Cover Page

PEERLESS Study

Statistical Analysis Plan v. 5.0

22MAR2024

NCT05111613

PEERLESS Study

Statistical Analysis Plan

Version 5.0

Study Treatment Device: FlowTriever® Retrieval/Aspiration System and CDT

Protocol Number: 21-002 version 5.0

Study Phase: Interventional

CONFIDENTIALITY STATEMENT

This Statistical Analysis Plan and all attachments are considered confidential information. The contents must not be disclosed, unless authorized in writing by Inari Medical, except that this document may be disclosed to the appropriate Institutional Review Board (IRB) with similar requests for maintenance of confidentiality. All data and study results are similarly subjected to the same disclosure restrictions as the investigational plan, as stated above. All copies of this analysis plan remain the property of Inari Medical.

APPROVAL

This document has been reviewed and approved by:		
Caitlin Mills Manager, Biostatistics	Signature	Date
Roger Chang Sr. Manager, Biostatistics	Signature	Date
Seth Swanlund Sr. Manager, Clinical Research Data	Signature	Date
Scott Skorupa Sr. Manager, Clinical Program	Signature	Date

VERSION HISTORY

SAP Version	Date	Primary Reasons for Amendment
0.1	12/OCT/2021	Initial version
0.2	01/NOV/2021	Incorporated internal reviewers' comments
1.0	02/NOV/2021	Incorporated internal reviewers' comments and format change
2.0	06/JAN/2023	Updated Inclusion/Exclusion criteria to align with study protocol v3.0
3.0	15/FEB/2023	Added section 6 covering data handling prior to database lock.
4.0	7/SEPT/2023	Updated to align with study protocol v4.0.
5.0	22/MAR/2024	Updated poolability section with further details.

TABLE OF CONTENTS

LISTING OF FIGURES

LIST OF ABBREVIATIONS AND DEFINITIONS

Abbreviation	Term
BP	Blood pressure
CDT	Catheter directed thrombolysis
CEC	Clinical Events Committee
CRNM	Clinically Relevant Non-Major
CT	Computed Tomography
CTED	Chronic thromboembolic disease
CTEPH	Chronic thromboembolic pulmonary hypertension
CTPA	Computed tomographic pulmonary angiography
ESC	European Society of Cardiology
EDC	Electronic Data Capture
eCRF	Electronic case report form
FT	FlowTriever
HIT	Heparin-induced thrombocytopenia
ICH	Intracranial hemorrhage
ICU	Intensive Care Unit
ISTH	International Society on Thrombosis and Haemostasis
ITT	Intention To Treat
LMWH	Low molecular weight heparin
LV	Left ventricle
mMRC	Modified Medical Research Council Dyspnea Scale
PA	Pulmonary Artery
PE	Pulmonary Embolism
PEmb-QOL	Pulmonary Embolism Quality of Life
PESI	Pulmonary Embolism Severity Index
PP	Per Protocol
RV	Right ventricle
RV/LV	Right ventricular to left ventricular diameter ratio
sPESI	Simplified Pulmonary Embolism Severity Index
VTE	Venous thromboembolism

1. Description of Study Objectives

The primary study objective is to compare the clinical outcomes of patients treated with the FlowTriever (FT) System versus catheter-directed thrombolysis (CDT) for use in the treatment of acute pulmonary embolism (PE).

2. Study Design

The PEERLESS study is a prospective, multicenter, open-label, randomized controlled trial of the FlowTriever System compared to CDT for acute PE, and includes a non-randomized cohort for subjects with an absolute contraindication to thrombolytics. The study will collect data on demographics, comorbidities, details from the PE diagnosis and treatment, and clinical outcomes through 30-day follow up. The follow up evaluations will include the 24-hour visit (24 hours ±8 hours), hospital discharge, and the 30-day/Exit visit (30 days +15 days). All windows reference time from Index Procedure completion (defined as time of exit from the procedure room).

2.1. Study Population

Randomized Controlled Trial Cohort (RCT Cohort):

Up to 550 subjects with acute PE will be enrolled and randomized at up to 60 study sites in US, EU, and/or UK. All subjects who sign informed consent and who meet all of the inclusion criteria and none of the exclusion criteria will be randomized (1:1, FlowTriever or CDT). The inclusion and exclusion criteria are described in sections 2.3 and 2.4.

- One-to-one (1:1) randomization will be stratified by bleeding risk, as measured by the VTE-BLEED score[1]. The detailed algorithm for computing VTE-BLEED score is described in Table 1 of the study protocol.
- Stratification by the VTE-BLEED algorithm will occur automatically in the Electronic Data Capture (EDC) system upon data entry, and randomization will be assigned accordingly.

Non-Randomized Absolute Contraindication to Thrombolytics Cohort (Contraindication Cohort):

Up to 150 additional subjects who meet study eligibility criteria and who have an absolute contraindication to thrombolytics, whose initial planned primary treatment strategy includes FlowTriever, will be evaluated as part of the Contraindication Cohort. The same RCT Cohort clinical assessments and follow up schedule will be administered in this Contraindication Cohort.

2.2. Point of Enrollment

To participate in the study, the patient must sign the informed consent. If a subject signs the informed consent but is later deemed NOT to meet the invasive systolic PA pressure eligibility criteria, the subject would be considered a screen failure and not enrolled in either the RCT Cohort or the Contraindication Cohort. Screen failures will be tracked in the EDC system with rationale for the screen failure. The point of enrollment is when the subject meets all eligibility criteria and study device enters the subject's body. If the index catheter-based intervention procedure is aborted before treatment catheter insertion or is not performed, the subject is considered a screen fail.

