ONLINE SUPPLEMENT

Supplement Figures

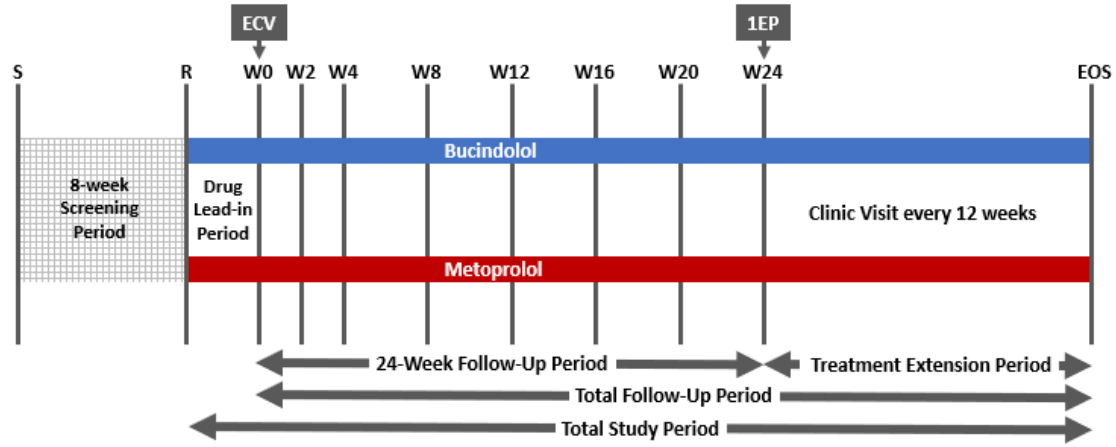
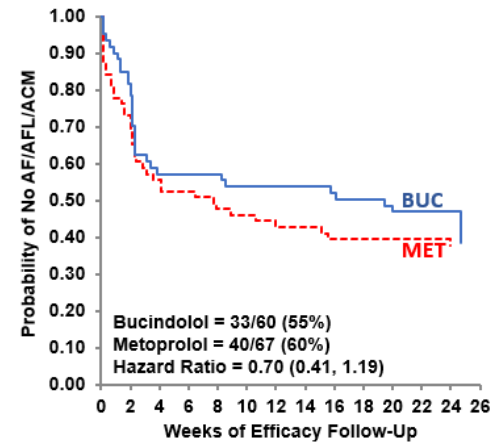
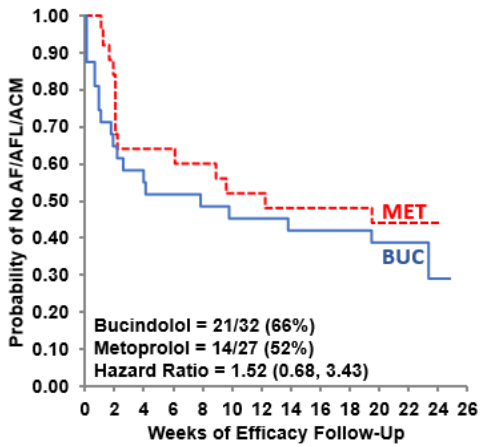


FIGURE 1. GENETIC-AF Study Visit Schedule
Note: ECV performed 3 weeks after randomization, if needed. Week 0 for patients in SR at randomization is 3 weeks (± 3 days). S = Screening Visit; R = Randomization Visit; W = week; ECV = electrical cardioversion; 1EP = primary endpoint; EOS = end of study.



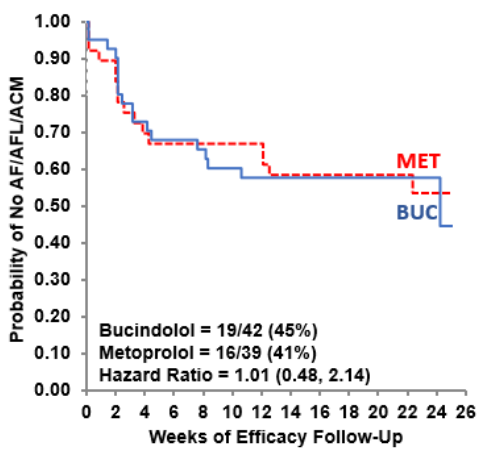
Number at Risk

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.

