Protocol Number: XG2114

Statistical Analysis Plan

A Clinical Study to Evaluate the Safety and Efficacy of Biologic Hollow Bone Screws

Name of Experimental Medical Device: Biologic Hollow Bone Screws

Clinical Trial Institution: The First Affiliated Hospital of PLA Air Force Medical University

The Sponsor: Jiangxi Sike Biotechnology, China

Collaborative Research Organization: Nanjing Sigma Medical Technology Co., LTD

NCT Number:

Document Date: January 21, 2024

Statistical Analysis

1. Statistical Design

A prospective, multicenter, randomized, single-blind, parallel-controlled clinical trial was conducted. The comparison type was non-inferiority design.

2. Analytical Methods

95% confidence interval method was used for the main efficacy indicators, and bilateral tests were used for all statistical tests, and $P \leq 0.05$ was considered to be statistically significant.

The description of quantitative indicators will calculate the mean, standard deviation, median, minimum, maximum. The classification index is described by the number of cases and percentage of each type. Statistical analysis was performed using SAS 9.4 software.

3. Calculation of Sample Size

A. Total sample size

With reference to the Guiding Principles for the Technical Review of Product Registration of Metal intramedullary nail System and the Guiding Principles for the Technical Review of Product Registration of Metal bone plate internal fixation System, the clinical healing rate of fractures at 24 weeks after surgery was used as the main therapeutic evaluation index to calculate the sample size.

According to the references, it is assumed that the clinical healing

rate of fracture PC at 24 weeks after operation of the control product is about 95.5%. The clinical fracture healing rate PT of the test product is expected to be comparable to that of the control product, according to the non-inferiority sample size formula:

$$n_T = n_C = \frac{\left(Z_{1-\alpha/2}+Z_{1-\beta}\right)^2\left[P_C\left(1-P_C\right)+P_T\left(1-P_T\right)\right]}{\left(|D|-\Delta\right)^2}$$

P_T and P_C were the expected event rates of the experimental group and the control group, respectively. |D| is the absolute value of the difference between the expected rates of the two groups, $|D|=| P_T - P_C |$. △ is the non-inferiority margin value, which is -10%. The type I error was a=0.05 (two-sided) and the type II error was B=0.20.There were 68 patients in each group, and a total of 136 patients in the two groups.Considering the dropout rate and block length, 160 cases were included in the experimental group and the control group, and the number of subjects in the experimental group and the control group was 1:1.

B. Number of clinical trials for each disease and reasons for their determination

According to the scope of application of the test instruments, subjects who meet the inclusion criteria and fail to meet the exclusion criteria will be strictly screened in this study. According to the statistical distribution, subjects with different diseases will be distributed in the experimental group and the control group, and the specific diseases will

no longer be restricted.

C. Minimum and maximum number of subjects and reasons for each clinical trial facility in multicenter clinical trials

This trial is expected to be conducted simultaneously in 6 clinical trial institutions, and in principle, the number of enrolled centers will be evenly distributed as far as possible to ensure adequate center representation. However, considering the feasibility and enrollment progress, the number of centers and the number of subjects in the centers will be adjusted according to the actual situation, so as to ensure that the enrollment scale of each center is relatively balanced, and the final enrollment scale of a particular center should not exceed 50% of the total number of cases.

3. The Significance Level and Degree of Assurance of Clinical Trials

A. Significance level: 0.05

B. Degree of confidence: 0.8

4. Expected Drop-out Rate

The expected shedding rate of this clinical trial is ≦15%.

5. Qualified/Unqualified Criteria for Clinical Trial Results

A. Eligibility Criteria: If the clinical healing rate of the experimental product is worse than that of the control product at 24 weeks after surgery (test group-control group) by 95%CI lower limit > -10%, then the experimental product is not inferior to the control product, and the

clinical verification can be considered successful.

B. Unqualified Criteria: If the clinical healing rate of fracture is different between the test product and the control product 24 weeks after surgery (experimental group-control group)95%CI lower limit ≦-10%, the clinical validation can be considered to have failed.

6. Criteria and Reasons for Terminating the Trial on Statistics

Not applicable

7. Statistical Methods for All Data, Along With Treatment of Missing, Unused, or Erroneous Data (Including Dropouts and Withdrawals) and Unreasonable Data.

A. Characteristics of the cases

Enrollment and completion: The number of enrolled and completed cases was summarized, and the list of dropped cases was listed.

General information and baseline characteristics: Demographic information and other disease history of the subjects were statistically described.

According to the numerical characteristics of the variables, T-test/corrected T-test was used to compare the age, blood pressure and other quantitative data of the two groups of subjects. Chi-square test/exact probability method was used to compare the qualitative variables such as gender and past medical history of the subjects between the groups. The Wilcoxon rank sum test was used to compare the ordered

classified data between groups.

Baseline: Baseline is defined as the basic situation of the subject before clinical operation.

B. Analysis of effectiveness evaluation indicators

Analysis of main evaluation indicators: The main evaluation index was the clinical healing rate of fracture at 24 weeks after surgery, and the difference of 95% confidence interval between the experimental group and the control group was calculated.

Analysis of secondary evaluation indicators: clinical healing rate of fractures and performance indicators of pain scoring instruments at 12, 48 and 72 weeks after surgery were statistically described by descriptive methods of qualitative or quantitative data according to the characteristics of variable types, and comparison between groups was performed by Chi-square test/exact probability /t test.

C. Analysis of security indicators

The occurrence of adverse events was clinically observed, event name, start time, end time, severity, corrective treatment, and relationship with the experimental product were recorded, and the incidence of adverse events was calculated and compared between groups. Study discontinuation due to adverse events and the occurrence of serious adverse events should be specifically noted.

For various safety indicators such as vital signs and laboratory tests,

T-test/corrected T-test /Wilcoxon rank sum test were used for comparison between groups according to the numerical characteristics of variables; Chi-square test/exact probability method was used for comparison between groups for qualitative data; Wilcoxon rank sum test or CMH chi square test was used for ordered classified data.

8. Procedures for Deviation from the Original Statistical Plan

The implementation shall be strictly in accordance with the requirements stipulated in **7.** If other statistical methods not specified in the program are needed, an application shall be submitted to the Ethics committee for approval before implementation.

9. Selection Criteria and Reasons for Subjects Included in the Analysis

Full Analysis Set (FAS) is a set of subjects that are as close as possible to the ideal in line with the principles of intentionality. The data set is derived from the elimination of subjects from all enrolled subjects by a minimum and reasonable method.

Per Protocol Set (PPS), Also known as a valid case, valid sample, or evaluable case sample. It is a dataset generated from a subset of cases that are fully compliant with the trial protocol and is a subset of the total analysis set. Adherence includes considerations such as the treatment received, the feasibility of the main measure, and the absence of significant violations of the protocol.

Safety Analysis Set (SS): When performing a safety evaluation, the set of subjects used for aggregation is called a safety analysis set. The safety data set shall include all subjects who have used this tester.