Enrollment considerations in study design

Any patient with a documented ***absolute contraindication*** to thrombolytics will be excluded from the RCT Cohort but may be enrolled in the Contraindication Cohort if treated with FlowTriever. Absolute contraindications include (per ESC Guidelines 2019[2] and AHA Scientific Statements 2019):

- History of hemorrhagic stroke or stroke of unknown origin
- Ischemic stroke in previous 6 months
- Presence of intracranial conditions that may increase the risk of bleeding, such as neoplasms, arteriovenous malformations, or aneurysms
- Recent (within 3 months) intracranial or intraspinal surgery or serious head trauma
- Bleeding diathesis
- Active internal bleeding, excluding menses
- Aortic dissection
- Severe uncontrolled hypertension
- Any other condition listed as an absolute contraindication on the product label for the thrombolytic agent planned for use by local standard and investigator discretion

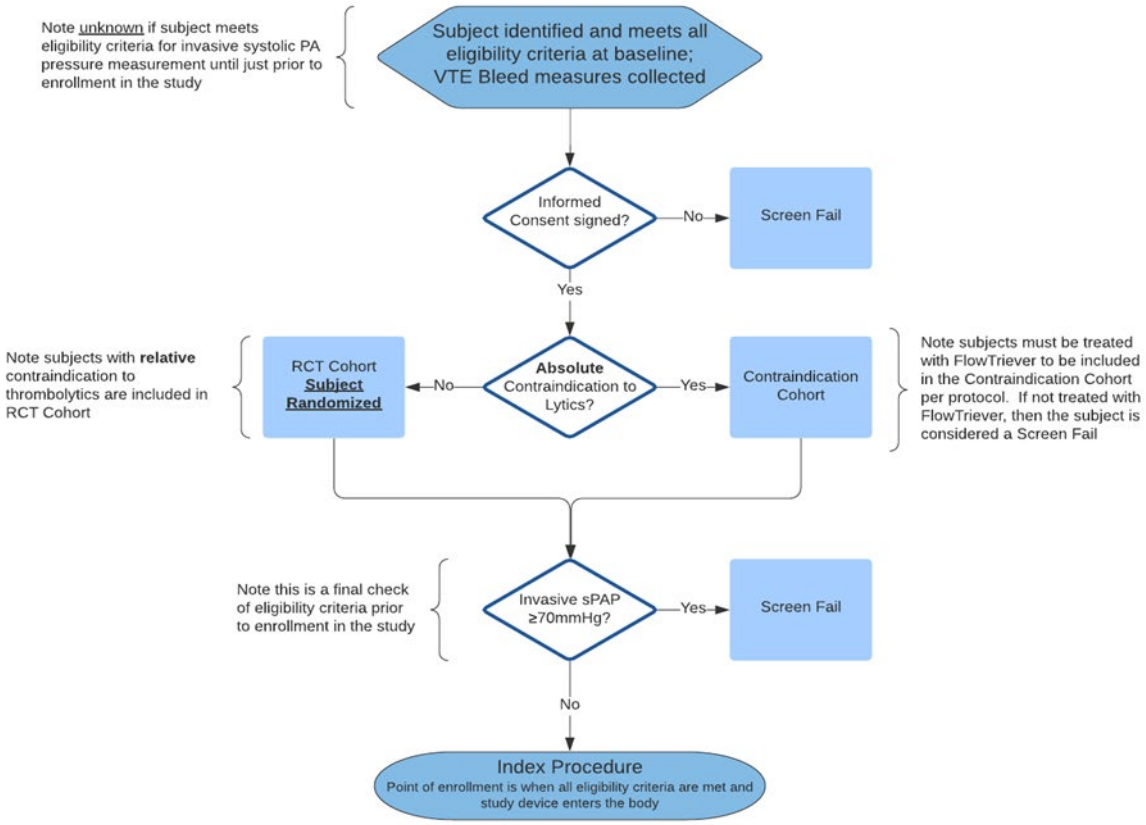
Subjects with a ***relative contraindication*** to thrombolytics are eligible for the RCT Cohort per protocol. Subjects with only relative contraindication(s) to thrombolytics are not eligible for the Contraindication Cohort. Relative contraindications include:

- Transient ischemic attack in previous 6 months
- Oral anticoagulation, except for aspirin
- Therapeutic LMWH within 24 hours
- Pregnancy or first post-partum week
- Non-compressible puncture sites
- Traumatic resuscitation, defined as prolonged (>10 min) cardiopulmonary resuscitation

- Refractory hypertension (systolic BP >180 mmHg or diastolic BP >110 mmHg) on two confirmed measurements
- Advanced liver disease
- Infective endocarditis
- Active peptic ulcer
- Recent administration of glycoprotein (GP) IIb/IIIa inhibitors
- Anemia (e.g. hemoglobin <10 g/dL)

Figure 1 summaries the enrollment process which will be implemented in this study.

Figure 1 Subject Enrollment Flowchart



2.3. Inclusion Criteria

Subjects must meet each of the following criteria to be included in the study:

1. Age ≥ 18 years
2. Echo, computed tomographic pulmonary angiography (CTPA), or pulmonary angiographic evidence of any proximal filling defect in at least one main or lobar pulmonary artery
3. Including ALL of the following:
    a. Clinical signs and symptoms consistent with acute PE, or PESI class III-V, or sPESI ≥1
    AND
    b. Hemodynamically stable
    AND
    c. RV dysfunction on echocardiography or CT
    AND
    d. Any one or more of the following present at the time of diagnosis:
        i. Elevated cardiac troponin levels
        ii. History of heart failure\
        iii. History of chronic lung disease
        iv. Heart rate ≥110 beats per minute
        v. SBP <100mmHg
        vi. Respiratory rate ≥30 breaths per minute
        vii. O2 saturation <90%
        viii. Syncope related to PE
        ix. Elevated lactate
4. Intervention planned to begin within 72 hours of the later of either
    a. Confirmed PE diagnosis
    OR
    b. If transferring from another hospital, arrival at the treating hospital
5. Symptom onset within 14 days of confirmed PE diagnosis

