Risk Management and Healthcare Policy

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

Challenges in Results Robustness of Trials with Missing Data for the Primary Endpoint: Insights from Coronary Balloon/Stent Trials

Xinyue Lang[1], Yanyan Zhao[2], Yingxuan Zhu[2], Lei Song[3], Chuangshi Wang[4], Duoer Wang[2], Chilie Danzeng[2], Yang Wang[2], Wei Li[2]

[1]Department of Pharmacy and Clinical Trial Unit, Beijing Anzhen Hospital, Capital Medical University, Beijing, 100029, People's Republic of China; [2]Medical Research & Biometrics Center, National Center for Cardiovascular Diseases, The National Clinical Research Center for Cardiovascular Diseases, Fuwai Hospital, Chinese Academy of Medical Sciences & Pe-king Union Medical College, Beijing, 102308, People's Republic of China; [3]Department of Cardiology, Fuwai Hospital, National Center for Cardiovascular Diseases, Chinese Academy of Medical Sciences, Beijing, 100037, People's Republic of China; [4]National Clinical Research Center for Cardiovascular Diseases, Fuwai Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, 100037, People's Republic of China

Correspondence: Yang Wang; Wei Li, Email wangyang@mrbc-nccd.com; liwei@mrbc-nccd.com

Background: To assess the consequence of different degrees of missing primary endpoint data for randomized controlled trials and to find the influence factors.

Methods: PubMed, Cochrane Library, EMBASE and ClinicalTrials.gov were searched up to Nov 30, 2023. We included trials of the drug-coated balloon/drug-eluted stent with angiographic outcomes as the primary endpoint. The tipping-point analysis was used to deal with the missing data for the primary endpoint. The inconsistency rate, tipping-point standardized effect size (SES) and tipping-point ratio were used to assess the result robustness.

Results: A total of 101 trials were included, which had 109 trial comparisons. Among them, 89 (81.7%) comparisons had superior/non-inferior conclusions (H_0 rejected); 85 (78.0%) comparisons had a missing rate of ≥10%, and 30 (27.5%) comparisons had a missing rate of ≥20%. For H_0 rejected comparisons with a missing rate of ≥10%, the median of inconsistency rate, tipping-point SES and tipping-point ratio was 32.2% (IQR 19.7%, 45.4%), 0.90 (IQR 0.17, 1.79) and −1.53 (IQR −2.43, −0.39). A higher missing rate and a larger (worse) observed-target SES were associated with a more unreliable result.

Conclusion: A high dropout rate and inflated target effect size could cause an unreliable result. We emphasize a robust evaluation of the results for clinical trials with missing data for the primary endpoint.

Plain Language Summary: Missing data for the primary outcome has a great impact on the interpretation of clinical trials. This study included 101 randomized controlled trials of the drug-coated balloon or drug-eluted stent with an angiographic primary endpoint and used three indicators to assess the result robustness. This study found that 78.0% trial comparisons had a missing rate of ≥10%, and 27.5% comparisons had a missing rate of ≥20%; The conclusions of some clinical trials may change after dealing with the missing data; A higher missing rate and an inflated target effect size were associated with a more unreliable result.

Keywords: missing data, tipping-point analysis, result robustness

Introduction

Patients withdrawing before study completion is an explicitly anticipated potential problem for trial sponsors and investigators despite a well-designed study protocol.[1] The reasons for the patient's withdrawal include side effects, intercurrent health problems, unpleasant treatment, lack of improvement, recovery, and external factors unrelated to the study.[2]

Drug-coated balloon (DCB) or drug-eluted stent (DES) have been shown to reduce the incidence of target vessel revascularization and become commonly used treatments for coronary artery disease.[3–5] For randomized controlled trials

Received: 16 December 2024
Accepted: 22 February 2025
Published: 27 March 2025

Graphical Abstract

(RCTs) involving these devices, clinical endpoints require a large sample size; therefore, quantitative coronary angiography (QCA) results were usually used as surrogate endpoints for its feasibility.[6] However, due to its invasiveness, these studies always have a high dropout rate.[7,8] Patients' dropout causes missing data which can produce bias for the study results and may even change the conclusions.[9]

Common imputation methods such as maximum likelihood, multiple imputation, and Bayesian methods are applied to data missing at random and rely on parametric assumption. Since the mechanism of missing data is always unknown and unverifiable, sensitivity analyses for handling missing data are required by regulatory agencies to test the robustness of the study result.[10] Tipping-point analysis was proposed by the US Food and Drug Administration (FDA)[10] and improved by Liublinska and Rubin[11] which did not need to postulate any missing data mechanism or model parameters. This method enumerates all possible results caused by missing data and then conducts hypothesis tests one by one to find the value that changes the conclusion and therefore obtains the most comprehensive results in the presence of missing data.

In the current study, we hypothesize that the results of RCTs with high dropout rates might be unreliable. We will systematically sort out the RCT focusing on DCB or DES with angiographic outcomes as the primary endpoints, and use the tipping-point analysis to deal with the missing data, to test the robustness of RCT results and their influence factors.

Methods

Study Subjects and Inclusion/Exclusion Criteria

The systematic literature review was carried out under the guidance of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement.[12] Four databases, PubMed, EMBASE, Cochrane Library and ClinicalTrials.gov were searched up to Nov 30, 2023, with the search terms ("drug-coated balloon" or "drug-eluting stent") and ("Randomized controlled trial"). The citations were imported into Endnote (Endnote X9.3.1; Thomson Reuters, San Francisco, CA) for screening. Two reviewers (L.X. and Z.Y.) independently reviewed the abstracts and confirmed the eligibility through full-text assessment. A third reviewer (W.Y.) made the final decision if disagreement existed (A detailed search strategy was shown in the Supplement method).

The following inclusion criteria were used to assess the eligibility: (1) patients with coronary artery lesions (native coronary vessel lesion or in-stent restenosis); (2) RCT with DCB arm or DES arm; (3) setting the angiographic outcomes (LLL, MLD, %DS, or restenosis) as the primary endpoint; (4) with a clear study hypothesis (superiority or non-inferiority) and sample size calculation. Exclusion criteria were: (1) conference abstract, protocol or review; (2) the same clinical trial reporting long-term follow-up results and (3) H0 not rejected trials with an inferior result.

