-normally distributed in both groups, and median values at baseline were similar (777 and 861 pg/ml, p = 0.38; **Online Table 6**). There was a significant decrease from baseline in the bucindolol group at week 4 (-96 pg/ml; p = 0.003) and week 12 (-96 pg/ml; p =0.002) that was not observed in the metoprolol group. At week 24, significant decreases relative to baseline values were observed in both the bucindolol (-197 pg/ml; p = 0.005) and metoprolol (-100 pg/ml; p = 0.014) groups, but the change from baseline was not significantly different between the two groups (p = 0.220).

Safety

The proportion of patients experiencing adverse events (AEs) was similar in the two groups (**Table 3**). More patients in the metoprolol group had symptomatic bradycardia or bradycardia leading to dose reduction or discontinuation of study drug compared to the bucindolol group (9.0% vs. 3.0%; p=0.042). Three (2.3%) patients in each group died while receiving study drug or within 30 days of their last dose. All deaths in the metoprolol group occurred during the primary endpoint period (worsening HF – day 25; sudden cardiac death – day 43; motor vehicle accident – day 77). All deaths in the bucindolol group occurred during the long-term extension period (respiratory failure – day 385; sudden death – day 535; cardiac tamponade – day 779). Rates of HF hospitalization (7.5% vs. 8.3%) and ACM/HF hospitalization (8.2% vs. 9.0%) were similar for the bucindolol and metoprolol groups, respectively. There were no strokes in either treatment group, with 93% of patients receiving oral anticoagulants prior to randomization.

Discussion

The GENETIC-AF trial was designed as an adaptive, randomized, controlled trial that was powered for a full Phase 3 investigational comparison if evidence from the Phase 2B study

suggested efficacy was likely on expansion to the Phase 3 sample size (9). In the Phase 2B analysis, pharmacogenetic-guided bucindolol did not reduce the recurrence of AF/AFL/ACM compared to metoprolol in the overall population. However, trends for bucindolol benefit were observed in key subgroups, particularly in those without long-standing and heavily treated AF prior to the development of HF. A lower proportion of patients with longstanding AF diagnosed prior to the development of HF likely contributed to the favorable bucindolol treatment effect in U.S. and device substudy patients, who were majority U.S. enrolled. In addition to the findings relevant to the investigational drug, this study also has several important findings relative to detection of AF in clinical trials.

GENETIC-AF also represents several firsts in the conduct of pharmacogenetic studies in cardiovascular disease and AF in particular. It is the first pharmacogenetically-targeted, randomized, controlled trial of rhythm control therapy in AF. Moreover, it is the first pharmacogenetic trial for prevention of recurrent AF in HFrEF, defined as HF with any decrease in LVEF (23). It is also the first study to compare AF burden to symptomatic AF/AFL as determined by adjudication of symptoms and ECG data. Finally, it represents the first comparative beta-blocker trial to include HF patients with mid-range ejection fraction (HFmrEF), defined as a LVEF ≥ 0.40 and < 0.50 (24).

There are several important findings from GENETIC-AF regarding AF in this HFrEF population. For example, nearly all patients who experienced AF recurrence had symptomatic AF, defined as new or worsening symptoms as adjudicated by a blinded clinical events committee. Recently, there has also been considerable interest in methods of AF diagnosis in clinical practice, including telemetry and device-based technologies (21,22). Our device substudy defined an AF/AFL event as AF burden ≥ 6 hours per day because this amount of

burden had previously been shown to be associated with an increased rate of hospitalization for HF (18). We found that AF burden ≥ 6 hours per day as recorded by continuous monitoring exhibited high predictive accuracy for clinically symptomatic AF/AFL and tended to identify these events earlier than intermittent ECG monitoring.

Approximately half of patients screened for this trial had the ADRB1 Arg389Arg genotype, consistent with previous findings (8-11). In this genotype only norepinephrine high affinity $beta_1$ Arg389 receptors are present, providing a substrate for the favorable effect of sympatholysis (9) that was again observed for bucindolol. Bucindolol lowered plasma norepinephrine levels after 4 weeks of treatment, which was not observed for metoprolol. Plasma NT-proBNP levels also decreased significantly with bucindolol treatment but not with metoprolol. These data indicate that the pharmacodynamic profile that contributes to the pharmacogenetic differentiation of bucindolol was operative in the trial.

It is also notable there were no safety concerns identified with bucindolol. Similar rates of death and hospitalization were observed in both treatment arms, though power was limited for detection of uncommon events. Interestingly, bradycardia was significantly lower in the bucindolol arm, suggesting that bucindolol may lead to less bradycardia than metoprolol in patients with the ADRB1 Arg389Arg genotype.

A major goal of a Phase 2 clinical trial is to further refine the study population that will be investigated in Phase 3. To this end we conducted an exercise in precision therapeutic phenotyping, or “individual treatment effect modeling” (23), designed to identify both prespecified obvious as well as nonobvious variables associated with a beneficial treatment effect of bucindolol. Exploration of factors contributing to the heterogeneity in response observed for regional subgroups led us to examine the timing of AF and HF onset prior to

randomization and relative to one another. This led us to identify two variables that were strongly associated with an attenuation of bucindolol response: 1) the interval of time from the initial diagnosis of HF and AF to randomization (i.e., DxT), and; 2) the onset of AF relative to initial HF diagnosis (i.e., DTRI). AF duration has previously been reported to modulate response for other drug therapies post-ECV (24) and for catheter ablation (25). Less well appreciated is how the HF duration may impact medical therapy, and how these two variables interact in HF patients with concomitant AF. It should also be noted that GENETIC-AF compared two members of a drug class that had been administered chronically to this population, in some cases for years, prior to randomization. As such, a survivor effect due to loss of patients who develop AF and HF within a few years of each other, potentially due to adverse effects on mortality with the combination (26), may be responsible for altering the composition of certain subpopulations (i.e., those with longstanding AF/HF DxT, **Online Figure 5B**) in a manner that influences treatment response (**Online Figure 6**). If a contemporaneous relationship between the onset of AF and HF is optimal for bucindolol to maintain sinus rhythm, potentially related to higher levels of adrenergic activity when both conditions manifest in some proximity (10, 26), then this would explain the phenotype identified in our analysis. Alternatively, or in addition, it is also possible that the DTRI effect has a biological origin based on differences in atrial and ventricular pathophysiology when AF precedes or dominates over HF, the major difference residing in chamber interstitial fibrosis being a more prominent feature in AF (27, 28).

For comparative efficacy studies that seek to observe a *differential response* between two drugs in the same drug class it is critical to identify a study population with high potential for overall response to the drug class. This is necessary because a differential response is, by definition, a fraction of the overall response to a specific drug and, therefore, is more difficult to

observe in a given study population. In this exploratory Phase 2 trial with limited sample size and statistical power, we identified HF populations who respond differentially to two beta-blockers based on genetic targeting. This approach circumvents potential issues associated with conventional subset analyses by evaluating monotonicity and consistency of trends across the full continuum of candidate variables such that the classifiers are readily conducive to numerical calibration (examples provided in **Online Supplement**). We propose that increasing the permissible limits of variation (i.e., tolerance) for the phenotype selection criteria increases the likelihood of reproducibility of these results in future studies.

Limitations

The results of this Phase 2B trial are best considered in light of its limitations. Given the conclusion of the study at Phase 2B, there was not adequate power to definitively test superiority. Although AF DxT and HF DxT were prespecified in the SAP prior to unblinding as potential predictors of treatment response, the onset relationship derived from these variables (i.e., DTRI) was retrospectively defined. Multiplicity via subgroup analysis can lead to false discovery, although this was tempered by examination for consistent trends across the entire dataset and other comparable datasets (i.e. BEST). Lastly, the selection of the precision therapeutic phenotype was based on response, but also considered the sample size needed to maintain feasibility for enrollment in future trials. As such, the treatment effect estimates derived from these analyses are hypothesis generating only and will need to be evaluated in a subsequent, prospectively-designed trial.

Conclusion

In the first trial of a pharmacogenetic-guided rhythm control intervention, bucindolol did not reduce the recurrence of AF/AFL or ACM compared to metoprolol in the overall population.

However, precision therapeutic phenotyping identified a large population of HF patients with an ADRB1 Arg389Arg genotype who display a differential response to bucindolol compared to metoprolol for the prevention of AF/AFL. This experience underscores the utility of performing relatively large Phase 2 studies comprised of heterogeneous populations in order to generate the data necessary to identify appropriate therapeutic phenotypes suitable for Phase 3 investigation.

Competency in Medical Knowledge

The intersection of atrial fibrillation (AF) and heart failure (HF) is common, worsens the prognosis of each disorder and lacks effective, easily administered and safe drug therapy. In the BEST trial pharmacogenetic substudy, against placebo in patients with an ADRB1 Arg389Arg genotype the 4th generation beta-blocker bucindolol reduced the risk of developing AF by 74%, leading to design and performance of the Phase 2 trial GENETIC-AF where 267 high AF risk HFrEF patients were randomized to bucindolol vs. the conventional, 2nd generation compound metoprolol succinate. Overall there was no difference in effectiveness (hazard ratio (HR) 1.01; 95% CI: 0.71, 1.42), but a trend for benefit with bucindolol was observed in the U.S. subgroup (N=127; HR=0.70; 95% CI: 0.41, 1.19) and in patients with implanted devices (N=69; HR=0.75; 95% CI: 0.43, 1.32). The trial exhibited marked regional heterogeneity, which was attributed to 2 countries predominately enrolling patients whose AF diagnosis preceded HF by many years; in countries that enrolled patients with a more contemporaneous presentation of AF and HF bucindolol was associated with a positive efficacy signal.

Translational Outlook

The theoretical basis for bucindolol's advantage over conventional beta-blockers for preventing AF and reducing HF events in HFrEF patients who are genotype ADRB1 Arg389Arg is its more powerful inhibition of the higher functioning Arg389 polymorphic variant of the $beta_1$-adrenergic receptor. The ADRB1 Arg389Gly polymorphism is not present in other species but can be and has been investigated by transgenic overexpression in mice. In terms of the potential for reverse translation, precision therapeutic phenotyping in GENETIC-AF identified a group of patients in whom AF developed many years prior to HF who did not respond favorably to

bucindolol, suggesting different pathophysiology compared to patients who develop AF and HF contemporaneously. This putative pathophysiologic difference and its impact on therapy, potentially related to a greater burden of atrial and ventricular fibrosis associated with longstanding AF, could be translationally investigated in animal models of AF and HF.

References

1. Trulock KM, Narayan SM, Piccini JP. Rhythm control in heart failure patients with atrial fibrillation: contemporary challenges including the role of ablation. J Am Coll Cardiol 2014;64:710-21.
2. Olsson LG, Swedberg K, Ducharme A et al. Atrial fibrillation and risk of clinical events in chronic heart failure with and without left ventricular systolic dysfunction: results from the Candesartan in Heart failure-Assessment of Reduction in Mortality and morbidity (CHARM) program. J Am Coll Cardiol 2006;47:1997-2004.
3. Wang TJ, Larson MG, Levy D et al. Temporal relations of atrial fibrillation and congestive heart failure and their joint influence on mortality: the Framingham Heart Study. Circulation 2003;107:2920-5.
4. Turagam MK, Garg J, Whang W et al. Catheter Ablation of Atrial Fibrillation in Patients With Heart Failure: A Meta-analysis of Randomized Controlled Trials. Ann Intern Med. 2019;170(1):41-50..
5. Marrouche NF, Brachmann J, Andresen D et al. Catheter Ablation for Atrial Fibrillation with Heart Failure. N Engl J Med 2018;378:417-427.
6. Nasr IA, Bouzamondo A, Hulot JS, Dubourg O, Le Heuzey JY, Lechat P. Prevention of atrial fibrillation onset by beta-blocker treatment in heart failure: a meta-analysis. Eur Heart J 2007;28:457-62.
7. Bristow MR, Krause-Steinrauf H, Nuzzo R et al. Effect of Baseline or changes in adrenergic activity on clinical outcomes in the beta-blocker evaluation of survival trial (BEST). Circulation 2004;110:1437-42.

8. Liggett SB, Mialet-Perez J, Thaneemit-Chen S et al. A polymorphism within a highly conserved β_1-adrenergic receptor motif alters beta-blocker response in multiple models and human heart failure. Proc Natl Acad Sci 2006;103:11288-93.

9. O'Connor CM, Fiuzat M, Carson PE et al. Combinatorial pharmacogenetic interactions of bucindolol and $beta_1$, $alpha_{2C}$ adrenergic receptor polymorphisms. PLoS One 2012;7:e44324

10. Aleong RG, Sauer WH, Davis G, et al. Prevention of atrial fibrillation by bucindolol is dependent on the beta-1 389 Arg/Gly adrenergic receptor polymorphism. JACC Heart Fail 2013;1:338-44.

11. Sehnert AJ, Daniels SE, Elashoff M et al. Lack of association between adrenergic receptor genotypes and survival in heart failure patients treated with carvedilol or metoprolol. J Am Coll Cardiol 2008;52:644-51.

12. White HL, de Boer RA, Maqbool A et al. An evaluation of the beta-1 adrenergic receptor Arg389Gly polymorphism in individuals with heart failure: a MERIT-HF sub-study. Eur J Heart Fail 2003;5:463-8.

13. Piccini JP, Connolly SJ, Abraham WT et al. A genotype-directed comparative effectiveness trial of Bucindolol and metoprolol succinate for prevention of symptomatic atrial fibrillation/atrial flutter in patients with heart failure: Rationale and design of the GENETIC-AF trial. Am Heart J 2018;199:51-58.