2.4. Exclusion Criteria

Subjects will be excluded from the study for any of the following criteria:

1. Unable to anticoagulate with heparin, enoxaparin or other parenteral antithrombin
2. Index presentation with hemodynamic instability that are part of the high-risk PE definition in the ESC Guidelines 2019[2], including ANY of the following:
    a. Cardiac arrest
    OR

   b. Systolic BP < 90 mmHg or vasopressors required to achieve a BP ≥90 mmHg despite adequate filling status, AND end-organ hypoperfusion
      OR
   c. Systolic BP < 90 mmHg or systolic BP drop ≥40 mmHg, lasting longer than 15 min and not caused by new-onset arrhythmia, hypovolemia, or sepsis
3. Known sensitivity to radiographic contrast agents that, in the Investigator's opinion, cannot be adequately pre-treated
4. Imaging evidence or other evidence that suggests, in the opinion of the Investigator, the patient is not appropriate for catheter-based intervention (e.g. inability to navigate to target location, clot limited to segmental/subsegmental distribution, predominately chronic clot)
5. Patient has right heart clot in transit identified at baseline screening
6. Life expectancy < 30 days (e.g. stage 4 cancer or severe COVID-19 infection), as determined by the Investigator
7. Current participation in another drug or device study that, in the investigator's opinion, would interfere with participation in this study
8. Current or history of chronic thromboembolic pulmonary hypertension (CTEPH) or chronic thromboembolic disease (CTED) diagnosis, per ESC 2019 guidelines[2]
9. Invasive systolic PA pressure ≥ 70mmHg prior to study device entering the body
10. Administration of bolus or drip/infusion thrombolytic therapy or mechanical thrombectomy for the index PE event within 48 hours prior to enrollment
11. Ventricular arrhythmias refractory to treatment at the time of enrollment
12. Known to have heparin-induced thrombocytopenia (HIT)
13. Subject has any condition for which, in the opinion of the investigator, participation would not be in the best interest of the subject (e.g., compromise the well-being or that could prevent, limit, or confound the protocol-specified assessments). This includes a contraindication to use of FlowTriever or CDT System (for example, EKOS System) per local approved labeling
14. Subject has previously completed or withdrawn from this study
15. Patient unwilling or unable to conduct the follow up visits per protocol

2.5. Primary Endpoint

The primary endpoint is a composite clinical endpoint constructed as a win ratio, a hierarchy of the following, which are assessed at hospital discharge or at 7 days after the index procedure, whichever is sooner:

1. All-cause mortality, or
2. Intracranial hemorrhage (ICH), or
3. Major bleeding per ISTH[3] definition, or

4. Clinical deterioration defined by hemodynamic or respiratory worsening, and/or escalation to a bailout therapy, or
5. ICU admission and post-procedure ICU length of stay

The first four endpoint events will be adjudicated by Clinical Events Committee (CEC). The definition of these endpoints can be found in the study protocol.

2.6. Secondary Endpoints

The secondary endpoints of the study will assess safety, effectiveness, and utility measures, as follows:

- Composite clinical endpoint constructed as a win ratio hierarchy of the following four components, assessed at hospital discharge or at 7 days after the index procedure, whichever is sooner:
  - All-cause mortality, or
  - Intracranial hemorrhage (ICH), or
  - Major bleeding per ISTH[3] definition, or
  - Clinical deterioration defined by hemodynamic or respiratory worsening, and/or escalation to a bailout therapy
- Individual components of the win ratio composite endpoint, assessed at hospital discharge or at 7 days after the index procedure, whichever is sooner:
  - All-cause mortality
  - Intracranial hemorrhage (ICH)
  - Major bleeding per ISTH[3] definition
  - Clinical deterioration defined by hemodynamic or respiratory worsening, and/or escalation to a bailout therapy
  - ICU admission and post index procedure ICU length of stay
- All-cause mortality within 30 days from index procedure
- PE-related and all-cause readmission within 30 days from index procedure
- Device and drug-related serious adverse events through the 30 day visit
- Clinically Relevant Non-Major (CRNM) and Minor bleeding events through hospital discharge or at 7 days after the index procedure, whichever is sooner
- Change in core-lab-adjudicated right-ventricular/left-ventricular (RV/LV) ratio from baseline to 24 hour visit, as measured by echocardiography or CT
- mMRC Dyspnea score at 24 hour visit and 30 day visit
- Length of post-index-procedure hospital stay (to a maximum of 30 days)
- Disease-specific and general health-related quality of life at the 30 day visit (PEmb-QoL and EQ-5D-5L)

3. Analysis Populations

The analysis populations are defined as follows.

3.1. Intention To Treat Population

All subjects who met the inclusion/exclusion criteria, provided informed consent, and who were randomized to the FlowTriever or CDT group will be included in ITT. To be included in ITT, subjects must receive some treatment for PE after randomization. This study will consist of up to 550 subjects with acute PE who meet the eligibility criteria, and are enrolled, consented, and randomized. Subjects will be excluded from ITT population if no treatments were given, or no data were available after randomization.

3.2. Per Protocol Population

Per Protocol (PP) population is a subset of ITT population that adhere to the protocol. The following are some of the possible reasons to exclude subjects from PP population.

- Non-compliant
- Protocol deviations
- Withdrew from study
- Received other PE treatments than the assigned

The primary and secondary endpoints will be analyzed separately for both ITT and PP populations.

3.3. Contraindication Cohort

Up to 150 additional subjects who meet study eligibility criteria and who have an absolute contraindication to thrombolytics, whose initial planned primary treatment strategy includes FlowTriever, will be evaluated as part of the Contraindication Cohort. The same RCT Cohort clinical assessments and follow up schedule will be administered in the Contraindication Cohort.