Data Collection and Definition

The following study information was extracted from the included trials, including first author, trial name, study design (noninferiority or superiority), interventions, sample size calculation parameter (power, alpha level, target standardized effect size and drop-out rate), number of patients/lesions at baseline, number of patients/lesions at follow-up, the primary endpoint definition and the observed standardized effect size.

The target/observed standardized effect size (*SES*) was created for the comparability of effect size among the different studies. The *SES* was calculated using formula (1) for continuous outcomes and (2) for binary outcomes,[13,14] where $\bar{x}_1$ and $\bar{x}_2$ were the mean value of the endpoints for the treatment and control, respectively. n_1 and n_2 were the sample size, s_1 and s_2 represented the standard deviation (SD) of each group, p_1 and represented the event rate of each group.

$$SES = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{(n_1-1){S_1}^2+(n_2-1){S_2}^2}{n_1+n_2-2}}} \tag{1}$$

$$SES = \frac{p_1 - p_2}{\sqrt{\frac{p_1(1-p_1)+p_2(1-p_2)}{2}}} \tag{2}$$

The primary endpoints in the included trials were mainly late lumen loss, percentage diameter stenosis, or binary angiographic restenosis, and a smaller value corresponds to better effectiveness. A smaller observed SES or target SES

indicated that the treatment group had a better effect than the control. So, we took the negative number of SES on minimal lumen diameter (MLD) to maintain the same direction of benefit with other endpoints.

Statistical Analysis

Tipping-point analysis was used to deal with the missing data of the primary endpoint.

We used three indexes to test the result robustness of the included trials, inconsistency rate, tipping-point SES and tipping-point ratio.

The inconsistency was considered if the outcome was different from the reported (For example, a study reported a superior/non-inferior result of the treatment group, but after the imputation, the result became inconclusive). The definition of the study result is plotted in Figure 1A.

For continuous outcomes, we imputed 150 times evenly for missing values in the treatment and control arm with a range of $\bar{x} \pm$ 3SD. A total of 22,500 possible outcomes were obtained. Welch's *t*-test method was then used to test the group difference. For binary outcomes, we enumerated all possible outcomes for the missing data, (from all failures to all successes) and compared the differences between groups using the chi-square test (See supplementary marital for the R code).

The inconsistency rate was calculated as follows:

$$\text{Inconsistency rate} = \text{the number of inconsistencies}/\text{the number of possible outcomes} * 100\%$$

The tipping point (tipping-point SES) was defined as the standardized difference of means or event rates between the treatment and control group in the missing cohort where the result changed when imputing.[10] For trials with superior results, the tipping point was considered as the result became "inconclusive". For trials with inconclusive results, two tipping points were considered where the result became "superior" or "inferior". We chose the one which was closer to the observed SES (See supplementary marital for the detailed description). The derived indicator (tipping-point ratio) was calculated as follows:

$$\text{Tipping} - \text{point ratio} = \text{tipping} - \text{point SES}/\text{observed SES}.$$

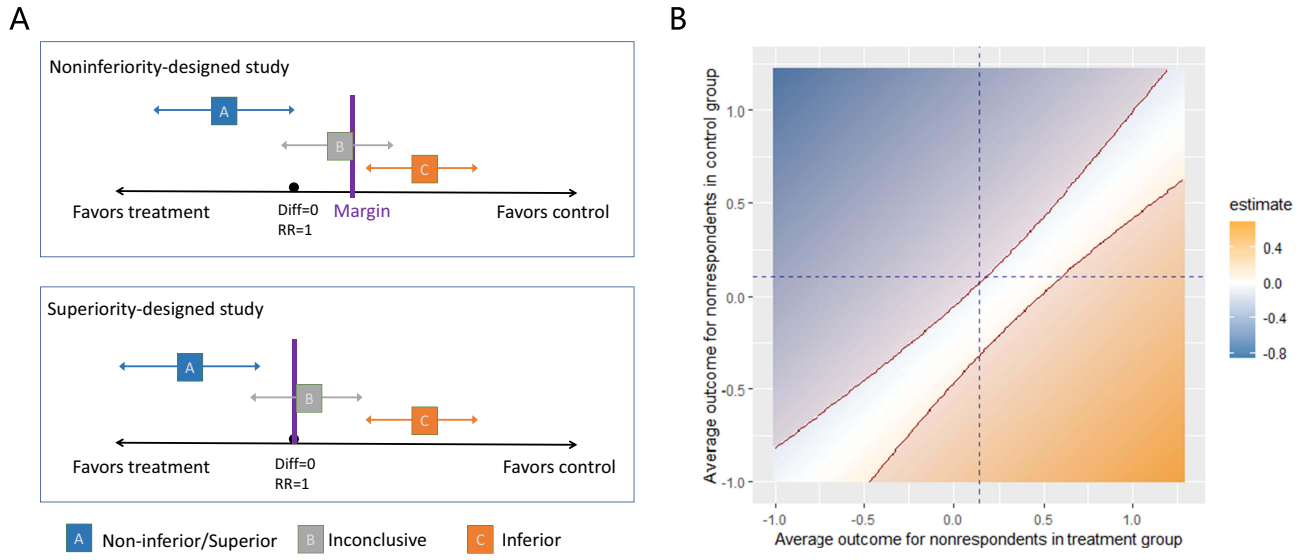

Figure 1 The definition of the study result and tipping-point analysis. (**A**) Forest plot with differences and 2-sided CIs showing the hypothetical outcomes from noninferiority-designed and superiority-designed studies. The blue one indicates a superior result, the grey one indicates an inconclusive result, and the Orange one indicates an inferior result. (**B**) An example of the result (LONG-DES V) after tipping point analysis. In the study of Lee 2014, The blue area means that the treatment group is noninferior or superior to the control. The white area indicates that the treatment group was inconclusive with the control. The orange area indicates that the treatment group is inferior to the control.

Abbreviations: CI, confidence interval; RR, relative risk.

