E7058

A Multi-Center, Prospective, Randomized Study Comparing Removable, Self-Expanding Metal Stents to Plastic Stents for the Treatment of Benign Biliary Strictures Secondary to Chronic Pancreatitis

WallFlex Biliary FC Chronic Pancreatitis RCT

NCT01543256

Statistical Analysis Plan

March 20, 2013

Statistical Analysis Plan

90913894 Version AA

A Multi-Center, Prospective, Randomized Study Comparing Removable, Self-Expanding Metal Stents to Plastic Stents for the Treatment of Benign Biliary Strictures Secondary to Chronic Pancreatitis

WallFlex Biliary FC Chronic Pancreatitis RCT

E7058

CONFIDENTIAL

DO NOT COPY OR DISTRIBUTE WITHOUT WRITTEN PERMISSION

Revision History

Version AB - Updated for Protocol Changes

Original Release: March 27, 2012

Revision History

Revision Number/Release Date	Section	Change	Reason for Change
AB	Secondary Endpoints, I/E Criteria	Added in new definition of tech success, I/E criteria to allow for patients with prior stenting, changed definition of stent removal technical success, and allowed for pre-baseline LFTs.	Updated to match new version of protocol

TABLE OF CONTENTS

1 PROTOCOL SUMMARY

Full Title	A Multi-Center, Prospective, Randomized Study Comparing Removable, Self-Expanding Metal Stents to Plastic Stents for the Treatment of Benign Biliary Strictures Secondary to Chronic Pancreatitis
Short Title	WallFlex Biliary FC Chronic Pancreatitis RCT
Primary Objective	To compare the use of Self Expanding Metal Stents (SEMS) to plastic stents for the treatment of benign biliary strictures secondary to chronic pancreatitis as it pertains to stricture resolution rates, complication rates and number of ERCP procedures during 24 months.
Study Devices	**Group A: Metal Stents – MS Arm** WallFlex™ Biliary RX Fully Covered Stent System RMV **Group B. Plastic Stents – PS Arm** Per Investigator preference
Study Design	Prospective, multi-center, randomized
Planned Number of Subjects	164
Planned Number of Sites	Up to 15
Primary Endpoint	Stricture resolution at 24 months
Secondary Endpoints	1. Occurrence of adverse events related to the stent and/or the stent placement or removal procedures 2. Number of ERCP procedures through 24 months after initial stent placement 3. Ability to deploy the stent(s) in satisfactory position 4. Stent Removal: • Ability to remove the stent(s) without serious stent removal related adverse events at each procedure involving removal of stent(s) (technical success at removal) or • Complete distal migration without serious stent removal related adverse events 5. Liver Function Tests (LFT's):

	a. LFT improvement at month 1 compared to baseline naïve stricture or to LFTs at time of plastic stent placement 6 months or fewer prior to enrollment in strictures treated with 1 prior plastic stent. b. LFTs at month 24 compared to removal of last stent (applicable for subjects who had not been re-stented at time of month 24 visit) 6. Health Economic Endpoints: 7. Number of outpatient procedures 8. Number of hospitalizations 9. Duration of hospitalizations 10. Length of procedures o Number of devices
Randomization	Subjects will be randomized at the time of the procedure to a 1:1 ratio between Metal Stent Arm (Group A – MS) and Plastic Stent Arm (Group B – PS).
Follow-Up Schedule	• Baseline: Subject screening, enrollment, LFT's and symptoms. • Study Treatment Procedure: o Group A (MS): Stent Placement o Group B (PS): Stent Placement: Two or more 8.5 Fr. or 10 Fr. PS whenever possible • 1 Month Follow-up: o Group A (MS): LFTs and symptoms o Group B (PS): LFTs and symptoms • Stent Exchange Follow-up: o Group A (MS): None o Group B (PS): Month 4 and Month 8 • Stent Removal: o Group A (MS): Removal at Month 12, LFTs and symptoms o Group B (PS): Removal of last stents at Month 12, LFTs and symptoms • Post-Stent Removal Follow-Up o Group A (MS): Month 24 – LFTs and symptoms o Group B (PS): Month 24 – LFTs and symptoms • Additional ERCP visits as needed Note: Recurrent strictures will be treated with a metal stent in Group A (MS) and with plastic stents in Group B (PS) – no cross-over. Re-stenting after the per-protocol 12 month stenting period will be considered primary endpoint failures. Follow-up, however, will