4. Incomplete Date Handling and Missing Data

Partial date/time input will not be collected in the electronic case report form (eCRF), i.e., date/time will be either non-missing or completely missing. On a case-by-case basis, when the missing date/time imputation is required, the most conservative date/time will be used.

In general, main analysis will be conducted in the complete case analysis. Additional sensitivity analysis will be conducted by imputing missing data using multiple imputation.

5. Statistical Methods and Analysis

5.1. Sample Size

By assuming an 80% power with one-sided alpha of 2.5%, the win ratio methodology is applied to the primary endpoint that consists of five components. The required sample size is 432 subjects (216 subjects per arm); and planning for follow up attrition and further describing secondary endpoints/exploratory analysis, the study will enroll up to a total of 550 randomized subjects (RCT Cohort). The ”event” proportions of the five components are based on Inari sponsored studies and a review of literature.

Up to 150 additional subjects who meet all eligibility criteria and who have an absolute contraindication to thrombolytics, whose initial planned primary treatment strategy includes FlowTriever, will be enrolled as part of the Contraindication Cohort. The same clinical assessments and follow up schedule will be administered in this cohort as is described for the RCT Cohort.

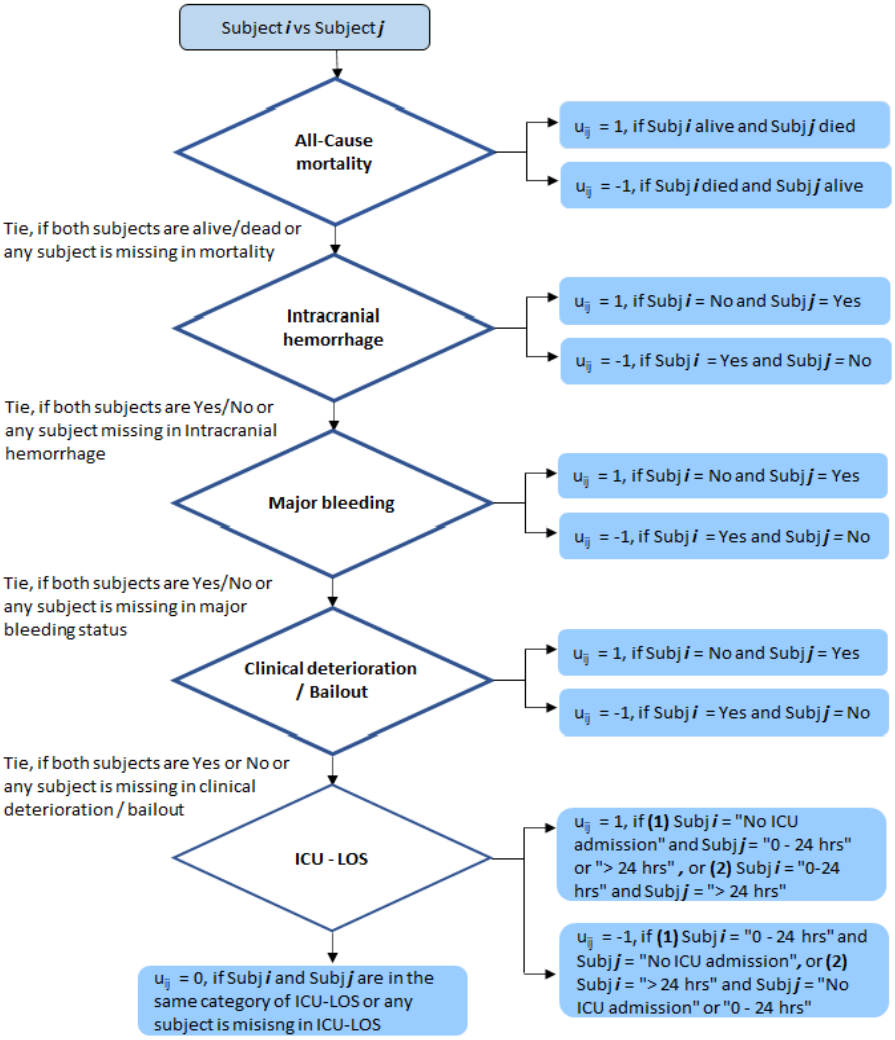
5.2. Statistical Methodology

The statistical design objective for this trial is to compare the clinical outcomes of the FlowTriever System versus CDT for use in the treatment of acute pulmonary embolism (PE). Since the primary endpoint is a composite clinical endpoint, a modified generalized Wilcoxon test (F-S test) proposed by Finkelstein & Schoenfeld[4] will be applied to examine the performance differences between the two treatment arms.

In this trial, the primary endpoint is a hierarchy of five clinical outcome components. Each subject in the study will be compared to each of the other subjects in a pairwise manner, regardless of which treatment arm the compared subject is in, and assigned a score, u_{ij}, of 1, -1, or 0, depending on whether the comparison has a favorable, unfavorable, or unsettled outcome in the hierarchy of the clinical events; *i* is the index subject to be compared with all other subjects in a pairwise fashion and *j* represents j^{th} subject comparison. For example, if a subject *i* is alive while a subject *j* died, the score is 1; if a subject *i* died while subject *j* is alive, the score is -1. If both subjects are alive/dead or any subject is missing in mortality status (i.e., a pair is unsettled/not comparable for the first outcome), the second clinical outcome (Intracranial hemorrhage) is compared and assigned a score of 1, -1, or unsettled in a similar comparison logic. Subsequently, within each pairwise comparison, the score is determined by comparing five clinical outcomes sequentially in the order of outcome priorities (see Figure 2). In summary, for each pair of subjects (*i*, *j*), the score is defined as

$$u_{ij} = \begin{cases} 1, \text{if subject } i \text{ does better than subject } j \\ -1, \text{if subject } j \text{ does better than subject } i \\ 0, \text{if it cannot be determined} \end{cases}$$