14, van Veldhuisen DJ, Aass H, El Allaf D, Dunselman PH, Gullestad L, Halinen M, Kjekshus J, Ohlsson L, Wedel H, Wikstrand J and Group M-HS. Presence and development of atrial fibrillation in chronic heart failure. Experiences from the MERIT-HF Study. Eur J Heart Fail 2006;8:539-46.

15. Nergårdh AK, Rosenqvist M, Nordlander R, Frick M. Maintenance of sinus rhythm with metoprolol CR initiated before cardioversion and repeated cardioversion of atrial fibrillation: a randomized double-blind placebo-controlled study. Eur Heart J 2007;28:1351-7.

16. BEST Investigators. Beta-Blocker Evaluation of Survival Trial I. A trial of the beta-blocker bucindolol in patients with advanced chronic heart failure. N Engl J Med 2001;344:1659-67.

17. MERIT-HF Investigators. Effect of metoprolol CR/XL in chronic heart failure: Metoprolol CR/XL Randomised Intervention Trial in Congestive Heart Failure (MERIT-HF). Lancet 1999;353:2001-7.

18. Sarkar S, Koehler J, Crossley GH et al. Burden of atrial fibrillation and poor rate control detected by continuous monitoring and the risk for heart failure hospitalization. Am Heart J 2012;164:616-24.

19. Spiegelhalter DJ, Freedman LS, Blackburn PR. Monitoring clinical trials: conditional or predictive power? Control Clin Trials. 1986;7(1):8-17.

20. Berry SM, Spinelli W, Littman GS, Liang JZ, Fardipour P, Berry DA, Lewis RJ, Krams M. A Bayesian dose-finding trial with adaptive dose expansion to flexibly assess efficacy and safety of an investigational drug. Clin Trials. 2010;7:121-35.

21. Steinhubl SR, Waalen J, Edwards AM et al. Effect of a Home-Based Wearable Continuous ECG Monitoring Patch on Detection of Undiagnosed Atrial Fibrillation: The mSToPS Randomized Clinical Trial. JAMA 2018;320:146-155.

22. Turakhia MP, Desai M, Hedlin H et al. Rationale and design of a large-scale, app-based study to identify cardiac arrhythmias using a smartwatch: The Apple Heart Study. Am Heart J 2019;207:66-75.

23, Wilson FP, Parikh CR. Translational Methods in Nephrology: Individual Treatment Effect Modeling. Am Soc Nephrol. 2018;29:2615-18.

24. Toso E, Blandino A, Sardi D, Battaglia A, Garberoglio L, Miceli S, Azzaro G, Capello AL, Gaita F. Electrical cardioversion of persistent atrial fibrillation: acute and long-term results stratified according to arrhythmia duration. Pacing Clin Electrophysiol 2012;35:1126-34.

25. Bunch TJ, May HT, Bair TL, et al. Increasing time between first diagnosis of atrial fibrillation and catheter ablation adversely affects long-term outcomes. Heart Rhythm 2013;10:1257-62.

26. Aleong RG, Sauer WH, MD, Davis G, Bristow MR. New onset atrial fibrillation predicts heart failure progression. Am J Med. 2014;127:963-71.

27. Li D, Fareh S, Leung TK, Nattel S. Promotion of atrial fibrillation by heart failure in dogs: atrial remodeling of a different sort. Circulation 1999;100:87-95.

28. Dzeshka MS, Lip GY, Snezhitskiy V, Shantsila E. Cardiac Fibrosis in Patients With Atrial Fibrillation: Mechanisms and Clinical Implications. J Am Coll Cardiol 2015;66:943-59.

Figure Legends

CENTRAL ILLUSTRATION Treatment Effect by Duration and Relative Onset of AF and HF prior to Randomization
A. 3-dimensional plot of HF DxT (x-axis) and AF DxT (y-axis) versus treatment effect (z-axis). B. 3-dimensional plot of AF onset prior to HF (x-axis) and HF onset prior to AF (y-axis) versus treatment effect (z-axis). Hazard ratio is for time to AF/AFL/ACM endpoint. HF DxT=time from initial HF diagnosis to randomization. AF DxT=time from initial AF diagnosis to randomization. DTRI (Diagnosis to Randomization Index) = HF DxT – AF DxT. AF onset prior to HF = absolute value of DTRI lower bound. HF onset prior to AF = DTRI upper bound.

FIGURE 1 Consort Diagram
Proportion of patients with the ADRB1 Arg389Arg genotype was consistent with previous findings (8-11)

FIGURE 2 Time to First AF/AFL/ACM Event
Cox proportional hazards model adjusted for the four randomization strata.
Non-stratified hazard ratio = 0.96 (95% CI: 0.69, 1.33). Stratified analysis including adjustment for previous use of class III anti-arrhythmic drugs (yes/no): HR = 0.92 (95% CI: 0.63, 1.33).

FIGURE 3 Time to First Event of AF/AFL/ACM in the Device Substudy
A. Device-based detection. B. ECG-based detection. For device-based detection an AF/AFL event was defined as AF burden ≥ 6 hours per day. Non-stratified Cox proportional hazards model.

FIGURE 4 Time to First Event of AF/AFL/ACM in the DxT12/DTRI-2 Cohort
A. ECG-based detection in the entire cohort. B. Device-based detection in the substudy cohort. For device-based detection an AF/AFL event=AF burden ≥6 hours per day. HR=hazard ratio. FU=follow-up.

Tables

TABLE 1 BASELINE CHARACTERISTICS

	Entire Study			Device Substudy		
Parameter	**All Patients N = 267**	**Bucindolol N = 134**	**Metoprolol N = 133**	**All Patients N = 69**	**Bucindolol N = 35**	**Metoprolol N = 34**
Age, years	65.6 ± 10.1	65.8 ± 10.3	65.5 ± 10.0	66.1 ± 10.7	65.5 ± 11.5	66.8 ± 9.9
Male/Female, %	82/18	83/17	81/19	93/7	94/6	91/9
Race: W/B/A/O, %	96/2/1/1	96/1/1/2	96/2/1/1	96/1/1/2	94/0/3/3	97/3/0/0
LVEF	0.36 ± 0.10	0.36 ± 0.10	0.36 ± 0.10	0.34 ± 0.08	0.33 ± 0.08	0.36 ± 0.09
NYHA I/II/III, %	28/57/15	30/60/10	26/54/20	23/57/20	29/49/23	18/65/18
Ischemic/Non-Ischemic HF, %	32/68	31/69	33/67	28/72	29/71	26/74
Randomized in AF/Not in AF, %	51/49	49/51	52/48	65/35	63/37	68/32
Persistent/Paroxysmal AF, %	51/49	51/49	51/49	64/36	63/37	65/35
HF DxT Duration, days	1153 ± 1909	1252 ± 2070	1054 ± 1733	1168 ± 1723	1208 ± 1880	1126 ± 1572
AF DxT Duration, days	1306 ± 2240	1431 ± 2271	1180 ± 2209	1355 ± 1984	1444 ± 1997	1263 ± 1995
Systolic blood pressure, mm Hg	123.3 ± 15.3	124.7 ± 14.9	121.8 ± 15.7	123.3 ± 15.1	122.4 ± 15.7	124.2 ± 14.5
Diastolic blood pressure, mmHg	75.3 ± 10.8	75.8 ± 11.0	74.8 ± 10.6	75.0 ± 10.1	73.7 ± 9.9	76.3 ± 10.3
Heart Rate, bpm	76.3 ± 17.8	76.5 ± 17.9	76.0 ± 17.7	78.4 ± 17.2	76.8 ± 16.4	80.1 ± 18.1
Previous ECV/AF Ablation/Type III AAD, %	49/21/48	49/21/50	50/20/46	55/13/54	57/17/57	53/9/50
Device Type: ICM/PM/ICD, %	16/17/15	17/15/18	15/20/12	62/22/16	66/20/14	59/24/18
Norepinephrine, pg/ml	673 ± 353	682 ± 348	664 ± 359	706 ± 368	710 ± 398	702 ± 339
NT-proBNP, pg/ml, median (IQR)	801 (384, 1420)	777 (355, 1326)	861 (420,1607)	996 (457, 1645)	923 (365, 1506)	1013 (537, 1806)

W/B/A/O=White/Black/Asian/Other. HF DxT Duration=time from HF diagnosis to randomization. AF DxT Duration=time from AF diagnosis to randomization. ECV=electrical cardioversion. AAD=antiarrhythmic drug. ICM=insertable cardiac monitor. ICD=implanted cardiac defibrillator. PM=pacemaker. IQR=interquartile range. Note: mean ± standard deviations are presented unless otherwise specified.

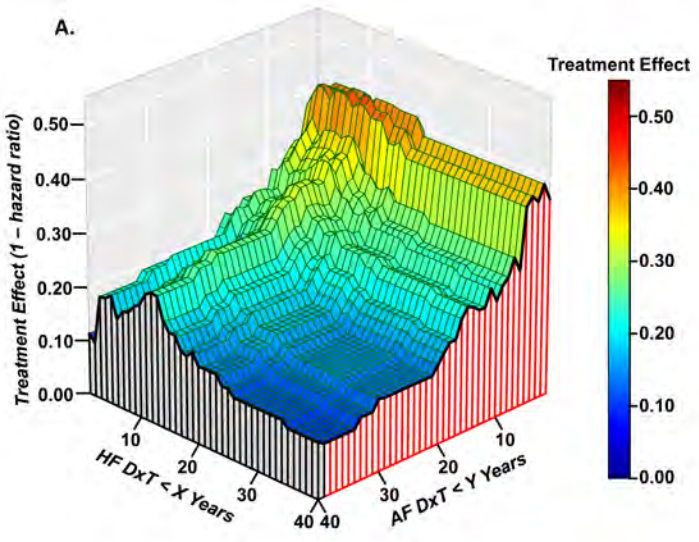
TABLE 2 Timing of HF and AF Onset Relative to Randomization

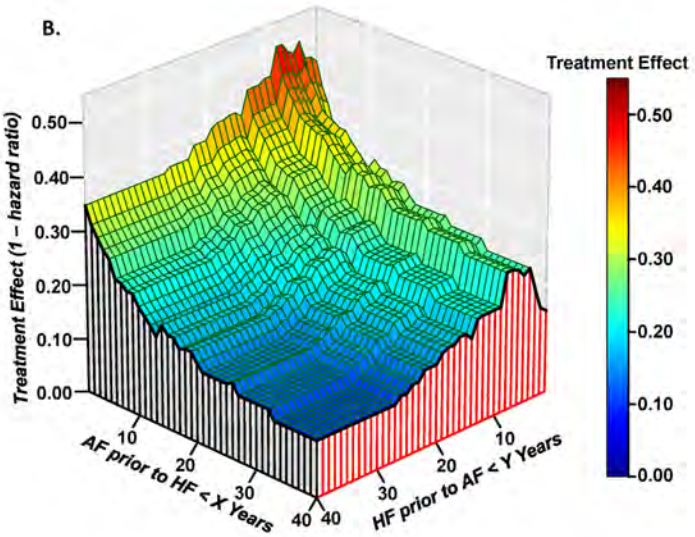
Cohort	HF DxT (years)		AF DxT (years)		DTRI (years)			Time to AF/AFL/ACM	
	Mean	Median	Mean	Median	Mean	Median	P value*	Stratified HR (95% CI)	Non-stratified HR (95% CI)
U.S. (N=127)	4.5	1.5	3.4	1.0	1.1	0.0	–	0.70 (0.41, 1.19)	0.77 (0.48, 1.22)
Non-U.S. (N=140)	2.0	0.4	3.8	0.9	-1.8	0.0	0.0005	1.34 (0.79, 2.28)	1.22 (0.76, 1.96)
Canada (N=59)	2.5	0.5	3.4	0.6	-0.9	0.0	0.024	1.52 (0.68, 3.43)	1.42 (0.72, 2.79)
Europe (N=81)	1.6	0.4	4.0	1.7	-2.4	0.0	0.0009	1.01 (0.48, 2.14)	1.06 (0.55, 2.07)
Hungary (N=33)	1.5	0.3	7.5	4.1	-5.9	-2.8	<0.0001	2.90 (0.71, 11.8)	3.57 (0.99, 12.9)
Poland (N=23)	1.6	0.9	1.4	0.7	0.3	0.0	0.590	0.25 (0.03, 2.22)	0.28 (0.07, 1.14)
Serbia (N=21)	0.4	0.3	0.9	0.4	-0.5	0.0	0.175	0.42 (0.08, 2.18)	0.59 (0.15, 2.36)
Netherlands (N=4)	8.0	7.1	6.4	3.8	1.6	-0.1	ND	ND	ND

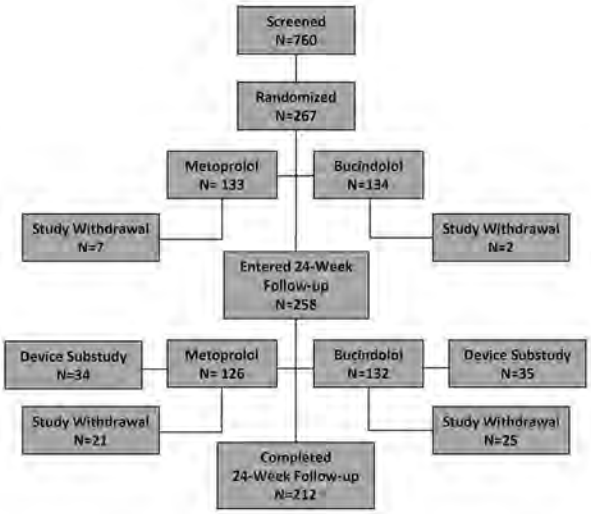
AF DxT=time from AF diagnosis to randomization. HF DxT=time from HF diagnosis to randomization.
DTRI=diagnosis to randomization index; DTRI=HF DxT – AF DxT.
*Wilcoxon rank sum test for comparison to U.S. Cohort.

