1.0 Title Page

1925-201-008

A Phase 2, Multi-Center, Randomized, Double-Blind, Placebo-Controlled, Dose-Ranging Study to Evaluate the Efficacy and Safety of Botulinum Toxin Type A (AGN-151607) Injections into the Epicardial Fat Pads to Prevent Post-Operative Atrial Fibrillation in Patients Undergoing Open-Chest Cardiac Surgery

STATISTICAL ANALYSIS PLAN

Final: 31 May 2022

2.0 Table of Contents

1.0 Title Page 1
2.0 Table of Contents 2
List of Tables 3
3.0 List of Abbreviations 5
4.0 Introduction 7
5.0 Objectives 11
5.1 Primary Efficacy Objectives 11
5.2 Secondary Efficacy Objectives 11
5.3 Safety Objectives 11
5.4 Additional Objectives 11
5.5 Exploratory Objectives 11
6.0 Patient Populations 11
6.1 Modified Intent-To-Treat Population 11
6.2 Safety Population 11
6.3 Supplementary ePatch Analysis Sets 12
6.4 Beta Blocking Agent Withdrawn ePatch Analysis Set 12
7.0 Participant Disposition 12
8.0 Demographics and Other Baseline Characteristics 13
9.0 Extent of Exposure 14
10.0 Efficacy Analyses 14
10.1 Primary Efficacy Variable 16
10.2 Secondary Efficacy Variables 17
10.3 Additional Efficacy Variables 19
10.4 Exploratory Efficacy Variables 23
10.4.1 Assessment of Fat Pads and Injection Procedure 23
10.4.2 Quality of Life Outcomes 24
10.4.3 Assessment of Biomarkers 33
11.0 Safety Analyses 33
11.1 Adverse Events 34
11.2 Clinical Laboratory Parameters 37

11.3 Vital Signs 40
11.4 Electrocardiogram 40
11.5 Other Safety Parameters 41
11.5.1 Immunogenicity Assessments 41
11.5.2 Pulmonary Function Test (PFT) 41
11.5.3 Time to Extubation 42
11.5.4 Physical Examination 42
12.0 Interim Analysis 42
13.0 Determination of Sample Size 43
14.0 Statistical Software 43
15.0 Data Handling Conventions 43
15.1 Visit Time Windows 43
15.2 Repeated or Unscheduled Assessments of Safety Parameters 44
15.3 Missing Severity Assessment for Adverse Events 45
15.4 Missing Causal Relationship to Study Drug for Adverse Events 45
15.5 Missing Date Information for Adverse Events 45
15.6 Missing Date Information for Prior or Concomitant Medications 46
15.6.1 Incomplete Start Date 46
15.6.2 Incomplete Stop Date 47
15.7 Character Values of Clinical Laboratory Parameters 48
15.8 Datetimes for ePatch Analyses 48
16.0 Changes to Analyses Specified in Protocol 49
17.0 Version History 49
18.0 References 51

List of Tables

Table 4–1 Schedule of Evaluations: Study 1925-201-008 8
Table 10–1 The University of Toronto Atrial Fibrillation Severity Scale Scoring Algorithm 25
Table 10–2 The University of Toronto Atrial Fibrillation Severity Scale Scoring Clarifications 26
Table 10–3 Duke Activity Status Index (DASI) Score 29
Table 11–1 Cardiovascular Adverse Events of Interest 35

Table 11–2 Criteria for Potentially Clinically Significant Laboratory Results 39
Table 11–3 Criteria for Potentially Clinically Significant Vital Signs 40
Table 11–4 Criteria for Potentially Clinically Significant Electrocardiograms.....41
Table 15–1 Analysis Visit Windows 44
Table 15–2 Examples of Coding Special Character Values for Clinical Laboratory Parameters.....48
Table 17–1 Final Analysis Plan Version History Summary.....49

3.0 List of Abbreviations

AE	adverse event
AESI	adverse event of special interest
AF	atrial fibrillation
ALT	alanine aminotransferase
AST	aspartate aminotransferase
BP	blood pressure
CABG	coronary artery bypass graft
CI	confidence interval
DSMB	data safety monitoring board
ECG	electrocardiogram
eCRF	electronic case report form
EQ-5D-5L	A measure of health-related quality of life developed by the EuroQol Group
FEV	forced expiratory volume
FEV_1	forced expiratory volume in 1 second
FEV_3	forced expiratory volume in 3 seconds
FEV_6	forced expiratory volume in 6 seconds
FVC	forced vital capacity
GCP	good clinical practice
HIPAA	Health Insurance Portability and Accountability Act of 1996
IB	investigator's brochure
ICF	informed consent form
ICH	International Conference on Harmonisation
ICU	intensive care unit
IDR	independent drug reconstitutor
ITT	intent-to-treat
LS	least squares
MedDRA	Medical Dictionary for Regulatory Activities
mITT	modified intent-to-treat
NB	negative binomial
PCS	potentially clinically significant
PDSOT	possible distant spread of toxin
POAF	post-operative atrial fibrillation
PRO	patient reported outcome
QTc	QT interval corrected for heart rate
SAE	serious adverse event
SAP	statistical analysis plan
SBP	systolic blood pressure
SF-12v2	Short-Form-12 Health Survey, Version 2
TEAE	treatment-emergent adverse event

WBC	white blood cell
ZINB	zero-inflated negative binomial

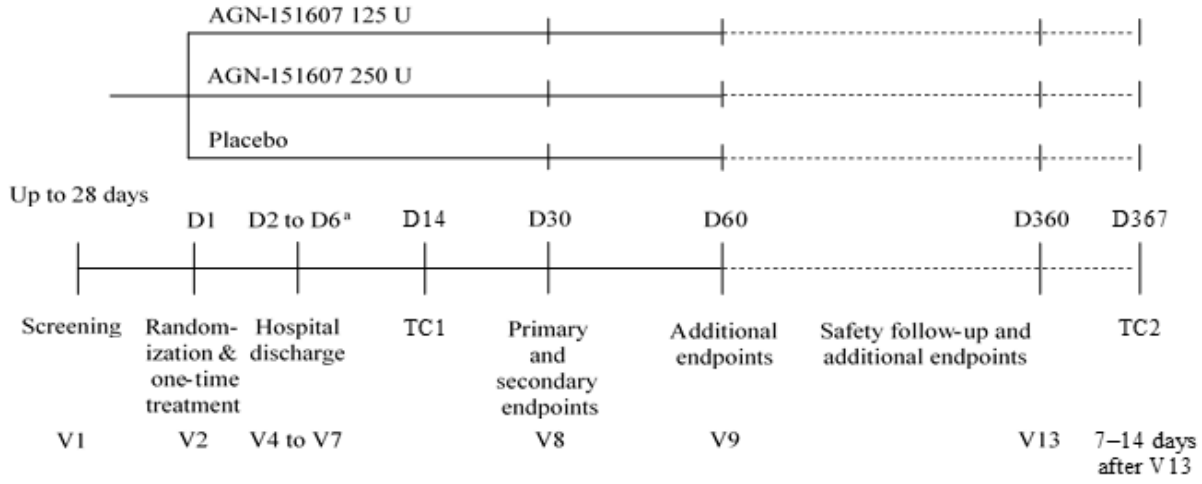
4.0 Introduction

This statistical analysis plan (SAP) provides a detailed elaboration of the statistical analyses of the efficacy and safety data as specified in the final protocol of Study 1925-201-008 (amendment 1 version dated 25 September 2020). Specifications of tables, figures, and data listings are contained in a separate document.

This is a phase 2, multi-center, randomized, double-blind, placebo-controlled, dose-ranging study in patients 55-90 years of age who are scheduled to undergo open-chest cardiac surgery.

The length of this study will be 367 days, not including the screening visit, which may occur up to 28 days before the first study day (randomization). Signed informed consent from the participant will be obtained before any study-related procedures are begun. Patients meeting the inclusion/exclusion criteria will be randomized (1:1:1 ratio) to one of three treatment arms at the end of the screening period. Primary and secondary efficacy assessments will be assessed for 30 days post-surgery. Additional efficacy and exploratory assessments will be assessed through Day 367 post-surgery. All participants will be followed through Day 367 for safety assessments. Every attempt will be made for participants who prematurely discontinue from the study, regardless of cause, to be seen for a final evaluation.

The study schema is presented below.



D = study day; TC = telephone call; U = units; V = Visit

a Hospital discharge could take more than 6 days.

The schedule of evaluations is presented in the Table 4–1 below.

Table 4–1 Schedule of Evaluations: Study 1925-201-008

Study Period	**Screening[a]**	**Days 1 to 60**									**Day 61 through Day 360**					**Early Disc[c]**
Visit Number	**V1**	**V2**	**V3**	**V4[b]**	**V5[b]**	**V6[b]**	**V7[b]**	**TC1**	**V8**	**V9**	**V10**	**V11**	**V12**	**V13**	**TC2**	
Day Number	**(up to 28 days)**	**D1**	**D2**	**D3**	**D4**	**D5**	**D6**	**D14**	**D30**	**D60**	**D90**	**D180**	**D270**	**D360**	**D367[d]**	
Visit Windows	-	-	-	-	-	-	-	**± 3d**	**± 3d**	**± 3d**	**± 7d**	**± 7d**	**± 7d**	**± 7d**	**V13 + 7d (+ 7d)**	-
Informed consent	X															
Inclusion and exclusion criteria	X[e]	X														
Demography	X															
Physical examination	X								X					X[f]		X
Medical history	X															
Urine pregnancy test (for women of childbearing potential)[e,g]	X	X														
Laboratory assessments[e,h]	X	X					X		X	X[i]	X[i]			X[i]		X
Sample collection for biomarker assessment[e]		X	X	X	X											
12-lead ECG[j]	X	X	X	X	X	X	X		X	X[k]	X[k]	X[k]	X[k]	X[k]		X
ECG patch		X	X	X	X	X	X	X	X[l]	X[m]	X[m]	X[m]	X[m]	X[m]	X	
Vital signs[e,n]	X	X	X	X	X	X	X		X	X	X	X	X	X		X
Height	X															
Body weight and abdominal circumference	X													X		X
Dispense AF symptom diary and train participant on proper use[o]	X															
AF symptom diary review[o]		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Randomization		X														
Study intervention[p]		X														
Assessment of fat pads and injection procedure		X														
AE/SAE review[q]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
PFT[r]	X								X							
Concomitant medication and concurrent procedure review[e]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

Study Period	**Screening[a]**	**Days 1 to 60**									**Day 61 through Day 360**					**Early Disc[c]**
Visit Number	**V1**	**V2**	**V3**	**V4[b]**	**V5[b]**	**V6[b]**	**V7[b]**	**TC1**	**V8**	**V9**	**V10**	**V11**	**V12**	**V13**	**TC2**	
Day Number	**(up to 28 days)**	**D1**	**D2**	**D3**	**D4**	**D5**	**D6**	**D14**	**D30**	**D60**	**D90**	**D180**	**D270**	**D360**	**D367[d]**	
Visit Windows	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**± 3d**	**± 3d**	**± 3d**	**± 7d**	**± 7d**	**± 7d**	**± 7d**	**V13 + 7d (+ 7d)**	**-**
AFEQT	X								X		X	X				X[s]
AFSS	X								X		X	X				X[s]
DASI	X								X		X	X				X[s]
EuroSCORE II[t]	X															
SF-12v2	X								X		X	X				X[s]
EQ-5D-5L	X								X		X	X				X[s]
Record discharge from ICU		X	X	X	X	X	X									
Record discharge from hospital			X	X	X	X	X									
Record any additional healthcare resource utilization								X	X	X	X	X	X	X	X	X
Serum sampling for immunogenicity assessment[u]	X								X		X					

