A Preliminary Analysis of Sample Size for Medical Device Registration Clinical Trials in China

Authors: Zeng Zhiyu, Han Lei, Zhang Xiaoxing, Peng Lin, Zeng Li, Zeng Zhiyu

Date: 2022-11-21T00:00:00+00:00

Abstract

Determination of clinical trial sample size is one of the core components of trial design. Based on the publicly disclosed medical device evaluation reports from the Center for Medical Device Evaluation (CMDE) of the National Medical Products Administration (NMPA) in recent years and related publicly available data, this study conducts a brief analysis of the current status of sample sizes for medical device registration clinical trials in China, covering aspects of study design, with comparisons made to the United States for select products. The analysis reveals that the median sample size for Class III medical device registration clinical trials in China is 120 (interquartile range 90~167.5), with sample sizes being significantly influenced by regulatory policies, and substantial differences observed in clinical trial sample sizes for some products compared to the United States. Public disclosure of evaluation reports represents a major advancement in medical device regulation in China; however, the content of publicly disclosed evaluation reports still requires further improvement.

Full Text

A Brief Analysis of Sample Sizes for Medical Device Registration Clinical Trials in China

Authors: Zeng Zhiyu, Han Lei, Zhang Xiaoxing, Peng Lin, Zeng Li
Affiliation: Boston Scientific International Medical Trading (Shanghai) Co., Ltd.
Contact: zhiyu.zeng@bsci.com

Abstract

Sample size determination for clinical trials is one of the key components of study design. Based on medical device registration review reports recently pub-

lished by the National Medical Products Administration Center for Medical Device Evaluation and other publicly available information, we conducted an analysis of sample sizes for medical device registration clinical trials in China, including study design considerations, with some products compared to those in the United States. Our results showed that the median sample size for Class III medical device registration trials is 120 (IQR 90~167.5). Sample size was significantly influenced by regulatory policies, and some differed markedly from those in the US. The public disclosure of registration review reports represents a major advancement for medical device regulation in China; however, the disclosed information needs to be further improved.

Keywords: Medical device, Registration, Clinical Trial, Sample Size

Introduction

With China' s economic development and increasing public demand for healthcare, China' s pharmaceutical industry, particularly the medical device sector, has experienced robust growth, with domestic medical devices showing explosive growth. How to expedite the registration and market launch of high-quality medical devices for clinical application to benefit patients while ensuring their safety and effectiveness has been a key focus for stakeholders including regulatory authorities, manufacturers, clinicians, and patients.

The 2021 revised *Regulation on the Supervision and Administration of Medical Devices* explicitly stipulates that clinical evaluation shall be conducted for medical device registration and filing, and that such evaluation may be performed through clinical trials or analysis of clinical literature and data from predicate devices to demonstrate safety and effectiveness. According to regulations from the State Council' s drug administration department, clinical trials must be conducted when existing clinical literature and data are insufficient to confirm product safety and effectiveness.

Therefore, clinical trials constitute an important basis for supporting medical device registration, particularly for innovative devices where existing clinical data or experience remains limited. Sample size determination is a critical aspect of clinical trials that often directly affects trial success. This paper reviews medical device evaluation reports publicly released by the National Medical Products Administration Center for Medical Device Evaluation (hereinafter referred to as the "Center") to provide a brief analysis of current sample sizes for medical device registration clinical trials in China. The views expressed herein are personal opinions of the authors and do not represent the position of their employer.

Current Status of Sample Sizes for Medical Device Registration Clinical Trials in China

The Center began publicly releasing evaluation reports for approved medical devices on its official website in November 2017. By the end of July 2021, a

total of 115 records had been published, including 33 for in vitro diagnostic reagents and corresponding equipment, and 15 for imported medical devices. After excluding these categories (with some overlap between the two groups), 72 records remained. Among these 72 records, one was clinical trial-exempt, and six involved predicate device comparisons, leaving 65 reports for analysis in this study, as detailed in and [Figure 1: see original paper].

Among the 65 evaluation reports analyzed, all medical devices were Class III. Twenty-one reports (32.3%) did not provide sample size information. For the 45 reports that included sample sizes, the median was 120 (IQR 90~167.5), with a minimum of 48 and a maximum of 1,663. As shown in [Figure 1: see original paper], the sample size distribution was markedly right-skewed. Excluding three trials with $ $1,000 participants, the remaining 41 trials had sample sizes between 48 and 274, with 13 trials having $ $100 participants, 24 trials having 101–200 participants, and 4 trials having 201–300 participants. Sample sizes were relatively concentrated, with limited variation across different product categories.