Figure 2 Pairwise Comparison Flowchart



Finkelstein & Schoenfeld[4] (F-S) then assigned a score $U_i = \sum_{i \neq j} u_{ij}$ to each subject *i*. Their proposed test is a score test based on the sum of the ranks for the treated group (see the following eqation).

$$T = \sum_{i=1}^{N} U_i D_i$$

where D_i= 1 for subjects in FlowTriever arm and D_i= 0 for subjects in CDT arm, and N is the number of total subjects in the trial. The proposed F-S statistic for the hypothesis of interest is $T/\sqrt{V}$, where $V = \frac{n_{FT} n_{CDT}}{N(N-1)} \sum_i U_i^2$ is the variance of T, n_{FT} and n_{CDT} are the number of subjects in FlowTriever and CDT

arms, respectively, and $n_{FT} + n_{CDT} = N$. The hypothesis is tested by comparing the F-S statistic to the standard normal distribution (Finkelstein & Schoenfeld[4]).

To determine the sample size required to achieve 80% power given one-sided alpha of 2.5%, we first will simulate subject level data per data replicate. For a data replicate and within a treatment arm, clinical outcomes with event proportions are sampled from binomial distribution; ICU is from a multinominal distribution with 3 categories, including no ICU admission, ≤ 24 hours, and > 24 hours. We will assume each of the clinical outcomes are independent of each other. For example, we may sample 100 subjects per treatment arm, and within each treatment arm each has different clinical outcome proportions or means. Once we generate 200 subjects in total (100 per treatment arm), we may derive the F-S test statistic and its p-value. We will then repeat the process for 2,500 data replicates and derive 2,500 p-values based on F-S test statistics; power is the proportion of p-values that are ≤ 0.05. Each given sample size number will thus lead to one power number to be calculated. As we alter the sample sizes in the simulation scenario, we can conduct a grid search to determine the minimal sample size required to achieve at least 80% power.

The primary and secondary endpoints will be analyzed for both ITT and PP populations, separately. Primary endpoint and secondary endpoints will lead to hypothesis testing whereas, only descriptive summary statistics will be provided for the Contraindication Cohort.

5.3. Primary Endpoint Analysis

The primary endpoint for this study will be assessed using the F-S statistic with the following null and alternative hypotheses:

H_0: $n_{FW} \leq n_{CW}$ versus H_A: $n_{FW} > n_{CW}$

where n_{FW} is the random variable of the number of winners for the FlowTriever System (i.e., counting the number of pairs in which the subject *i* in FlowTriever System does better than those in CDT arms) while n_{CW} is the random variable of the number of winners for the CDT (i.e., counting the number of pairs in which the subject *i* in FlowTriever System does worse than those in CDT arms).

To separately quantify the treatment effect of interest, the win ratio of the treatment effect is defined by Pocock el al.[5], where the performance differences between the two treatment arms is defined as $\frac{N_{FW}}{N_{CW}}$, where N_{FW} is the number of winners for the FlowTriever System and N_{CW} is the number of winners for CDT in the dataset; the 95% CI of the win ratio estimate can be derived via a bootstrap method, a method whereby the original data is sampled with replacement with a large number of times, e.g. 1000 samples being drawn to create 1000 datasets. From each bootstrap sample, the win ratio of the treatment effect is computed and an empirical distribution of the win ratio is determined from the total number of sample datasets. The lower and upper limits of the 95% CI are the 2.5%ile and 97.5%ile

of this empirical distribution of the win ratio out of 1000 samples. In this study, FT and CDT groups are labelled ***F*** and ***C*** with n_F and n_C subjects, respectively. The following are the five steps of calculating 95% CI via a bootstrap approach:

1. Draw a random sample S_F of size N_F (275) with replacement from the original FT group and a random sample S_C of N_C with replacement from original CDT group. In this study, $S_F = S_C$.
2. Perform the unmatched analysis, all possible pairs comparison, on the samples S_F, S_C and calculate the win ratio of the treatment effect.
3. Repeat Steps 1 and 2 1000 times.
4. Determine the empirical bootstrap distribution of the win ratio from the 1000 bootstrap values.
5. Obtain the 2.5%ile and 97.5%ile of the bootstrap distribution which are the estimated limits of the 95% confidence interval of the win ratio.

The hypothesis testing using the F-S statistic and the estimation of the win ratio of the treatment effect of the primary endpoint will be assessed in both ITT and PP populations. If the results of the ITT and PP populations are similar, it indicates robust confidence on the treatment effect of FT relative to CDT.

5.4. Secondary Endpoints Analysis

5.4.1. Composite Endpoint of Four Clinical Events

Similar to the primary endpoint for this study, the composite clinical endpoint constructed as a win ratio hierarchy of the four components (see section 2.6) will be assessed using the F-S statistic with the following null and alternative hypotheses:

H_0: $n'_{FW} \leq n'_{CW}$ versus H_A: $n'_{FW} > n'_{CW}$

where n'_{FW} is the random variable of the number of winners for the FlowTriever System (i.e., counting the number of pairs in which the subject *i* in FlowTriever System does better than those in CDT arms) while n'_{CW} is the random variable of the number of winners for CDT (i.e., counting the number of pairs in which the subject *i* in FlowTriever System does worse than those in CDT arms). Additionally, the 95%CI of the win ratio of the treatment effect of this endpoint will be presented.

This composite endpoint will be assessed in both ITT and PP populations. If the results of the ITT and PP populations are similar, it indicates robust confidence on the treatment effect of FT relative to CDT.