	continue until Month 24 after initial stent placement for all subjects, in order to assess all secondary endpoints in a comparative fashion for the MS and PS arms.
Key Inclusion Criteria	• Age 18 or older • Willing and able to comply with the study procedures and provide written informed consent to participate in the study • Chronic pancreatitis • Stricture not previously dilated or stricture previously dilated with no more than one plastic stent of 10 Fr or less for 6 months or fewer • Symptomatic bile duct stricture (defined by cholangitis or persistent jaundice for at least one month or cholestasis associated with at least 3 times normal alkaline phosphatase levels) documented at time of enrollment for naïve stricture or at the time of one prior plastic stent placement up to 6 months before enrollment in strictures that had one prior plastic stent inserted.[12] • Common bile duct stricture based on imaging assessment of dilatation of the common and/or intrahepatic bile ducts
Key Exclusion Criteria	• Biliary stricture of benign etiology other than chronic pancreatitis • Any prior biliary metal stent or any plastic stenting other than one plastic stent of 10 Fr or less for 6 months or fewer • Developing obstructive biliary symptoms associated with an attack of acute pancreatitis • Biliary stricture of malignant etiology • Stricture within 2 cm of common bile duct bifurcation • Known bile duct fistula or leak • Subjects for whom endoscopic techniques are contraindicated • Known sensitivity to any components of the stent or delivery system • Symptomatic duodenal stenosis (with gastric stasis) • Participation in another investigational study within 90 days prior to consent • Investigator Discretion
Statistical Hypothesis	A literature search of metal and plastic stenting for treatment of benign biliary strictures secondary to chronic pancreatitis yielded 4 articles representing 70 subjects treated with metal stenting (MS) [1,9-11] and 3 articles

	representing 60 subjects treated with plastic stenting (PS) [3,4,8]. The following meta-analysis was conducted of the probability of stricture resolution: • Metal Stenting: A meta-analysis of the stricture resolution rate during the reported follow-up after initial stent placement yields a proportion of 0.762 [95% CI: 0.593 – 0.895] [1,9-11]. • Plastic Stenting: A meta-analysis of the stricture resolution rate during the reported follow-up after initial stent placement yields a proportion of 0.611 [95% CI: 0.311 – 0.870] [3,4,8]. Statistical testing will be performed to determine if the rate of stricture resolution for the metal stent is non-inferior to the plastic stent group. The null hypothesis is that the stricture resolution rate is non-inferior in the Metal Stent Arm versus the Plastic Stent Arm: $Ho: \pi_{test} - \pi_{control} \geq \Delta$ (Inferior) $Ha: \pi_{test} - \pi_{control} < \Delta$ (Non-inferior) where π_{test} and $\pi_{control}$ are the probabilities of having a stricture resolution in the metal stent arm and the plastic stent arm respectively, and Δ is defined as the non-inferiority margin. The sample size was calculated for a one-sided 0.050 Farrington-Manning test using SAS 9.2®. If the P value from the Farrington-Manning test is <0.05 then the metal stent group will be considered non-inferior to the plastic stent group. The expected probability of stricture resolution in the metal stent arm and plastic stent arm is 66.0%, which was taken from the 95% CIs from the meta-analysis above. The non-inferiority margin (Δ) is 20%. Given these assumptions and a one-sided 5% significance level, 2 x 74 = 148 subjects will provide 80% power to reject the null hypothesis, that the metal stent group is inferior to the plastic stent group. To compensate for possible loss of subjects after enrollment and complete assessment of inclusion/exclusion criteria, an additional 10% of subjects will be enrolled, for a total of 2 x 82 = 164.

2 INTRODUCTION

This statistical plan addresses the planned analyses for the WallFlex Biliary FC Chronic Pancreatitis RCT based on the protocol dated 17 January 2012, Version AB. All of the specified analyses may not be provided in reports to Competent Authorities but may be used for scientific presentations and/or manuscripts. The primary analysis will be based on the data through 24 months post-procedure.

3 ENDPOINT ANALYSIS

3.1 Primary Endpoint

The rate of stricture resolution is the primary endpoint for the study.

3.1.1 Hypotheses

Statistical testing will be performed to determine if the rate of stricture resolution for the metal stent is non-inferior to the plastic stent group. The null hypothesis is that the rate of stricture resolution is non-inferior in the Metal Stent Arm versus the Plastic Stent Arm:

$Ho: \pi_{test} - \pi_{control} \geq \Delta$ (Inferior)

$Ha: \pi_{test} - \pi_{control} < \Delta$ (Non-inferior)

where π_{test} and $\pi_{control}$ are the probabilities of having a stricture resolution in the metal stent arm and the plastic stent arm respectively, and Δ is defined as the non-inferiority margin.