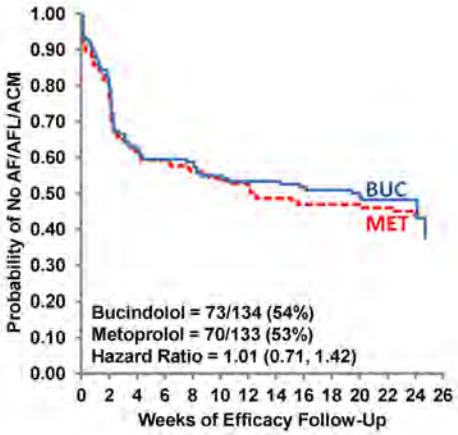
TABLE 3 Treatment Emergent Adverse Events

Endpoint	Bucindolol (N=134)	Metoprolol (N=133)
Any adverse event (AE)	100 (74.6%)	95 (71.4%)
AE possible/probably related to study drug	32 (23.9%)	40 (30.1%)
AE leading to permanent study drug discontinuation	11 (8.2%)	11 (8.3%)
AE leading to study withdrawal (excluding death)	2 (1.5%)	2 (1.5%)
AE of symptomatic bradycardia or bradycardia leading to dose reduction or discontinuation of study drug	4 (3.0%)	12 (9.0%)
Any serious adverse event	34 (25.4%)	27 (20.3%)
AE leading to death	3 (2.3%)	3 (2.3%)
Data presented from randomization through 30 days after last dose of study drug.		



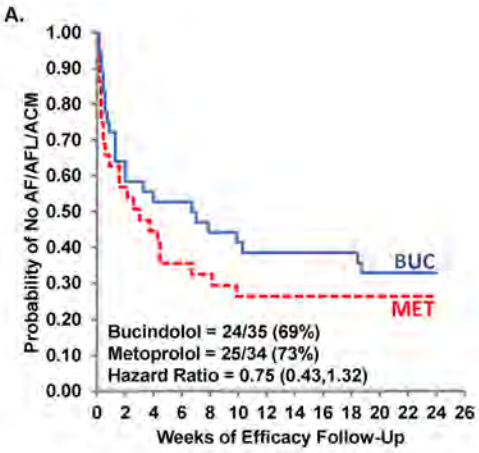






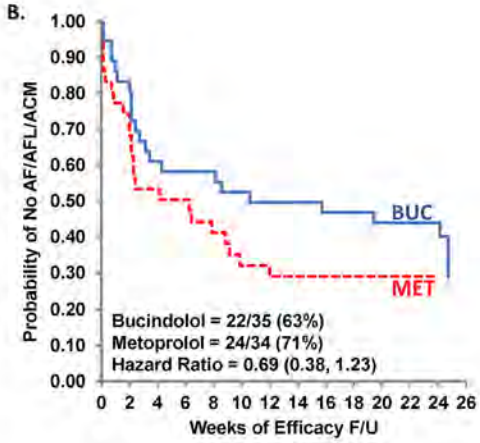
Number at Risk

Week	0	8	16	24
BUC	134	75	64	39
MET	133	68	55	41



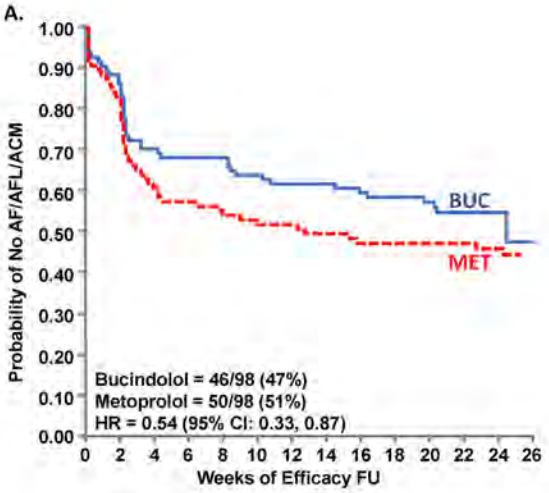
Number at Risk

Week	0	8	16	24
BUC	35	15	13	8
MET	34	10	8	6



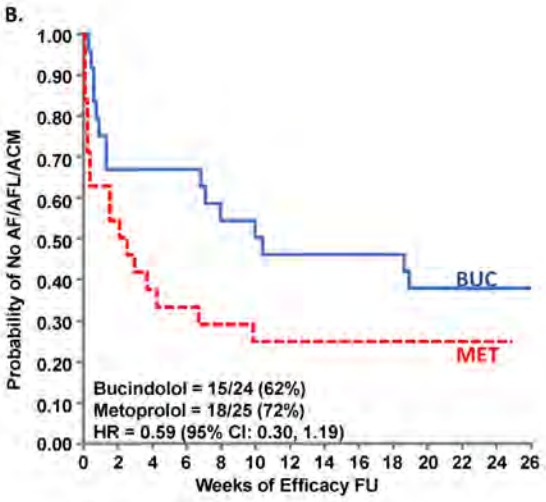
Number at Risk

Week	0	8	16	24
BUC	35	20	16	12
MET	34	13	9	7



Number at Risk

Week	0	8	16	24
BUC	98	65	55	34
MET	98	49	41	32



Number at Risk

Week	0	8	16	24
BUC	24	13	11	7
MET	25	7	6	5

ONLINE SUPPLEMENT

Supplement Figures

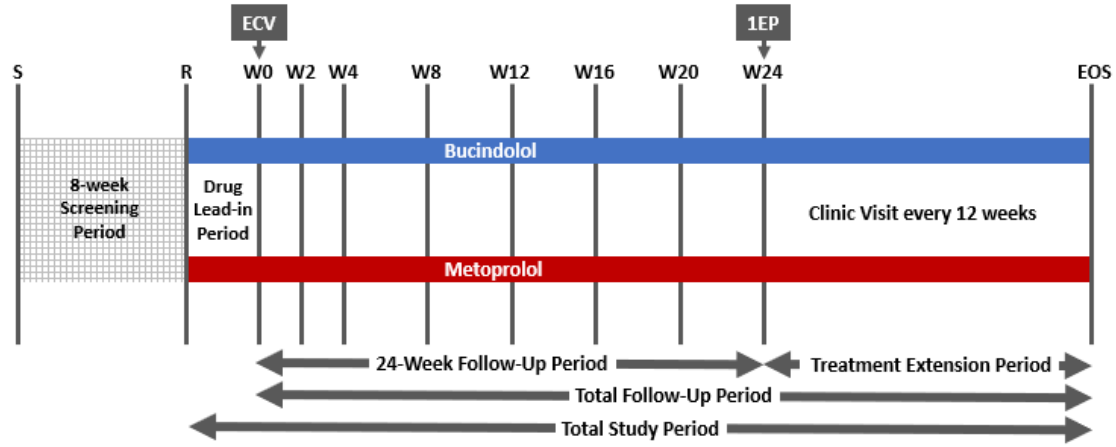
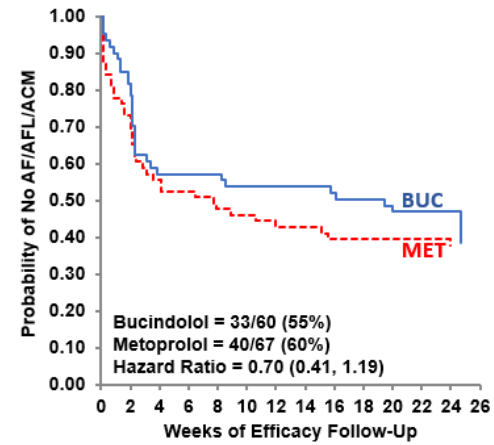
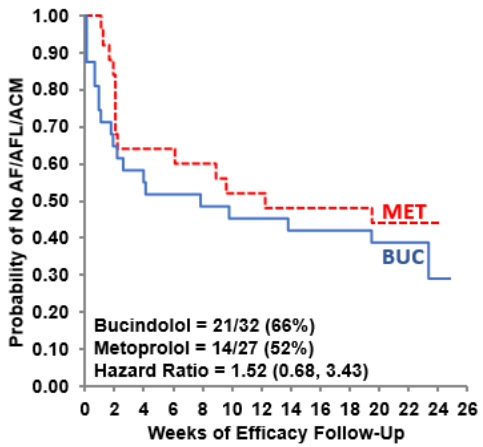


FIGURE 1. GENETIC-AF Study Visit Schedule
Note: ECV performed 3 weeks after randomization, if needed. Week 0 for patients in SR at randomization is 3 weeks (± 3 days). S = Screening Visit; R = Randomization Visit; W = week; ECV = electrical cardioversion; 1EP = primary endpoint; EOS = end of study.



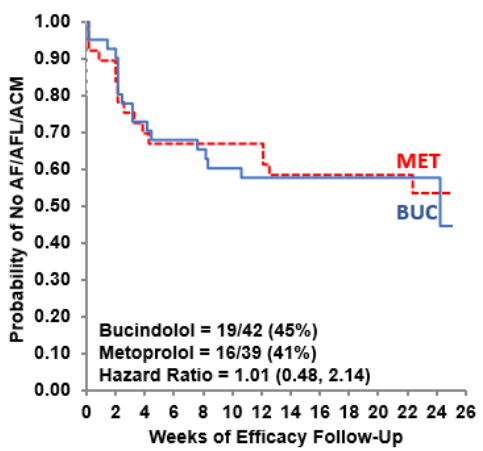
Number at Risk

Week	0	8	16	24
BUC	60	33	30	20
MET	67	29	23	21



Number at Risk

Week	0	8	16	24
BUC	32	16	13	8
MET	27	15	12	10



Number at Risk

Week	0	8	16	24
BUC	42	26	21	11
MET	39	24	20	10

FIGURE 2 Time to First AF/AFL/ACM Event by Region
A., U.S. cohort; B., Canada cohort; C., Europe cohort.
Cox proportional hazards model adjusted for the four randomization strata.

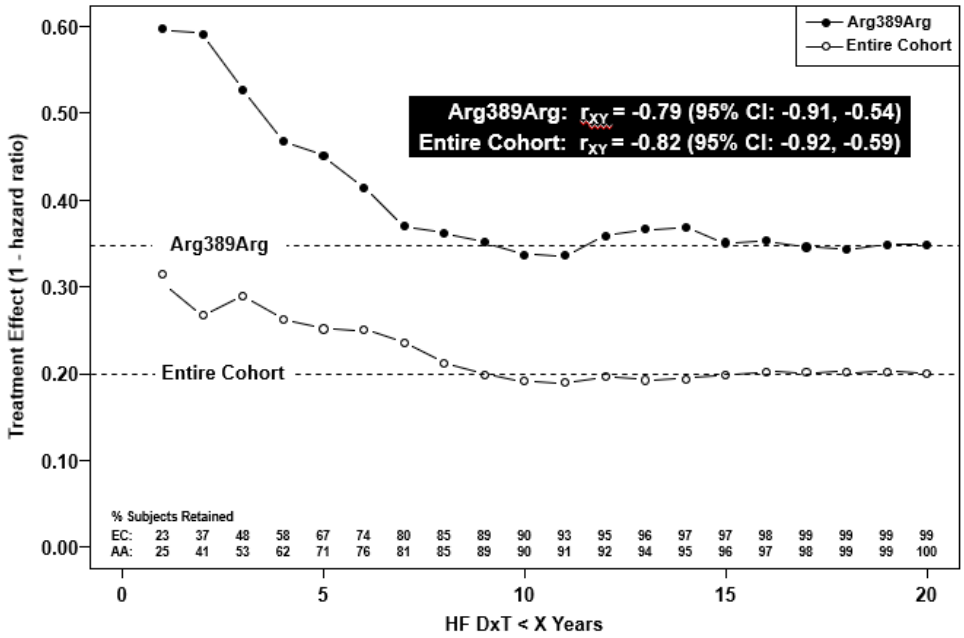


FIGURE 3 Treatment Effect and the Duration of HF in the BEST HF Trial
Entire cohort (open circles, n=2708) and ADRB1 Arg389Arg subgroup (closed circles, n=493). Hazard ratio is for time to first heart failure hospitalization or death for bucindolol and placebo. HF DxT=time from initial HF diagnosis to randomization. R_{xy}=correlation coefficient. Arg389Arg=patients homozygous for ADRB1 Arg389.