Continuous variables were described using means (standard deviations) and median (Interquartile range) and compared by Student's *t*-test or Mann–Whitney *U*-test for two groups, Analysis of variance or Kruskal–Wallis *H*-test for three groups. Categorical variables were described using frequencies (percentages) and compared by the chi-squared test or Fisher's exact test. Multivariable generalized linear regression models were used to assess the influence factors of the inconsistency rate, tipping-point SES and tipping-point ratio. Covariates were selected in the models according to the clinical experience, including study design (superiority or noninferiority), primary endpoint (LLL or other), time to primary endpoint analysis (≤6m, 8m–9m or ≥12m), sample size, power, 1-side alpha (0.050 or 0.025), missing rate (<10%, 10–19% or ≥20%) and observed-target SES. Pearson correlation analysis was used to test the correlation among inconsistency rate, tipping-point SES and tipping-point ratio. All analyses were conducted using R version 4.0.3 and SAS version 9.4. A 2-sided $p < 0.05$ was considered significant.

Results

Study Selection

Of 1057 articles, 808 papers were excluded after the titles and abstracts screening. Through full texts reading, 147 studies were further deleted for the following reasons: result from the same study (n = 53), without angiographic outcomes (n = 42), without dropout (n = 8), without DCB or DES as the study arm (n = 18), without study hypothesis (n = 26), with inferior result (n = 1). Finally, 101 RCTs with 31862 patients were enrolled (Figure S1). Specific information on the included publications is available in Table S1.

Summary of the Included Trials

Among the included 101 trials, 93 trials had both a single control and treatment arm. The ISAR-DESIRE 3,[15] TAX-001[16] and STRESSED[17] trials had 1 active treatment arm and 2 control arms, the ISAR-DESIRE,[18] PAINT,[19] ISAR-TEST 3,[20] and EVOLVE[21] had 2 active treatment arms and 1 control arm, and Song[22] conducted the trial under two conditions with two comparisons. Finally, we got 109 powered trial comparisons.

Of the 109 trial comparisons, 67 (61.5%) had a noninferiority hypothesis, and 42 (38.5%) had a superiority hypothesis. 86 (78.9%) trials used LLL as the primary endpoint.74 (66.7%) studies used 8–9 months as the time for the primary endpoint comparison (range, 6–36 months). The median sample size was 230 (IQR 148, 397) and the projected power was 80% (IQR 80%, 90%). One-sided alpha of 0.025 or 0.05 was used in 78 (71.6%) trials, and 31 (28.4%) trials, separately. The median missing rate for the primary endpoint was 16.1% (IQR 10.6%, 20.2%). And the missing rate of 24 (22.0%) trials was <10%, of 55 (50.5%) trials was 10–19%, and of 30 (27.5%) trials was >20%. The median target SES and observed SES were −0.4 (IQR −0.6, −0.4) and −0.6 (IQR −0.8, −0.3), respectively. Forty-four (40.4%) trials did not reach the target SES (target SES minus margin for noninferiority trials). The median observed-target SES was −0.1 (IQR −0.3, 0.1). 89 (81.7%) trial comparisons rejected the H_0 hypothesis, and 20 (18.3%) failed and gave inconclusive results. Except for a larger power, a larger observed SES and a larger observed-target SES, H_0 not rejected trials had no difference from H_0 rejected trials (Table 1).

Tipping-Point Analysis for the Primary Endpoints

Figure 1B plots the imputation result of a typical trial (the LONG-DES V trial) after tipping-point analysis. A 51.3% inconsistency rate was shown (the white area plus the yellow area). The tipping-point SES and tipping-point ratio were −0.10 and 0.48, respectively. The LONG-DES V trial assumed that the 9m in-segment LLL of the biolimus A9-eluting stent were not inferior to those of the durable PtCr-EES and observed their noninferiority. However, this trial had a high drop-out rate of the angiographic outcomes, 33.1% for the treatment arm, and 35.7% for the control arm. And its observed effect size did not meet the target effect size.

After conducting a tipping-point analysis, for 20 H_0 not rejected trial comparisons, the median of the inconsistency rate and tipping-point SES was 56.9% (IQR 37.8%, 68.9%) and −0.5 (IQR −0.7, 0.6), separately. The tipping-point ratio is not available because H_0 not rejected trial comparisons always have a small observed SES, which causes the tipping-point ratio extremely large. For 89 H_0 rejected trial comparisons, the median of the inconsistency rate, tipping-point SES,

Table 1 Studies Characteristics of Included Studies by Study Original Result

Variables	Overall(N=109)	H_0 Rejected[a](N=89)	H_0 not Rejected[b](N=20)	P value
Study design				
Noninferiority	67 (61.5)	56 (62.9)	11 (55.0)	0.511
Superiority	42 (38.5)	33 (37.1)	9 (45.0)	
Primary endpoint				
LLL	86 (78.9)	70 (78.7)	16 (80.0)	0.339
MLD	7 (6.4)	5 (5.6)	2 (10.0)	
%DS	6 (5.5)	6 (6.7)	0 (0.0)	
Restenosis	5 (4.6)	3 (3.4)	2 (10.0)	
Other	5 (4.6)	5 (5.6)	0 (0.0)	
Time to the primary endpoint analysis				
≤6m	28 (25.7)	26 (29.2)	2 (10.0)	0.202
8m–9m	73 (67.0)	57 (64.0)	16 (80.0)	
≥12m	8 (7.3)	6 (6.7)	2 (10.0)	
Sample size	290.9±240.4/230.0 (148.0, 397.0)	269.1±188.9/221.0 (135.0, 372.0)	388.3±388.5/363.0 (176.0, 450.5)	0.129
Power	84.8±5.6/80.0 (80.0, 90.0)	84.2±5.3/80.0 (80.0, 90.0)	87.4±6.3/90.0 (80.0, 90.0)	0.038
1-side alpha				
0.025	78 (71.6)	66 (74.2)	12 (60.0)	0.205
0.050	31 (28.4)	23 (25.8)	8 (40.0)	
Missing rate				
<10%	24 (22.0)	21 (23.6)	3 (15.0)	0.354
10–19%	55 (50.5)	46 (51.7)	9 (45.0)	
≥20%	30 (27.5)	22 (24.7)	8 (40.0)	
Target SES[c]	−0.5±0.2/−0.4 (−0.6, −0.4)	−0.5±0.2/−0.4 (−0.6, −0.4)	−0.4±0.3/−0.4 (−0.6, −0.3)	0.202
Observed SES[d]	−0.6±0.4/−0.6 (−0.8, −0.3)	−0.7±0.3/−0.7 (−0.9, −0.5)	−0.0±0.2/−0.0 (−0.1, 0.1)	<0.001
Observed-target SES[e]	−0.1±0.4/−0.1 (−0.3, 0.1)	−0.2±0.3/−0.1 (−0.4, 0.0)	0.4±0.3/0.4 (0.2, 0.5)	<0.001