3.1.2 Sample Size

A literature search of metal and or plastic stenting for treatment of benign biliary strictures secondary to chronic pancreatitis yielded 4 articles representing 70 subjects treated with metal stenting (MS) [2-7] and 3 articles representing 60 subjects treated with plastic stenting (PS) [8-13].

The following meta-analysis was conducted of the probability of stricture resolution:

- Metal Stenting: A meta-analysis of the stricture resolution rate during the reported follow-up after initial stent placement yields a proportion of 0.762 [95% CI: 0.593 – 0.895] [2-7].
- Plastic Stenting: A meta-analysis of the stricture resolution rate during the reported follow-up after initial stent placement yields a proportion of 0.611 [95% CI: 0.311 – 0.870] [8-13].

The sample size was calculated for a one-sided 0.050 Farrington-Manning test using SAS 9.2®. If the P value from the Farrington-Manning test is <0.05 then the metal stent group will be considered non-inferior to the plastic stent group. The expected probability of a stricture resolution in the metal stent arm and plastic stent arm is 66.0%, which was taken from the 95% CIs from the meta-analysis above. The non-inferiority margin (Δ) is 20%. Given these assumptions, and a one-sided 5% significance level, 2 x 74 = 148 subjects will provide 80% power to reject the null hypothesis, that the metal stent group is inferior to the plastic stent group. To compensate for possible loss of subjects between enrollment and complete assessment of inclusion/exclusion criteria and subsequent randomization, an additional 10% of subjects will be enrolled for a total of 2 x 82 = 164.

3.1.3 Statistical Methods

Eligibility of Subjects, Exclusions, and Missing Data

The distribution of prognostic factors between patients with and without data will be examined. Statistical models that account for censored data will be employed in appropriate circumstances, e.g. for time-to-event outcomes.

For calculating the rate of stricture resolution at 24 months after initial stent placement, patients with insufficient follow-up will be excluded from the calculation; that is, only patients that have at least follow-up through 24 month visit will be included. Sensitivity analyses, such as a tipping point analysis will be performed to assess the impact of the missing data, if the data support such analysis.

When calculating rates of adverse events, missing and partial dates will be handled as follows:

Partial Date Description	Action Taken
Entire onset date is missing	The procedure date will be used for the onset date.
The month and the day of the month are missing but the year is available	January 1 will be used for the month and day of the onset date. However, if the imputed date falls before the procedure date, then the procedure date will be used for the onset date.
Day is missing, but the month and year are available	The 1st will be used as the day of the onset date. However, if the imputed date falls before the procedure date, then the procedure date will be used for the onset date.

Baseline Comparability

Baseline data will be analyzed to assess the comparability of the treatment groups. Patient demographics, clinical history, risk factors, and procedure characteristics will be summarized using descriptive statistics (mean, standard deviation, n, minimum, maximum) for continuous variables and frequency statistics for discrete variables. Statistical testing will be performed using appropriate methods.

Post-Procedure Endpoints

The treatment effect due to important prognostic factors at the baseline level for both groups will be assessed using the logistic regression model for categorical outcomes, analysis of covariance model for continuous outcomes, or stratified methods, as appropriate.

Overall safety and effectiveness will be reported based on patients with sufficient follow-up by Metal Stent Arm and Plastic Stent Arm using the logistic regression or analysis of covariance model adjusted for baseline covariates. The Kaplan-Meier product-limit method will also be used to estimate event-free rates for time-to-event outcomes and compared using log-rank and Wilcoxon tests. Additional analyses will be performed as appropriate.

4 GENERAL STATISTICAL METHODS

4.1 Analysis Sets

Primary endpoint and selected secondary endpoints will be done for the following cohorts.

Enrolled Cohort

A subject is considered "enrolled" after signing the study-specific ICF. Subjects who sign the ICF but subsequently do not meet one or more of the selection criteria will be considered screen failures and excluded from the study.

Intent-to-Treat Cohort (ITT)

This cohort consists of those "enrolled" subjects who meet all inclusion/exclusion criteria and are subsequently randomized. Subjects in this cohort who do not receive a study stent(s) will be counted towards the enrollment ceiling. Any adverse events occurring or resulting from a treatment attempt will be collected. Protocol deviations will be collected as necessary.

Per-Protocol Cohort (PP)

The per-protocol cohort is a subset of the ITT subjects who receive the study stent(s) as randomized and no major protocol deviations (Per ICH E9).