Among the 65 evaluation reports, 25 were randomized controlled trials (RCTs; 10 superiority and 15 non-inferiority), 20 employed single-arm target value designs, and 20 were self-control studies (primarily imaging diagnostic devices; 3 explicitly non-inferiority, 1 target value, and the remaining 16 consistency comparisons. Some reports without specified comparison types were also categorized as consistency comparisons).

Evaluation reports rarely included sample size calculation explanations. For instance, among the 12 reports published in 2021, only one provided sample size calculation parameters. Some reports also omitted clinical trial results, with five of the 12 reports published in 2021 failing to provide trial outcomes.

Another notable phenomenon is that most evaluation reports mentioned the concept of Full Analysis Set (FAS), regardless of whether they were RCTs or single-arm studies. For example, six of the 12 reports from 2021 mentioned FAS. The ICH (International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use) E6 guideline from 1998 provides detailed descriptions of FAS, primarily for use in RCTs. When intention-to-treat (ITT) analysis is difficult to implement (e.g., due to major protocol violations or lack of post-randomization data), some subjects may be excluded with justification to form the FAS. Therefore, the FAS is not necessarily "full," which can cause misunderstanding. ICH guidelines primarily target drug clinical trials. Current authoritative literature on clinical trials, including medical device trials [1,2], published papers, and US FDA Summaries of Safety and Effectiveness Data (SSEDs) predominantly use ITT and mITT (modified ITT). Conceptually, FAS is more consistent with mITT.

Impact of Regulatory Policies on Clinical Trial Sample Sizes

Numerous factors influence clinical trial sample sizes, with regulatory policies being an important consideration alongside product characteristics and clinical factors. Several guidance documents issued by Chinese authorities in recent years have provided specific sample size recommendations.

The Center' s 2016 *Technical Review Guidance for Imaging Ultrasound Diagnostic Equipment (Class III)* [3] states: Based on clinical experience, the image consistency rate between the test probe and control probe should be at least 85%, establishing 85% as the target value for final evaluation. If the expected overall image consistency rate between test and control probes is assumed to be 95%, then 80 subjects are required at a significance level of 5% (two-sided) and power of 80%. Assuming a control intracavitary probe image excellence rate of 96% and a non-inferiority margin of 10% (10% of the control excellence rate), each group (test and control) requires 61 subjects, totaling 122 cases. Sample sizes for imaging products in this study (not limited to ultrasound) were largely consistent with these calculations.

Two of the three products with sample sizes exceeding 1,000 in this study were bioresorbable coronary drug-eluting stents. The Center' s 2018 *Guidance for Clinical Trials of Coronary Drug-Eluting Stents* [4] recommends that confirmatory trials consist of two studies: one RCT and one single-arm target value trial. The RCT should have no fewer than 200 pairs, and the single-arm target value trial should have no fewer than 800 subjects (some cases may be derived from the test group of the RCT). The total sample size should be no fewer than 1,000 subjects based on statistical considerations. The 2019 *Guidance for Clinical Trials of Bioresorbable Coronary Drug-Eluting Stents* [5] similarly stipulates: RCT sample size no fewer than 200 pairs, single-arm trial sample size no fewer than 1,000 subjects (some cases may be derived from the RCT test group), with total sample size no fewer than 1,200 subjects based on statistical considerations. The sample sizes for the two bioresorbable coronary drug-eluting stents in this study were largely consistent with these requirements.

The Center's 2019 *Guidance for Clinical Trials of Transcatheter Implantable Aortic Valve Prostheses* [6] states that the primary endpoint should be a clinically meaningful evaluation indicator, currently recommending 12-month cumulative all-cause mortality as the primary endpoint for sample size estimation. This guidance does not recommend specific sample sizes. The three transcatheter aortic valve trials in this study all employed single-arm target value designs with sample sizes of 81, 110, and 120.

Comparison of Sample Sizes Between China and the US for Selected Medical Devices

The US FDA (Food and Drug Administration) is generally considered to have stringent review requirements for Class III medical devices, although the proportion of Class III devices in the US is far smaller than in China. Transcatheter

aortic valves and left atrial appendage closure devices have been particularly popular in recent years. We briefly compare sample sizes for Class III device registration trials between China and the US using these two products as examples.