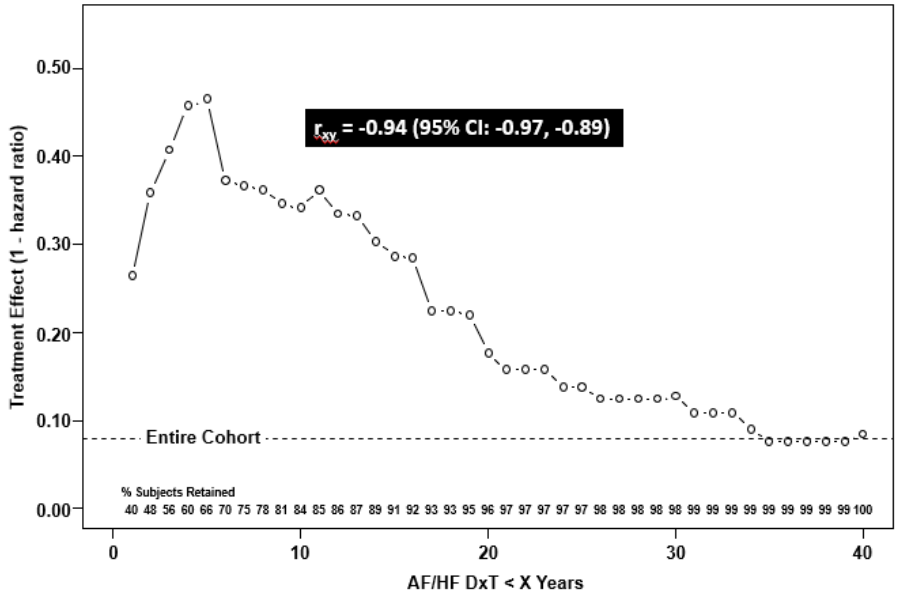
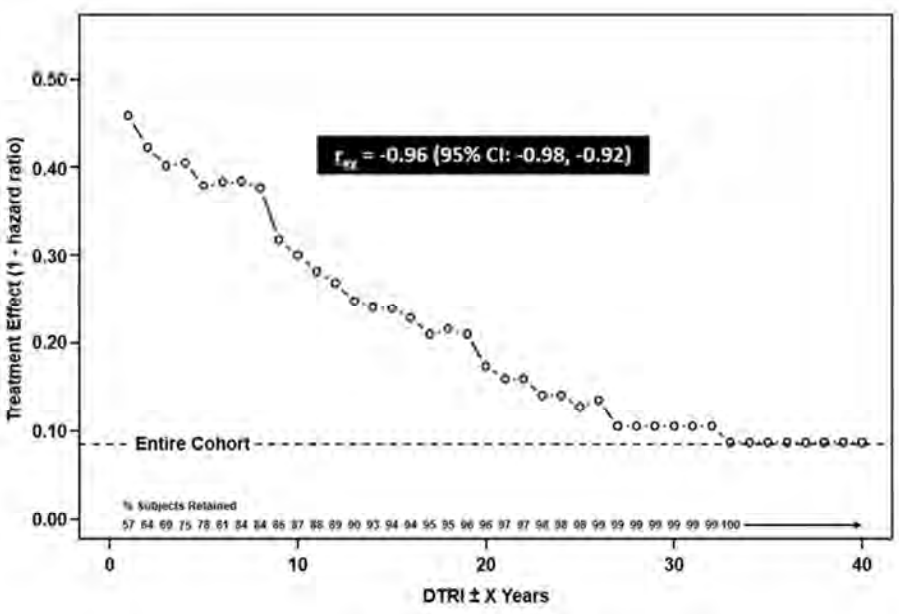


FIGURE 4 Treatment Effect by AF and HF Duration
Treatment effect versus AF/HF DxT (i.e., both HF DxT and AF DxT< X years).
Hazard ratio is for time to AF/AFL/ACM endpoint. AF/HF DxT= time from initial AF and HF diagnosis to randomization.

A.



B.

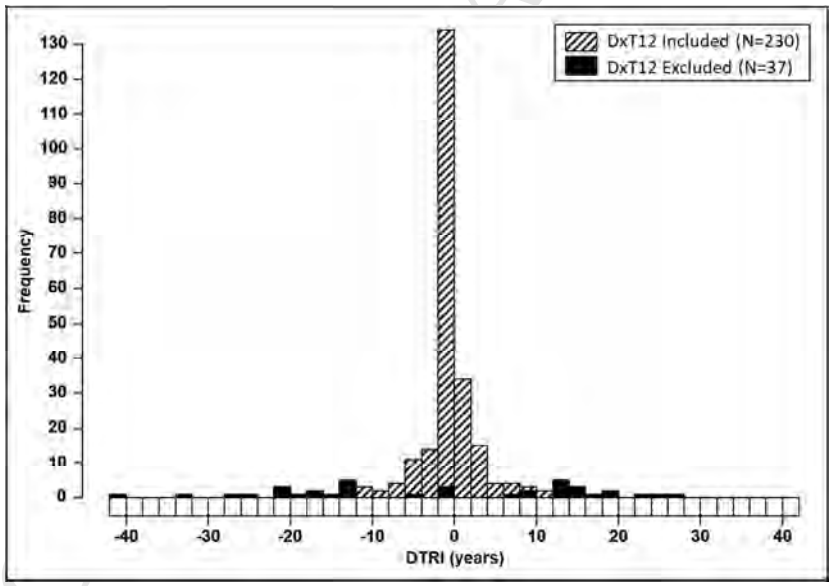


FIGURE 5 Treatment Effect Relationship to Relative Onset of AF and HF (DTRI)
A. Treatment effect versus absolute value of DTRI upper and lower bounds.
B. Histogram of DTRI distribution for DxT12 cohort and cohort excluded by DxT12 criteria.
Hazard ratio is for time to AF/AFL/ACM endpoint. DTRI=Diagnosis to Randomization Index.
DxT12 = cohort with <12 years of AF and HF prior to randomization. X-axis is in 2-year intervals.

A.

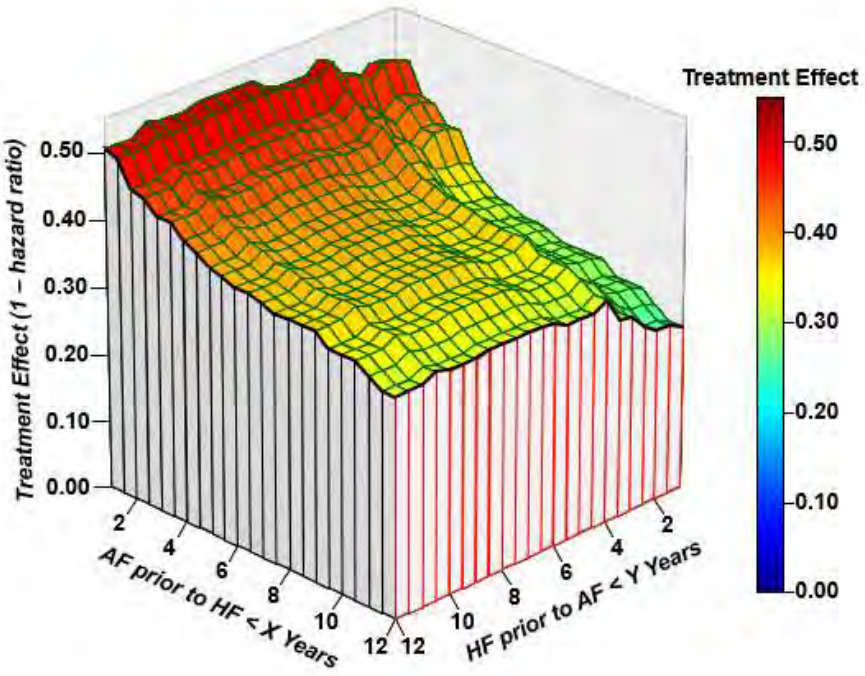


FIGURE 6 Treatment Effect and the Relative Onset of AF and HF in DxT12 Cohort
3-dimensional plot of AF onset prior to HF (x-axis) and HF onset prior to AF (y-axis) versus treatment effect (z-axis) in DxT12 Cohort. Hazard ratio is for time to AF/AFL/ACM endpoint. DTRI (Diagnosis to Randomization Index) = HF DxT – AF DxT. AF onset prior to HF = absolute value of DTRI lower bound. HF onset prior to AF = DTRI upper bound. DxT12 = cohort with <12 years of AF and HF prior to randomization.

Supplement Tables

TABLE 1. Study Drug Titration Schedule

Previous Commercial Beta-blocker Dose[1]												Randomized Beta-blocker Dose	
Metoprolol XL/CR (mg QD)		Metoprolol IR (mg BID)		Carvedilol CR (mg QD)		Carvedilol IR (mg BID)		Bisoprolol (mg QD)		Nebivolol (mg QD)		Metoprolol XL (mg QD)	Bucindolol (mg BID)
>	≤	>	≤	>	≤	>	≤	>	≤	>	≤	=	=
-	50	-	25	-	20		6.25	-	2.5	-	1.25	25	6.25
50	100	25	50	20	40	6.25	12.5	2.5	5	1.25	2.5	50	12.5
100	200	50	100	40	80	12.5	25	5	10	2.5	5	100	25
200[3]	-	100[3]	-	80[3]	-	25[3]	-	10[3]	-	5	10[3]	200	50
-	-	-	-			-	-			-	-	200	100[2]
Transition to Starting Dose of Study Drug ➡➡➡												Up-titration ⬇	

[1]Transition from β-blockers other than those above requires approval from the Sponsor or its designee prior to randomization.

[2]Patients who weigh < 75 kg at randomization will receive a maximum bucindolol dose of 50 mg BID.

[3]Patients receiving commercial β-blocker doses higher than those currently approved will require pre-approval from the Sponsor or its designee prior to randomization.

TABLE 2. Baseline Characteristics by Region

Parameter	U.S. Cohort N = 127	Non-U.S. Cohort N = 140	P-value
Age, years	66.3 ± 10.7	65.1 ± 9.5	0.516
Male/Female, %	87/13	78/22	0.079
Race: W/B/A/O, %	93/4/1/2	99/0/1/0	0.017
LVEF	0.33 ± 0.09	0.39 ± 0.09	<0.001
NYHA I/II/III, %	17/57/26	39/56/5	<0.001
Ischemic/Non-Ischemic HF, %	31/69	33/67	0.896
Randomized in AF/Not in AF, %	59/41	43/57	0.010
Persistent/Paroxysmal AF, %	52/48	50/50	0.807
AF DxT Duration, days	1236 ± 2192	1370 ± 2288	0.517
HF DxT Duration, days	1627 ± 2306	724 ± 1326	<0.001
Systolic blood pressure, mm Hg	119.9 ± 15.7	126.3 ± 14.4	0.001
Diastolic blood pressure, mmHg	73.8 ± 11.3	76.6 ± 10.2	0.024
Heart Rate, bpm	78.4 ± 19.4	74.4 ± 16.0	0.118
Previous ECV, %	55	44	0.041
Previous AF Ablation, %	17	24	0.373
Previous Type III AAD use, %	47	49	0.902
Device Type: ICM/PM/ICD, %	19/15/21	14/20/9	0.002
Norepinephrine, pg/ml	657 ± 373	687 ± 335	0.389
NT-proBNP, pg/ml, median (IQR)	953 (488, 1506)	678 (143, 1252)	0.045

W/B/A/O = White/Black/Asian/Other. AF DxT = time from AF diagnosis to randomization. HF DxT = time from HF diagnosis to randomization. ECV = electrical cardioversion. AADs = antiarrhythmic drugs. ICM = insertable cardiac monitor. ICD = implanted cardiac defibrillator. PM = pacemaker. IQR = interquartile range. Note: mean ± standard deviations are presented unless otherwise specified. Wilcoxon Rank Sum Test for continuous values and Fishers Exact Test for categorical values.

TABLE 3. Cox Proportional Hazards Regression Modeling for Time to First AF/AFL/ACM Event

Predictor	Two Predictor Model		Three Predictor Model		
	Treatment	Predictor	Treatment	Predictor	Treatment x Predictor
Rhythm at randomization†	0.83	<0.001*	0.66	<0.001*	0.51
Baseline heart rate†	0.80	<0.001*	0.96	0.042*	0.99
AF type	0.72	0.001*	0.77	0.06	0.49
Baseline systolic blood pressure	0.84	0.006*	0.15	0.63	0.15
HF DxT	0.77	0.007*	0.66	0.63	0.73
Initial study dose	0.39	0.017*	0.79	0.89	0.35
Prior ECV count	0.76	0.018*	0.37	0.78	0.30
HF etiology	0.81	0.023*	0.91	0.04*	0.53
Baseline NT-proBNP	0.91	0.040*	0.48	0.75	0.28
Baseline NYHA class	0.99	0.043*	0.59	0.91	0.57
AF DxT	0.83	0.07	0.18	0.14	0.025**
Device strata	0.72	0.11	0.98	0.77	0.77
Prior ECV or ablation	0.79	0.13	0.51	0.13	0.52
Region	0.82	0.09	0.87	0.16	0.33
Baseline diastolic blood pressure	0.71	0.28	0.18	0.09	0.16
Previous use of class III AAR†	0.76	0.35	0.58	0.32	0.64
Beta blocker prior to randomization	0.84	0.42	0.66	0.68	0.98
Baseline creatinine	0.82	0.48	0.30	0.19	0.26
Total prior ECV or ablation	0.74	0.52	0.75	0.64	0.93
Prior ablation	0.78	0.62	0.83	0.14	0.19
LVEF	0.80	0.66	0.79	0.96	0.84
LVEF strata	0.80	0.68	0.74	0.89	0.82
CYP2D6	0.98	0.93	0.21	0.29	0.17
Baseline norepinephrine	0.73	0.99	0.63	0.73	0.72

*P<0.05 for prediction of primary endpoint. **P<0.05 for treatment x predictor interaction. †Violation of proportionality of hazards assumption (p<0.05). AF DxT=time from initial AF diagnosis to randomization. HF DxT=time from initial HF diagnosis to randomization. ECV=electrical cardioversion. AAR=antiarrhythmic drug. LVEF=left ventricular ejection fraction. CYP=cytochrome p450.