5.4.2. All-cause Mortality at Discharge/at 7 Days after the Index Procedure

All-cause mortality, assessed at hospital discharge or at 7 days after the index procedure, whichever is sooner, will be analyzed as with the following null and alternative hypotheses:

H_0: $p_{DF} \geq p_{DC}$ versus H_A: $p_{DF} < p_{DC}$

where p_{DF} represents the proportion of mortalities in the FlowTriever System group while p_{DC} represents the proportion of mortalies in the CDT group. The hypothesis will be tested using a Fisher's exact test.

5.4.3. All-cause Mortality within 30 Days from Index Procedure

All-cause mortality within 30 days will be analyzed as with the following null and alternative hypotheses:

H_0: $p_{D30F} \geq p_{D30C}$ versus H_A: $p_{D30F} < p_{D30C}$

where p_{D30F} represents the proportion of mortalities within 30 days in the FlowTriever System group while p_{D30C} represents the proportion of mortalies within 30 days in the CDT group. The hypothesis will be tested using a Fisher's exact test.

5.4.4. Intracranial Hemorrhage

Intracranial hemorrhage (ICH), assessed at hospital discharge or at 7 days after the index procedure, whichever is sooner, will be analyzed as with the following null and alternative hypotheses:

H_0: $p_{ICHF} \geq p_{ICHC}$ versus H_A: $p_{ICHF} < p_{ICHC}$

where p_{ICHF} represents the proportion of subjects with ICH in the FlowTriever System group while p_{ICHC} represents the proportion of subjects with ICH in the CDT group. The hypothesis will be tested using a Fisher's exact test.

5.4.5. Major Bleeding per ISTH[3] Definition

Major bleeding, per ISTH[3] definition, assessed at hospital discharge or at 7 days after the index procedure, whichever is sooner, will be analyzed as with the following null and alternative hypotheses:

H_0: $p_{BF} \geq p_{BC}$ versus H_A: $p_{BF} < p_{BC}$

where p_{BF} represents the proportion of subjects with major bleeding in the FlowTriever System group while p_{BC} represents the proportion of subjects with major bleeding in the CDT group. The hypothesis will be tested using a Fisher's exact test.

5.4.6. Clinical Deterioration and / or Escalation to a Bailout Therapy

Clinical deterioration defined by hemodynamic or respiratory worsening, and/or escalation to a bailout therapy, assessed at hospital discharge or at 7 days after the index procedure, whichever is sooner, will be analyzed as with the following null and alternative hypotheses:

H_0: $p_{CF} \geq p_{CC}$ versus H_A: $p_{CF} < p_{CC}$

where p_{CF} represents the proportion of subjects with clinical deterioration / a bailout therapy in the FlowTriever System group while p_{CC} represents the proportion of subjects with clinical deterioration / a bailout therapy in the CDT group. The hypothesis will be tested using a Fisher's exact test.

5.4.7. ICU admission and ICU length of stay

ICU admission and ICU length of stay, post index procedure and following the index procedure, will be analyzed as with the following null and alternative hypotheses:

H_0: $p_{IF1} \leq p_{IC1}$ and $p_{IF2} \geq p_{IC2}$ and $p_{IF3} \geq p_{IC3}$ versus H_A: $p_{IF1} > p_{IC1}$ or $p_{IF2} < p_{IC2}$ or $p_{IF3} < p_{IC3}$

where p_{IF1} represents the proportion of subjects without ICU admission in the FlowTriever System group, p_{IF2} represents the proportion of subjects with 0-24 hours ICU stay in the FlowTriever System group, p_{IF3} represents the proportion of subjects with > 24 hours ICU stay in the FlowTriever System group, p_{IC1} represents the proportion of subjects without ICU admission in the CDT group, p_{IC2} represents the proportion of subjects with 0-24 hours ICU stay in the CDT group, p_{IC3} represents the proportion of subjects with > 24 hours ICU stay in the CDT group. The hypothesis will be tested using a Fisher's exact test.

5.4.8. PE-related Readmission within 30 Days from Index Procedure

PE-related readmission within 30 days from index procedure will be analyzed as with the following null and alternative hypotheses:

H_0: $p_{PRF} \geq p_{PRC}$ versus H_A: $p_{PRF} < p_{PRC}$

where p_{PRF} represents the proportion of subjects had PE-related readmission within 30 days in the FlowTriever System group while p_{PRC} represents the proportion of subjects had PE-related readmission within 30 days in the CDT group. The hypothesis will be tested using a Fisher's exact test.

5.4.9. All-cause Readmission within 30 Days from Index Procedure

All-cause readmission within 30 days from index procedure will be analyzed as with the following null and alternative hypotheses:

H_0: $p_{ARF} \geq p_{ARC}$ versus H_A: $p_{ARF} < p_{ARC}$

where p_{ARF} represents the proportion of subjects had all-cause readmission within 30 days in the FlowTriever System group while p_{ARC} represents the proportion of subjects had all-caused readmission within 30 days in the CDT group. The hypothesis will be tested using a Fisher's exact test.

5.4.10. Clinically Relevant Non-Major Events at Discharge/at 7 Days after the Index Procedure

Clinically relevant non-major (CRNM) events through hospital discharge or at 7 days after the index procedure, whichever is sooner, will be analyzed as with the following null and alternative hypotheses:

H_0: $p_{CRF} \geq p_{CRC}$ versus H_A: $p_{CRF} < p_{CRC}$

where p_{CRF} represents the proportion of CRNM events in the FlowTriever System group while p_{CRC} represents the proportion of CRNM events in the CDT group. The hypothesis will be tested using a Fisher's exact test.

5.4.11. Minor Bleeding Events at Discharge/at 7 Days after the Index Procedure

Minor bleeding events through hospital discharge or at 7 days after the index procedure, whichever is sooner, will be analyzed as with the following null and alternative hypotheses:

H_0: $p_{MBF} \geq p_{MBC}$ versus H_A: $p_{MBF} < p_{MBC}$

where p_{MBF} represents the proportion of minor bleeding events in the FlowTriever System group while p_{MBC} represents the proportion of minor bleeding events in the CDT group. The hypothesis will be tested using a Fisher's exact test.

