JAMA Internal Medicine | Original Investigation

# Characteristics of Contemporary Randomized Clinical Trials and Their Association With the Trial Funding Source in Invasive Cardiovascular Interventions

Mario Gaudino, MD; Irbaz Hameed, MD; Mohamed Rahouma, MD; Faiza M. Khan, MD; Derrick Y. Tam, MD; Giuseppe Biondi-Zoccai, MD, MStat; Michelle Demetres, MLIS; Mary E. Charlson, MD; Marc Ruel, MD; Filippo Crea, MD; Volkmar Falk, MD; Leonard N. Girardi, MD; Stephen Fremes, MD; Joanna Chikwe, MD

**IMPORTANCE** Changes in evidence-based practice and guideline recommendations depend on high-quality randomized clinical trials (RCTs). Commercial device and pharmaceutical manufacturers are frequently involved in the funding, design, conduct, and reporting of trials, the implications of which have not been recently analyzed.

**OBJECTIVE** To evaluate the design, outcomes, and reporting of contemporary randomized clinical trials of invasive cardiovascular interventions and their association with the funding source.

**DESIGN, SETTING, AND PARTICIPANTS** This cross-sectional study analyzed published RCTs between January 1, 2008, to May 31, 2019. The trials included those involving coronary, vascular and structural interventional cardiology, and vascular and cardiac surgical procedures.

**MAIN OUTCOMES AND MEASURES** We assessed (1) trial characteristics, (2) finding of a statistically significant difference in the primary end point favoring the experimental intervention, (3) reporting of implied treatment advantage in trials without significant differences in primary end point, (4) existence of major discrepancies between registered and published primary outcomes, (5) number of patients whose outcomes would need to switch from a nonevent to an event to convert a significant difference in primary end point to nonsignificant, and (6) association with funding source.

**RESULTS** Of the 216 RCTs analyzed, 115 (53.2%) reported having commercial sponsorship. Most trials had 80% power to detect an estimated treatment effect of 30%, and 128 trials (59.3%) used composite primary end points. The median (interquartile range [IQR]) sample size was 502 (204-1702) patients, and the median (IQR) follow-up duration was 12 (1.0-14.4) months. Overall, 123 trials (57.0%) reported a statistically significant difference in the primary outcome favoring the experimental intervention; reporting strategies that implied an advantage were identified in 55 (65.5%) of 84 trials that reported nonsignificant differences. Commercial sponsorship was associated with a statistically significantly greater likelihood of favorable outcomes reporting (exponent of regression coefficient β, 2.80; 95% CI, 1.09-7.18; $P = .03$) and with the reporting of findings that are inconsistent with the trial results. Discrepancies between the registered and published primary outcomes were found in 82 trials (38.0%), without differences in trial sponsorship. A median (IQR) number of 5 (2.8-12.5) patients experiencing a different outcome would have change statistically significant results to nonsignificant. Commercial sponsorship was associated with a greater number of patients (exponent of regression coefficient β, 1.29; 95% CI, 1.00-1.66; $P = .04$).

**CONCLUSIONS AND RELEVANCE** These results suggest that contemporary RCTs of invasive cardiovascular interventions are relatively small and fragile, have short follow-up, and have limited power to detect large treatment effects. Commercial support appeared to be associated with differences in trial design, results, and reporting.

Invited Commentary
Related article
Supplemental content

**Author Affiliations:** Author affiliations are listed at the end of this article.

**Corresponding Author:** Mario Gaudino, MD, Department of Cardiothoracic Surgery, Weill Cornell Medicine, 525 E 68th St, New York, NY 10065 (mfg9004@med.cornell.edu).

*JAMA Intern Med*. doi:10.1001/jamainternmed.2020.1670
Published online June 1, 2020.