TABLE 4. Baseline Characteristics for Selected Phenotypes

Parameter	AF12/HF12			AF12/HF12/DTRI-2		
	Included N=230	Excluded N=37	p-value	Included N=196	Excluded N=34	p-value
Age, years	64.9 ± 10.2	70.1 ± 8.4	0.012	65.2 ± 9.9	63.1 ± 11.8	0.435
Male/Female, %	80/20	95/5	0.036	80/20	79/21	1.000
Race: W/B/A/O, %	97/2/0/1	95/0/0/5	0.087	96/2/1/1	97/3/0/0	0.728
LVEF	36.6 ± 9.4	33.4 ± 10.5	0.104	36.0 ± 9.3	39.8 ± 9.6	0.010
NYHA I/II/III, %	30/57/13	6/59/24	0.099	28/57/15	41/56/3	0.074
Ischemic/Non-Ischemic HF, %	30/70	43/57	0.132	32/68	21/79	0.227
Randomized in AF/Not in AF, %	47/53	73/27	0.004	48/52	41/59	0.577
Persistent/Paroxysmal AF, %	49/51	62/38	0.159	48/52	56/44	0.459
AF DxT, days	770 ± 983	4642 ± 4201	<0.001	539 ± 787	2098 ± 955	<0.001
HF DxT, days	698 ± 1012	3988 ± 3289	<0.001	778 ± 1064	231 ± 402	<0.001
Systolic blood pressure, mm Hg	124.0 ± 15.0	118.9 ± 16.7	0.094	123.9 ± 15.4	124.5 ± 13.1	0.827
Diastolic blood pressure, mmHg	75.7 ± 10.2	72.6 ± 13.7	0.090	75.3 ± 10.4	78.0 ± 9.3	0.093
Heart rate, bpm	76.2 ± 18.3	76.6 ± 14.3	0.61	75.7 ± 18.5	79.4 ± 16.9	0.223
Previous ECV (0, 1, 2+), %	51/28/20	46/22/32	0.263	52/31/18	50/15/35	0.032
Previous AF ablation (0, 1, 2+), %	82/13/5	62/27/11	0.017	85/11/4	65/24/12	0.010
Previous class I AAD use: Y/N, %	8/92	8/92	1.000	6/94	21/79	0.008
Previous class III AAD use: Y/N, %	46/54	59/41	0.157	42/58	71/29	0.003
Device type: None/ILR/TD, %	55/18/27	32/3/65	<0.001	55/17/28	53/26/21	0.347
Norepinephrine, pg/ml	646 ± 311	839 ± 519	0.030	656 ± 316	585 ± 278	0.243
NT-proBNP, pg/ml, median (IQR)	769 (372, 1338)	1044 (528, 1983)	0.043	790 (392, 1387)	588 (263, 1147)	0.266

AF12/HF12=AF DxT and HF DxT< 12 years. AF12/HF12/DTRI-2=AF12/HF12 and DTRI > -2 years. W/B/A/O=White/Black/Asian/Other. ECV=electrical cardioversion. AAD=antiarrhythmic drug. ILR=implanted loop recorder. TD=therapeutic device (implanted cardiac defibrillator or pacemaker). IQR=interquartile range. AF DxT=time from initial AF diagnosis to randomization. HF DxT=time from initial HF diagnosis to randomization. DTRI=Diagnosis to Randomization Index. Note: mean±standard deviations are presented unless otherwise specified. Wilcoxon Rank Sum Test for continuous values and Fishers Exact Test for categorical values.

TABLE 5. Time to First Event of AF/AFL/ACM for Subgroups by LVEF

Cohort	HFrEF LVEF < 0.50		HFmrEF 0.40 ≤ LVEF < 0.50		HFlrEF LVEF < 0.40	
	N (%)	HR (95% CI)	N (%) {% of Cohort}	HR (95% CI)	N (%) {% of Cohort}	HR (95% CI)
All Patients	267 (100)	0.92 (0.63, 1.33)	128 (100) {48}	0.78 (0.45, 1.33)	139 (100) {52}	1.03 (0.58, 1.83)
AF12/HF12	230 (86)	0.68 (0.45, 1.02)	113 (88) {49}	0.61 (0.34, 1.10)	117 (84) {51}	0.74 (0.38, 1.44)
AF12/HF12/DTRI-2	196 (73)	0.54 (0.33, 0.87)	91 (71) {46}	0.42 (0.21, 0.86)	107 (77) {54}	0.69 (0.33, 1.43)

AF12/HF12=AF/HF DxT < 12 years; 12/12/DTRI-2=AF/HF DxT < 12 years and DTRI > -2 years.
HFrEF=HF with reduced LVEF; HFmrEF=HF with mid-range LVEF; HFlrEF=HF with lower-range LVEF. DTRI=Diagnosis to Randomization Index.

Table 6. NT-proBNP values (pg/ml)‡

Parameter	Metoprolol N = 123	Bucindolol N = 125
Baseline	861 (420, 1607)	777 (355, 1326)
P value vs. Met†	NA	0.378
ΔWeek 4	-35 (-384, 246)	-96 (-431, 70)
P value vs. Bsl*	0.320	0.003
P value vs. Met†	NA	0.300
ΔWeek 12	-50 (-610, 303)	-96 (-482, 69)
P value vs. Bsl*	0.198	0.002
P value vs. Met†	NA	0.051
ΔWeek 24	-100 (-634, 117)	-197 (-613, 115)
P value vs. Bsl*	0.014	0.005
P value vs. Met†	NA	0.220
‡Median and interquartile range presented due to non-normal distribution; *Wilcoxon signed rank test; †Wilcoxon rank sum test; Δ = change from baseline.		

Composition of Oversight Committees

GENETIC-AF Steering Committee

Stuart J. Connolly, MD – Population Health Research Institute, McMaster University (Chair)

William T. Abraham, MD – Ohio State University Medical Center (Co-Chair)

Jonathan P. Piccini, MD – Duke Clinical Research Institute and Duke University Medical Center

Jeff S. Healey, MD – Population Health Research Institute, McMaster University

Inder S. Anand, MD – U.S. Department of Veterans Affairs / University of Minnesota

D.J. van Veldhuisen, MD – University of Groningen, University Medical Center Groningen, The Netherlands

Michel White, MD – Montreal Heart Institute

Stephen B. Wilton, MD – Libin Cardiovascular Institute of Alberta, University of Calgary

William H. Sauer, MD – University of Colorado

David Haines, MD – Beaumont Health Systems

Michael R. Bristow, MD, PhD, University of Colorado and ARCA biopharma Inc.

GENETIC-AF Data Safety Monitoring Committee

Voting Members:

Christopher O'Connor, MD – Inova Heart and Vascular Institute (Chair)

Jonathan Steinberg, MD – University of Rochester School of Medicine & Dentistry

Victor Hasselblad, PhD – Duke University School of Medicine

Non-Voting Members:

Hussein Al-Khalidi, PhD – Duke Clinical Research Institute

Joan Gu, MS – Duke Clinical Research Institute

GENETIC-AF Clinical Events Committee

James P. Daubert, MD (Co-Chair) – Duke Clinical Research Institute and Duke University Medical Center

Albert Y. Sun, MD (Co-Chair) – Duke Clinical Research Institute and Duke University Medical Center

Sean D. Pokorney, MD, MBA – Duke Clinical Research Institute and Duke University Medical Center

Daniel J. Friedman, MD – Duke Clinical Research Institute and Duke University Medical Center

Andrew Ambrosy, MD – Duke Clinical Research Institute and Duke University Medical Center

Adam DeVore, MD – Duke Clinical Research Institute and Duke University Medical Center

Marat Fudim, MD – Duke Clinical Research Institute and Duke University Medical Center

Bayesian Statistical Modeling

Bayesian modeling of interim analysis data on 230 patients was performed by:

Ben Saville, PhD – Berry Consultants, Austin, TX.

Trial Operational Management

The trial was managed by ARCA with the assistance of three research organizations:

Duke Clinical Research Institute, Durham, NC

Population Health Research Institute, Hamilton, ON

Argint International, Budapest, Hungary.

GENETIC-AF Investigators at Sites who Screened and/or Randomized Patients

Canada: F Ayala-Paredes, A Bakbak, ML Bernier, DH Birnie, SJ Connolly, B Coutu, E Crystal, MW Deyell, KM Dyrda, MC Hartleib, Y Khaykin, ZW Laksman, P Leong-Sit, CA Morillo, AS Pandey, F Philippon, S Vizel, SB Wilton; **Hungary**: P Andréka, Z Csanadi, GZ Duray, T Forster, G Kerkovits, B Merkely, AC Nagy, T Simor; **Poland**: D Czarnecka, JD Kasprzak, WJ Musial, G Raczak, J Szachniewicz, JK Wranicz; **Serbia**: S Apostolović, S Hinić, V Miloradović, D Simić; **The Netherlands**: GJ Milhous, A Oomen, M Rienstra, TJ Romer, LM van Vijk; **United States**: PB Adamson, RG Aleong, JD Allred, N Amjadi, MM Bahu, AJ Bank, AE Berman, MA Bernabei, RS Bhagwat, L Borgatta, AJ Buda, RT Cole, JL Collier, SJ Compton, O Costantini, MR Costanzo, IM Dauber, MP Donahue, I Dor, GF Egnacyzk, EJ Eichhorn, CC Eiswirth, S Emani, GA Ewald, RC Forde-McLean, MD Gelernt, DE Haines, CA Henrikson, JM Herre, B Herweg, L Ilkhanoff, LR Jackson 2nd, SK Krueger, A Lala, R Lo, B London, BD Lowes, JA Mackall, V Malhotra, FA McGrew, S Murali, A Natale, KR Nilsson, J Okolo, MV Perez, RS Phang, R Ranjan, MY Rashtian, MJ Ross, SM Samii, T Shinn, MB Shoemaker, SA Strickberger, VN Tholakanahalli, A Tzur, PJ Wang, LT Younis.

Statistical Analysis Plans

STATISTICAL ANALYSIS PLAN

Study Title:	GENETIC-AF – A **Gen**otype-Directed Comparative **E**ffectiveness **Tri**al of Bu**c**indolol and Toprol-XL for Prevention of Symptomatic **A**trial **F**ibrillation/Atrial Flutter in Patients with Heart Failure
Sponsor:	ARCA biopharma, Inc. 11080 CirclePoint Road, Suite 140 Westminster, Colorado 80020 Phone: 720.940.2100
Study Drug:	Bucindolol hydrochloride (bucindolol)
Comparator:	Metoprolol succinate (Toprol-XL, metoprolol)
IND No.:	118,935
Indication:	Atrial Fibrillation
Protocol ID:	BUC-CLIN-303
Date:	15 February 2017

Note: The interim analysis methodology is not included in this plan. That methodology can be found in the DSMB Charter and DSMB Statistical Analysis Plan documents.

DEFINITIONS OF ANALYSIS POPULATIONS AND ENDPOINT FOLLOW-UP PERIODS

The efficacy analysis will follow the intent-to-treat (ITT) principle and all patients randomized to study treatment will be included regardless of (1) the success of the treatment titration process and (2) result of electrical cardioversion (ECV) aimed at converting atrial fibrillation (AF) to sinus rhythm (SR). As an additional sensitivity analysis, testing of the primary and secondary endpoints will be repeated on a protocol-compliant subpopulation. Further sensitivity analyses specific to endpoints are described below. The safety analyses will include all patients that received at least one dose of blinded study treatment. The screened population includes any patient who signs informed consent for the study. The screen failure population is a subpopulation of the screened population who are not randomized to study drug for any reason.

Four follow-up periods will be defined for inclusion of each patient's results in endpoint calculations:

- Drug Titration Period: starts on the day of randomized treatment initiation and extends for six weeks after randomization.
- 24-Week Follow-up Period: starts on the day of 1) the first ECG that establishes stable SR (defined in Section 3.2.1), or; 2) the last ECV attempt for patients who fail to convert to stable SR, or; 3) the Week 0 Visit, for patients in AF who do not undergo ECV for any reason. Ends on the day of the Week 24 Visit or the End of Study (EOS) Visit, if patient discontinues prior to Week 24 Visit.
- Total Follow-up Period: starts on the same day as the 24-Week Follow-up Period and extends until the EOS Visit.
- Total Study Period: starts on the day of the Randomization Visit and extends until the EOS Visit.

PATIENT CHARACTERISTICS

Screen Failure

Screen failure reasons will be tabulated in order of frequency. These reasons are collected on the eCRF DEMOG form.

Randomization

Randomized treatment assignment is centralized and in versions 1 and 2 of the protocol was stratified by: 1) HF etiology (ischemic/non-ischemic); 2) LVEF ($< 0.35/ \geq 0.35$) and; 3) type of Medtronic device (Reveal/Non-Reveal/No Device). In protocol version 3 a fourth strata was added: rhythm status at randomization: (SR vs AF). The count of patients randomized by strata within each treatment group will be tabulated by site and overall. The randomization process will be described in full detail.

Baseline Characteristics

The treatment groups will be examined for comparability with respect to demographics, cardiovascular history, AF risk factors, current disease state, HF and AF therapies, physical exam abnormalities, CYP2D6 and α_{2C} genotyping, vital signs, ECG and laboratory parameters

using descriptive statistics. Continuous variables will be analyzed with a mean, standard error, standard deviation, median, minimum, maximum and n=count of results available. Categorical variables will be described with n=count of results available and percentage of study population, with a clear explanation of the denominators provided in footnotes when necessary.

Treatment Exposure and Compliance

The treatment groups will be examined for comparability with respect to the outcome of the titration period (broken down by pre-study beta blocker usage), the attainment of target dose and the days of double blind treatment by dose level and overall. Elapsed days and days of treatment exposure during the four follow-up periods will also be described by treatment group.