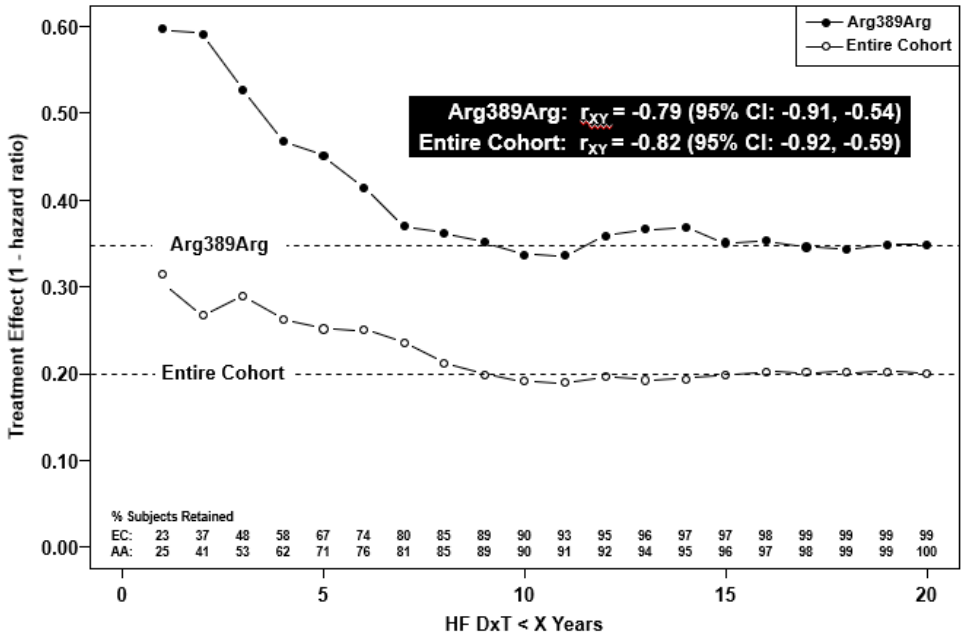


FIGURE 3 Treatment Effect and the Duration of HF in the BEST HF Trial
Entire cohort (open circles, n=2708) and ADRB1 Arg389Arg subgroup (closed circles, n=493). Hazard ratio is for time to first heart failure hospitalization or death for bucindolol and placebo. HF DxT=time from initial HF diagnosis to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