Evidence-based practice relies on high-quality randomized clinical trials (RCTs).[1] However, past analyses of RCTs in different fields have shown major limitations in study design and reporting.[2-4] Randomized clinical trials require substantial funding that is often provided by commercial drug and device manufacturers, which may incur large financial penalties from negative or neutral study findings.[5] Systematic reviews of RCTs in several fields, including cardiovascular medicine, have reported an association between commercially sponsored trials and a greater likelihood of reporting results that favor the sponsor-related experimental intervention.[5-7]At the individual-study level, this favorable finding has been associated with mechanisms such as unrepresentative patient selection, use of surrogate end points, and follow-up intervals that maximize the outcome of the sponsored intervention.[5-8]

For trials that find no difference between the treatment cohorts, biased reporting of nonsignificant outcomes that intentionally or unintentionally favor a specific treatment option has occurred,[8,9] although the relationship of such reports with the trial sponsors has not been investigated.

Most attributes of RCT design, conduct, results, and reporting can be quantified and analyzed using statistical methods. For example, the fragility of study findings (ie, Fragility Index) may be defined by the minimum number of patients whose outcomes, if switched from a nonevent to an event, would convert a statistically significant result to a nonsignificant result.[10,11] Lower values indicate less robust results. Similarly, the reporting of implied advantages of the experimental intervention despite nonsignificant differences in the primary outcome (ie, spin) may also be measured and quantified.[3,12,13]

The aim of this cross-sectional study was to describe the main characteristics of contemporary RCTs of invasive cardiovascular interventions, including study design, outcomes, and reporting, and to evaluate their association with the funding source.

We focused on cardiovascular research because of the public health burden presented by cardiovascular disease, which accounts for approximately 800 000 deaths annually and 1 in every 6 health care dollars spent in the US.[14] We also examined commercial sponsorship, which has been reported in more than 44% of cardiovascular trials.[5]

### Key Points

**Questions** What are the characteristics, including design, outcomes, and reporting, of contemporary randomized clinical trials of invasive cardiovascular interventions, and what is their association with the funding source?

**Findings** In this cross-sectional analysis of 216 trials of invasive cardiovascular interventions, most trials reported a statistically significant difference in the primary outcome favoring the experimental therapy and used reporting strategies that implied an advantage with nonsignificant differences in primary outcome (*spin*). Discrepancies between the registered and published primary outcomes were found in 38% of trials; a median of 5 patients experiencing a different outcome would have changed statistically significant results to nonsignificant; and commercially sponsored trials appeared to be associated with differences in trial design, results, and reporting.

**Meaning** Findings of this study suggest that contemporary trials in invasive cardiovascular treatments may be small and fragile, have short follow-up, and have limited power to detect large treatment effects.

## Methods

### Search Strategy

A comprehensive literature search was performed to identify coronary, vascular, and structural interventional cardiology and vascular and cardiac surgical RCTs published between January 1, 2008, to May 31, 2019. The search was performed by the medical librarian on our team (M.D.), and the search strategy, full set of key words, and MeSH (Medical Subject Headings) terms used are provided in eTable 1 in the Supplement. We used the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) reporting guideline (eAppendix in the Supplement). This study (PROSPERO registration: CRD42019133404) was approved by the National Institute of Health Research.

### Extraction of Trial Data

The following data were recorded for each RCT: information on the trial sponsors (names and affiliations), declared conflict of interest of first and last authors, journal of publication and impact factor (according to Clarivate Analytics[15]), year of publication, type of intervention, single-center or multicenter study, geographic locations of the participating centers, details of the primary outcome (definition of the outcome and composite or noncomposite end points), number of screened patients and percentage of screened patients enrolled in the trial, sample size, statistical power, treatment effect (relative risk reduction) size estimation used for sample size calculation, length or duration of the follow-up, number of events, number of patients lost to follow-up, number of citations on Web of Science[15] (until July 2019), and details of the primary analysis (intention to treat, as treated or per protocol, superiority, and equivalence or noninferiority).

To identify the primary outcomes, we examined the following sequentially: trial methods, design, primary aim, and outcome used in the sample size calculation. If no primary outcome was clearly identified (ie, explicitly specified in the article, in a sample size calculation, or in the primary study objectives), we deemed the trial ineligible for subsequent analyses and excluded it. We classified primary outcomes as major or minor clinical events using a classification scheme (eTable 2 in the Supplement). We classified composite outcomes that included 1 or more minor clinical events as minor clinical events.