Compliance since the previous visit is reported by the sites on the VISREC eCRF form. Overall compliance rates for the 24-Week Follow-up Period and the Total Study Period will be calculated for each patient and compared between the two treatment groups with descriptive statistics. Note that if a patient discontinues study treatment, compliance is calculated through the date of discontinuation.

Concomitant Medications

Patients must be receiving optimal anticoagulation therapy for stroke prevention. A tabulation of anticoagulant drug usage by treatment group will be generated. For warfarin users, INR is collected on the LAB eCRF as the following ranges: < 1, ≥ 1 and < 2, ≥ 2 and < 3, ≥ 3 and < 4, ≥ 4. A tabulation of these reported ranges by treatment group will be generated for each of the study visits in which reporting is required.

Reported usage of all concomitant medications during the study will be standardized with preferred name and Anatomical Therapeutic Classification (ATC) using the WHODrug dictionary for tabulation by treatment group.

Metrics for Key Study Procedures

Metrics for the following study procedures and medical interventions will be presented with descriptive statistics by randomized treatment group:

- The cardiac rhythm status of every patient at both the Randomization Visit and at the start of the 24-Week Follow-up Period will be tabulated as follows.
  - Patients in Stable SR at Week 0 who did not require ECV
    - Pts in SR at Randomization
    - Pts in AF at Randomization
  - Patients in Stable SR at Week 0 who did require ECV
    - Pts in SR at Randomization
    - Pts in AF at Randomization
  - Patients in AF/AFL at Week 0
    - Pts in SR at Randomization

        - Pts in AF at Randomization
    - Death/Loss to Follow-up (LTF) prior to Week 0
        - Pts in SR at Randomization
        - Pts in AF at Randomization

- Elapsed days on treatment prior to ECV.
- Outcome of ECV.
- Compliance with procedures for collection of transtelephonic monitoring (TTM) results, and
- Compliance with procedures for collection of Medtronic device results.

Final Study Disposition

The disposition of patients screened and randomized into the study will be tabulated by treatment group and displayed with a flow diagram. This will include the counts of screens, screen failures, re-screens, randomizations, completion of the Week 24 Visit, reasons for permanent discontinuation of study treatment and reasons for discontinuation of study follow-up (broken down by pre/post Week 24 Visit). Note that all patients classified as completing the Week 24 Visit will have all components of the primary and secondary endpoints ascertained through the entire 24-Week Follow-up Period.

Protocol Deviations

ARCA Clinical Operations maintains an Excel spreadsheet of protocol deviations reported during the study. Each protocol deviation is classified as being Major or Minor, based on its potential impact on clinical results per ARCA SOP CLIN-005. Tabulations and listings of the reported protocol deviations will be provided for both treatment groups.

EFFICACY ANALYSIS

General Methodology

Time-to-Event Analysis Methodology

Time-to-event is calculated as the date of the event minus the date of initiation of efficacy follow-up, with 1 added in order to include both the start date and end date of the interval.

For all endpoints, follow-up will be censored when a patient receives a cardiac transplant, is declared to be permanently lost to follow-up or withdraws consent. The follow-up periods and specific censoring rules are identified in the endpoint descriptions.

These analyses will be a two-tailed comparison of bucindolol and metoprolol, using the log rank statistic with the exact variance calculation stratified by the randomized treatment assignment strata: 1) HF etiology (ischemic/non-ischemic); 2) LVEF ($< 0.35 / \geq 0.35$); 3) type of Medtronic device (Reveal/Non-Reveal/No Device); and 4) rhythm status at randomization: (SR vs AF). Note that patients enrolled under versions 1 and 2 of the protocol were not stratified by rhythm status however their rhythm status is known due to inclusion criteria (all were in AF). The calculations will be performed with the SAS® LIFETEST procedure, with the stratification

variables specified in the STRATA statement and the TEST statement used to specify the treatment group comparator and any covariates being examined. Cox's proportional hazards model will be used to calculate estimated hazard ratios and 95% confidence intervals. The calculations will be performed with the SAS PHREG procedure, with the stratification variables specified in the STRATA statement and the treatment group comparator and any covariates being examined specified in the MODEL statement. For the primary endpoint, the appropriateness of assuming proportional hazards will be explored by the graphing of log (-log(survival function)) over follow-up for each treatment group.

Where appropriate, Kaplan-Meier survival curves for bucindolol versus metoprolol will be generated to provide a graphical comparison of the two treatment groups.

Follow-up for the time-to-event endpoints will generally end either at the Week 24 Visit or the EOS Visit for the Total Follow-up Period or Total Study Period endpoints. If the Week 24 Visit falls later than day 180, follow-up will be censored on day 180.

Components of Combined Endpoints

This report will contain many endpoints that involve the time to the first occurrence of multiple events, such as AF/AFL onset, mortality or hospitalization. For these endpoints, the count of first events provided by each component will be tabulated. In addition, each component of the combined endpoints will be analyzed separately with a time-to-first-event analysis following the same methodologies used for the combined statistic.

Adjudication

A Clinical Events Classification (CEC) group will adjudicate the primary endpoint, first symptomatic AF/AFL event or death during the 24-Week Follow-up Period. As part of the adjudication process for the primary endpoint, the CEC will also evaluate the secondary endpoint of first AF/AFL event (i.e., symptomatic or asymptomatic). Specifically, the ECGs for the first report of AF/AFL will be reviewed and adjudicated for the presence of AF/AFL regardless of the symptom status. If the first protocol-defined AF/AFL event is not considered a symptomatic AF/AFL event, the triggering process will continue for that patient until the first symptomatic AF/AFL event is identified for the primary endpoint. The CEC over-read of ECG tracings will be used in the calculation of other pertinent study endpoints (such as non-symptomatic AF/AFL within the 24-Week Follow-up Period). More details can be found in the CEC Charter.

Core Lab and Transtelephonic Monitoring

In the original study protocol, an Electrophysiology Core Lab (Agility Centralized Research Services) provided a centralized ECG interpretation of the individual ECGs performed at the clinic site and the transtelephonic monitors (TTM) worn by the patients, both during the 24-Week Follow-up Period. In version 4 of the protocol, the collection of these two sources of data was discontinued. The CEC adjudication process was not in production mode at that time point, so it was decided the CEC would perform their own interpretation (over-reads) of the site ECG tracings and not use any of the Core Lab interpretations. Further, the CEC adjudication would make use of available TTM data.

Hospitalization

Many of the efficacy endpoints involve hospitalization. Only non-voluntary, overnight hospital admissions will be included in these endpoints; emergency room visits will not be included. Patients in this study will often have scheduled hospital admissions for treatment of their AF and/or HF. Examples include ablation procedures, Tikosyn induction, placement/replacement of implanted devices, and IV drug treatment. These will not be included in the endpoints. The eCRF specifically collects the investigator's assessment of hospitalization causation, which includes assessments of non-CV, CV and HF hospitalizations. In addition to the investigator assessment of causation, the data will be reviewed by the Sponsor via a blinded listing review prior to database lock to confirm which hospitalizations are considered voluntary, overnight admissions.

Data Collection Cut-off at End of Study

The protocol states the study will end with approximately 620 randomized patients and accrual of at least 330 primary endpoint events, presuming the sample size and target event counts are not altered due to the Phase 3 interim analysis (see DSMB Charter). At this point, any patients still participating in the 24-Week Follow-up Period will remain on blinded study treatment until they complete the Week 24 Visit. Those patients in the Extension Period will be called in for an EOS Visit.

Missing Data Due to Withdrawal or Loss to Follow-up

The rate of withdrawal or loss to follow-up prior to the Week 24 Visit is expected to be low. If a withdrawal or loss to follow-up occurs prior to the Week 24 Visit, all time-to-event endpoints will be censored as of the last completed visit. Note that patients that withdraw from the study will be requested to consent to have their vital status checked via phone calls. If deaths are detected by this procedure the date of death will be incorporated into the efficacy and safety datasets and analyses.

P-value Adjustment for Interim Analysis

The goals and operational details for the interim efficacy analysis and ongoing safety monitoring can be found in the DSMB Charter and the DSMB SAP.

At the end of Phase 3, the alpha level for the primary endpoint will be reduced to 0.04989 to adjust for the Phase 2B ($\alpha = 0.00001$) and Phase 3 ($\alpha = 0.0001$) interim analyses.

Efficacy Endpoints

Primary Efficacy Endpoint

The primary endpoint is elapsed time-to-first-event of symptomatic AF/AFL or all-cause mortality (ACM) during the 24-Week Follow-up Period. This is a time-to-event endpoint censored at the end of the 24-Week Follow-up Period. The identification of first event of symptomatic AF/AFL or death is provided by the CEC. The CEC does not distinguish between the presence of AF or AFL so a component analysis will not be possible.

The following definitions apply to this endpoint:

- Stable SR on study drug is defined as any of the following:
  - SR confirmed $\geq$ 1 hour after ECV.

  - SR confirmed ≥ 1 hour after spontaneous conversion from AF/AFL.
  - SR confirmed ≥ 1 hour at the Week 0 Visit for patients randomized in SR.

- An AF/AFL event is defined as AF or AFL observed on two consecutive measures separated by at least 10 minutes as assessed by ECG/TTM.

- A symptomatic AF/AFL event is defined as an AF/AFL event that is associated with a clinically relevant change in patient-reported symptoms, as determined by the CEC examination of blinded data.

The CEC charter and associated documents describe the "triggers" that are established to identify events for their consideration and the data sources to be used in their adjudication proceedings. The charter also describes their approach for identifying an AF event as symptomatic and for identifying the onset date and time of the event since that is needed for this time-to-event endpoint. Note that version 3 of the protocol involved a comprehensive change to the symptoms collected, with 6 of the original 8 symptoms having their descriptions modified and 2 new symptoms being added. Also the symptom characteristics were clarified with addition of a 'frequency' field to the collection form. All of these changes were made to give the CEC more specific information to support their identifying symptoms that were new or worsened in association with AF onset. Since these changes were implemented after only 12 patients were randomized (2% of the planned 620) and the identification of overall symptom onset/worsening is an adjudicated decision, no modification of analysis methodology is planned.

AF/AFL will be assessed at scheduled and unscheduled clinic visits via 12-lead ECG. Patients will be queried at the time of each ECG assessment to determine if they have experienced any change in symptoms that could be potentially related to AF.

The vast majority of patients will either be in SR or successfully convert from AF to SR after one or two ECV procedures around three weeks after they begin randomized treatment. However, there are several scenarios that depart from this norm and the methodology for establishing the start of efficacy follow-up and censoring for the primary endpoint is described below:

1. Spontaneous conversion to stable SR prior to the planned cardioversion. For these patients, the day of the first ECG assessment that meets the definition of stable SR, as defined above, will be designated as Day 1 of the 24-Week Follow-up Period.

2. Failure to attain stable SR because the ECV procedure was not performed due to drop out or any reason other than those described below. These patients will be included in the analysis as censored on Day 1 of the 24-Week Follow-up Period.

3. Failure to attain stable SR, either spontaneously or following ECV. These patients will be included in the endpoint calculation as experiencing the event on Day 1 of the 24-Week Follow-up Period.

4. Deaths occurring after randomization and prior to conversion to stable SR will be counted as events on Day 1 of the 24-Week Follow-up Period.

5. Patients with AF/AFL stopped at the Week 0 visit by any means other than ECV will be censored on Day 1 of the 24-Week Follow-up Period. An example is the performance of AV nodal ablation at the Week 0 visit.

The primary endpoint analysis will also be performed within the following prospectively identified subgroups based on pathophysiological or clinical importance:

1) Started the 24-Week Follow-up Period in SR vs AF
2) LVEF strata at randomization: ≤0.35 vs. >35
3) Gender
4) Ischemic etiology vs. nonischemic
5) Age above/below median
6) Duration of AF diagnosis above and below median.
7) Baseline norepinephrine above and below median
8) Baseline NT-proBNP
9) α_{2C} AR polymorphisms (i.e., Del carriers vs. α_{2C} wild type homozygotes).

In exploratory analyses, the following covariates will be included as potentially relevant explanatory variables in the Cox regression models:

1. Initial study treatment dose level.
2. Baseline NYHA Class.
3. Gender.
4. Race.
5. Age.
6. Baseline serum creatinine.
7. Baseline norepinephrine level.
8. Baseline heart rate.
9. Baseline systolic blood pressure.
10. History of diabetes.
11. Duration of AF diagnosis.
12. Previous amiodarone use (both historical and stopped just prior to randomization).
13. Ablation procedure prior to study.
14. Therapeutic device type: CRT, ICD, single ventricular lead pacemaker.
15. For the subset of patients in AF at baseline, type of rhythm abnormality: (paroxysmal AF or persistent AF).
16. For the subset of patients in SR at baseline: the time since last attaining SR, the type of previous rhythm abnormality, and the intervention that ended the previous AF episode.
17. Elapsed days of treatment from randomization date to start of the 24-Week Follow-up Period.
18. CYP2D6 metabolizer status.
19. α_{2C} AR polymorphisms (i.e., Del carriers vs. α_{2C} wild type homozygotes).
20. Country in which clinic site is located.
21. Other clinically significant AF risk factors.

Additional exploratory analyses will include the following:

- A qualitative analysis of the symptoms associated with the primary endpoint events. The symptoms will be classified as arrhythmia-related (palpitations or lightheadedness/dizziness) HF-related (fatigue or tiredness, weakness or problems exercise, weight gain or swelling of both legs and/or feet), or both.

- For patients with primary endpoint events of symptomatic AF/AFL, how many had prior events of asymptomatic AF/AF that progressed into symptomatic.