AE = adverse event; AF = atrial fibrillation; AFEQT = Atrial Fibrillation Effect on Quality-of-life Questionnaire; AFSS = University of Toronto Atrial Fibrillation Severity Scale; D(d) = day; DASI = Duke Activity Status Index; Disc = discontinuation; ECG = electrocardiogram; EQ-5D-5L = A measure of health-related quality of life developed by the EuroQol Group; FU = follow-up; ICU = intensive care unit; PFT = pulmonary function test; SAE = serious adverse event; SF-12v2 = Short Form-12 Health Survey, version 2; TC = telephone call; V = visit

a Screening and randomization activities can occur on the same day. If screening and randomization occur on the same day, all assessments for Visit 1 and Visit 2 should be completed. Duplicate assessments (eg, laboratory assessments) do not have to be completed twice; please note a 12-lead ECG is required both before and after surgery.

b These visits are optional and are only required if the participant is still hospitalized. Participants may remain in the hospital beyond Day 6, based on the recovery from surgery. The study visit procedures will have to be performed each day until discharge.

c In case of early discontinuation, attempts will be made to follow-up with the participant.

d Day 367 must be at least 7 days and should be no more than 14 days after Visit 13 (Day 360).

e On Day 1, assessments to be performed prior to surgery.

f Physical examinations should be conducted where feasible. If the participant cannot attend the study visit, then the physical examination can be omitted, after sponsor approval.

g Local urine testing will be standard for the protocol unless serum testing is required by local regulation or IRB/IEC.

h Laboratory assessments (hematology, chemistry, coagulation and urinalysis) to be performed at screening, Day 1 and day of discharge. The day of discharge can be from Day 3 onwards.

i Laboratory assessments should be conducted where feasible. If the participant cannot attend the study visit and/or an in-home assessment is not feasible, then the laboratory assessments can be omitted, after sponsor approval.

j On Visit 2, Day 1, two 12-lead ECGs will be performed; 1 before surgery and 1 after surgery. PFT should be performed at a reasonable time apart from vital signs and ECG to avoid confounding ECG and vital signs data.

k 12-lead ECGs should be conducted where feasible. If the participant cannot attend the study visit and/or an in-home assessment is not feasible, then the 12-lead ECG can be omitted, after sponsor approval.

l ECG patch must be worn for a full 30 days post-surgery, even if Day 30 visit occurs prior to Day 30 due to window allowance.

m ECG patch must be used during the first 7 days following Day 60, Day 90, Day 180, Day 270 and Day 360 visits.

n Vital signs taken after a participant has been in a resting position for a minimum of 5 minutes; include blood pressure, pulse, respiration rate, and body temperature. On Visit 2, Day 1, vital signs will be measured before and after surgery.

o AF symptom diary must be used for first 30 days after surgery (or until early discontinuation if before Day 30) and during the first 7 days following Day 60, Day 90, Day 180, Day 270 and Day 360 visits. The diary should be returned to the site and be part of the medical records.

p Unblinded site pharmacist or designee will prepare drug; injected by investigator who is blinded to study intervention.

q Method of detecting AEs and SAEs: care will be taken not to introduce bias when detecting AEs and/or SAEs. Open-ended and nonleading verbal questioning of the participant is the preferred method to inquire about AE occurrences.

r Pulmonary function to be assessed using a spirometer and will include forced vital capacity (FVC) and the forced expiratory volume in 1 second (FEV1), 3 seconds (FEV3) and 6 seconds (FEV6). Note, if local policies contraindicate or prohibit PFT assessments due to potential for viral spread or microbial precautions, then the PFT is not required after explicit approval from the sponsor.

s If participant withdraws after Day 180, these early discontinuation procedures will not be performed.

t Impaired prognosis defined as EuroSCORE II > 7% perioperative mortality at screening is exclusionary.

u Serum collection should be conducted where feasible. If the participant cannot attend the study visit and/or an in-home assessment is not feasible, then the serum collection can be omitted, after sponsor approval.

5.0 Objectives

5.1 Primary Efficacy Objectives

- To compare the efficacy of AGN-151607 with placebo to prevent post-operative atrial fibrillation (POAF) in participants who are undergoing open-chest cardiac surgery.

5.2 Secondary Efficacy Objectives

- To compare the efficacy of AGN-151607 with placebo to reduce atrial fibrillation (AF) burden in participants who are undergoing open-chest cardiac surgery.
- To compare the efficacy of AGN-151607 with placebo to prevent POAF using alternative definitions for POAF in participants who are undergoing open-chest cardiac surgery.

5.3 Safety Objectives

- To compare the safety of AGN-151607 with placebo in participants undergoing open-chest cardiac surgery.

5.4 Additional Objectives

- To compare the clinical benefit of AGN-151607 with placebo in participants who are undergoing open-chest cardiac surgery.

5.5 Exploratory Objectives

- Assessment of fat pads and injection procedure
- Quality of life outcomes
- Assessment of biomarkers

6.0 Patient Populations

6.1 Modified Intent-To-Treat Population

The Modified Intent-to-Treat (mITT) Population will consist of all randomized participants who received study treatment and had at least 1 post-dose ECG (by ePatch) by Day 30 post-surgery. Analyses will be based on randomized treatment.

6.2 Safety Population

The safety population will consist of all participants who received study treatment. Analyses will be based on actual treatment received.

6.3 Supplementary ePatch Analysis Sets

To conduct supplementary analyses of the primary efficacy and specified secondary efficacy variables, two Supplementary ePatch Analysis (SEA) Sets will be defined. Literature suggests that a majority of POAF occurs within one week post-operation (Aranki *et. al*., 1996). The first SEA set will analyze participants without substantial missingness for their ePatch data during the first week post-surgery. The second SEA set will analyze participants that additionally are without substantial missingness for their ePatch data during the first 30 days post-surgery. A threshold of 80% of the expected analyzable time will be used to define "without substantial missingness". The two SEA sets are defined as follows:

- SEA-7 consists of all randomized participants who received study treatment and have at least 134.4 hours (5.6 days) of analyzable time of ePatch data for patch "2.1" (the patch expected to be worn for the first 7 days post-surgery).
- SEA-30 consists of all randomized participants who received study treatment and have at least 134.4 hours (5.6 days) of analyzable time of ePatch data for patch "2.1" and have at least 576 hours (24 days) of analyzable time during the first 30 days post-surgery.

Analyses will be based on randomized treatment.

6.4 Beta Blocking Agent Withdrawn ePatch Analysis Set

To conduct supplementary analyses of the primary efficacy and specified secondary efficacy variables within the subgroup of patients that had a beta blocking agent withdrawn post-surgery, the Beta Blocking Agent Withdrawn ePatch Analysis Set will be defined. The Beta Blocking Agent Withdrawn ePatch Analysis Set consists of all randomized participants who received study treatment, have at least 1 post-dose ECG (by ePatch) by Day 30 post-surgery, and had a beta blocking agent at baseline or after surgery that was withdrawn within 3 days post-surgery and no replacement (no new beta blocker prescription) within 7 days post-surgery. Analyses will be based on randomized treatment.

7.0 Participant Disposition

The number and percentage of participants in the two study populations (Safety and mITT) will be summarized by treatment group; the overall number of patients screened will be presented.

The number and percentage of participants who complete the study and of participants who prematurely discontinue the study will be presented for each treatment group and pooled across treatment groups for the mITT population. The reasons for premature discontinuation from the study as recorded on the termination pages of the electronic case report forms will be summarized (number and percentage) by treatment group for the mITT population. All

participants who prematurely discontinue during the study will be listed by discontinuation reason for the mITT population.

8.0 Demographics and Other Baseline Characteristics

Demographic parameters (age; age group; race; ethnicity; sex), baseline characteristics (weight; height; and body mass index, calculated as weight [kg]/(height [m])^2), waist circumference (m), and other disease characteristics (type of surgery and number of grafts) will be summarized descriptively by treatment group for the safety population. Continuous variables will be summarized by number of participants and mean, SD, first quartile (Q1), median, minimum, third quartile (Q3), and maximum values. Categorical variables will be summarized by number and percentage of participants.

Medical history data will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). The number and percentage of participants with medical history ongoing at screening will be summarized by system organ class, preferred term, and treatment group for the safety population. Data for medical history that is not ongoing at screening will be summarized and presented in a similar manner. An overall total will be provided for the study as well as by treatment group.

Prior medication is defined as any medication taken before the study treatment. Concomitant medication is defined as any medication taken on or after the day of study treatment.

Both prior and concomitant medications will be coded using the Anatomical-Therapeutic-Chemical classification text from World Health Organization (WHO) Drug Dictionary. The use of prior and concomitant medications will be summarized by the number and percentage of subjects in each treatment group for the safety population. If a participant took a specific medication multiple times or took multiple medications within a specific therapeutic class, that participant would be counted only once for the coded drug name or therapeutic class.

In addition to summarizing all medications as discussed above, a separate tabulation will also be provided for the following medications of interest:

- The following antiarrhythmics: quinidine, procainamide, disopyramide, mexiletine, flecainide, propafenone, amiodarone (including amiodarone hydrochloride, amiodarone hydrochloride w/ glucose and amiodarone;glucose), dronedarone, dofetilide, sotalol (including sotalol hydrochloride), and ibutilide.

- The following anticoagulants: warfarin (including warfarin sodium), dabigatran, apixaban, rivaroxaban, and edoxaban (including edoxaban tosilate).

Concurrent procedures will be presented in a listing. A tabulation will be provided for concurrent procedures performed due to atrial fibrillation or atrial flutter. Concurrent procedures performed

due to atrial fibrillation or atrial flutter may include: alcohol septal ablation, cardiac ablation, cardioversion, and additional cardiac surgery. See Section 10.3 below for additional discussion of concurrent procedures performed due to atrial fibrillation or atrial flutter.

Additional medications or concurrent procedures meeting the classifications above may be added at the discretion of the clinical team prior to the interim database lock for the primary analysis.

9.0 Extent of Exposure

All study participants will be treated only once on their day of surgery. Study treatment exposure data will be listed for the safety population.

Descriptive statistics of the study duration, calculated for each participant as the number of days from the date of the study treatment to the date the participant exits the study, will be presented by treatment group. In addition, the number and percentage of participants whose study duration belong to each of the following intervals will also be presented by treatment group: ≥ 6 days, ≥ 14 days, ≥ 30 days, ≥ 60 days, ≥ 90 days, ≥ 180 days ≥ 270 days, ≥ 360 days. These summaries will be based on the safety population.

10.0 Efficacy Analyses

All efficacy analyses will be based on the mITT Population, unless otherwise specified. Note that the use of the phrase "AF" or "AF episode" in the section and subsections below implies an occurrence of either atrial fibrillation or atrial flutter. Other analyses will be conducted to assess only atrial fibrillation, only atrial flutter, only atrial tachycardia, and a combination of either atrial fibrillation, atrial flutter, or atrial tachycardia.

Each participant will be continuously monitored for the first 30 days post-surgery, and during the first 7 days following Day 60, Day 90, Day 180, Day 270, and Day 360 visits, for all occurrences of AF episodes as well as the length of time spent in each of the corresponding AF episodes using ECG patches (ePatch) placed on the participant's chest.