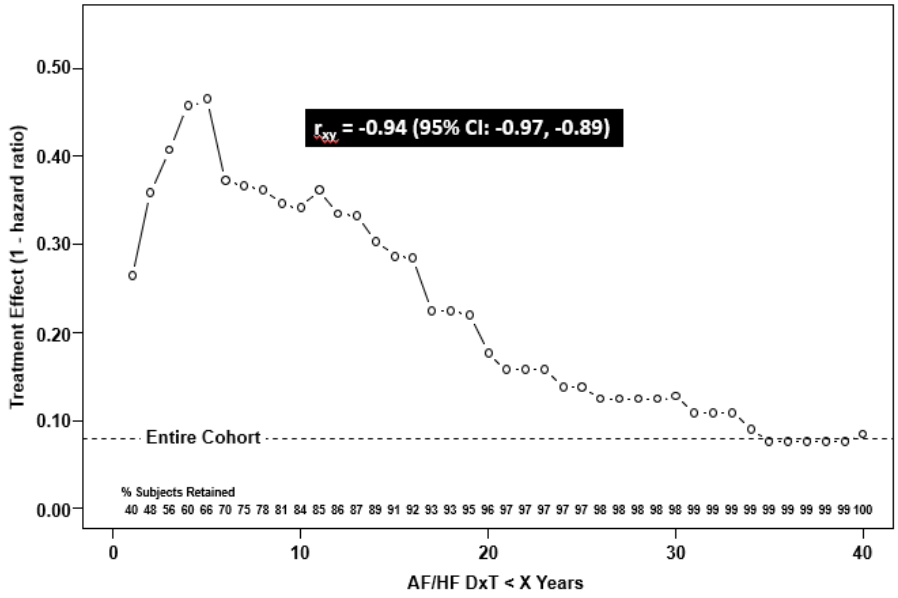
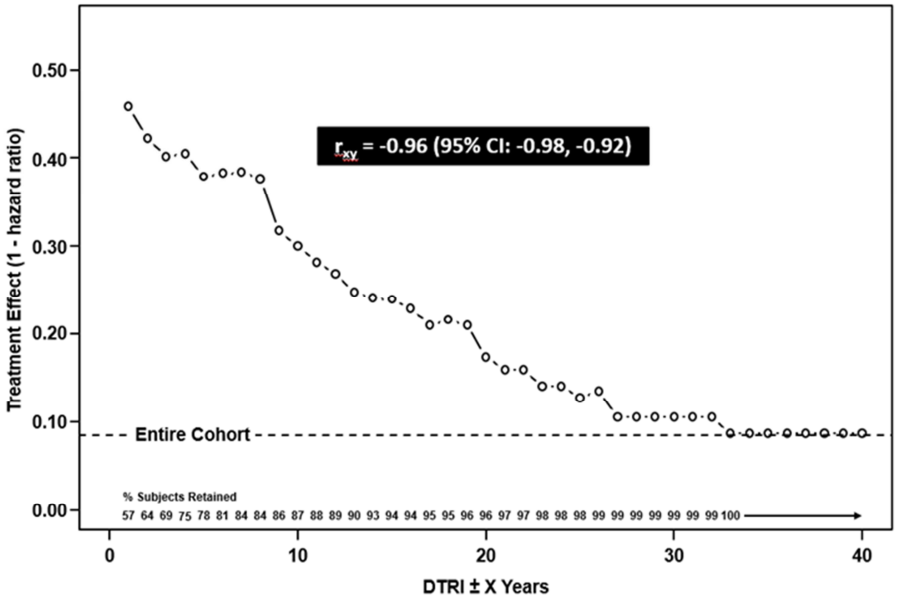


FIGURE 4 Treatment Effect by AF and HF Duration
Treatment effect versus AF/HF DxT (i.e., both HF DxT and AF DxT< X years).
Hazard ratio is for time to AF/AFL/ACM endpoint. AF/HF DxT= time from initial AF and HF diagnosis to randomization.

A.



B.

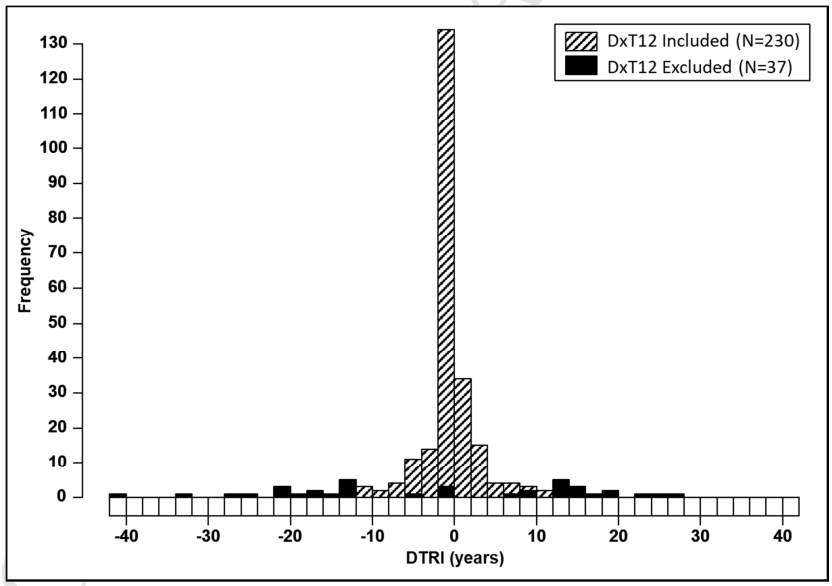


FIGURE 5 Treatment Effect Relationship to Relative Onset of AF and HF (DTRI)
A. Treatment effect versus absolute value of DTRI upper and lower bounds.
B. Histogram of DTRI distribution for DxT12 cohort and cohort excluded by DxT12 criteria.
Hazard ratio is for time to AF/AFL/ACM endpoint. DTRI=Diagnosis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

A.