The following sensitivity analyses will be performed:

- A subpopulation analysis including only those patients beginning the 24-Week Follow-up Period in SR.
- In the per-protocol analysis, endpoint events and deaths that occur more than 30 days after permanent discontinuation of study treatment are omitted.
- All Week 24 Visits included (ie - no exclusion of events observed at Week 24 Visits after day 180).
- Patients that have not previously reverted to AF/AFL that withdraw or are lost to follow-up prior to the Week 24 Visit, will be assigned an AF/AFL event at the first missed clinic visit or scheduled TTM.
- Patients that withdraw or are lost to follow-up prior to the Week 24 Visit are omitted from the analysis.

Secondary Efficacy Endpoints

The following endpoints will be tested for superiority of bucindolol benefit relative to metoprolol by fixed sequence provided that bucindolol is found to be significantly superior in the primary endpoint. The time-to-event endpoint methodology described in Sections 3.1.1 and 3.2.1 for events involving AF/AFL recurrence will be used unless otherwise noted:

- Time-to-first-event of AF/AFL (i.e., symptomatic or asymptomatic) or ACM during the 24-Week Follow-up Period.

  Supportive Analyses:

  Events accrued during the Total Follow-up Period.

  For patients with events based on symptomatic AFL, the rate of patients subsequently progressing to AF. Also for these patients, the elapsed time from symptomatic AFL to AF.

  Data Source:

  ECG (over-read by CEC for first 24 weeks)

  TTM (first 24 weeks only)

- Proportion of patients with VT, VF, or symptomatic supraventricular tachycardia (SVT) during the 24-Week Follow-up Period. Includes VF and symptomatic SVT events of any duration, VT events ≥ 15 seconds, and VT events that result in appropriate firing of an ICD. It will be tested with a Cochran-Mantel-Haenszel statistic to control for the four stratification variables.

  Supportive Analyses:

  Events accrued during the Total Follow-up Period.

Data Source:

The CVEVENT eCRF form is the source of all components of these compound endpoints.

- Total all-cause hospitalization days per patient during the Total Study Period. The count of hospitalization days will be normalized for the total number of days of follow-up prior to testing with the Wilcoxon Rank Sum statistic.

  Supportive Analyses:

  Number of heart failure hospitalization days per patient.

  All-cause hospitalization days through first recurrence of AF/AFL versus days after recurrence, normalized for days of follow-up within each period. The comparison will take place within treatment group and across treatment.

  All-cause hospitalization days for patients with ventricular rate control (VRR) control compared to those without VRR control. The comparison will take place within treatment group and across treatment.

  Data Source:

  The HOSP eCRF form provides the number of hospitalization days and the reason for hospitalization.

  The ECG and AE eCRF will be used to identify the patients in AF with VRR control at the end of the study.

- Time-to-first-event of AF/AFL (i.e., symptomatic or asymptomatic), HF hospitalization (as assessed by the Investigator), or ACM during the Total Follow-up Period. As in the primary endpoint, any incidence of ACM prior to start of the 24-Week Follow-up Period will be analyzed as an event on Day 1. Hospitalization prior to Week 0 are not included, but those are included in the safety analyses.

  Supportive Analyses:

  - Events accrued during the 24-Week Follow-up Period.
  - Combinations of each component ((i.e., AF/AFL+ACM, AF/AFL+HFH, HFH+ACM).

  Data Source:

  - ECG (over-read by CEC for first 24 weeks), HOSP and DEATH eCRF forms.
  - TTM (first 24 weeks only).

- Proportion of patients with adequate ventricular rate control (VRR) in the setting of AF/AFL. Adequate VRR in setting of AF/AFL is defined as follows: 1) the presence of AF or AFL; 2) a VRR between 40 and 80 beats per minute (bpm) at rest; and 3) the absence of symptoms associated with bradycardia. Thus this is a subset analysis only involving patients with AF/AFL recurrence. The endpoint is evaluated for the last tracing demonstrating AF/AFL during the 24-Week Follow-up Period prior to intervention (eg:

ablation, ECV, initiation of anti-arrhythmic drugs). Will be tested with a Cochran-Mantel-Haenszel statistic to control for the four stratification variables.

Supportive Analyses:

- Evaluated for the last tracing demonstrating AF/AFL when the patient is still on study treatment during the 24-Week Follow-up Period.

Data Source:

- ECG and AE eCRF form (for symptomatic bradycardia).

Tertiary Efficacy Endpoints

The following endpoints will be tested for superiority of bucindolol benefit relative to metoprolol. The time-to-event endpoint methodology described in Section 3.1.1 and 3.2.1 for events involving AF/AFL recurrence will be used unless otherwise noted:

- Time-to-first-event of VT/VF or ACM during the Total Follow-up Period. Includes VF events of any duration, VT events of ≥ 15 seconds, and VT events that result in appropriate firing of an ICD.

  Supportive Analyses:

  - Events accrued during the 24-Week Follow-up Period.

  Data Source:

  - CVEVENT and DEATH eCRF forms.

- Time-to-first-event of AF/AFL (i.e., symptomatic or asymptomatic), CV-related hospitalization (as assessed by the Investigator), or ACM during the Total Study Follow-up Period.

  Supportive Analyses:

  - Events accrued during the 24-Week Follow-up Period.
  - Combinations of each component (i.e., AF/AFL+ACM, AF/AFL+CVH, CVH+ACM).

  Data Source:

  - ECG (over-read by CEC during the 24-Week Follow-up Period), HOSP and DEATH eCRF forms.
  - TTM (24-Week Follow-up Period).

- Proportion of patients with stroke or systemic embolism during the Total Follow-up Period. Stroke is defined as a focal neurologic deficit from a non-traumatic ischemic, hemorrhagic, or uncertain cause lasting at least 24 hours (as assessed by the Investigator). Tested with a Cochran-Mantel-Haenszel statistic to control for the four stratification variables.

  Data Source:

- CVEVENT eCRF form.

- Proportion of patients randomized with AF/AFL who convert to stable SR (spontaneous or post-ECV) and enter the 24-Week Follow-up Period. Tested with a Cochran-Mantel-Haenszel statistic to control for the four stratification variables.

  Supportive Analyses:

  Subset of patients with spontaneous conversion.

  Data Source:

  FUSTART eCRF form.

- Total number of ECV procedures per patient during the Total Study Period. This count will be normalized for the total number of days of follow-up prior to testing with the Wilcoxon Rank Sum statistic.

  Data Source:

  ECV eCRF form.

- Proportion of patients at Week 24 Visit who are receiving study drug and have not had an AF/AFL event. Tested with a Cochran-Mantel-Haenszel statistic to control for the four stratification variables.

  Data Source:

  ECG (over-read by CEC), DRUGLOG and EOT eCRF forms.

  TTM (24-Week Follow-up Period).

- Change in NT-proBNP, assessed relative to baseline (Randomization Visit). Change from baseline will be tested for greater reduction in the bucindolol treatment group with the Wilcoxon Rank Sum test because of the expected lack of normality of this measure.

  Data source:

  LabCorp vendor dataset.

- Change in norepinephrine, assessed relative to baseline (Randomization Visit). Change from baseline will be tested for greater reduction in the bucindolol treatment group with the Wilcoxon Rank Sum test because of the expected lack of normality of this measure.

  Data source:

  LabCorp vendor dataset.

- The EQ-5D questionnaire has 5 dimensions (mobility, self-care, usual activities, pain/discomfort and anxiety/depression) and each is self-rated by the patient as no problems, some problems, or severe problems. The results for each dimension will be analyzed individually at both time points. The change from randomization to each visit will be categorized as improved or no change/worsened and the proportions of these categories in both treatment groups will be tabulated with a 2 by 2 table. The bucindolol

treatment group will be tested for superior response using a Cochran-Mantel-Haenszel statistic to control for the four stratification variables.

Data source:

EQ-5D eCRF form.

- Pharmacoeconomic modeling of healthcare utilization. Details of this analysis will be prespecified in a separate analysis plan.

SAFETY ANALYSIS

The following four periods are established for analysis of safety endpoints:

- 24-Week On-Drug Period: starts at day of randomization and extends to latest visit attended through Week 24 Visit. For patients that discontinue treatment early, data collected through 30 days after the final dose of study treatment are included.
- 24-Week On-Study Period: starts at day of randomization and extends to latest visit attended through Week 24 Visit. For patients that discontinue the study prior to Week 24, data collected through 30 days after the final study visit are included. Study treatment status is not considered for data inclusion.
- Total Study On-Drug Period: starts at day of randomization and extends through 30 days after the final dose of study treatment.
- Total Study On-Study Period: starts at day of randomization and extends through 30 days after final clinic visit attended. Study treatment status is not considered for data inclusion.

Analysis of SAEs will be performed for all four timeframes. For the other safety endpoints, the 24-Week On-Study and Total Study On-Study Periods will be used. If treatment group imbalances are observed for an endpoint, it will be further analyzed with the other data inclusion timeframes.

The results for the following safety endpoints will be compared with descriptive statistics between the treatment groups for all patients receiving study treatment. Results collected from first dose of study drug to 30 days after the last dose for each patient will be included in the assessments of safety. Results specific to scheduled visits will be included in the by-visit analyses if they were collected within a ± 7-day window for the prescribed visit study day.

- Incidence of ACM during the Total Study Period.

  Supportive Analyses:

  The association of VRR control with mortality will be examined using the final heart rate measurement available for each patient (comparisons will be within the treatment groups).

  Data Source:

  DEATH eCRF form.

- Incidence of ACM, CV-related hospitalization (as assessed by the Investigator), or withdrawal of study drug due to an AE during the Drug Titration Period.

  Data source:

  DEATH, HOSP, EOS and AE eCRF forms.

- Incidence of symptomatic heart block during the Total Study Period. Symptomatic Heart Block is defined as the first of any of the following: 1) 3rd degree heart block (complete heart block); 2) any 2nd degree heart block with the presence of symptoms attributable to, and temporally correlated with the occurrence of heart block which include any of the following: Near-fainting or fainting (syncope) / Dizziness; Weakness or Fatigue; Shortness of breath; Chest pain; or 3) 2nd or 3rd degree heart block requiring implantation of a permanent pacemaker (with or without defibrillator).

  Data source:

  CVEVENT and AE eCRF forms.

- Overall incidence and severity of treatment-emergent AEs/SAEs over time during the Total Study Period. Also events associated with device implantation. The events will have standardized MedDRA preferred terms and System Organ Classes assigned to them for tabulation.

  Supportive analyses:

  Incidence of AEs leading to reduction, interruption or permanent discontinuation of study treatment.

  Incidence of AEs associated with device implantation.

  Incidence of AEs by CYP2D6 metabolizer status.

  Incidence of AEs by α2C AR polymorphisms.

  Data source:

  AE eCRF form.

- Incidence of neoplasm-related AEs during the Total Study Period. The AEs of special interest will be tabulated according to the following characteristics.

  Development of treatment-emergent neoplastic conditions.

  Progression or worsening of pre-study neoplastic conditions.

  Progression or worsening of treatment-emergent neoplastic conditions.

  Data source:

  - AE, NEOPLHX and NEOPLAS eCRF forms.

- Clinical Chemistry and Hematology.

  Visit collection: screen, start of follow-up Week 0 (protocol versions 1 and 2), Week 4 (protocol versions 3 and 4), Week 12 (protocol versions 3 and 4), Week 24,

every 24 weeks during extension, end of treatment and end of study. Screen results will serve as the pre-treatment baseline.

Change from baseline to each planned study visit of collection will be calculated and analyzed with descriptive statistics.

The numbers and percentages of patients with values exceeding the bounds of normal ranges will be tabulated for scheduled visits.

The numbers and percentages of patients with values exceeding the panic bounds each visit.

Data source:

LabCorp vendor-supplied dataset.

- ECG quantitative parameters.

  Measured at every visit. Randomization Visit measurement prior to first dose will serve as the baseline. Will be analyzed at Week 0, 4, 12 and 24 visits as well as end of treatment and end of study.

  Change from baseline to each analysis visit will be calculated and analyzed with descriptive statistics.

  The numbers and percentages of patients with QTc increase from baseline exceeding 60 ms at any time point during the study.

  Data source:

  - ECG eCRF form.

- Vital signs and weight (data source: VITALS eCRF form).

  Measured at every in-clinic visit. Randomization Visit measurement prior to first dose will serve as the baseline. Will be analyzed at Week 0, 4, 12 and 24 visits as well as end of treatment and end of study.

  Change from baseline to each analysis visit will be calculated and analyzed with descriptive statistics.

  Data source:

  VITALS eCRF form.

- Proportion of patients attaining target study drug dose during the Drug Titration Period. Will be calculated for all patients, those receiving β-blocker therapy prior to randomization and those not previously receiving β-blocker therapy.

  Data Sources:

  VISREC and DRUGLOG eCRF forms.

MEASUREMENTS OF INTEREST AND SUBSTUDIES

- AF Burden (AFB) Substudy.

In this optional substudy, AFB, defined as the amount of time per day that a patient is in AF/AFL, is measured by implanted Medtronic devices, including cardiac monitors, pacemakers, cardioverter-defibrillators, and cardiac resynchronization therapy. These devices also measure VRR during periods of AF. Approximately 50% of the study participants are expected to participate in the AFB substudy.

The distribution of device types will be presented by treatment group, by patient baseline characteristics, by disease severity, by treatment exposure prior to device implantation and elapsed days to start of the 24-Week Follow-up Period. AFB will be presented as hours/day in graphical displays for each patient with the dates of randomization and initial ECV and other interventions annotated.