Even though each participant is expected to wear the epatch for the first 30 days post-surgery, and during the first 7 days following Day 60, Day 90, Day 180, Day 270 and Day 360 visits, all analyzable data (interpretable) collected from the epatch will be included in the analysis regardless of any deviations from these timeframes.

The ECG measurements collected for AF episodes post-surgery will be used to provide 10 different definitions for POAF for use in statistical analyses, namely:

1. at least 1 continuous AF episode ≥ 30 seconds
2. at least 1 continuous AF episode ≥ 2 minutes

3. at least 1 continuous AF episode ≥ 5 minutes
4. at least 1 continuous AF episode ≥ 6 minutes
5. at least 1 continuous AF episode ≥ 30 minutes
6. at least 1 continuous AF episode ≥ 1 hour
7. at least 1 continuous AF episode ≥ 4 hours
8. at least 1 continuous AF episode ≥ 6 hours
9. at least 1 continuous AF episode ≥ 12 hours
10. at least 1 continuous AF episode ≥ 24 hours

Based on information collected from continuous monitoring for all occurrences of AF during the first 30 days post-surgery, the percent of time spent in AF (hereinafter referred to as AF Burden) during the first 30 days post-surgery will be calculated as (the total time spent in AF during the first 30 days post-surgery divided by the total time of analyzable data obtained from the ECG patch during the first 30 days post- surgery) multiplied by 100.

In addition, similar definitions of POAF as well as the derivation of the AF Burden will be obtained based on continuous monitoring of all occurrences of AF during the first 7 days following Day 60, Day 90, Day 180, Day 270, and Day 360 visits.

All analyses using ECG patch data will be based on observed data; no imputation will be performed on missing data from the ECG patch device. Descriptive statistics of the total time of analyzable data obtained from the ECG patch during the first 30 days post- surgery will be presented by treatment group. Similar statistics will be presented for each of the first 7 days following Day 60, Day 90, Day 180, Day 270, and Day 360 visits.

Unless otherwise stated, the ICU admission date for the index surgery (initial care of indication under study) will be used as reference timepoint to establish if an event occurred within a specified timeframe “post-surgery”. For example, the occurrence of AF during the first 30 days post-surgery will imply the occurrence of AF during the first 30 days following the initial ICU admission date.

In general, continuous and ordinal variables will be summarized by descriptive statistics and will include sample size, mean, standard deviation, median, minimum, and maximum. Nominal variables will be summarized by frequency and percentage.

Inferential analyses will include 3 treatment groups: Placebo, 125 U, and 250 U of AGN-151607. Each dose level of AGN-151607 (125 U and 250 U) will be separately compared to placebo. All treatment comparisons will be made at the $\alpha = 0.05$ level (2-sided). To control for multiplicity in comparing 2 doses of AGN-151607 with placebo, the primary efficacy variable will be tested sequentially by first comparing the high dose to placebo. The low dose will be compared to

placebo only if the comparison to the high dose is found to be statistically significant. There will be no adjustment for multiple comparisons for all other efficacy variables.

Mis-stratified participants, if any, will be analyzed based on the correct stratification group [type of surgery (presence or absence of valve surgery) and age group (< 65 or ≥ 65 years)], not the stratification group to which the participants were assigned at the time of randomization.

Unless otherwise stated, the use of the phrase "percentage of participants" for any efficacy variable in the sections below assumes the variable is a binary categorical response outcome.

10.1 Primary Efficacy Variable

For primary efficacy consideration, POAF will be defined as any continuous episode of AF lasting 30 seconds or more during the first 30 days post-surgery. The primary efficacy variable is the percentage of participants with at least 1 continuous AF episode ≥ 30 seconds during the first 30 days post-surgery.

For each of the AGN-151607 doses of 125 U and 250 U, the null hypothesis is that there is no difference between that dose group and placebo in the percentage of participants with at least 1 continuous AF episode ≥ 30 seconds during the first 30 days post-surgery. The alternative hypothesis is that there is a difference between the AGN-151607 dose group under consideration and placebo.

The proportion of participants with at least 1 continuous AF episode ≥ 30 seconds during the first 30 days post-surgery will be presented for each treatment group. Between-group comparisons will be performed using the Cochran-Mantel-Haenszel (CMH) test for general association stratified by type of surgery (presence or absence of valve surgery) and age group (< 65 or ≥ 65 years). P-values, adjusted Mantel-Haenszel (MH) odds ratios and relative risks (and their corresponding 95% confidence intervals) will be presented for each comparison of AGN-151607 versus placebo. Exact (Clopper-Pearson) 95% confidence intervals of the observed proportion will be presented for each treatment group. In addition, exact unconditional 95% confidence interval (using the score method) for the difference in observed proportion between each AGN-151607 dose group versus placebo will also be presented.

For the primary efficacy variable, supplementary analyses will be performed based on the Supplementary ePatch Analysis Sets (SEA-7 and SEA-30) and the Beta Blocking Agent Withdrawn ePatch Analysis Set. The supplementary analyses will use the same methods described for the primary analysis.

For the primary efficacy variable, subgroup analyses will be performed using the same methods described for the primary analysis based on the correct stratification group (with the other stratification factor used for the stratified CMH test):

- type of surgery (presence or absence of valve surgery)
- age group (< 65 or ≥ 65 years)]

10.2 Secondary Efficacy Variables

The secondary efficacy variables are:

- Percentage of time spent in AF (AF Burden) during the first 30 days post-surgery
- Percentage of participants with at least 1 event of symptomatic AF (symptoms occurring within 2 hours of an AF episode) during the first 30 days post-surgery
- Time to first occurrence of AF during the first 30 days post-surgery
- Percentage of participants with at least 1 continuous AF episode ≥ 2 minutes during the first 30 days post-surgery
- Percentage of participants with at least 1 continuous AF episode ≥ 5 minutes during the first 30 days post-surgery
- Percentage of participants with at least 1 continuous AF episode ≥ 6 minutes during the first 30 days post-surgery
- Percentage of participants with at least 1 continuous AF episode ≥ 30 minutes during the first 30 days post-surgery
- Percentage of participants with at least 1 continuous AF episode ≥ 1 hour during the first 30 days post-surgery
- Percentage of participants with at least 1 continuous AF episode ≥ 4 hours during the first 30 days post-surgery
- Percentage of participants with at least 1 continuous AF episode ≥ 6 hours during the first 30 days post-surgery
- Percentage of participants with at least 1 continuous AF episode ≥ 12 hours during the first 30 days post-surgery
- Percentage of participants with at least 1 continuous AF episode ≥ 24 hours during the first 30 days post-surgery
- Percentage of participants with at least 1 continuous atrial fibrillation (excluding atrial flutter) episode ≥ 30 seconds during the first 30 days post-surgery
- Percentage of participants with at least 1 continuous atrial flutter episode ≥ 30 seconds during the first 30 days post-surgery

- Percentage of participants with at least 1 continuous atrial tachycardia episode (defined as the duration of the longest SVT run) ≥ 30 seconds during the first 30 days post-surgery

- Percentage of participants with at least 1 continuous episode of either atrial fibrillation or atrial flutter or atrial tachycardia ≥ 30 seconds during the first 30 days post-surgery

Between-group treatment comparisons in the percentage of time spent in AF (AF burden) during the first 30 days post-surgery will be performed for each AGN-151607 treatment arm versus placebo using the stratified Wilcoxon (Van Elteren) test stratified by age group and type of surgery. Calculation of time spent in AF (AF burden) will exclude continuous AF episodes < 30 seconds in duration. For participants with no AF during the first 30 days post-surgery, the AF burden will be set to 0.

For each efficacy variable with binary categorical response outcome, the response proportions will be analyzed using the same methods used to analyze the primary efficacy variable. For symptomatic AF, symptoms that occur in the interval that starts two hours prior the onset of the AF episode and ends two hours after the conclusion of the AF episode will meet the definition of "within 2 hours of an AF episode". As noted in the protocol, participants will mark the date and time that a symptom is experienced by pressing a button on the ECG patch itself.

For each treatment group, the Kaplan-Meier (KM) survival analysis methodology will be used to estimate the median time (days) to first occurrence of AF during the first 30 days post-surgery using each of the 9 definitions of POAF stated in Section 10.0 above. Furthermore, the 25th and 75th percentiles of the time to the first occurrence of AF during the first 30 days post-surgery will also be estimated. The associated 95% CIs calculated using Greenwood's formula for the variance of the KM estimate will also be presented for each treatment group. Pairwise treatment comparisons of the KM curves will be performed primarily by using the log rank test stratified by type of surgery and age group. Time to event data will also be compared between treatment groups using Cox proportional hazards models adjusting for baseline risk factors of POAF (type of surgery, and age group), with hazard ratios and their 95% CIs presented for each pairwise comparison. Cox proportional hazards models additionally adjusting for total analyzable time during the first 30 days post-surgery will also be analyzed. The difference in restricted mean survival time between treatment groups for time to first occurrence of AF during the first 30 days post-surgery (i.e., restricted time $\tau = 30$ days) will be compared using a nonparametric *K*-sample chi-square test stratified by type of surgery and age group. Participants who do not experience an AF episode during the first 30 days post-surgery will be censored at the maximum timepoint during the first 30 days post-surgery for which they have recorded data.

For the time to first occurrence of AF during the first 30 days post-surgery, a supplementary analysis will be performed based on the Supplementary ePatch Analysis Sets (SEA-7 and SEA-30). The supplementary analysis will use the same methods described above.

For each minimum duration specified above for the percentage of participants with at least 1 continuous AF episode during the first 30 days post-surgery, a supplementary analysis will be performed based on the Supplementary ePatch Analysis Sets (SEA-7 and SEA-30). The supplementary analysis will use the same methods used to analyze the primary efficacy variable. For these endpoints, subgroup analyses will also be performed using the same methods described above based on the correct stratification group [type of surgery (presence or absence of valve surgery) and age group (< 65 or ≥ 65 years)].

10.3 Additional Efficacy Variables

Hospital and ICU Length of Stay

Hospital length of stay for the initial care of indication under study is defined as the time (days) from ICU admission date/time to the date/time of discharge from the hospital. ICU length of stay is defined as the time (hours) from the ICU admission date/time to the date/time of discharge from the ICU. For participants not admitted to the ICU, ICU length of stay will be set to zero and hospital length of stay will be defined as the time (days) from end date/time of study drug administration to date/time of discharge from hospital. For subjects missing ICU admission date/time, ICU admission date/time will be imputed to the end date/time of study drug administration. The date/time of death will be used to derive hospital length of stay for participants who died at the hospital (and ICU length of stay, for participants who die in the ICU).

Descriptive statistics (n, mean, SD, Q1, median, Q3, min, max) for hospital and ICU length of stay will be presented by:

- treatment group
- region (U.S. versus non-U.S.) and treatment group
- country and treatment group
- sites and treatment group

Between-group treatment comparisons in hospital and ICU length of stay will be performed for each AGN-151607 treatment arm versus placebo using the following methods:

- An ANOVA model with treatment group, type of surgery, age group, and region as factors
- An ANOVA model with treatment group and region as factors
- An ANOVA model with treatment group and country as factors

- A stratified Wilcoxon (Van Elteren) test, stratified by type of surgery, age group, and region

- A stratified Wilcoxon (Van Elteren) test, stratified by type of surgery and age group

Descriptive statistics of hospital length of stay measured from the time of initial hospital admission to discharge will be presented by:

- treatment group
- region (U.S. versus non-U.S.) and treatment group
- country and treatment group
- sites and treatment group

ANOVA models, as described above, will be used to estimate confidence intervals for difference in mean hospital length of stay measured from the time of initial hospital admission between each AGN-151607 treatment arm and placebo.