As mentioned above, based on evaluation report data, sample sizes for the three approved domestic transcatheter aortic valves in China were 81, 110, and 120. Two additional approved domestic transcatheter aortic valves without publicly available evaluation reports had sample sizes of 1,007 and 1,018, also using single-arm target value designs. In contrast, Edwards Lifesciences' transcatheter aortic valve trials in the US for inoperable/surgical high-risk, intermediate-risk, and low-risk indications had sample sizes of 3,589/6,9910, 2,03211, and 1,00012, respectively. Similarly, Medtronic's trials for high-risk, intermediate-risk, and low-risk indications had sample sizes of 79513, 1,74614, and 1,46815, respectively. Boston Scientific is currently conducting a registry trial covering all surgical risk levels with a sample size of 1,67016. All US transcatheter aortic valve trials listed here were RCTs. Thus, sample sizes differ substantially between China and the US for transcatheter aortic valve registration trials. Additionally, the Center' s guidance for transcatheter aortic valves states: For devices already marketed overseas that have completed well-designed prospective clinical trials meeting Chinese requirements, in addition to submitting overseas clinical data as required, prospective clinical trial data from China must also be submitted; for patients with symptomatic severe aortic stenosis who are unsuitable for conventional surgery or at high surgical risk, separate clinical trials should be conducted with no fewer than 50 subjects each. This guidance appears to impose relatively high requirements for imported products. Edwards Lifesciences' SAPIEN 3 valve has been approved in China, with the public evaluation report mentioning a domestic trial sample size of 58, while publicly available information indicates Edwards conducted another SAPIEN XT trial in China17 with 67 subjects, which was not reflected in the evaluation report.

Based on evaluation report data, sample sizes for approved domestic left atrial appendage closure devices in China were 71 and 175. Two additional approved domestic devices without publicly available evaluation reports had sample sizes of 15418 and 15819. Currently, only Boston Scientific' s WATCHMAN left atrial appendage closure device is marketed in the US. The FDA required two trials for this product: PROTECT AF20 and PREVAIL21, with sample sizes of 707 and 407, respectively, totaling 1,114. Abbott and Johnson & Johnson are currently conducting registration trials for their left atrial appendage closure devices in the US, both comparing against WATCHMAN, with sample sizes of 1,87822 and 1,25023, respectively. Sample sizes thus also differ substantially between China and the US for left atrial appendage closure device registration trials. Similarly, the Chinese domestic trials mentioned above all employed single-arm target value designs (although some current domestic trials have adopted RCT designs), while US registration trials were all RCTs.

Discussion

This paper provides a brief analysis of sample sizes for medical device registration clinical trials in China based on evaluation reports and other publicly available information, including study design considerations and simple comparisons with US trials for selected products.

The public disclosure of evaluation reports represents tremendous progress in China' s medical device regulation, demonstrating a move toward transparency in administrative affairs. However, current evaluation reports still have issues, such as limited disclosed content that remains incomplete, with some reports even omitting critical information like sample sizes and study results.

Overall, sample sizes for medical device registration clinical trials in China are relatively small, with more than half between 101–200 participants and a median of 120—approximately one-tenth of corresponding US trial sample sizes. Nevertheless, domestic clinical trials are often required for medical devices already marketed overseas, as recommended in the *Guidance for Clinical Trials of Transcatheter Implantable Aortic Valve Prostheses*. This situation may improve as regulators increasingly recognize overseas data. In April 2021, the Center' s review forum published an article stating: "When conducting clinical evaluation through clinical trials, if guidance requires submission of overseas clinical trial data plus domestic trials, must applicants conduct domestic trials?" The article clarified that Article 5 of the *Guidance for Acceptance of Overseas Clinical Trial Data for Medical Devices* states: "If specific technical review guidance contains relevant clinical trial requirements, overseas clinical trials should consider such requirements, and any inconsistencies should be accompanied by adequate and reasonable justification." Therefore, if applicants have submitted overseas clinical trial data that complies with ethical, legal, and scientific principles and adequately addresses differences in technical review requirements, subject populations, and clinical trial conditions, additional domestic trials may not be necessary. We welcome this interpretation and believe that, considering the fundamental mechanistic differences between medical devices and drugs, subject population difference analyses should not be overly stringent.

This study has several limitations. First, constrained by the content of publicly available materials, our analysis of sample sizes lacks depth, with no evaluation of sample size calculations or their appropriateness, nor discussion of endpoint selection and result analysis, despite their importance. Additionally, because sample size distributions across different trials were relatively concentrated, we did not conduct further comparisons across different device categories. Furthermore, some data were inferred by the authors from context due to lack of explicit notation in evaluation reports.