5.4.12. Change in RV/LV Ratio

Change in right-ventricular/left-ventricular (RV/LV) ratio from baseline to 24 hour visit, as measured by echocardiography or CT, will be analyzed as with the following null and alternative hypotheses:

H_0: $\mu_{RF} \leq \mu_{RC}$ versus H_A: $\mu_{RF} > \mu_{RC}$

where μ_{RF} represents the mean reduction in the RV/LV ratio treated from baseline to 24 hour visit in the FlowTriever System group, and μ_{RC} represents the mean reduction in the RV/LV ratio treated from baseline to 24 hour visit in the CDT group. The hypothesis will be tested using a Wilcoxon Rank Sum test.

5.4.13. mMRC Dyspnea Score at 24 Hour Visit

The mMRC Dyspnea score at 24 hour visit will be analyzed as with the following null and alternative hypotheses:

H_0: $p_{mD24hF0} \leq p_{mD24hC0}$ and $p_{mD24hF1} \geq p_{mD24hC1}$ and $p_{mD24hF2} \geq p_{mD24hC2}$ and $p_{mD24hF3} \geq p_{mD24hC3}$ and $p_{mD24hF4} \geq p_{mD24hC4}$

versus

H_A: $p_{mD24hF0} > p_{mD24hC0}$ or $p_{mD24hF1} < p_{mD24hC1}$ or $p_{mD24hF2} < p_{mD24hC2}$ or $p_{mD24hF3} < p_{mD24hC3}$ or $p_{mD24hF4} < p_{mD24hC4}$

where $p_{mD24hFz}$ represents the proportion of subjects with mMRC Dyspnea Score equal to ***z*** at 24 hour visit in the FlowTriever System group, where z = 0, 1, 2, 3, and 4; and, $p_{mD24hCz}$ represents the proportion of subjects with mMRC Dyspnea Score equal to ***z*** at 24 hour visit in the CDT group, where z = 0, 1, 2, 3, and 4. The hypothesis will be tested using a Fisher's exact test.

5.4.14. mMRC Dyspnea Score at 30 Day Visit

The mMRC Dyspnea score at 30 day visit will be analyzed as with the following null and alternative hypotheses:

H_0: $p_{mD30dF0} \leq p_{mD30dC0}$ and $p_{mD30dF1} \geq p_{mD30dC1}$ and $p_{mD30dF2} \geq p_{mD30dC2}$ and $p_{mD30dF3} \geq p_{mD30dC3}$ and $p_{mD30dF4} \geq p_{mD30dC4}$

versus

H_A: $p_{mD30dF0} > p_{mD30dC0}$ or $p_{mD30dF1} < p_{mD30dC1}$ or $p_{mD30dF2} < p_{mD30dC2}$ or $p_{mD30dF3} < p_{mD30dC3}$ or $p_{mD30dF4} < p_{mD30dC4}$

where $p_{mD30dFz}$ represents the proportion of subjects with mMRC Dyspnea Score equal to ***z*** at 30 day visit in the FlowTriever System group, where z = 0, 1, 2, 3, and 4; and, $p_{mD30dCz}$ represents the proportion of subjects with mMRC Dyspnea Score equal to ***z*** at 30 day visit in the CDT group, where z = 0, 1, 2, 3, and 4. The hypothesis will be tested using a Fisher's exact test.

5.4.15. Length of Post-Index-Procedure Hospital Stay

The length of post-index-procedure hospital stay (to a maximum of 30 days) will be analyzed as with the following null and alternative hypotheses:

H_0: $\mu_{PHSF} \geq \mu_{PHSC}$ versus H_A: $\mu_{PHSF} < \mu_{PHSC}$

where μ_{PHSF} represents the mean length of post-index-procedure hospital stay in the FlowTriever System group, and μ_{PHSC} represents the mean length of post-index-procedure hospital stay in the CDT group. The hypothesis will be tested using a Wilcoxon Rank Sum test.

5.4.16. Disease-Specific Quality of Life at 30 Day Visit – PEmb-QoL

The PEmb-QoL at 30 days visit will be analyzed as with the following null and alternative hypotheses:

H_0: $\mu_{PQF} \geq \mu_{PQC}$ versus H_A: $\mu_{PQF} < \mu_{PQC}$

where μ_{PQF} represents the mean PEmb-QoL score at 30 day visit in the FlowTriever System group, and μ_{PQC} represents the mean PEmb-QoL score at 30 day visit in the CDT group. The hypothesis will be tested using a Wilcoxon Rank Sum test.

5.4.17. General Health-Related Quality of Life at 30 Day Visit – EQ-5D-5L

The EQ-5D-5L at 30 days visit will be analyzed as with the following null and alternative hypotheses:

H_0: $\mu_{EQF} \leq \mu_{EQC}$ versus H_A: $\mu_{EQF} > \mu_{EQC}$

where μ_{EQF} represents the mean EQ-5D-5L score at 30 day visit in the FlowTriever System group, and μ_{EQC} represents the mean EQ-5D-5L score at 30 day visit in the CDT group. The hypothesis will be tested using a Wilcoxon Rank Sum test.

The secondary endpoints listed in this section will be assessed for patients in ITT and PP populations with descriptive statistics and hypothesis testing. Continuous variables will report any combination of mean, SD, median, Q1, Q3, IQR, minimum and maximum, while the categorical variables will report frequency count and percentages (%).

5.5. Exploratory Analysis

Exploratory analysis may be, but not limited to, performed on additional outcomes of interest with descriptive statistics and hypothesis testing. The same approach described in Section 5.6 will be applied to control for overall α. Continuous variables will report any combination of mean, SD, median, Q1, Q3, IQR, minimum and maximum, while the categorical variables will report frequency count and percentages (%).