Descriptive statistics of number of days spent in the hospital within 30 days post-discharge and within 60 days post-surgery will be presented by treatment group. Summary statistics will also be presented by treatment group for the reasons for hospital and ICU admissions.

Rehospitalization within 30 Days Post-Discharge and 60 Days Post-Surgery

The percentage of participants with at least one all-cause rehospitalization within 30 days post-discharge and percentage of participants with at least one all-cause rehospitalization 60 days post-surgery will each be presented by treatment group. Similar summaries will be presented for non-arrhythmia cardiovascular rehospitalization and arrhythmia cardiovascular rehospitalization. These variables will be analyzed using the same method presented for the primary efficacy variable in Section 10.1. Descriptive statistics of the total number of rehospitalizations within 30 days post-discharge and 60 days post-surgery as well as the length of rehospitalization stay will also be presented by treatment group for each of the different causes of rehospitalizations (all-cause, non-arrhythmia cardiovascular and arrhythmia cardiovascular).

The time (days) to the first all-cause rehospitalization will be calculated as the date/time of first re-hospitalization minus the date/time of hospital discharge after surgery. This will be analyzed using the KM method discussed in Section 10.2. Participants with no rehospitalizations will be censored at their respective study exit day. Similar analyses will be performed for non-arrhythmia cardiovascular rehospitalization and arrhythmia rehospitalizations.

Healthcare Resource Utilization within 30 Days, 60 Days, and 365 days Post-Discharge

The proportion of participants with any healthcare resource utilization (outside of any planned in-patient hospitalization, procedures and medications) within 30 days, 60 days, and 365 days post-discharge will be presented by treatment group. The type of resource utilized (e.g., Outpatient hospital, Physician's office, etc.) as well as the reason for the visit will also be tabulated in a similar manner. Descriptive statistics of the number of times each participant utilizes a healthcare resource within each of the specified timeframe will also be presented by treatment group. Note that any unplanned inpatient hospitalization will also be included in these summaries.

Need for Medical or Procedural Intervention Due to AF

The following five binary (yes/no) variables will be used to analyze the need for pharmacologic and/or procedural intervention due to AF during the first 30 days post-surgery:

1. at least one pharmacologic intervention (i.e., anticoagulation or antiarrhythmic drugs)?
2. at least one anticoagulation pharmacologic intervention?
3. at least one antiarrhythmic pharmacologic intervention?
4. at least one procedural intervention (see list in Section 8.0)?
5. at least one pharmacologic or procedural intervention?

In addition, time to event variables will also be created for each of the corresponding 5 variables above. The time to the first event will be used for analysis. All these variables will be analyzed using the same methods for similar variables in Sections 10.1 and 10.2.

The need for pharmacologic and/or procedural intervention due to AF from Day 31 to Day 360 post-surgery will be analyzed in a similar fashion. Participants with no events who exit the study prior to Day 360 will be censored at their respective study exit day minus 30 days, while participants with no events who exit the study on or beyond day 360 will be censored at day 330 (360 – 30 = 330). Pharmacologic and/or procedural interventions due to AF with onset dates prior to day 31 that are ongoing at day 31 will count as events for these endpoints, with a time-to-event equal to zero.

Descriptive statistics will be presented by treatment group for each of the following:

- time to POAF resolution from first dose of prescription of pharmacologic intervention
- time to POAF resolution from first procedural intervention
- types of pharmacologic interventions used due to POAF

- average duration of pharmacologic interventions due to POAF. In addition, this summary will be further categorized by "medication of interest" (antiarrhythmics versus anticoagulants) as defined in Section 8.0 above.

- yes/no: Beta blocking agent (ATC code: C07A) withdrawn within 3 days post-surgery and no replacement (no new beta blocker prescription) within 7 days post-surgery

Other Efficacy Variables

- Percentage of participants with clinically important tachycardia in AF (defined as heart rate ≥ 100 bpm for at least 2 minutes) within 30 days post-surgery

- Percentage of participants with any occurrence of AF as defined in the primary and secondary endpoints during the first 7 days following Day 60, Day 90, Day 180, Day 270, and Day 360 visits

- Percentage of participants with at least 1 event of symptomatic AF (symptoms occurring within 2 hours of an AF episode) during the first 7 days following Day 60, Day 90, Day 180, Day 270, and Day 360 visits

- Percentage of participants with actionable AF within 30 days post-surgery. Actionable AF is defined as meeting at least one of the following conditions: hospitalization due to AF, or medicinal or procedural intervention due to AF

- Percentage of participants discharged from hospital (after initial care of indication under study) without prescription of any of the medication of interest listed in Section 8.0

- Percentage of participants free from AF through 12 months post-surgery and not taking any antiarrhythmic drugs (as defined in Section 8.0)

- Percentage of participants free from AF through 12 months post-surgery and not taking any anticoagulant drugs (as defined in Section 8.0)

- Percentage of time spent in AF (AF burden) during the first 7 days following Day 60, Day 90, Day 180, Day 270, and Day 360 visits

The first 6 variables above will be analyzed using the same method used to analyze the primary efficacy variable in Section 10.1. Similarly, the AF burden will be analyzed using the same method discussed in Section 10.2. In addition, descriptive statistics of the AF Burden, including tertiles, will be presented by treatment group. For symptomatic AF, symptoms that occur in the interval that starts two hours prior the onset of the AF episode and ends two hours after the conclusion of the AF episode will meet the definition of "within 2 hours of an AF episode". The date and time when a symptom is experienced is marked by a participant pressing a button on the ePatch itself.

- A negative binomial (NB) regression model will be used to model the number of AF events (using ≥ 30 seconds definition of AF) within 30 days post-surgery, with treatment, type of surgery, and age group as factors and log(total analyzable time) as the offset in the model. A zero-inflated negative binomial (ZINB) regression model will also be estimated with treatment, type of surgery, and age group as factors in both the main and zero-count models and log(total analyzable time) as the offset. If $p < 0.05$ for the Scaled Pearson chi-square goodness of fit statistic for the NB model (suggesting overdispersion may be present), then the ZINB model will be considered the primary analysis; otherwise, the NB model will be considered the primary analysis. Sensitivity analyses of the number of AF events using Poisson regression and zero-inflated Poisson regression using similar models will be performed. Least squares (LS) mean differences and incidence rate ratios with 95% CIs will be presented for all models. A supplementary analysis of the number of AF events will also be performed based on the SEA-30 set using the same models.

- The win ratio unmatched approach (Pocock et. al., 2012) will be used to analyze the following composite endpoint:

    1. Death
    2. Stroke
    3. All-cause 30-day rehospitalization
    4. Use of pharmacologic or procedural intervention of interest due to AF
    5. Hospital length of stay

    The components of the endpoints are listed in order of clinical importance. The win ratio, with the associated 95% CI and p-value will be presented for the composite endpoint, as well as for each of the component endpoints. Two win ratios analyses will be performed. The first will analyze binary outcomes for each of the components (with the fifth component being "Hospital length of stay > median hospital length of stay for the total study population"). The second will analyze time-to-event outcomes for each of the components (with the time-to-first-event used for components with multiple events experienced, and with the fifth component being time-to-hospital-discharge from initial care). A list of Medical Dictionary for Regulatory Activities (MedDRA) preferred terms for stroke will be provided by the clinical team. All-cause 30-day rehospitalization includes any hospitalization within 30 days of (first) discharge from the hospitalization for initial care.

10.4 Exploratory Efficacy Variables

10.4.1 Assessment of Fat Pads and Injection Procedure

[redacted]

[redacted] Descriptive or summary statistics will be presented by treatment groups for each of these assessments as appropriate.

10.4.2 Quality of Life Outcomes

Each quality-of-life questionnaire discussed below will be completed by the participants at Screening (baseline), Day 30, Day 90, Day 180 or at early discontinuation visit if necessary.

Descriptive statistics for each of the aggregate measures and their corresponding change from baseline scores will be presented by treatment group at each visit.

The change from baseline in each of the aggregate scores will be analyzed using a mixed model for repeated measures (MMRM). The model will include treatment group, visit, age group, type of surgery, and treatment group by visit interaction as categorical fixed effects, and baseline score as a continuous fixed effect. Restricted maximum likelihood method will be used. The within-patient correlation will be modeled using the unstructured covariance matrix. The Kenward-Roger approximation will be used to estimate denominator degrees of freedom. Contrasts will be constructed to obtain the treatment effects at each postbaseline visit (i.e., Day 30, Day 90, and Day 180) to compare each AGN-151607 treatment group versus the placebo group. Each treatment effect and treatment comparisons will be estimated by the LS Means and their differences in LS Means, along with their SE and 95% confidence intervals, and the p-value corresponding to the between-treatment group difference.

10.4.2.1 University of Toronto Atrial Fibrillation Severity Scale (AFSS)

A self-reported, disease-specific quality of life instrument used to capture subjective and objective ratings of disease burden in patients with AF. The AFSS consists of 19 items. It contains a visual analogue scale, 4 items on AF frequency, duration and severity that are used to calculate AF burden, 4 items on health care utilization, and 7 items regarding the severity of specific symptoms. See Appendix 10.9 of study protocol for more details.

The tables below show the scoring algorithm for some of the items on the questionnaire as well as the derivation of the following aggregate measures:

- Total AF Burden
- AF Symptoms Score

Table 10–1 The University of Toronto Atrial Fibrillation Severity Scale Scoring Algorithm

Question Number (range)	Scoring
Q4 (1-10) Global well-being	Raw score
Q5 (1-11) AF frequency *lower scores denote more frequent AF	Continuous = 1 More than twice a day = 2 Daily or almost daily = 3 4 - 5 times a week = 4 2 - 3 times a week = 5 About once a week = 6 About twice a month = 7 About once a month = 8 About 2-4 times a year = 9 About once a year = 10 Less than once a year = 11
Q6 (1-8) AF duration *lower scores denote AF of longer duration	Continuous = 1 Several days = 2 All day = 3 Several hours = 4 About 1 hour = 5 30 - 45 minutes = 6 < 30 minutes = 7 A few minutes = 8
Q7, Q8 (1-10) AF severity *higher scores denote more severe AF	Arithmetic mean of 2 scores
Q10 (0-7) ER visits	0 = 0 1 = 1 2 = 2 3 = 3 4 = 4 5 to ≤ 10 = 5 > 10 to ≤ 15 = 6 > 15 = 7
Q11 (0-7) Hospitalizations	0 = 0 1 = 1 2 = 2 3 = 3 4 = 4 5 to ≤ 10 = 5 > 10 to ≤ 15 = 6 > 15 = 7

Question Number (range)	Scoring
Q12 (0-7) Specialist visits	0 = 0 1 = 1 2 = 2 3 = 3 4 = 4 5 to ≤ 10 = 5 > 10 to ≤ 15 = 6 > 15 = 7

Table 10–2 The University of Toronto Atrial Fibrillation Severity Scale Scoring Clarifications

CLARIFICATIONS FOR PART A	
Total AF Burden The Total AF Burden has been historically obtained by combining measures of *frequency* (Q5), *duration* (Q6) and patient perceived *severity* (the mean of Q7 & Q8) equally to obtain a score that ranges from 3-30. Though Total AF Burden may be reported as an aggregate measure, we recommend reporting frequency, duration and severity separately as current literature has suggested both frequency and duration are different from severity in that frequency and duration are not well correlated to quality of life however quality of life is driven by patient perceived severity.	
Clarifications for Q5 (AF frequency)	Due to the fact that Q5 has responses ranging from 1-11, for patients who respond "Less than once a year", please give them a score of 10 instead of 11 (the response, "About once a year" is also scored as 10). Patients who respond, "Not applicable, I have never had an irregular heart rhythm" will not contribute a frequency score. Suggest reporting the % of N/A responses separately with the assumption that these patients may not be aware of their AF frequency. **Reverse code scores so that 1=continuously now becomes 10, etc.**
Clarifications for Q6 (AF duration)	Take the score of each patient, divide it by 8 and multiply it by 10 so that the score will fall in the range of 1-10. Round to the nearest whole number. Patients who respond, "Not applicable, I have never had an irregular heart rhythm" separately will not contribute a duration score. Suggest reporting the % of N/A responses separately with the assumption that these patients may not be aware of their AF duration. **Reverse code scores so that 1=continuously now becomes 10, etc.**
Clarifications for Q7 & Q8 (AF severity)	Add Q7 and Q8 and divide by 2 to obtain the mean. Round to the nearest whole number.