Notes: [a]Trial comparisons had superior or non-inferior results. [b]Trial comparisons had inconclusive results. [c]Calculated based on the study design, the standardized difference of the expected efficacy of the treatment and control groups. A higher MLD indicates a better status, so we took the opposite number of SES on MLD to maintain the same direction of benefit with other endpoints (such as LLL and %DS). [d]Calculated based on the observed result, the standardized difference of the observed efficacy of the treatment and control groups. [e]The absolute difference between observed SES and target SES. The values in the table are presented as n(%) or mean ± sd/median (IQR).

Abbreviations: %DS, percentage diameter stenosis; LLL, late lumen loss; MLD, minimal lumen diameter; SES, standardized effect size.

and tipping-point ratio was 23.2% (IQR 1.8%, 38.4%), 1.3 (IQR 0.5, 3.7) and −2.1 (IQR −4.3, −0.8), separately. H0 rejected comparisons were significantly more robust than H0 not rejected comparisons. For 89 H_0 rejected trial comparisons, the study result became less robust with increasing missing rate (Table 2).

Table 2 Tipping-Point Analysis for the Primary Endpoints

	H_0 not Rejected (N=20)	H_0 Rejected(N=89)		P value
Inconsistency rate[a]	50.6±22.7/56.9 (37.8, 68.9)	23.1±18.3/23.2 (1.8, 38.4)		<0.001
Tipping-point SES[b]	−0.4±1.1/−0.5 (−0.7, 0.6)	2.9±3.6/1.3 (0.5, 3.7)		<0.001
Tipping-point ratio[c]	NA[d]	−3.3±3.8/−2.1 (−4.3, −0.8)		NA
	H_0 Rejected Comparisons			
	Missing Rate <10%(N=21)	Missing Rate 10–19%(N=46)	Missing Rate ≥20%(N=22)	P for Trend
Inconsistency rate	1.0±2.6/0.0 (0.0, 0.0)	26.6±15.9/26.5 (16.3, 39.3)	36.7±11.7/37.3 (28.5, 47.5)	<0.001
Tipping-point SES	8.3±3.5/8.1 (5.5, 11.8)	1.7±1.8/1.3 (0.5, 2.1)	0.6±0.7/0.6 (−0.1, 1.3)	<0.001
Tipping-point ratio	−8.9±4.1/−6.7 (−12.4, −6.4)	−2.2±1.6/−2.2 (−3.4, −0.8)	−0.8±1.0/−1.0 (−1.5, 0.2)	<0.001

Notes: [a]Inconsistency rate was defined as the rate of inconsistencies of the possible outcomes after filling. [b]Tipping-point SES was defined as the standardized difference of means or event rates between the treatment and control group in the missing cohort, where the result changed when imputing. [c]Tipping-point ratio was defined as the ratio of tipping-point SES and observed SES. [d]H_0 not rejected trial comparisons always have a small observed SES, which causes the tipping-point ratio extremely large. The values in the table are presented as mean ± sd/ median (IQR).

Abbreviation: SES, standardized effect size.

Figure 2A–C plots the imputation results after tipping-point analysis in the H_0 rejected trial comparisons with a missing rate of ≥10%. In 68 trial comparations, the median of the inconsistency rate, tipping-point SES, and tipping-point ratio was 32.2% (IQR 19.7%, 45.4%), 0.90 (IQR 0.17, 1.79) and −1.53 (IQR −2.43, −0.39), separately.