The conflicts of interest of the first and last authors were identified from the disclosure statement published in the article or its supplementary material. For trials that listed co-first authors, the disclosures of both authors were considered. Author conflicts of interest were defined as any report

of consulting, advisory, or speaking fees or honoraria, stock ownership, affiliation, or employment by the study sponsor.

Two of us (I.H. and M.R.), blinded to the trial funding sources and study characteristics, independently screened the citations and extracted data using a method that has been described previously.[16-18] One of us (M.G.) resolved any discrepancies that emerged.

## Analysis of Trials

Randomized clinical trials were classified as commercially sponsored if they were industry initiated and sponsored or if they were investigator initiated that received commercial support. Trials were classified as noncommercially sponsored if they were investigator initiated and reported sponsorship from a local or federal government or hospital or if they had no sponsors. For commercially sponsored trials, the body of the articles, supplementary materials, and original trial designs were also analyzed for report of commercial or sponsor involvement in the trial design, conduct, analysis, or reporting.

### Classification of Trial Results

Consistent with previous reports,[6,19,20] the trials in this study were classified as favorable or unfavorable for the experimental therapy according to their results. A trial was classified as favorable if, for at least 1 primary outcome among those defined in the protocol, the experimental therapy was statistically significantly ($P < .05$ or a 95% CI that excluded the null value) better than the control therapy (in superiority trials), the experimental therapy was not substantially worse than the control therapy (in noninferiority trials), or the effects of the treatments differed by no more than the equivalence margin (in equivalence trials).

### Appraisal of Spin

In trials that reported a statistically nonsignificant difference in the primary outcome, we assessed the spin, the presence and amount of distortion or misrepresentation of value or advantage.[3,13] *Spin* is the use of specific reporting strategies to suggest that the experimental treatment is advantageous or noninferior despite a statistically nonsignificant difference in the primary outcome, or to distract the reader from statistically nonsignificant results.

For each selected article, 2 of us (I.H. and F.K.), blinded to the funding sources and study characteristics, independently read the full manuscripts and the online appendices to assess their contents, using a pretested and standardized data abstraction form as previously described.[3] One of us (M.G.) resolved any discrepancies. The presence of spin was assessed in the following sections of the manuscript: title; abstract results and abstract conclusions; and main text results, discussion, and conclusions.

On the basis of the described methods, the following strategies of spin were considered for superiority design trials: (1) focusing on secondary statistically significant results (within-group comparison, secondary outcomes, subgroup analyses, and modified population of analyses), (2) interpreting statistically nonsignificant results for the primary outcomes as showing treatment equivalence or comparable effectiveness, and (3) claiming or emphasizing the advantage of the experimental treatment despite statistically nonsignificant results.[3] For noninferiority design trials, spin was considered when trials (1) claimed or emphasized noninferiority despite not establishing noninferiority boundaries or when data were inconclusive, and (2) focused on other results (such as secondary outcomes or information from other studies) when noninferiority was not established, inconclusive, or unclear.[13] Other spin strategies that could not be classified according to this scheme were systematically recorded and classified as others.[3,13] (eTable 3 in the Supplement). The extent of spin is defined as the number of sections in the entire article with spin.

### Assessment of Discrepancy Between the Registered and Published Primary Outcomes

For each RCT, we identified the registration number listed in the article (ie, registration with ClinicalTrials.gov, ISRCTN register, or country-specific registries). We considered only trials that were prospectively registered and with a clear description of the primary outcome. A major discrepancy between the registered and the published primary outcomes was identified if the outcomes were different or assessed at different time points. Consistent with previous definitions,[21,22] we defined major discrepancies as (1) a prespecified primary outcome in the trial registration protocol that was reported as a secondary outcome in the final published article, (2) a published primary outcome that was described as a secondary outcome in the registry, (3) prespecified primary outcomes in the trial registration that were not reported in the published article, (4) a new primary outcome that was introduced in the published article, and (5) a difference between the timing of assessment of the primary outcome in the registered protocol and in the published article.[22]

Two of us (I.H. and M.R.), blinded to the trial funding sources, analyzed the trials. All discrepancies were discussed to obtain consensus, and if needed, the article in question was discussed with our first author (M.G.).