Clarifications to obtain Total AF Burden (AF frequency + AF duration + AF severity)	Total AF Burden=AF frequency + AF duration + AF severity. Each of the 3 measures contributes equally to the Total AF Burden score, and each measure varies from 1-10 to yield Total AF Burden scores ranging from 3-30 where higher scores indicate greater AF burden.
CLARIFICATIONS FOR PART B	
Health Care Utilization No clarifications for this section.	
CLARIFICATIONS FOR PART C	
AF Symptoms Score How affected patients are by specific symptoms related to AF such as palpitations, shortness of breath, exercise intolerance, fatigue, etc.	
Clarifications to obtain AF Symptoms Score	The AF Symptoms Score is scored in a strictly additive manner. For each of the 7 questions (answers vary from 0=I have not had this symptom in the past 4 weeks to 5=A great deal), the values are determined and summed together to form a total score (from 0-35). This score assumes linearity of severity (i.e., if patient X has a total score of 7 and patient Y has a total score of 35, patient Y has 5x greater severity in symptoms than patient X). Note: although the stem question says "How often…" , in practice this is similar to "How severe were the symptoms when they occurred".

10.4.2.2 Atrial Fibrillation Effect on Quality-of-life (AFEQT) Questionnaire

A questionnaire to assess Health Related Quality of Life (HRQoL) in patients with atrial fibrillation (AF). AFEQT evaluates QoL across 3 domains: symptoms, daily activities and treatment concerns. The questionnaire uses 20 questions on a 7-point Likert scale to measure the overall AFEQT score and treatment satisfaction.

The responses on the AFEQT are scored on a 1 to 7 Likert scale, where for questions 1-18, 1="Not at all…" to 7 = "Extremely...". Questions 19- 20 relate to patients' satisfaction with treatment and are not included in HRQoL score of the AFEQT questionnaire.

Calculation of the overall AFEQT score is based on the following formula:
Overall AFEQT score = 100 – [((sum of severity for all questions answered - number of questions answered) X 100) / (total number questions answered X 6)].

Note that questions 19 and 20 are not included in the derivation of the overall AFEQT score.

Subscale scores are computed similarly to the overall score from each subscale used to generate its own score. The 18 questions are grouped into 3 functional subscales as described below:

Subscales	**Questions**
Symptoms	1, 2, 3, and 4
Daily Activities	5, 6, 7, 8, 9, 10, 11, and 12
Treatment Concern	13, 14, 15, 16, 17, and 18

The 2 treatment satisfaction questions (Questions 19 and 20) are not part of the overall AFEQT score but are computed as its own subscale similarly to the overall score.

Subscale	**Questions**
Treatment Satisfaction	19, 20

Overall or subscale scores range from 0 to 100. A score of 0 corresponds to complete disability, while a score of 100 corresponds to no disability.

Separate analyses will be performed for each of the following scores:

- Overall AFEQT score
- Symptoms subscale score
- Daily activities subscale score
- Treatment concern subscale score
- Treatment satisfaction subscale score

10.4.2.3 Duke Activity Status Index (DASI)

A 12-item, self-administered questionnaire which provides a standardized assessment of functional status that uses the patient's ability to perform a set of common activities of daily living to gauge functional capacity. Scoring of the Activity Status Index is shown on the table below.

Table 10–3 Duke Activity Status Index (DASI) Score

Questions	Scores (only for response "Yes, with no difficulty")[a]
1	2.75
2	1.75
3	2.75
4	5.50
5	8.00
6	2.70
7	3.50
8	8.00
9	4.50
10	5.25
11	6.00
12	7.50

a All other responses gets a score of zero for that question.

The DASI score is obtained by summing all the scores from the 12 items on the questionnaire. The higher the score (maximum 58.2), the higher the functional status.

10.4.2.4 European Quality of Life - 5-Dimensional - 5-Level (EQ-5D-5L)

The EQ-5D-5L is made up of two components: health state description and evaluation. The description component consists of five dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. The mobility dimension queries the participant's walking ability. The self-care dimension queries the participant's ability to wash or dress by himself. The usual activities dimension assesses the participant's performance in "work, study, housework, family or leisure activities". The pain/discomfort dimension measures how much pain or discomfort a participant has. The anxiety/depression dimension assesses how anxious or depressed a participant is. The respondents rate their level of severity for each dimension using a 5-level scale (EQ-5D-5L) by ticking (or placing a cross) in the box against the most appropriate statement in each of the 5 dimensions. The second component of the EQ-5D-5L is a visual analogue scale (EQ-VAS) by which participants can rate their overall health from 0 (worst imaginable health state) to 100 (best imaginable health state).

With the EQ-5D-5L, rating levels can be coded as numbers 1, 2, 3, 4 or 5 which correspond to "have no problems", "have slight problems", "have moderate problems", "have severe problems", and "unable to do/have extreme problems", respectively. As a result, a participant's health state can be defined by a 5-digit number by combining the numeric levels from the 5 dimensions, ranging from 11111 ("have no problems" in all 5 dimensions) to 55555 ("unable to

do/have extreme problems" in all 5 dimensions). The US-based value set for the EQ-5D-5L will be derived using an international standardized protocol (Pickard *et. al*., 2019).

The EQ-5D-5L will be completed by the participants at Screening, Day 30, Day 90, Day 180 or at early discontinuation visit if necessary.

10.4.2.5 Short Form-12 Health Survey (SF-12v2)

The SF-12v2 health survey is a generic assessment of health-related quality of life from patient's perspective. It consists of 12 items that measures 8 health domain scales: physical functioning (PF), role limitations due to physical health problems (RP), bodily pain (BP), general health (GH), vitality (energy/fatigue) (VT), social functioning (SF), role limitations due to emotional problems (RE), and mental health (physiologic distress and physiologic well-being) (MH). The 8 scales are aggregated into 2 summary measures: the Physical Component Summary (PCS) and the Mental Component Summary (MCS) scores, which range from 0 to 100, with higher scores indicating better quality of life.

Within each scale, if there are less than or equal to 50% of the item scores missing, the missing item score(s) will be imputed using the mean score of the non-missing items and used to calculate the scale score. If more than 50% of the item scores are missing, no imputation will be performed, and the scale score will be set to missing.

The component summary scores will be derived using the following scoring algorithm.

Step 1: Item coding

Code the response to each item as follows such that higher score indicates better health state.

Response to Item 1

Response Choices	Precoded Item Value	Final Item Value
Excellent	1	5.0
Very Good	2	4.4
Good	3	3.4
Fair	4	2.0
Poor	5	1.0

Response to Items 2a and 2b

Respond Choices	Precoded Item Value	Final Item Value
Yes, limited a lot	1	1
Yes, limited a little	2	2
No, not limited at all	3	3

Response to Items 3a, 3b, 4a, 4b and 7

Response Choices	Precoded Item Value	Final Item Value
All of the time	1	1
Most of the time	2	2
Some of the time	3	3
A little of the time	4	4
None of the time	5	5

Response to Item 5

Response Choices	Precoded Item Value	Final Item Value
Not at all	1	5
A little bit	2	4
Moderately	3	3
Quite a bit	4	2
Extremely	5	1

Response to Items 6a, 6b and 6c

Response Choices	Precoded Item Value	Final Item Value
All of the time	1	5
Most of the time	2	4
Some of the time	3	3
A little of the time	4	2
None of the time	5	1

Step 2: Computing raw scale scores

After item recoding, a raw score is computed for each scale which is the simple algebraic sum of the final item values for all items in that scale as shown in the below table. For example, the raw score for PF scale is the sum of the final item values of items 2a and 2b.

Scale	Sum Final Item Values	Lowest and Highest Possible Raw Scores	Possible Raw Score Range
PF	Items 2a + 2b	2, 6	4
RP	Items 3a + 3b	2, 10	8
BP	Item 5	1, 5	4
GH	Item 1	1, 5	4
VT	Item 6b	1, 5	4
SF	Item 7	1, 5	4
RE	Items 4a + 4b	2, 10	8
MH	Items 6a + 6c	2,10	8

Step 3: Transforming raw scale scores

Transforming the raw scale score to a 0-100 scale using the formula below. The table above include the lowest possible and possible range of raw score for each scale.

$$Transformed\ scale = \frac{Actual\ raw\ score - Lowest\ possible\ raw\ score}{Possible\ raw\ score\ range} \times 100$$

For example, the transformed score for PF scale is (Raw PF score – 2) / 4.

Step 4: Standardizing transformed scale scores

Each transformed scale is standardized to a z-score using scale means and standard deviations from the 1998 general United States population as shown in the below table.

Scale	Mean	Standard Deviation
PF	81.18122	29.10558
RP	80.52856	27.13526
BP	81.74015	24.53019
GH	72.19795	23.19041
VT	55.59090	24.84380
SF	83.73973	24.75775
RE	86.41051	22.35543
MH	70.18217	20.50597

The formula for z-score transformation is:

Z-score standardized scale = (Transformed scale score – Mean) / Standard Deviation

For example, the z-score of PF scale = (Transform PF score – 81.18122) / 29.10558.

Step 5: Aggregating scale scores to compute raw component summary scores

The standardized scales are aggregated using weights (factor score coefficients) from the 1998 general United States population to derive raw component summary scores (PCS and MCS). Formulae for aggregating z-score standardized scales are:

$$Raw\ PCS = PF \times .42402 + RP \times .35119 + BP \times .31754 + GH \times .24954 +$$
$$VT \times .02877 + SF \times -.00753 + RE \times -.19206 + MH \times .22069$$

$$Raw\ MCS = PF \times -.22999 + RP \times -.12329 + BP \times -.09731 + GH \times -.01571 +$$
$$VT \times .23534 + SF \times .26876 + RE \times .43407 + MH \times .48581$$

Step 6: Standardizing composite summary scores

The raw PCS and MCS scores are standardized as follows such that the final PCS and MCS scores have values with mean 50 and standard deviation 10.

$$PCS = 50 + Raw\ PCS \times 10$$

$$MCS = 50 + Raw\ MCS \times 10$$

10.4.3 Assessment of Biomarkers

Blood samples will be collected from all participants in this study prior to dosing on day 1 and at days 2, 3, and 4. These samples will be used to test for biomarkers involved in the inflammatory process and in the autonomic nervous system, including the inflammatory panel (IFNγ, IL1β, IL2, IL4, IL6, IL8, IL10, IL13, and TNFα), Catecholamines (Epinephrine, norepinephrine and dopamine) and hsCRP.