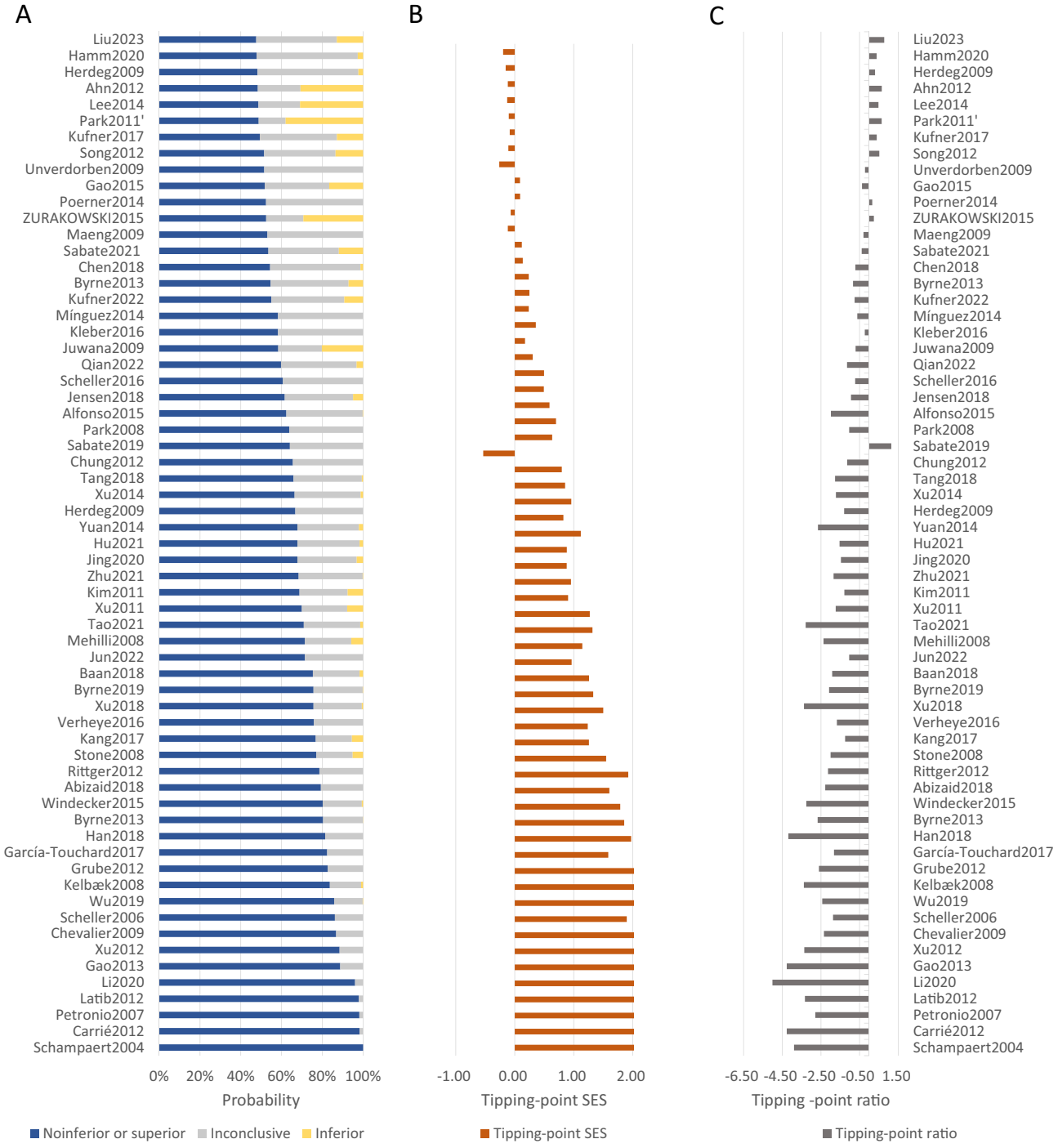

Figure 2 Tipping-point analysis for the primary endpoints of the H_0 rejected comparisons with a missing rate of no less than 10%. (**A**) The proportion of the H_0 rejected comparisons with a missing rate of no less than 10% after tipping point analysis (the inconsistency rate was the grey area plus the yellow area). (**B**) The tipping-point SES. (**C**) The tipping-point ratio.

Abbreviation: SES, standardized effect size.

https://doi.org/10.2147/RMHP.S511449

Table 3 Multivariable Generalized Linear Regression Analysis of Study Characteristics in the H_0 Rejected Comparisons

Variables	Inconsistency Rate		Tipping-point SES		Tipping-point Ratio	
	β (95% CI)	P value	*β* (95% CI)	P value	*β* (95% CI)	P value
Study design						
Superiority	ref.		ref.		ref.	
Noninferiority	0.023 (−0.019, 0.065)	0.286	−0.306 (−0.684, 0.072)	0.112	0.197 (−0.325, 0.720)	0.459
Primary endpoint						
Other	ref.		ref.		ref.	
LLL	0.013 (−0.035, 0.060)	0.605	−0.047 (−0.494, 0.399)	0.835	0.171 (−0.446, 0.789)	0.586
Time to primary endpoint analysis						
≤6m	ref.		ref.		ref.	
8m–9m	0.021 (−0.021, 0.064)	0.322	−1.227 (−2.803, 0.348)	0.127	0.309 (−0.763, 1.381)	0.572
≥12m	−0.001 (−0.080, 0.079)	0.99	−1.795 (−4.740, 1.151)	0.233	0.966 (−1.008, 2.939)	0.337
Sample size (per 100)	−0.001 (−0.012, 0.010)	0.856	0.044 (−0.153, 0.241)	0.662	−0.176 (−0.448, 0.097)	0.206
Power	−0.005 (−0.008, −0.001)	0.012	0.011 (−0.054, 0.076)	0.736	−0.012 (−0.101, 0.078)	0.794
1-side alpha						
0.050	ref.		ref.		ref.	
0.025	−0.023 (−0.068, 0.021)	0.305	0.017 (−0.386, 0.421)	0.933	0.218 (−0.340, 0.776)	0.443
Missing rate						
<10%	ref.		ref.		ref.	
10%-19%	0.165 (0.113, 0.216)	<0.001	−5.236 (−6.202, −4.270)	<0.001	6.291 (4.954, 7.627)	<0.001
≥20%	0.251 (0.194, 0.309)	<0.001	−6.187 (−7.245, −5.128)	<0.001	7.747 (6.284, 9.211)	<0.001
Observed-target SES	0.327 (0.264, 0.390)	<0.001	−4.559 (−5.699, −3.419)	<0.001	2.007 (0.430, 3.584)	0.013

Abbreviations: CI, confidence interval; SES, standardized effect size.

Association Between Study Characteristics and Inconsistency Rate, Tipping-Point SES and Tipping-Point Ratio

The Pearson correlation analysis showed the association of the missing rate with the inconsistency rate, tipping-point SES and tipping-point ratio (Figure S2). The inconsistency rate, tipping-point SES and tipping-point ratio were correlated with each other (Figure S3).

The multivariable generalized linear regression model demonstrated that missing rate (<10%, 10–19%, ≥20%) and observed-target SES were associated with inconsistency rate, tipping-point SES and tipping-point ratio for the H_0 rejected studies after adjusting other study characteristics.

Compared with missing rate <10%, the missing rate at 10–19% increased 16.5% of the inconsistency rate (95% CI, 0.194 to 0.309 $p < 0.001$), decreased −5.236 of the tipping-point SES (95% CI, −6.202 to −4.270, $p < 0.001$), increased 6.291 of the tipping-point ratio (95% CI, 4.954 to 7.627, $p < 0.001$), and missing rate ≥20% increased 25.1% of the inconsistency rate (95% CI, 0.194 to 0.309, $p < 0.001$), decreased −6.187 of the tipping-point SES (95% CI, −7.245 to −5.128, $p < 0.001$), increased 7.747 of the tipping-point ratio (95% CI, 6.284 to 9.211, $p < 0.001$).