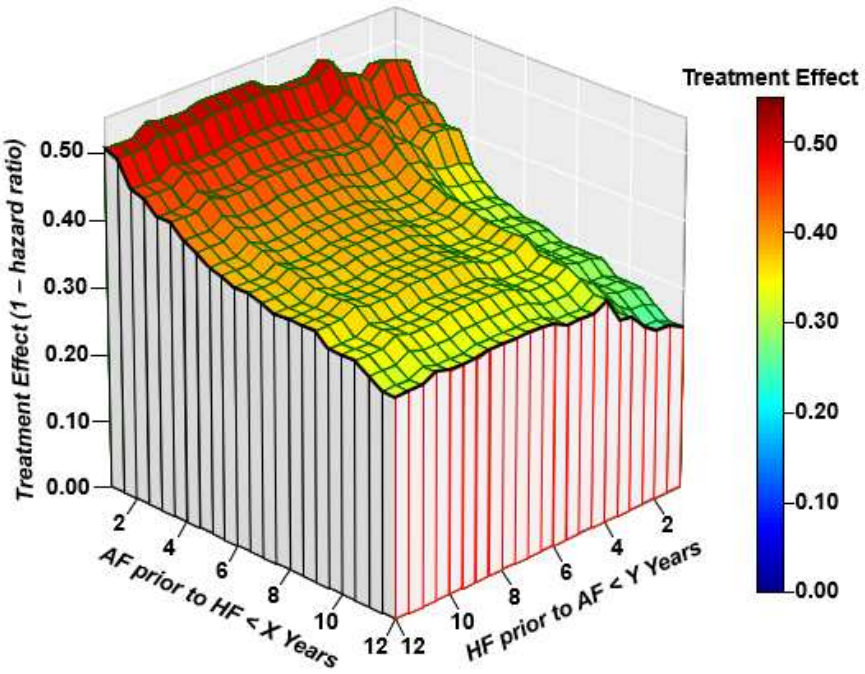


FIGURE 6 Treatment Effect and the Relative Onset of AF and HF in DxT12 Cohort
3-dimensional plot of AF onset prior to HF (x-axis) and HF onset prior to AF (y-axis) versus treatment effect (z-axis) in DxT12 Cohort. Hazard ratio is for time to AF/AFL/ACM endpoint. DTRI (Diagnosis to Randomization Index) = 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.

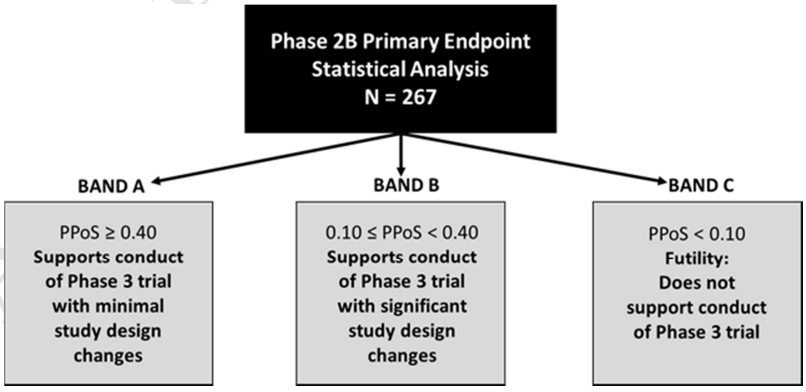


Figure 1. Predictive Probability of Success (PPoS) bands and boundary designations for Bayesian modeling of the GENETIC-AF Phase 2B primary endpoint of time to any AF/AFL or ACM.

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]