Descriptive statistics and graphs will be presented by treatment group for each biomarker of the following:

- Biomarker concentration at baseline and each timepoint post treatment
- Change from baseline (CFB) in biomarker concentration.

Exploratory analysis of biomarker concentration and their CFB following treatment will be conducted in correlation with the observed clinical response to study treatment.

11.0 Safety Analyses

The safety analysis will be performed using the safety population. The safety parameters will include adverse events (AEs), physical examinations, clinical laboratory test results, vital sign

measurements, electrocardiograph (ECG) results, and pulmonary function testing. For each of the physical examinations, clinical laboratory test results, vital sign measurements, and ECG parameters, the last non-missing safety assessment before the study treatment will be used as the baseline for all analyses of that safety parameter. Continuous variables will be summarized by number of participants and mean, SD, first quartile (Q1), median, minimum, third quartile (Q3), and maximum values. Categorical variables will be summarized by number and percentage of participants.

11.1 Adverse Events

Adverse events will be coded by system organ class and preferred term using the Medical Dictionary for Regulatory Activities (MedDRA).

An AE will be considered a treatment-emergent adverse event (TEAE) if the AE began or worsened (increased in severity or became serious) on or after the date/time of the study injection. Per case report form instructions, a new AE record will be created with a new AE onset date/time for any AE that worsens. Therefore, TEAEs can simply be identified as those AEs with recorded onset date/time on or after the date/time of the study injection.

The number and percentage of participants reporting TEAEs in each study treatment group will be tabulated by descending percentage of the total across treatment groups, by system organ class and preferred term, and further categorized by severity and relationship to the study treatment. If more than 1 TEAE is coded to the same preferred term for the same participant, the participant will be counted only once for that preferred term. For tables presented by severity, the greatest severity observed will be used.

The number and percentage of participants who have treatment-emergent serious adverse events (TESAE) will be summarized by system organ class, preferred term and treatment group. A similar summary will be presented for subjects with TEAEs leading to discontinuation from the study. In addition, separate subject listings will be generated for deaths, serious adverse events (SAEs), and TEAEs leading to study discontinuation.

Cardiovascular events of interest in this study are: heart failure or congestive heart failure, stroke or transient ischemic attack, myocardial infarction, renal failure, respiratory failure, ventricular arrhythmias, and supraventricular arrhythmias. The number and percentage of participants who have treatment-emergent cardiovascular event of interest will be presented by treatment group. The list of preferred terms within each of the cardiovascular events of interest is provided below:

Table 11–1 Cardiovascular Adverse Events of Interest

Heart failure or Congestive heart failure		
Cardiac failure Cardiac failure acute Cardiac failure chronic Cardiac failure congestive Cardiac failure high output Cardiopulmonary failure Acute left ventricular failure	Acute right ventricular failure Chronic left ventricular failure Chronic right ventricular failure Left ventricular failure Right ventricular failure Ventricular failure Cardiogenic shock	Acute pulmonary oedema Pulmonary oedema Cor pulmonale Cor pulmonale acute Cor pulmonale chronic Low cardiac output syndrome
Stroke or Transient ischemic attack		
Cerebrovascular accident Cerebral infarction Embolic cerebral infarction Thrombotic cerebral infarction Cerebellar infarction Haemorrhagic infarction Haemorrhagic cerebral infarction Ischaemic cerebral infarction Infarction Lacunar infarction Basal ganglia infarction Basal ganglia stroke Brain stem infarction Brain stem stroke Cerebellar stroke Embolic stroke Haemorrhagic stroke Haemorrhagic transformation stroke Ischaemic stroke Lacunar stroke Cerebral septic infarct Thalamic infarction Post procedural stroke Pseudostroke Stroke in evolution Thrombotic stroke	Vertebrobasilar stroke Basilar artery thrombosis Brain stem embolism Brain stem thrombosis Cavernous sinus thrombosis Cerebellar artery thrombosis Cerebellar embolism Cerebral artery embolism Cerebral artery thrombosis Cerebral gas embolism Cerebral microembolism Cerebral thrombosis Cerebral venous thrombosis Intracranial venous sinus thrombosis Precerebral artery thrombosis Superior sagittal sinus thrombosis Transverse sinus thrombosis Vertebral artery thrombosis Basal ganglia haemorrhage Brain stem haemorrhage Brain stem microhaemorrhage Central nervous system haemorrhage Cerebellar haemorrhage Cerebellar microhaemorrhage Cerebral arteriovenous malformation haemorrhagic	Cerebral cyst haemorrhage Cerebral haemorrhage Cerebral microhaemorrhage Epidural haemorrhage Extra-axial haemorrhage Haemorrhage intracranial Intracranial tumour haemorrhage Intraventricular haemorrhage Meningorrhagia Pituitary haemorrhage Subarachnoid haemorrhage Subdural haemorrhage Thalamus haemorrhage Traumatic intracranial haemorrhage Transient ischaemic attack Brain stem ischaemia Cerebral ischaemia Cerebellar ischaemia Cerebral small vessel ischaemic disease Reversible ischaemic neurological deficit Delayed ischaemic neurological deficit
Myocardial infarction		
Acute myocardial infarction ECG signs of myocardial infarction Myocardial infarction Periprocedural myocardial infarction	Post procedural myocardial infarction Silent myocardial infarction Papillary muscle infarction	Myocardial necrosis Myocardial ischaemia Subendocardial ischaemia
Renal failure		
Renal failure	Postrenal failure	Acute kidney injury

Hepatorenal failure Postoperative renal failure	Prerenal failure	Chronic kidney disease
Ventricular arrhythmias		
Ventricular arrhythmia Ventricular tachyarrhythmia Arrhythmia Accelerated idioventricular rhythm Cardiac fibrillation Cardiac flutter Extrasystoles	Parasystole Rhythm idioventricular Tachyarrhythmia Torsade de pointes Ventricular extrasystoles Ventricular fibrillation Ventricular flutter	Ventricular parasystole Ventricular pre-excitation Ventricular tachycardia Anomalous atrioventricular excitation Ventricular asystole
Supraventricular arrhythmias		
Arrhythmia supraventricular Atrial fibrillation Atrial flutter Atrial parasystole Atrial tachycardia Junctional ectopic tachycardia	Sinus tachycardia Supraventricular extrasystoles Supraventricular tachyarrhythmia Supraventricular tachycardia Paroxysmal arrhythmia Bradyarrhythmia	Nodal arrhythmia Nodal rhythm Sinus arrest Sinus arrhythmia Sinus bradycardia Sinus node dysfunction

To assess possible distant spread of toxin (PDSOT), 40 MedDRA preferred terms that may be associated with botulinum toxin effects have been identified. All TEAEs associated with PDSOT will be tabulated by SOC, preferred term and treatment group; in addition, all PDSOT TEAEs will be listed by subject. The 40 terms are listed below.

MedDRA Preferred Terms Evaluated for Possible Distant Spread of Toxin

Cardiac Disorders	**Nervous System Disorders**
Bradycardia	Bell's palsy
	Bulbar palsy
Eye Disorders	Cranial nerve palsies multiple
Accommodation disorder	Cranial nerve paralysis
Diplopia	Dysarthria
Opthalmoplegia	Facial paralysis
Eyelid function disorder	Facial paresis
Eyelid ptosis	Hyporeflexia
Pupillary reflex impaired	Hypotonia
Vision blurred	Paralysis
	Paresis cranial nerve
Gastrointestinal Disorders	Peripheral nerve palsy
Constipation	Peripheral paralysis
Dry mouth	Speech disorder
Dysphagia	Vocal cord paralysis
Ileus paralytic	Vocal cord paresis
Infections and Infestations	**Renal and Urinary Disorders**
Botulism	Urinary retention
Musculoskeletal and Connective Tissue Disorders	**Respiratory, Thoracic and Mediastinal Disorders**
Muscular weakness	Aspiration
	Diaphragmatic paralysis
	Dysphonia
	Dyspnoea
	Pneumonia aspiration
	Respiratory arrest
	Respiratory depression
	Respiratory failure
	Reproductive System and Breast Disorders
	Pelvic floor muscle weakness

11.2 Clinical Laboratory Parameters

Descriptive statistics for clinical laboratory values (in SI units) and changes from the baseline values at each assessment time point will be presented by visit and treatment group for all continuous laboratory parameters. Categorical laboratory parameters will be presented in listings. The laboratory parameters are:

Hematology: Basophils absolute cell count, eosinophils absolute cell count, hematocrit, hemoglobin, lymphocytes absolute cell count, monocytes absolute cell count, neutrophil absolute cell count, mean corpuscular hemoglobin (MCH), mean corpuscular hemoglobin concentration (MCHC), mean corpuscular volume (MCV), platelet count, percentage reticulocytes, red blood cell (RBC) count, and white blood cell (WBC) count.

Chemistry: Albumin, alanine aminotransferase (ALT), alkaline phosphatase, aspartate aminotransferase (AST), direct bilirubin, indirect bilirubin, total bilirubin, blood urea nitrogen (BUN), calcium, chloride, total cholesterol, creatinine, estimated glomerular filtration rate (eGFR), glucose non-fasting, triglycerides, HDL, LDL, Hemoglobin A1c, potassium, total protein, and sodium.

Urinalysis: Bilirubin, blood, glucose, ketones, leukocyte esterase, nitrite, pH, protein, specific gravity, urobilinogen, and microscopic examination (if blood or protein is abnormal)

Coagulation: Prothrombin time (PT), activated partial thromboplastin time (aPTT), and thrombin time (TT).

Clinical laboratory test values will be considered potentially clinically significant (PCS) if they meet either the lower-limit or higher-limit PCS criteria listed in Table 11–2. The number and percentage of participants who have PCS postbaseline clinical laboratory values will be tabulated by treatment group at each assessment. The percentages will be calculated relative to the number of participants with available non-PCS baseline values and at least 1 postbaseline assessment. The numerator will be the total number of participants with available non-PCS baseline values and at least 1 PCS postbaseline value. A supportive listing of participants with PCS postbaseline values will be provided, including the PID number, baseline and all postbaseline (including non-PCS) values.

Shift tables from baseline to each post-baseline visit for clinical laboratory parameters will be presented by treatment group for the following categories: low, normal, and high, which are provided by lab vendor.