As for observed-target SES, 32.7% increased in inconsistency rate per unit increased in observed-target SES (95% CI, 0.264 to 0.390, $p < 0.001$), 4.559 decreased in tipping-point SES per unit increased in observed-target SES (95% CI, −5.699 to −3.419, $p < 0.001$), 2.007 increased in tipping-point ratio per unit increased in observed-target SES (95% CI, 0.430 to 3.584, $p < 0.001$) (Table 3). The results were stable in the sensitivity analysis of all 109 trial comparations (Table S2).

Discussion

Main Findings

This is the first study, to our knowledge evaluating the robustness of the coronary drug-eluted stent/balloon study results using tipping-point analysis. We used three indicators to assess the robustness, inconsistency rate, tipping-point SES and tipping-point ratio. We found that (1) for trials with an angiographic primary endpoint, only 22.0% of trials with a missing rate of less than 10%; (2) the conclusions of some clinical trials may change after tipping-point analysis. (3)

a higher missing rate and a larger observed-target SES were associated with a more unreliable result (a higher inconsistency rate, a lower tipping-point SES and a higher tipping-point ratio).

Except for the first proposed tipping-point SES,[10] we posed another two indexes (inconsistency rate and tipping-point ratio) to test the robustness of the results. They all had a good correlation with each other except for trials with high robustness. Since mean±3SD was considered when filling, for studies with an inconsistency rate of 0, tipping-point SES and tipping-point ratio could distinguish these studies by filling in values beyond this range. However, there is little need to distinguish these situations. The inconsistency rate could more intuitively display all results after filling than the tipping-point SES. As for the tipping-point ratio, which compares the tipping-point SES with the observed SES, the value would help to make a judgment of whether such a difference was implausibly unfavourable. For example, if a tipping-point ratio was less than −1, it would indicate a robust result (because the tipping-point SES was less likely to be the opposite number of the observed SES, or even farther). These indexes all have their advantages, Evaluation together allows for a more comprehensive presentation of the results.

Relation to Other Literature

Our findings were primarily based on the H_0 rejected trial comparisons. H_0 not rejected comparisons had a significantly not robust result compared with H_0 rejected comparisons and were considered in the sensitivity analysis. For H_0 rejected trial comparisons with missing rate ≥ 10%, the median of inconsistency rate, tipping-point SES and tipping-point ratio was 32.2% (IQR 19.7%, 45.4%), 0.90 (IQR 0.17, 1.79) and −1.53 (IQR −2.43, −0.39), separately. The results emphasize a sensitivity analysis for trials with missing data. For the clinical trials, some patients might not be missing at random, and the reason included poor curative effect, adverse reactions, etc. In addition, it is recommended by National Research Council (US) Panel on Handling Missing Data in Clinical Trials that not only a primary analysis that assumes data are missing at random, but also sensitivity analyses which allow for data not missing at random are needed.[23] However, a review of 77 RCTs at the top medical journals showed that only 27 (35%) trials performed a sensitivity analysis, among them 10 (37%) trials weakened the assumptions regarding missing data from their primary analysis.[24]

In our study, we demonstrated that the more missing, the less robust the findings. For H_0 rejected trials with a missing rate of ≥20%, the median inconsistency rate was even 37.3% with an interquartile range from 28.5% to 47.5%. As described by Dong et al,[25] they used a listwise deletion method, multiple imputations, full information maximum likelihood, and expectation-maximization algorithm to deal with a complete real-world data set, and indicated that any level of missing data was a potential risk to the validity of a trial, and missing over 20% would give a significant risk. A systematic review of RCTs testing palliative interventions also demonstrated that missing data reduced the power and potentially became a source of bias for trials.[26]

We also found that the observed-target SES was the stable influence factor of the inconsistency rate, tipping-point SES and tipping-point ratio instead of the study power, alpha level or sample size. The larger the observed-target SES (the observed SES not reached and is worse than the target SES), the more unstable the study result. A large observed-target SES indicates a poor observed SES, or an inflated target SES. An excessive target SES corresponds to arbitrarily large margins or overestimated mean values/event rates for non-inferiority designed trials and an inflated expected difference for superiority designed trials. Previous studies have demonstrated that the target SES always determined the sample size and the study costs, hence, it was tempting for researchers to inflate the target SES to reduce the sample size, which would cause an inappropriate trial design and introduce bias.[27–30]

Take ISAR-DESIRE 4 trial[31] for example, it reported that neointimal modification with scoring balloon pre-dilation before a drug-coated balloon showed superior results of the primary endpoint (in-segment percentage diameter stenosis) than a drug-coated balloon. Its observed SES was −0.281 (treatment vs control: 35.0±16.8 vs 40.4±21.4), which did not reach the target SES (−0.398, treatment vs control: 26.25±22 vs 35±22). Although the authors used a Student's *t*-test to calculate the difference and found the superiority, the 95% confidence intervals of the primary endpoint of the treatment and the control were overlapped (treatment, 35.0, 95% CI 31.8 to 38.2; control, 40.4, 95% CI 36.2 to 44.6). Thus, the result was unstable and it was not easy to pass the sensitivity analysis.

For clinicians, it is important to use clinical endpoints, such as target vessel failure, to assess the safety and effectiveness of new devices.[32,33] When surrogate endpoints are used due to feasibility, missing data for the primary endpoint should be considered. Clinicians could use our R codes provided in the supplementary material to calculate the inconsistency rate of their research, thereby providing more rigorous evidence for their results.

Strengths and Limitations

This study has some limitations. Firstly, our study was limited to coronary drug-eluting stent or drug-coated balloon trials using angiographic outcomes as the primary endpoint. This kind of trial always has a higher dropout rate due to the invasive examination. However, drop-out is a common problem for RCTs, and the tipping-point analysis is applicable for other studies, not only RCT but also cohort, case-control and cross-sectional studies.[34–39] Additionally, we did not have the original datasets of the eligible trials, so we filled them based on the normal distribution according to the reported mean and SD. Third, although we conducted a comprehensive search, some unpublished articles were not included due to negative results. To avoid publication bias, we also included H_0 not rejected trials. According to the analysis of the characteristics of included studies, H_0 rejected and H_0 not rejected had similar study characteristics. We consider that the publication bias may have little impact on the results. Last but not least, we only used tipping-point analysis to fill in the missing data, and the indexes we proposed still need further sensitivity and specificity tests.