Clinical trials require appropriate sample sizes—neither excessive nor insufficient. Current sample sizes for medical device registration clinical trials in China have positively encouraged innovation in medical device research and development, though their true adequacy remains to be determined through long-term obser-

vation.

References

1. Friedman LM, Furberg CD, DeMets DL, Reboussin DM, Granger CB: *Fundamentals of Clinical Trials*. Fifth edition. 2015.
2. GALIN JI, OGNIBENE FP, JOHNSON LL: *PRINCIPLES AND PRACTICE OF CLINICAL RESEARCH*. Fourth edition. 2018.
3. National Medical Products Administration: *Technical Review Guidance for Imaging Ultrasound Diagnostic Equipment (Class III)*. 2016.
4. National Medical Products Administration: *Guidance for Clinical Trials of Coronary Drug-Eluting Stents*. 2018.
5. National Medical Products Administration: *Guidance for Clinical Trials of Bioresorbable Coronary Drug-Eluting Stents*. 2019.
6. National Medical Products Administration: *Guidance for Clinical Trials of Transcatheter Implantable Aortic Valve Prostheses*. 2019.
7. Artificial Heart Bioprosthetic Valve [Internet]. [cited 2021 Aug 6]. Available from: https://www.chictr.org.cn/showproj.aspx?proj=10929
8. Song Guangyuan, Wang Moyang, Wang Yuan, et al.: Effectiveness of Venus-A Aortic Valve Intervention for Severe Aortic Stenosis. *Chinese Journal of Cardiology* 2017; 45:843–847.
9. Leon MB, Smith CR, Mack M, et al.: Transcatheter Aortic-Valve Implantation for Aortic Stenosis in Patients Who Cannot Undergo Surgery. *New England Journal of Medicine* 2010; 363.
10. CR S, MB L, MJ M, et al.: Transcatheter versus surgical aortic-valve replacement in high-risk patients. *The New England Journal of Medicine* [Internet] 2011 [cited 2021 Aug 10]; 364:2187–2198. Available from: https://pubmed.ncbi.nlm.nih.gov/21639811/
11. Leon MB, Smith CR, Mack MJ, et al.: Transcatheter or Surgical Aortic-Valve Replacement in Intermediate-Risk Patients. *New England Journal of Medicine* 2016; 374.
12. Mack MJ, Leon MB, Thourani VH, et al.: Transcatheter Aortic-Valve Replacement with a Balloon-Expandable Valve in Low-Risk Patients. *New England Journal of Medicine* 2019; 380.
13. Adams DH, Popma JJ, Reardon MJ, et al.: Transcatheter Aortic-Valve Replacement with a Self-Expanding Prosthesis. *New England Journal of Medicine* 2014; 370.
14. Reardon MJ, van Mieghem NM, Popma JJ, et al.: Surgical or Transcatheter Aortic-Valve Replacement in Intermediate-Risk Patients. *New*

England Journal of Medicine 2017; 376.

15. Popma JJ, Deeb GM, Yakubov SJ, et al.: Transcatheter Aortic-Valve Replacement with a Self-Expanding Valve in Low-Risk Patients. *New England Journal of Medicine* 2019; 380.
16. ACURATE IDE [Internet]. [cited 2021 Aug 6]. Available from: https://clinicaltrials.gov/ct2/show/NCT03735667
17. China XT [Internet]. [cited 2021 Aug 6]. Available from: https://clinicaltrials.gov/ct2/show/NCT0331485
18. LAmbre [Internet]. [cited 2021 Aug 6]. Available from: https://clinicaltrials.gov/ct2/show/NCT02029014
19. MemoLefort [Internet]. [cited 2021 Aug 6]. Available from: https://www.sohu.com/a/432481067_{72747
20. Reddy VY, Sievert H, Halperin J, et al.: Percutaneous Left Atrial Appendage Closure vs Warfarin for Atrial Fibrillation. *JAMA* [Internet] 2014; 312:1988. Available from: http://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.2014.15192
21. Holmes DR, Kar S, Price MJ, et al.: Prospective randomized evaluation of the Watchman Left Atrial Appendage Closure device in patients with atrial fibrillation versus long-term warfarin therapy: the PREVAIL trial. *Journal of the American College of Cardiology* [Internet] 2014; 64:1–12. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24998121
22. Amulet IDE [Internet]. [cited 2021 Aug 6]. Available from: https://clinicaltrials.gov/ct2/show/NCT02879
23. WAVECREST2 [Internet]. [cited 2021 Aug 6]. Available from: https://clinicaltrials.gov/ct2/show/NCT03302494

Note: Figure translations are in progress. See original paper for figures.