Table 11–2 Criteria for Potentially Clinically Significant Laboratory Results

Parameter	***SI Unit***	***Lower Limit***	***Higher Limit***
CHEMISTRY			
Albumin	g/L	< 0.9 × LLN	> 1.1 × ULN
Alanine aminotransferase	U/L	—	≥ 3 × ULN
Alkaline phosphatase	U/L	—	≥ 3 × ULN
Aspartate aminotransferase	U/L	—	≥ 3 × ULN
Bilirubin, total	µmol/L	—	> 1.5 × ULN
Calcium	mmol/L	< 0.9 × LLN	> 1.1 × ULN
Chloride	mmol/L	< 0.9 × LLN	> 1.1 × ULN
Cholesterol	mmol/L	—	> 1.3 × ULN
Creatinine	µmol/L	—	> 1.3 × ULN
Glucose, fasting	mmol/L	< 0.8 × LLN	> 1.2 × ULN
Glucose, nonfasting	mmol/L	< 0.8 × LLN	> 1.4 × ULN
Potassium	mmol/L	< 0.9 × LLN	> 1.1 × ULN
Protein, total	g/L	< 0.9 × LLN	> 1.1 × ULN
Sodium	mmol/L	< 0.9 × LLN	> 1.1 × ULN
Triglycerides	mmol/L	—	> 2 × ULN
Urea nitrogen	mmol/L	—	> 1.2 × ULN
Uric acid	µmol/L	—	> 1.2 × ULN
HEMATOLOGY			
Basophils, absolute cell count	10^9/L	—	> 3 × ULN
Neutrophils, absolute cell count	10^9/L	< 0.8 × LLN	> 1.5 × ULN
Hematocrit	Ratio	< 0.9 × LLN	> 1.1 × ULN
Hemoglobin	g/L	< 0.9 × LLN	> 1.1 × ULN
Platelet count	10^9/L	≤ 0.5 × LLN	≥ 1.5 × ULN
Red blood cell count	10^{12}/L	< 0.9 × LLN	> 1.1 × ULN
White blood cell count	10^9/L	≤ 0.7 × LLN	≥ 1.5 × ULN
URINALYSIS			
pH	—	< 0.9 × LLN	> 1.1 × ULN
Specific gravity	—	—	> 1.1 × ULN

LLN = lower limit of normal value provided by the laboratory; SI = *Le Système International d'Unités* (International System of Units); ULN = upper limit of normal value provided by the laboratory.

11.3 Vital Signs

Descriptive statistics for vital signs (systolic and diastolic blood pressures, pulse rate, respiration rate, temperature, body weight, abdominal circumference, and body mass index) and changes from baseline values at each visit will be presented by treatment group.

Vital sign values will be considered PCS if they meet both the observed-value criteria and the change-from-baseline criteria listed in Table 11–3. The number and percentage of participants with PCS postbaseline values will be tabulated by treatment group. The percentages will be calculated relative to the number of participants with available baseline values and at least 1 postbaseline assessment. The numerator will be the total number of patients with available baseline values and at least 1 PCS postbaseline value. A supportive listing of participants with PCS postbaseline values will be provided, including the PID number, baseline and all postbaseline (including non-PCS) values.

Table 11–3 Criteria for Potentially Clinically Significant Vital Signs

Parameter	***Flag***	***Criteria[a]***	
		Observed Value	***Change From Baseline***
Sitting systolic blood pressure, mm Hg	High	≥ 160	Increase of ≥ 20
	Low	≤ 90	Decrease of ≥ 20
Sitting diastolic blood pressure, mm Hg	High	≥ 105	Increase of ≥ 15
	Low	≤ 50	Decrease of ≥ 15
Sitting pulse rate, bpm	High	≥ 110	Increase of ≥ 15
	Low	≤ 50	Decrease of ≥ 15
Weight, kg	High	—	Increase of ≥ 7%
	Low	—	Decrease of ≥ 7%

a A postbaseline value is considered potentially clinically significant if it meets both the observed-value and the change-from-baseline criteria.
bpm = beats per minute.

11.4 Electrocardiogram

Descriptive statistics for ECG parameters (heart rate, RR interval, PR interval, QRS interval, QT interval, and QTc) and changes from baseline values at each assessment time point to the end of study will be presented by treatment group. The QTc will be calculated using both the Bazett and Fridericia corrections (if the vendor does not provide).

Electrocardiographic parameter values are considered PCS if they meet or exceed the higher-limit PCS criteria listed in Table 11–4. The number and percentage of participants with PCS postbaseline ECG values will be tabulated by treatment group. The percentages will be

calculated relative to the number of participants with available baseline values and at least 1 postbaseline assessment. The numerator is the total number of participants with available baseline values and at least 1 PCS postbaseline value. A supportive listing of participants with PCS postbaseline values will be provided, including the PID number, baseline, all postbaseline (including non-PCS) values, and change from baseline.

Table 11–4 Criteria for Potentially Clinically Significant Electrocardiograms

Parameter	***Unit***	***Higher Limit***
QRS interval	msec	≥ 120
QRS interval	msec	>120 and increase from baseline > 25%
PR interval	msec	≥ 230
PR interval	msec	>230 and increase from baseline > 25%
QTc	msec	>480 or increase from baseline > 60
QTc	msec	increase from baseline > 30

QTc = QT interval corrected for heart rate.

11.5 Other Safety Parameters

11.5.1 Immunogenicity Assessments

Blood samples will be collected from all participants prior to dosing on day 1 and at the day 30 and day 90 follow-up visits. A 2-tier assay approach will be used for the detection of binding and neutralizing antibodies to AGN-151607 in human serum.

Immunogenicity findings (positive) will be tabulated with the number and percentage of participants at each visit separately for each AGN-151607 treatment group. Percentages will be based on the number of treated participants with interpretable antibody assays in each treatment group at the specified visit. A cumulative summary based on positive findings at any visit will be done for each AGN-151607 treatment group.

11.5.2 Pulmonary Function Test (PFT)

Study participant pulmonary function will be assessed at the screening and day 30 visits using forced procedures and compared across all 3 study groups. To demonstrate reproducibility of results, participants will be required to perform at least 3 acceptable spirometry maneuvers up to a maximum of 8 attempts. The results of all respiratory measures will be reviewed for clinically significant findings.

The assessments will include the following:

1. Forced vital capacity (FVC) measured from a maximal forced exhalation
2. Forced expiratory volume (FEV) in 1 second (FEV_1), 3 seconds (FEV_3) and 6 seconds (FEV_6)
3. FEV_1/FVC ratio

The average and the maximum of the measurements recorded on the eCRF for each of the PFT parameters (FVC, FEV1, FEV3, FEV6, and FEV_1/FVC ratio) at each visit will each be used in the data analysis. Descriptive statistics for each PFT parameter and changes from baseline values at each assessment time point will be presented by treatment group.

11.5.3 Time to Extubation

Descriptive statistics for time to extubation will be presented by treatment group. Time to extubation is defined as the time (hours) from the date/time of intubation to the date/time of extubation.

11.5.4 Physical Examination

Physical examination will be performed at screening, Day 30, Day 360 and at early discontinuation visits. All findings at the screening visit will be recorded on the medical history eCRF page. Any new or worsening condition observed after study injection will be recorded on the adverse event eCRF page. Hence summaries for physical examinations will be captured in summaries presented for medical history and adverse events.

12.0 Interim Analysis

An interim analysis of all efficacy data will be performed when all randomized participants have completed the Day 60 visit or exited the study earlier. The purpose of this analysis is to identify early trends in the data for administrative planning of following studies. There is no plan to stop the study or drop a treatment arm based on the results from the interim analysis. The interim analysis will serve as the primary analysis for the study.

Specifications of tables, figures, and data listings to be generated during the interim analysis are provided in a separate document. On completion of the database lock for the interim/primary analysis, the study will be unblinded to AbbVie study personnel. However, investigators and study sites personnel will remain blinded until completion of the final analysis. The final analysis will occur when all participants have completed the long-term safety follow-up or have exited the study earlier.

An independent data safety monitoring board (DSMB) will hold meetings after 5%, 25%, 50% and 75% of the total planned study population has completed Day 14 to evaluate partially blinded safety data for possible harmful effects of AGN-151607. More details regarding the policies, procedures, and composition of the DSMB, including precautions that will be employed to assure that inadvertent unblinding of parties does not occur; are described in the DSMB charter for this study. In addition, a separate SAP will be finalized before the first DSMB meeting.

13.0 Determination of Sample Size

A sample size of 100 randomized participants per treatment group will provide approximately 80% power to detect the treatment difference between each of the AGN-151607 doses (assumed equally effective) and placebo for the primary efficacy endpoint. The power calculations are based on the following assumptions:

- The expected incidence of early postoperative AF within 30 days post-surgery will be 47.8% in the placebo group and 27.8% in each of the AGN-151607 groups. This assumption is based on the results observed in the Pokushalov study (Pokushalov 2015) and the Duke University Medical Center clinical trial (Waldron et. al., 2019).

- Two-sided Mantel-Haenszel test statistic and a significance level of 0.05 for the test.

Calculations were performed using commercial software PASS 2008 (Hintze 2008). Assuming that approximately 10% in each intervention group will drop out early from the study, a total of approximately 330 participants will be randomly assigned to study intervention, in order to have approximately 300 participants complete all the efficacy assessments up to Day 60 visit.

14.0 Statistical Software

Statistical analyses will be performed using version 9.2 (or newer) of SAS on a Linux operating system.

15.0 Data Handling Conventions

15.1 Visit Time Windows

Table 15–1 presents the visits assigned for safety analyses and the corresponding range of treatment days (window) during which an actual visit may occur. Unless otherwise specified, efficacy parameters collected by visit (e.g., Quality of Life Outcomes) will use the same windows during which an actual visit may occur. All analyzable ePatch data collected from ePatches with a start date within the visit window will be included in the analysis for the derived visit for Days 60, 90, 180, 270, and 360.

Table 15–1 Analysis Visit Windows

Scheduled Visit	***Study Day[a]***	***Time Window (Study Day Range)***
Baseline	1	1
Day 2	2	2
Day 3	3	3
Day 4	4	4
Day 5	5	5
Day 6	6	[6, 10]
Day 14	14	[11, 21]
Day 30	30	[22, 45]
Day 60	60	[46, 75]
Day 90	90	[76, 135]
Day 180	180	[136, 225]
Day 270	270	[226, 315]
Day 360	360	> 315 days

a Relative to the date of the study treatment. Day 1 = the date of the study treatment.

If the assessment date (if the assessment date is unavailable, use visit date instead) is on or after the date of the study treatment, the study day is calculated by assessment date – date of the study treatment + 1. If the assessment date is before the date of the study treatment, the study day is calculated by assessment date – date of the study treatment. Therefore, a negative day indicates a day before the start of the study treatment.

If a subject has 2 or more visits within the same window, the last visit with a non-missing value will be used for analysis. All postbaseline assessments will be considered for PCS categorization. All assessments will be included in respective listings.

15.2 Repeated or Unscheduled Assessments of Safety Parameters

If a patient has repeated assessments before the start of the study treatment, the results from the last non-missing assessment made prior to the start of the study treatment will be used as baseline. If end-of-study assessments are repeated or if unscheduled visits occur, the last non-missing postbaseline assessment will be used as the end-of-study assessment for generating summary statistics. However, all postbaseline assessments will be used for PCS value determinations, and all assessments will be presented in the data listings.

15.3 Missing Severity Assessment for Adverse Events

If severity is missing for an AE that started before the date of the study treatment, an intensity of mild will be assigned. If severity is missing for an AE that started on or after the date of the study treatment, an intensity of severe will be assigned. The imputed values for severity assessment will be used for the incidence summary; the values will be shown as missing in the data listings.

15.4 Missing Causal Relationship to Study Drug for Adverse Events

If the causal relationship to the study treatment is missing for an AE that started on or after the date of the study treatment, a causality of yes will be assigned. The imputed values for causal relationship to study treatment will be used for the incidence summary; the values will be shown as missing in the data listings.

15.5 Missing Date Information for Adverse Events

The following imputation rules only apply to cases in which the start date for AEs is incomplete (i.e., partly missing).