The treatment efficacy endpoint will be the time to first device-detected event or ACM, with an event defined as at least 6 hours of AFB in a single day. This endpoint will be analyzed through the Week 24 Visit with the same methodology used for the study primary endpoint. Patients with no AFB data available after the start of the 24-Week Follow-up Period will be excluded. Patients with an implanted therapeutic device that produces paced rhythm which confounds the measurement of AFB will also be excluded.

- Supportive Analyses:

  Time to device detected AF/AFL event during the Total Follow-up Period.

  The proportion of patients with VRR on the last day demonstrating AF/AFL during the 24-Week Follow-up Period. Will be tested using a Cochran-Mantel-Haenszel statistic to control for the four stratification variables.

  The percent of follow-up days in AFB, calculated as the number of days with AFB of at least six hours divided by the total number of days in the 24-Week Follow-up Period. Statistical testing will be performed with the Wilcoxon Rank Sum Statistic. A sensitivity analysis will be performed on the subset of patients beginning the 24-Week Follow-up Period in SR.

- Data Sources:

  Medtronic vendor-supplied dataset.

- DNA Bank, with collection at time of screening, for patients who agree to participate in the substudy. No analysis of these data have been pre-planned.

- Sparse sampling of bucindolol hydrochloride plasma concentrations for population pharmacokinetic analysis. The analysis plan for the substudy will be prepared separately prior to unblinding.

GENETIC-AF Clinical Trial

Phase 2B Statistical Analysis Plan Amendment

RATIONALE FOR PHASE 2B STATISTICAL ANALYSIS PLAN

On the pre-specified first interim analysis of the GENETIC-AF trial conducted on August 7, 2017, based on application of pre-defined Bayesian predictive probability of success (PPoS) modeling of the "modified primary endpoint" data, the GENETIC-AF Data and Safety Monitoring Board (DSMB) recommended completing the trial in Phase 2B rather than immediately stopping for futility or "seamlessly" transitioning to Phase 3. Shortly thereafter, the Sponsor (ARCA biopharma) informed the trial investigators of the DSMB decision and instructed sites to complete follow-up of all randomized patients by December 31, 2017. This implies that 267 patients will constitute the final Phase 2B population, with nearly all of them having completed the planned 24 weeks of follow-up or having reached the Phase 2B modified primary endpoint (hereafter referred to as the Phase 2B primary endpoint) of time to symptomatic or asymptomatic atrial fibrillation/atrial flutter (AF/AFL) or all-cause mortality (ACM).

The DSMB Phase 2B interim analysis, conducted and reported to the Sponsor on August 7, 2017 was based on 103 AF/AFL/ACM events from 215 patients randomized through June 19, 2017 including 162 who had attained full follow-up or experienced the Phase 2B primary endpoint. In contrast, the completed Phase 2B dataset on 267 patients will likely include approximately 50% more Phase 2B primary endpoint events. Currently the patients are attending final study visits and all data are being subjected to full monitoring QA during close-out of each site. ARCA expects to receive the final data and treatment assignments in February of 2018.

The GENETIC-AF Statistical Analysis Plan (SAP)[1], which focused primarily on analyses pertinent to the Phase 3 population, was completed on March 15, 2017 and submitted to FDA on March 30, 2017. In the Phase 3 SAP, the primary efficacy endpoint is time to symptomatic AF/AFL or ACM, which was powered based on an expectation of 330 events from a total of approximately 620 patients. As this study is now stopping at Phase 2B, ARCA estimates that the total number of events will be less than half of what was planned for the full Phase 3 study. As such, the prespecified analysis described in the SAP for the Phase 3 primary endpoint is not expected to provide adequate guidance to the Sponsor regarding the utility of conducting a reasonably sized Phase 3 trial based on a time to AF/AFL/ACM primary endpoint.

The DSMB charter[2] was approved on October 13, 2015 and submitted to FDA on October 16, 2015. In the charter, the DSMB acknowledges that a traditional time-to-first AF/AFL/ACM event analysis would have very low statistical power for a population of 200-250 patients; therefore, the DSMB charter and an accompanying white paper[3] outlined a Bayesian methodology for the interim analysis that would be more informative for the Phase 2B population. More specifically, the DSMB charter identified time to first event of symptomatic or asymptomatic AF/AFL or ACM as the primary efficacy endpoint for the Phase 2B interim analysis, since this more inclusive endpoint was expected to have significantly more events than the Phase 3 primary endpoint (i.e., symptomatic AF/AFL or ACM). ARCA's ongoing review of blinded data supports this conclusion, with approximately 75% of first AF/AFL events being adjudicated as symptomatic and 25% of events being adjudicated as asymptomatic.

Therefore, ARCA plans to conduct the primary efficacy analysis of this Phase 2B study in a similar manner, following the Bayesian methodology that was prespecified in the DSMB charter for the Phase 2B interim analysis. As described below, these analyses will model the Phase 2B data to generate Bayesian predictive probability of success (PPoS) values for a discrete Phase 3 trial with 620 randomized patients who have accrued 330 events (i.e., symptomatic or asymptomatic AF/AFL or ACM). Additional Bayesian modeling will also be performed for Phase 3 planning purposes but these analyses will be secondary to the Phase 2B primary efficacy analysis described above. ARCA will also perform all analyses described in the GENETIC-AF SAP, recognizing that most of these endpoints (e.g., symptomatic AF/AFL, hospitalizations, mortality) will be significantly under powered and primarily hypothesis-generating in nature.

DESCRIPTION OF PHASE 2B STATISTICAL ANALYSES

As described in the DSMB Charter[2], the of time to first event of AF/AFL or ACM endpoint will be subjected to Bayesian modeling for derivation of PPoS estimates by Berry Consultants, Austin TX (Dr. Ben Saville, Project Lead). The PPoS bands and boundaries, identical to those described in the first interim analysis, are given in Figure 1 and will be used to inform/guide the Sponsor. The primary efficacy analysis will be based on Bayesian modeling of the Phase 2B data assuming a discrete Phase 3 population of 620 patients with 330 events (i.e., symptomatic or asymptomatic AF/AFL or ACM).

A secondary analysis will also be performed based on Bayesian modeling of the Phase 2B data assuming a discrete Phase 3 population of 820 patients with 440 events (i.e., symptomatic or asymptomatic AF/AFL or ACM). This secondary analysis reflects what ARCA believes is the approximate upper bounds of clinical feasibility for a Phase 3 trial, and was the final sample size planned for the current study if the second (Phase 3) interim analysis described in the DSMB Charter[2] indicated that the data was in the "promising zone"[4].

As described in Section 3.2.1 of the GENETIC-AF SAP[1], sensitivity analyses will be performed on both the primary and secondary models described above for the subset of patients who began the 24-week Follow-up Period in sinus rhythm. Additional exploratory analyses may also be performed with other sample sizes and event rates, as necessary.

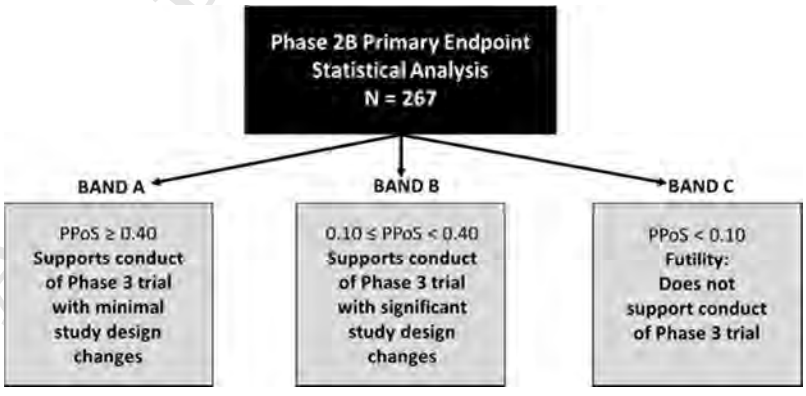


Figure 1. Predictive Probability of Success (PPoS) bands and boundary designations for Bayesian modeling of the GENETIC-AF Phase 2B primary endpoint of time to any AF/AFL or ACM.

All analyses described above will also be repeated for the symptomatic AF/AFL or ACM endpoint; however, since there are significantly fewer events for this endpoint these analyses are considered exploratory and the PPoS boundaries in Figure 1 do not directly apply.

To determine if modification of inclusion/exclusion criteria could improve the design of a future Phase 3 trial, exploratory Bayesian analyses will be conducted following the primary (i.e., 620 patients/330 events) and secondary (i.e., 820 patients/440 events) models described above to explore treatment effects in various subgroups.

1 Subgroups of interests are prespecified in Section 3.2.1 of the GENETIC-AF SAP[1]. For the Phase 2B analysis, the following subgroups have been prioritized in order of importance based on pathophysiological and/or clinical relevance:

1) Randomized in sinus rhythm versus AF/AFL
2) LVEF at randomization: ≤ 0.35 versus > 0.35
3) History of persistent AF versus paroxysmal AF
4) Geographic region (USA, Canada, or Europe)

Due to well-known issues associated with inflated false positive rates with subgroup analyses, these analyses will focus on estimation rather than hypothesis testing, and will incorporate Bayesian hierarchical methods to shrink estimated treatment effects in subgroups toward the respective estimate in the overall study population. The GENETIC-AF Steering Committee, which consists of AF and heart failure experts will review the subgroup analyses and determine whether there exists sufficient biologic or clinical plausibility to support further development in any of the subgroups.

REFERENCES

1. GENETIC-AF Phase 3 Statistical Analysis Plan submitted to FDA on March 30, 2017
2. DSMB Charter Version 2.0 submitted to FDA on October 16, 2015
3. DSMB Charter White Paper submitted to FDA on October 16, 2015
4. Chen YHJ, DeMets DL and Lan KKG. Increasing the sample size when the unblinded interim result is promising. Stat Med. 2004; 23:1023-38.

Classification of Heart Failure by LVEF

The definition of heart failure with reduced LV ejection fraction based on a lower limit of normal of 0.50 (1, 2) was used to define HFrEF (LVEF < 0.50 and a history of HF). HFrEF patients were subdivided into HFmrEF (HF with mid-range LVEF) according to Ponikowski et al. as HF with an LVEF ≥ 0.40 and < 0.50 (3), and HFlrEF (HF with "lower-range" LVEFs < 0.40).

References.

1. Davis BR, Kostis JB, Simpson LM, Black HR, Cushman WC, Einhorn PT, et al. and the ALLHAT Collaborative Research Group. Heart failure with preserved and reduced left ventricular ejection fraction in the antihypertensive and lipid-lowering treatment to prevent heart attack trial. Circulation. 2008;118:2259-67.

2. Gupta DK, Shah AM, Castagno D, Takeuchi M, Loehr LR, Fox ER, Butler KR, Mosley TH, Kitzman DW, Solomon SD. Heart failure with preserved ejection fraction in African Americans: The ARIC (Atherosclerosis Risk In Communities) study. JACC Heart Fail. 2013 Apr;1(2):156-63.

3. Ponikowski P, Voors AA, Anker SD et al. 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: The Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC)Developed with the special contribution of the Heart Failure Association (HFA) of the ESC. Eur Heart J 2016; 37:2129-200.

Modeling of Variables and Selection of Optimal Boundaries for Therapeutic Phenotypes

In this exploratory Phase 2 trial with limited sample size and statistical power, we employed precision therapeutic phenotyping to identify HF populations who respond differentially to two beta-blockers based on genetic targeting. This approach circumvents potential issues associated with conventional subset analyses by evaluating monotonicity and consistency of trends across the full continuum of candidate variables. The benefit of deriving these therapeutic phenotype characteristics from continuous variables is that the classifiers are readily conducive to numerical calibration. With discrete and/or categorical classifiers, a hypothetical predictor variable is either correct or not, with limited or no gradation possible as a hedge against spuriousness. For the calibration of the continuous variable DxT and DTRI, one could select more restrictive criteria such as DxT10/DTRI-1 (i.e., < 10 years of AF and HF with AF not preceding HF by more than 1 year), which yields a similar treatment effect estimate (HR = 0.51; 95% CI: 0.30, 0.85) compared to DxT12/DTRI-2 (HR = 0.54; 95% CI: 0.33, 0.87); whereas, more inclusive criteria such as DxT15/DTRI-3 results in only a slight loss of signal (HR = 0.63; 95% CI: 0.40, 0.98). We propose that increasing the permissible limits of variation (i.e., tolerance) for the phenotype selection criteria increases the likelihood of reproducibility of these results in future studies.

AF Symptoms Questionnaire (AFSQ)

1. Since your last clinic visit, have you experienced any of the following:
   a) Heart palpitations (pounding, racing or irregular heart beat)? [Yes/No]
   b) Shortness of breath? [Yes/No]
   c) Chest pain or pressure? [Yes/No]
   d) Fatigue or tiredness? [Yes/No]
   e) Weakness or problems exercising? [Yes/No]
   f) Lightheadedness, dizziness or fainting? [Yes/No]
   g) Confusion/trouble concentrating? [Yes/No]
   h) Sweating unrelated to physical activity? [Yes/No]
   i) Weight gain greater than 2 pounds? [Yes/No]
   j) Swelling of both legs and/or feet? [Yes/No]

2. Which symptom do you consider the predominant or worst symptom?
   [choose only one from above, or 'NA' if no symptom experienced]

3. For questions 1a-j, if "yes" collect the following:
   a) How frequently have you experienced this symptom? [rarely, sometimes, often, always]
   b) How would you rate the intensity/discomfort of this symptom? [mild, moderate, severe]
   c) When did you first experience this symptom during this reporting period? [MM/DD/YYYY]
   d) When did you last experience this symptom during this reporting period? [MM/DD/YYYY]