5 ADDITIONAL DATA ANALYSES

5.1 Secondary Endpoints

1. Occurrence of adverse events related to the stent and/or the stent placement or removal procedures
2. Number of ERCP procedures through 24 months after initial stent placement
3. Ability to deploy the stent(s) in satisfactory position
4. Stent Removal:

a. Ability to remove the stent(s) without serious stent removal related adverse events at each procedure involving removal of stent(s) (technical success at removal) or

b. Complete distal migration without serious stent removal related adverse events

5. Liver Function Tests (LFT's)

a. LFT improvement at month 1 compared to baseline naïve stricture or to LFTs at time of plastic stent placement 6 months or fewer prior to enrollment in strictures treated with 1 prior plastic stent.

b. LFTs at month 24 compared to removal of last stent (applicable for subjects who had not been re-stented at time of month 24 visit)

6. Health Economic Endpoints:

- Number of outpatient procedures
- Number of hospitalizations
- Duration of hospitalizations
- Length of procedures

The secondary endpoints will be summarized using descriptive statistics (mean, standard deviation, n, minimum, maximum) for continuous variables and frequency statistics for discrete variables. Statistical testing will be performed using appropriate methods. For analysis of LFT's, McNemar's Test or repeated measures regression model may be used.

5.2 Interim Analysis

Informal Interim Analysis once the first 50 subjects have reached 24 months post initial stent placement. This analysis will consist of the primary and some or all of the secondary endpoints and demographics if deemed necessary. This analysis will consist of only descriptive statistics (mean, standard deviation, n, minimum, maximum) for continuous variables and frequency statistics for discrete variables. There will be no testing done for this interim analysis.

5.3 Subgroup Analyses

The subgroup analyses will include analyzing the primary endpoint and select secondary endpoints by gender.

5.4 Justification of Pooling

The analyses will be presented using pooled data across institutions. An analysis of the poolability will be made using logistic regression for binary outcomes, proportional hazards regression for time-to-event outcomes, or analysis of variance for continuous outcomes, to assess differences between study institutions and to justify pooling data across institutions.

5.5 Multivariable Analysis

Univariate and multivariate analyses will be performed to assess possible predictors of the stricture resolution, removability of stents, and distal migration. Possible predictors may include any but not limited to demographic/baseline data and medical history data. Factors from the univariate model with $p \leq 0.20$ will also be modeled multivariately using a stepwise procedure in a logistic regression model or Cox proportional hazards regression model. The significance thresholds for entry and exit into the model will be set to $p < 0.10$.

5.6 Changes to Planned Analyses

Any changes to the planned statistical analyses made prior to performing the analyses will be documented in a Statistical Analysis Plan approved prior to performing the analyses.

6 VALIDATION

All clinical data reports generated per this plan will be validated per Global WI: Clinical Data Reporting Validation.

7 PROGRAMMING CONSIDERATIONS

7.1 Statistical Software

All statistical analyses will be done using The SAS System software, version 8 or higher (Copyright © 2000 SAS Institute Inc., SAS Campus Drive, Cary, North Carolina 27513, USA. All rights reserved). Other statistical software (e.g. StatXact) may be used if necessary.

7.2 Format of output

Results of analysis will be output programmatically to Word documents from SAS with no manual intervention. All output for the final statistical report will be in the form of a Word document containing tables, figures, graphs, and listings, as appropriate.

7.3 Rules and Definitions

Binary event rates (proportions) will be reported on a per patient basis.

The last follow-up date will be the latest of the following dates for each patient: date of an adverse event, index procedure date, follow-up visit date, any stent or ECRP procedure date, stent removal date, and device event date.

Stricture resolution at 24 months is defined by the following two criteria being met:

- Absence of re-stenting after the per-protocol stenting period through the 24 month visit

- Absence of cholestasis at the 24 month visit, defined as alkaline phosphatase level not exceeding 2 times the level at completion of the per-protocol stenting period

Primary endpoint failures are 1) subjects who are re-stented during follow-up after the per-protocol stenting period and 2) subjects who have not been re-stented at month 24 but have alkaline phosphatase level exceeding 2 times the level at completion of the per-protocol stenting period.

Subjects who experience early stent removal or complete distal stent migration without subsequent re-stenting will not be considered failures. If re-stenting occurs, but the cumulative stenting period does not exceed 12 months, the subject remains eligible for primary endpoint assessment.

For the final tables, two-sided exact tests will be done to test for differences between the two groups for removability and AEs. A one-sided non-inferiority Farrington-Manning test will be done to see if stricture resolution is non-inferior in the Metal group compared to the Plastic group. A two-sided student's t-test will done to test for differences between the two groups for the number of ERCP procedures done.