Conclusions

Rigorous sensitivity analysis (such as tipping-point analysis) adds more difficulty to the trials, which may restrict some new technologies entering the market. Having said this, the changes in results after filling in the missing data presented from the stent/balloon trials should raise concerns about the high dropout rate and inflated target effect size and their consequences. We emphasize a reliability evaluation of the results for further clinical trials with missing data for the primary endpoint.

Abbreviations

DCB, drug-coated balloons; DES, drug-eluting stents; LLL, late lumen loss; MLD, minimal lumen diameter; RCT, randomized controlled trials; SES, standardized effect size.

Data Sharing Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Research Ethics Approval

This study was based on published papers, we analyzed the results from 101 randomized controlled trials (109 trial comparisons).

This study does not involve human participants, so we don't have ethical approval.

Patient and Public Involvement

Patients or the public were not involved in the design, or conduct, or reporting, or dissemination plans of our research.

Acknowledgments

Thanks for the help of the members of the Medical Research and Biometrics Center.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

Fuwai Hospital, Chinese Academy of Medical Sciences, Special Project for the Development of Characteristic Disciplines. Construction of clinical trial sample database and statistical methodology service platform (Grant No. 2022-FWTS08); Fuwai Hospital, Chinese Academy of Medical Sciences (Grant No. 2023-GSP-GG-1); Fuwai Shenzhen Hospital, Chinese Academy of Medical Sciences, Shenzhen High-level Hospital Construction Fund (Grant No. NCRCSZ-2023-012); Guangdong Basic and Applied Basic Research Foundation (Grant No.2024A1515220161).

Disclosure

The authors report no conflicts of interest in this work.

References

1. Ware JH, Harrington D, Hunter DJ, D'Agostino RB. Missing data. *N Engl J Med*. 2012;367(14):1353–1354. doi:10.1056/NEJMsm1210043
2. Heyting A, Tolboom JT, Essers JG. Statistical handling of drop-outs in longitudinal clinical trials. *Stat Med*. 1992;11(16):2043–2061. doi:10.1002/sim.4780111603
3. Holmes DR, Teirstein P, Satler L, et al. Sirolimus-eluting stents vs vascular brachytherapy for in-stent restenosis within bare-metal stents: the SISR randomized trial. *JAMA*. 2006;295(11):1264-1273. doi:10.1001/jama.295.11.1264
4. Spaulding C, Henry P, Teiger E, et al. Sirolimus-eluting versus uncoated stents in acute myocardial infarction. *N Engl J Med*. 2006;355(11):1093-1104. doi:10.1056/NEJMoa062006
5. Habara S, Iwabuchi M, Inoue N, et al. A multicenter randomized comparison of paclitaxel-coated balloon catheter with conventional balloon angioplasty in patients with bare-metal stent restenosis and drug-eluting stent restenosis. *Am Heart J*. 2013;166(3):527-533. doi:10.1016/j.ahj.2013.07.002
6. Ueki Y, Räber L. Late lumen loss in the era of new generation drug-eluting stents: perspective on a quarter century companion. *Eur Heart J*. 2018;39(36):3390–3392. doi:10.1093/eurheartj/ehy582
7. Hu P, Sun Y, Li CL, et al. A randomized comparison of two paclitaxel-coated balloons for the treatment of in-stent restenosis: the LONGTY ISR China randomized trial (LONGTY DCB vs. SeQuent Please DCB). *Catheter Cardiovasc Interv*. 2021;97(S2):988–995. doi:10.1002/ccd.29589
8. Wu Y, Shen L, Yin J, et al. Twelve-month angiographic and clinical outcomes of the XINSORB bioresorbable sirolimus-eluting scaffold and a metallic stent in patients with coronary artery disease. *Int J Cardiol*. 2019;293:61-66. doi:10.1016/j.ijcard.2019.06.053
9. van der Heijde D, Fleischmann R, Wollenhaupt J, et al. Effect of different imputation approaches on the evaluation of radiographic progression in patients with psoriatic arthritis: results of the RAPID-PsA 24-week Phase III double-blind randomised placebo-controlled study of certolizumab pegol. *Ann Rheum Dis*. 2014;73(1):233–237. doi:10.1136/annrheumdis-2013-203697
10. Yan X, Lee S, Li N. Missing data handling methods in medical device clinical trials. *J Biopharm Stat*. 2009;19(6):1085–1098. doi:10.1080/10543400903243009
11. Liublinska V, Rubin DB. Sensitivity analysis for a partially missing binary outcome in a two-arm randomized clinical trial. *Stat Med*. 2014;33(24):4170–4185. doi:10.1002/sim.6197
12. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med*. 2009;151(4):264–269, w264.
13. Flury BK, Riedwyl H. Standard distance in univariate and multivariate analysis. *Am Stat*. 1986;40(3):249–251.
14. Killeen PR. An alternative to null-hypothesis significance tests. *Psychol Sci*. 2005;16(5):345–353. doi:10.1111/j.0956-7976.2005.01538.x
15. Byrne RA, Neumann FJ, Mehilli J, et al. Paclitaxel-eluting balloons, paclitaxel-eluting stents, and balloon angioplasty in patients with restenosis after implantation of a drug-eluting stent (ISAR-DESIRE 3): a randomised, open-label trial. *Lancet*. 2013;381(9865):461–467. doi:10.1016/S0140-6736(12)61964-3
16. Herdeg C, Göhring-Frischholz K, Haase KK, et al. Catheter-based delivery of fluid paclitaxel for prevention of restenosis in native coronary artery lesions after stent implantation. *Circulation*. 2009;2(4):294-301. doi:10.1161/CIRCINTERVENTIONS.108.827865.108.827865
17. Remkes WS, Somi S, Roolvink V, et al. Direct drug-eluting stenting to reduce stent restenosis: a randomized comparison of direct stent implantation to conventional stenting with pre-dilation or provisional stenting in elective PCI patients. *JACC Cardiovasc Interv*. 2014;7(7):751-758. doi:10.1016/j.jcin.2014.02.012
18. Kastrati A, Mehilli J, von Beckerath N, et al. Sirolimus-eluting stent or paclitaxel-eluting stent vs balloon angioplasty for prevention of recurrences in patients with coronary in-stent restenosis: a randomized controlled trial. *JAMA*. 2005;293(2):165-171. doi:10.1001/jama.293.2.165
19. Lemos PA, Moulin B, Perin MA, et al. Randomized evaluation of two drug-eluting stents with identical metallic platform and biodegradable polymer but different agents (paclitaxel or sirolimus) compared against bare stents: 1-year results of the PAINT trial. *Catheter Cardiovasc Interv*. 2009;74(5):665-673. doi:10.1002/ccd.22166
20. Mehilli J, Byrne RA, Wieczorek A, et al. Randomized trial of three rapamycin-eluting stents with different coating strategies for the reduction of coronary restenosis. *Eur Heart J*. 2008;29(16):1975-1982. doi:10.1093/eurheartj/ehn253
21. Meredith IT, Verheye S, Dubois CL, et al. Primary endpoint results of the EVOLVE trial: a randomized evaluation of a novel bioabsorbable polymer-coated, everolimus-eluting coronary stent. *J Am Coll Cardiol*. 2012;59(15):1362-1370. doi:10.1016/j.jacc.2011.12.016
22. Song HG, Park DW, Kim YH, et al. Randomized trial of optimal treatment strategies for in-stent restenosis after drug-eluting stent implantation. *J Am Coll Cardiol*. 2012;59(12):1093-1100. doi:10.1016/j.jacc.2011.11.047
23. Roderick, JA, Ralph, D, Kay, D et al Introduction and Background.. In: *The Prevention and Treatment of Missing Data in Clinical Trials*. Washington (DC): National Academies Press (US); 2010.
24. Bell ML, Fiero M, Horton NJ, Hsu CH. Handling missing data in RCTs; a review of the top medical journals. *BMC Med Res Methodol*. 2014;14:118. doi:10.1186/1471-2288-14-118