Missing month and day

- If the year of the incomplete start date is the same as the year of the study treatment, the month and day of the study treatment will be assigned to the missing fields
- If the year of the incomplete start date is before the year of the study treatment, *December 31* will be assigned to the missing fields
- If the year of the incomplete start date is after the year of the study treatment, *January 1* will be assigned to the missing fields

Missing month only

- If only the month is missing, the day will be treated as missing and both the month and the day will be replaced according to the above procedure

Missing day only

- If the month and year of the incomplete start date are the same as the month and year of the study treatment, the day of the study treatment will be assigned to the missing day
- If either the year of the incomplete start date is before the year of the date of the study treatment or if both years are the same, but the month of the incomplete start date is before the month of the date of the study treatment, the last day of the month will be assigned to the missing day

- If either the year of the incomplete start date is after the year of the date of the study treatment or if both years are the same, but the month of the incomplete start date is after the month of the date of the study treatment, the first day of the month will be assigned to the missing day

If the stop date is complete and the imputed start date as above is after the stop date, the start date will be imputed by the stop date.

If the start date is completely missing and the stop date is complete, the following algorithm will be used to impute the start date:

- If the stop date is after the date of the study treatment, the date of the study treatment will be assigned to the missing start date
- If the stop date is before the date of the study treatment, the stop date will be assigned to the missing start date

15.6 Missing Date Information for Prior or Concomitant Medications

For prior or concomitant medications, including rescue medications, incomplete (i.e., partly missing) start dates and/or stop dates will be imputed. When the start date and the stop date are both incomplete for a patient, the start date will be imputed first.

15.6.1 Incomplete Start Date

The following rules will be applied to impute the missing numeric fields for an incomplete prior or concomitant medication start date. If the stop date is complete (or imputed) and the imputed start date is after the stop date, the start date will be imputed using the stop date.

Missing month and day

- If the year of the incomplete start date is the same as the year of the study treatment, the month and day of the study treatment will be assigned to the missing fields
- If the year of the incomplete start date is before the year of the study treatment, *December 31* will be assigned to the missing fields
- If the year of the incomplete start date is after the year of the study treatment, *January 1* will be assigned to the missing fields

Missing month only

- If only the month is missing, the day will be treated as missing and both the month and the day will be replaced according to the above procedure

Missing day only

- If the month and year of the incomplete start date are the same as the month and year of the study treatment, the day of the study treatment will be assigned to the missing day
- If either the year of the incomplete start date is before the year of the date of the study treatment or if both years are the same, but the month of the incomplete start date is before the month of the date of the study treatment, the last day of the month will be assigned to the missing day.
- If either the year of the incomplete start date is after the year of the date of the study treatment or if both years are the same, but the month of the incomplete start date is after the month of the date of the study treatment, the first day of the month will be assigned to the missing day

15.6.2 Incomplete Stop Date

The following rules will be applied to impute the missing numeric fields for an incomplete prior or concomitant medication stop date. If the date of the study exit is missing, the last available study visit date will be used as the study exit date.

Missing month and day

- If the year of the incomplete stop date is the same as the year of the study exit, the month and day of the study exit will be assigned to the missing fields
- If the year of the incomplete stop date is before the year of the study exit, *December 31* will be assigned to the missing fields
- If the year of the incomplete stop date is after the year of the study exit, *January 1* will be assigned to the missing fields

Missing month only

- If only the month is missing, the day will be treated as missing and both the month and the day will be replaced according to the above procedure

Missing day only

- If the month and year of the incomplete stop date are the same as the month and year of the study exit, the day of the study exit will be assigned to the missing day
- If either the year of the incomplete stop date is before the year of the date of the study exit or if both years are the same, but the month of the incomplete stop date is before the month of the date of the study exit, the last day of the month will be assigned to the missing day

- If either the year of the incomplete stop date is after the year of the date of the study exit or if both years are the same, but the month of the incomplete stop date is after the month of the date of the study exit, the first day of the month will be assigned to the missing day

15.7 Character Values of Clinical Laboratory Parameters

If the reported value of a clinical laboratory parameter cannot be used in a statistical summary table because, for example, a character string is reported for a parameter of the numeric type, a coded value must be appropriately determined for use in the statistical analyses. The actual values, however, as reported in the database will be presented in the data listings.

Table 15–2 shows examples of how some possible laboratory results should be coded for the analysis.

Table 15–2 Examples of Coding Special Character Values for Clinical Laboratory Parameters

Laboratory Test, SI Unit	*Possible Laboratory Results*	*Coded Value for Analysis*
CHEMISTRY		
ALT, U/L	< 5	5
AST, U/L	< 5	5
Bilirubin, total, µmol/L	< 2	2
URINALYSIS		
Glucose, mmol/L	= OR > 55, ≥ 55, > 0	Positive
	≤ 0, negative	Negative
pH	> 8.0, ≥ 8.0	8.0
	≥ 8.5	8.5
Protein	= OR > 3.0, ≥ 3.0, > 0	Positive
	≤ 0	Negative

15.8 Datetimes for ePatch Analyses

For analyses of time to first occurrence of AF, datetimes from the internal clock on the ePatch device will be used to determine the onset of event occurrence. During study conduct, it was discovered that some ePatches experienced an internal clock reset prior to collecting data, such that the internal clock was reset to a much earlier calendar date (e.g., a date in the year 1999 or 2000). But, the measurements of sequential time relative to the reset calendar datetime was determined to be accurate.

To correct for this clock reset issue and still utilize all collected data, the recording start time of each ePatch was compared to the application time reported by the site on the eCRF. For each

ePatch such that the application time reported on the eCRF – ePatch-reported recording start time > 30 days, the following adjustment was made to all times on that ePatch prior to analysis:

$$\text{adjusted time} = \text{ePatch-reported time} + \text{offset}$$

where 'offset' was calculated as: eCRF application time – ePatch-reported recording start time. In other words, when the ePatch-reported recording start time was more than 30 days earlier than the application time reported on the eCRF, all times on the ePatch were adjusted by the difference between the application time reported on the eCRF and the ePatch-reported recording start time. No adjustments were made to ePatch-reported times when the application time reported on the eCRF – ePatch-recording start time ≤ 30 days.

16.0 Changes to Analyses Specified in Protocol

The following changes were made to analyses specified in the protocol:

- Logistic regression models adjusting for baseline risk factors of POAF (type of surgery, and age group) will not be performed for the primary endpoints of the proportion of participants with at least one continuous AF Episode during the first 30 days post-surgery.
    - Rationale: adjusted odds ratios, stratified by type of surgery (presence or absence of valve surgery) and age group (< 65 or ≥ 65 years), with associated 95% CIs and P-values, are already presented using the CMH test.
- In lieu of t-tests, ANOVA models will be used to analyze hospital and ICU length of stay. Stratified Wilcoxon (Van Elteren) tests will be used for supplementary analyses for these variables, regardless of skewedness.
    - Rationale: ANOVA models are commonly used for stratified analyses of continuous outcomes. The stratified Wilcoxon (Van Elteren) test may provide value as a supplementary analysis even if the distribution is not skewed.
- Percentage of participants with at least 1 continuous AF episode ≥ 6 minutes during the first 30 days post-surgery was added as a secondary efficacy variable based on input from cardiology experts.

17.0 Version History

DOCUMENT HISTORY	
Document	**Date**
Original Final Analysis Plan	26 Apr 2021
Amendment 1	19 Nov 2021
Amendment 2	31 May 2022

Table 17–1 Final Analysis Plan Version History Summary

Amendment	Date	Summary
	26 Apr 2021	Original version
1	19 Nov 2021	• Added supplementary analyses using Supplementary ePatch Analysis Sets, described in Sections 6.3, Section 10.1 and Section 10.2;

Amendment	Date	Summary
		• Added list of antiarrythmics and concurrent procedures due to atrial fibrillation or atrial flutter to Section 8.0; • Edited so that mis-stratified participants, if any, will be analyzed based on the correct stratification group in Section 10.0; • Added percentage of patients at least 1 continuous AF episode ≥ 6 minutes during the first 30 days post-surgery as a secondary efficacy variable in Section 10.0 and Section 10.2; • Clarified definition of "symptomatic AF" in Section 10.2; • Added Cox proportional hazards models as additional analyses for time-to-event endpoints in Section 10.2; • Added imputation rules for missing ICU admission date/times in Section 10.3; • Added stratified Wilcoxon (Van Elteren) tests for length of stay endpoints in Section 10.3; • Added hospital length of stay measured from the time of initial hospital admission to discharge as an endpoint in Section 10.3; • Added clarification for Poisson regression model and an additional analysis using total analyzed time in the model in Section 10.3; • Added components to the composite endpoint, specified binary and time-to-event versions, and clarified definitions of terms for the Win Ratio analysis in Section 10.3; • Added changes to analyses specified in protocol in Section 16.0.
2	31 May 2022	• Added supplementary analyses using Beta Blocking Agent Withdrawn ePatch Analysis Set, described in Sections 6.4 and Section 10.1; • Clarified that calculation of time spent in AF (AF burden) will exclude continuous AF episodes < 30 seconds in duration in Section 10.2; • Added restricted mean survival time (RMST) to time-to-event analyses described in Section 10.2; • Added negative binomial, zero-inflated negative binomial and zero-inflated Poisson models to analysis of number of AF events in Section 10.3; • Added 'Percentage of participants free from AF through 12 months post-surgery and not taking any anticoagulant drugs' as Additional Efficacy variable in Section 10.3; • Updated that TEAEs would be sorted based on descending frequency of total percentage across treatment groups and not just descending frequency in the AGN-151607 250 U arm; • Added FEV_1/FVC ratio and maximum of measurements to analysis of PFT parameters in Section 11.5.2 to reflect parameters analyzed in other studies of botulinum toxins; • Removed cardiac pacemaker insertion from list of procedures of interest due to atrial fibrillation or atrial flutter in Section 8.0 (at time of SAP amendment, no such procedures were marked as due to atrial fibrillation or atrial flutter); • Added description of datetime adjustment for ePatch analyses to correct for instances of internal clock resets as Section 15.8.

18.0 References

Pokushalov E, Kozlov B, Romanov A, Strelnikov A, Bayramova S, Sergeevichev D, et al. Long-term suppression of atrial fibrillation by botulinum toxin injection into epicardial fat pads in patients undergoing cardiac surgery: one-year follow-up of a randomized pilot study. Circ Arrhythm Electrophysiol. 2015;8:1334-1341.

Waldron NH, Cooter M, Haney JC, Schroder JN, Gaca JG, Lin SS, Sigurdsson MI, Fudim M, Podgoreanu MV, Stafford-Smith M, Milano CA, Piccini JP, Mathew JP. Temporary autonomic modulation with botulinum toxin type A to reduce atrial fibrillation after cardiac surgery. Heart Rhythm. 2019 Feb;16(2):178-184.

Stuart J. Pocock, Cono A. Ariti, Timothy J. Collier, and Duolao Wang. The win ratio: a new approach to the analysis of composite endpoints in clinical trials based on clinical priorities. European Heart Journal (2012) 33, 176–182

Kenward MG, Roger JH. Small sample inference for fixed effects from restricted maximum likelihood. Biometrics 1997; 53:983-97.

Pickard AS, Law EH, Jiang R, Pullenayegum E, Shaw JW, Xie F, Oppe M, Boye KS, Chapman RH, Gong CL, Balch A, Busschbach JJV, United States Valuation of EQ-5D-5L Health States Using an International Protocol. *Value Health* 2019; 22(8):931-941.

Aranki SF, Shaw DP, Adams DH, Rizzo RJ, Couper GS, VanderVliet M, Collins JJ Jr, Cohn LH, Burstin HR. Predictors of atrial fibrillation after coronary artery surgery. Current trends and impact on hospital resources. *Circulation*. 1996 Aug;94(3):390-7.