### Calculation of the Fragility Index

For trials with a superiority design that reported at least 1 statistically significant dichotomous primary outcome ($P < .05$ or a 95% CI that excluded the null value), we quantified how robust the results were by using the Fragility Index described by Walsh et al[11] and applied by Gaudino et al[23] for cardiovascular clinical trials. The Fragility Index is defined as the number of patients whose status would need to switch from a nonevent to an event to render a statistically significant difference not significant. The results for each outcome were entered in a 2 × 2 contingency table, and then the *P* value for each outcome was calculated with a 2-sided Fisher exact test. Single participants were iteratively shifted 1 at a time in the lower-incidence treatment group from nonevent to event, and the *P* value for the 2 × 2 table was recalculated. The Fragility Index for an outcome equaled the smallest number of patients required to turn the recalculated *P* value nonsignificant ($P \geq .05$). Lower values indicate less robust results.

### Statistical Analysis

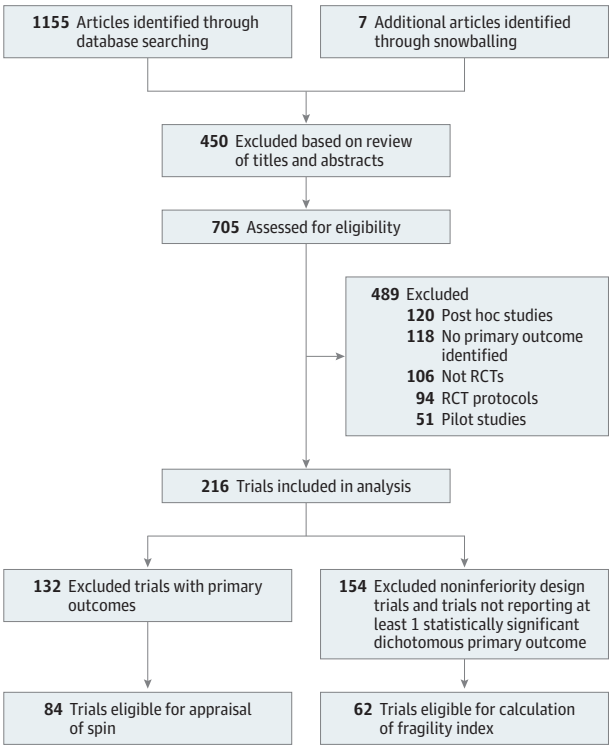
Categorical variables were reported as counts and percentages. All continuous variables were not normally distributed by visual inspection of the data and Shapiro-Wilk normality test and were reported as medians with interquartile ranges (IQRs).

Descriptive analyses were performed to compare commercially and noncommercially sponsored groups. For continuous variables, the Mann-Whitney test was used to compare the 2 groups. Categorical variables were compared using $\chi^2$ and Fisher exact tests. Regression analyses were conducted to identify the factors associated with the (1) selection of major clinical events as the primary end point (using binary logistic regression), (2) favorable trial outcomes (using binary logistic regression), (3) extent of spin (using multivariable ordinal logistic regression), (4) discrepancy between registered and published primary outcomes (using binary logistic regression), and (5) Fragility Index (using Poisson regression). Covariates for the models were selected according to their relevance to trial methods and included trial sponsorship, author conflict of interest, type of intervention in experimental intervention (nonsurgical vs surgical), power, sample size, year of trial publication, scale of the trial (single center vs multicenter), and type of end points. In the multivariable regression models, each variable was adjusted for covariates that were statistically significantly associated with the variable on univariate analyses. Multicollinearity was assessed using the variance inflation factor. Nonlinearity was assessed using the residual plot of fitted values vs the residuals, and by testing quadratic terms of the continuous variables in the models. Nonlinear continuous variables were categorized on the basis of their median value.

Results were reported as the exponent of the regression coefficient (Exp [β]) and its 95% Wald CI (ie, interpreted as the odds ratio for logistic regression, with values >1 suggesting greater association of the covariate with the outcome). Sensitivity analysis was performed to compare trials that were industry initiated and supported with trials that were noncommercially sponsored, excluding investigator-initiated studies that received commercial support.