5.6. Controlling for Multiplicity

For the primary and secondary endpoints related to safety, α is set to 0.05, to conservatively include all potential safety signals. The secondary endpoints listed below are considered as safety endpoints:

- Composite clinical endpoint constructed as a win ratio hierarchy of the following four components, assessed at hospital discharge or at 7 days after the index procedure, whichever is sooner:
    - All-cause mortality, or
    - Intracranial hemorrhage (ICH), or
    - Major bleeding per ISTH[3] definition, or
    - Clinical deterioration defined by hemodynamic or respiratory worsening, and/or escalation to a bailout therapy
- Four Individual components of the win ratio composite endpoint, assessed at hospital discharge or at 7 days after the index procedure, whichever is sooner:
    - All-cause mortality
    - Intracranial hemorrhage (ICH)
    - Major bleeding per ISTH[3] definition

    - Clinical deterioration defined by hemodynamic or respiratory worsening, and/or escalation to a bailout therapy
- All-cause mortality within 30 days from index procedure
- PE-related and all-cause readmission within 30 days from index procedure
- Device and drug-related serious adverse events through the 30 day visit
- Clinically Relevant Non-Major (CRNM) and Minor bleeding events through hospital discharge or at 7 days after the index procedure, whichever is sooner

For the secondary endpoints not related to safety, the family-wide type I error rate will be controlled at one-sided α=0.025 using the Bonferroni correction with an effective number of independent tests. The typical Bonferroni correction assumes all tests are independent and is therefore extremely conservative. An alternative approach to estimate the effective number of independent tests for the correlated tests has been proposed by Gao et al (2008)[9]. Therefore, the α to adjust for multiplicity is simply

$$\frac{familywide\ type\ I\ error}{effective\ number\ of\ independent\ tests}$$

The secondary endpoints listed below are considered as non-safety endpoints:

- ICU admission and ICU length of stay post index procedure and following the index procedure
- Change in right-ventricular/left-ventricular (RV/LV) ratio from baseline to 24 hour visit, as measured by echocardiography or CT
- mMRC Dyspnea score at 24 hour visit and 30 day visit
- Length of post-index-procedure hospital stay (to a maximum of 30 days)
- Disease-specific and general health-related quality of life at the 30 day visit (PEmb-QoL and EQ-5D-5L)

5.7. Data Poolability Assessment

The planned analysis for this study will pool data by treatment arms across clinical study sites since the study is randomized at the study level and not at the site level; it is possible that certain sites may have heavier proportions of one type of treatment vs the other. Efforts were made to ensure that consistent procedures were used across study sites, including use of the same study protocol, Sponsor monitoring the sites for compliance, and use of identical data-gathering instruments. The appropriateness of pooling the data by treatment arms across sites will be evaluated[7,8].

For each treatment arm, it may be necessary to combine two or more low enrolling study sites into pseudo-sites to allow for these analyses. Sites with fewer than six subjects will be ranked by enrollment from low to high. Starting from the lowest enrollment site, sites will be combined into a pseudo site until the combined size reaches the median enrollment among all sites. This process will be repeated until all

resulting sites have enrollment equal to or greater than six subjects for each treatment arms. This will be done in a manner to preserve the structure of the study and prevent bias.

Poolability analysis will be performed on the primary endpoint comparing across sites within each treatment arm. Random effects modeling using the inverse variance method will be used to assess heterogeneity between pseudo-sites. This is done by using pseudo-sites as a random effect and further quantifying the heterogeneity in terms of Higgin & Thompson's I^2 index for each arm, where higher I^2 values indicate higher levels of heterogeneity. Poolability analysis will only be performed in the ITT population.

6. Data Handling

Prior to database lock, data extracts may be provided to an unblinded member of the Biostatistics and Programming team. The unblinded team member is responsible for masked treatment assignments prior to handing the data over to the blinded Biostatistics and Programming team. Details of the masking process are provided in the Data Masking Plan.

7. References

1. Klok FA, Hösel V, Clemens A, Yollo WD, Tilke C, Schulman S, et al. Prediction of bleeding events in patients with venous thromboembolism on stable anticoagulation treatment. Eur Respir J. 2016;48(5):1369-76.
2. Konstantinides SV, Meyer G, Becattini C, Bueno H, Geersing GJ, Harjola VP, et al. 2019 ESC Guidelines for the diagnosis and management of acute pulmonary embolism developed in collaboration with the European Respiratory Society (ERS). Eur Heart J. 2019;41(4):543-603.
3. Schulman S, Kearon C. Definition of major bleeding in clinical investigations of antihemostatic medicinal products in non-surgical patients. J Thromb Haemost. 2005;3(4):692-4.
4. Finkelstein DM, Schoenfeld DA. Combining mortality and longitudinal measures in clinical trials. Stat Med. 1999;18(11):1341-54.
5. Pocock SJ, Ariti CA, Collier TJ, Wang D. The win ratio: a new approach to the analysis of composite endpoints in clinical trials based on clinical priorities. Eur Heart J. 2012;33(2):176-82.19.
6. Gao X, Starmer J, Martin ER. A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. Genet Epidemiol. 2008 May;32(4):361-9. doi: 10.1002/gepi.20310. PMID: 18271029.
7. Localio AR, Berlin JA, Ten Have TR, Kimmel SE. Adjustments for center in multicenter studies: an overview. Ann Intern Med. 2001 Jul 17;135(2):112-23. doi: 10.7326/0003-4819-135-2-200107170-00012. PMID: 11453711.
8. Gallo PP. Center-weighting issues in multicenter clinical trials. J Biopharm Stat. 2000;10:2, 145-163, DOI: 10.1081/BIP-100101019.