25. Dong Y, Peng CY. Principled missing data methods for researchers. *Springerplus*. 2013;2(1):222. doi:10.1186/2193-1801-2-222
26. Hussain JA, White IR, Langan D, et al. Missing data in randomized controlled trials testing palliative interventions pose a significant risk of bias and loss of power: a systematic review and meta-analyses. *J Clin Epidemiol*. 2016;74:57–65. doi:10.1016/j.jclinepi.2015.12.003
27. Head SJ, Kaul S, Bogers AJ, Kappetein AP. Non-inferiority study design: lessons to be learned from cardiovascular trials. *Eur Heart J*. 2012;33 (11):1318–1324. doi:10.1093/eurheartj/ehs099
28. Pocock SJ, Clayton TC, Stone GW. Challenging issues in clinical trial design: part 4 of a 4-part series on statistics for clinical trials. *J Am Coll Cardiol*. 2015;66(25):2886–2898. doi:10.1016/j.jacc.2015.10.051
29. Macaya F, Ryan N, Salinas P, Pocock SJ. Challenges in the design and interpretation of noninferiority trials: insights from recent stent trials. *J Am Coll Cardiol*. 2017;70(7):894–903. doi:10.1016/j.jacc.2017.06.039
30. Simonato M, Ben-Yehuda O, Vincent F, Zhang Z, Redfors B. Consequences of inaccurate assumptions in coronary stent noninferiority trials: a systematic review and meta-analysis. *JAMA Cardiol*. 2022;7(3):320–327. doi:10.1001/jamacardio.2021.5724
31. Kufner S, Joner M, Schneider S, et al. Neointimal modification with scoring balloon and efficacy of drug-coated balloon therapy in patients with restenosis in drug-eluting coronary stents: a randomized controlled trial. *JACC Cardiovasc Interv*. 2017;10(13):1332-1340. doi:10.1016/j.jcin.2017.04.024
32. Stone GW, Ellis SG, Gori T, et al. Blinded outcomes and angina assessment of coronary bioresorbable scaffolds: 30-day and 1-year results from the ABSORB IV randomised trial. *Lancet*. 2018;392(10157):1530–1540. doi:10.1016/S0140-6736(18)32283-9
33. Lansky A, Wijns W, Xu B, et al. Targeted therapy with a localised abluminal groove, low-dose sirolimus-eluting, biodegradable polymer coronary stent (TARGET All Comers): a multicentre, open-label, randomised non-inferiority trial. *Lancet*. 2018;392(10153):1117-1126. doi:10.1016/S0140-6736(18)31649-0
34. Goudman L, Molenberghs G, Duarte RV, Moens M. The influence of missing data on disabilities in patients treated with high-dose spinal cord stimulation: a tipping point sensitivity analysis. *J Clin Med*. 2021;10(21):4897. doi:10.3390/jcm10214897
35. Jin M, Liu R, Robieson W. Modified reference based imputation and tipping point analysis in the presence of missing data due to COVID-19. *Contemp Clin Trials*. 2021;110:106575. doi:10.1016/j.cct.2021.106575
36. Pigeon WR, Crean HF, Cerulli C, Gallegos AM, Bishop TM, Heffner KL. A randomized clinical trial of cognitive-behavioral therapy for insomnia to augment posttraumatic stress disorder treatment in survivors of interpersonal violence. *Psychother Psychosom*. 2022;91(1):50–62. doi:10.1159/000517862
37. Lipkovich I, Ratitch B, O'Kelly M. Sensitivity to censored-at-random assumption in the analysis of time-to-event endpoints. *Pharm Stat*. 2016;15 (3):216–229. doi:10.1002/pst.1738
38. Mehrotra DV, Liu F, Permutt T. Missing data in clinical trials: control-based mean imputation and sensitivity analysis. *Pharm Stat*. 2017;16 (5):378–392. doi:10.1002/pst.1817
39. Ouyang J, Carroll KJ, Koch G, Li J. Coping with missing data in phase III pivotal registration trials: tolvaptan in subjects with kidney disease, a case study. *Pharm Stat*. 2017;16(4):250–266. doi:10.1002/pst.1808