Two-sided significance testing was used, and $P < .05$ was considered statistically significant without adjustment for multiple testing. All analyses were performed with SPSS, version 24 (IBM), and R, version 3.4.2 within RStudio (R Foundation for Statistical Computing).

Figure 1. CONSORT Flow Diagram




RCT indicates randomized clinical trial.

## Results

Of the 1155 articles screened, 216 were eligible for analysis (**Figure 1** and **Table 1**). The outcomes analyzed for the trials are summarized in **Table 2**. Of the 216 articles, 170 (78.7%) were multicenter trials and 125 (57.9%) originated from Europe, 26 (12.0%) from North America, and 23 (10.6%) from Asia. Most trials were designed to have 80% power to detect a treatment effect of 30%. Most trials used a superiority design (n = 171 [79.2%]) and intention to treat as the main analysis (n = 196 [90.7%]). One hundred eighty-one trials (83.8%) were registered before the start of enrollment, and 35 (16.2%) were not prospectively registered. One hundred twenty-eight trials (59.3%) used composite end points as the primary outcome, and 115 (53.2%) used major clinical events as the primary outcome. The median (IQR) percentage of screened patients enrolled was 62.4% (26.9%-95.5%), and the median (IQR) sample size was 502 (204-1702) patients. The median (IQR) follow-up duration was 12 (1.0-14.4) months, and the median (IQR) percentage of patients lost to follow-up was 0.5% (0.0%-2.5%) (Table 1).

More than half of the trials (n = 115 [53.2%]) were commercially sponsored, of which 37 (17.1%) were investigator-initiated studies that received commercial support; 21 trials (9.7%) specified commercial involvement in study design, conduct, or reporting (Table 1). Sixty-seven (31.0%) of the commercially sponsored trials reported a conflict of interest of the first and/or last author with the study sponsors.

Compared with noncommercially sponsored studies, commercially sponsored trials were more often multicenter (101 of 115 [87.8%] vs 69 of 101 [68.3%]; $P = .001$) and used a noninferiority design (30 of 115 [26.1%] vs 15 of 101 [14.9%]; $P = .04$). No significant difference was found in the treatment effect size used for sample size calculation (median [IQR], 33.0% [25.0%-50.0%] vs 40.0% [25.0%-50.8%]; $P = .18$) and the use of composite outcomes (median [IQR], 75 of 115 [65.2%] vs 53 of 101 [52.5%]; $P = .07$) between commercially sponsored and noncommercially sponsored trials (Table 1).

Table 1. Summary of the Characteristics of the Trials Analyzed

| Characteristic | No. (%) | | | |
|---|---|---|---|---|
| | | RCT | | |
| | Overall | Commercially sponsored | Noncommercially sponsored | *P* value |
| Total No. of trials | 216 | 115 (53.2) | 101 (46.8) | NA |
| Prospective registration of trial | 181 (83.8) | 102 (88.7)[a] | 79 (78.2)[a] | .04 |
| Superiority design | 171 (79.2) | 85 (73.9)[a] | 86 (85.1)[a] | .04 |
| Noninferiority design | 45 (20.8) | 30 (26.1)[a] | 15 (14.9)[a] | .04 |
| Intention-to-treat analysis | 196 (90.7) | 104 (90.4) | 92 (91.1) | .45 |
| Multicenter trial | 170 (78.7) | 101 (87.8)[a] | 69 (68.3)[a] | .001 |
| Use of composite primary outcome | 128 (59.3) | 75 (65.2) | 53 (52.5) | .07 |
| Estimated treatment effect, median (IQR), % | 33.0 (25.0-50.0) | 33.0 (25.0-50.0) | 40.0 (25.0-50.8) | .18 |
| Power, median (IQR), % | 80 (80-90) | 80 (80-88) | 80 (80-90) | .34 |
| No. of screened patients, median (IQR) | 1710 (447-3891) | 1817 (447-3740) | 1576 (439-4113) | .82 |
| Sample size, median (IQR) | 502 (204-1702) | 800 (353-2032)[a] | 302 (140-788)[a] | <.001 |
| Screened patients included in sample size, median (IQR), % | 62.4 (26.9-95.5) | 75.4 (33.3-97.1)[a] | 48.8 (19.0-81.9)[a] | .02 |
| Sample size lost to follow-up, median (IQR), % | 0.5 (0.0-2.5) | 1.0 (0.0-2.5)[a] | 0.1 (0.0-2.4)[a] | .01 |
| Duration of follow-up, median (IQR), mo | 12.0 (1.0-14.4) | 12.0 (1.0-12.0) | 12.0 (1.0-18.4) | .67 |
| Commercial involvement in trial conduct and reporting | 21 (9.7) | 21 (18.3) | NA | NA |
| Conflict of interest of first or last author | 67 (31.0) | 67 (58.2)[a] | NA | NA |
| Consulting fees | 16 (23.9) | 16 (23.9) | NA | NA |
| Advisory fees | 7 (10.4) | 7 (10.4) | NA | NA |
| Speaking fees | 12 (17.9) | 12 (17.9) | NA | NA |
| Honoraria | 10 (14.9) | 10 (14.9) | NA | NA |
| Stock ownership | 4 (6.0) | 4 (6.0) | NA | NA |
| Affiliation with the study sponsor | 29 (43.4) | 29 (43.4) | NA | NA |
| Employment by the study sponsor | 3 (4.5) | 3 (4.5) | NA | NA |
| Grant support | 25 (37.3) | 25 (37.3) | NA | NA |
| Impact factor of journal, median (IQR) | 18.8 (9.9-53.3) | 27.1 (16.8-79.3)[a] | 16.8 (4.9-53.3)[a] | .001 |
| Year of publication, median (IQR) | 2013 (2011-2015) | 2013 (2011-2015) | 2013 (2011-2015) | .46 |
| No. of citations, median (IQR) | 112 (38-240) | 192 (65-349)[a] | 100 (32-251)[a] | .001 |
| Location | | | | |
| Multicontinental | 36 (16.7) | 28 (24.3)[a] | 8 (7.9)[a] | .001 |
| Asia | 23 (10.6) | 9 (7.8) | 14 (13.9) | .15 |
| Europe | 125 (57.9) | 58 (50.4)[a] | 67 (66.3)[a] | .02 |
| North America | 26 (12.0) | 17 (14.7) | 9 (8.9) | .18 |
| South America | 4 (1.9) | 2 (1.7) | 2 (2.0) | .89 |
| Not reported | 2 (0.9) | 1 (0.9) | 1 (1.0) | .93 |
| Surgical trial | 64 (29.6) | 23 (20.0)[a] | 41 (40.6)[a] | <.001 |
| Medical trial | 18 (8.3) | 0 (0.0)[a] | 18 (17.8)[a] | <.001 |
| Interventional cardiology trial | 175 (81.0) | 99 (78.2)[a] | 76 (75.2)[a] | .04 |
| Percutaneous coronary intervention | 127 (58.8) | 69 (60.0) | 58 (57.4) | .70 |
| Transcatheter aortic valve replacement | 6 (2.7) | 5 (4.3) | 1 (1.0) | .13 |
| Carotid interventions | 26 (12.0) | 16 (13.9) | 10 (9.9) | .36 |
| Lower-limb interventions | 5 (2.3) | 3 (2.6) | 2 (2.0) | .89 |
| Percutaneous mitral interventions | 3 (1.4) | 3 (2.6) | 0 (0.0) | .37 |
| Endovascular aneurysm repair | 8 (3.7) | 3 (2.6) | 5 (5.0) | .36 |

Abbreviations: IQR, interquartile range; NA, not applicable; RCT, randomized clinical trials.

[a] $P < .05$ for commercially vs noncommercially sponsored trials.

Commercially sponsored trials vs noncommercially sponsored trials screened similar numbers of patients (median [IQR], 1817 [447-3740] vs 1576 [439-4113]; $P = .82$), included more screened patients (median [IQR], 75.4% [33.3%-97.1%] vs 48.8% [19.0%-81.9%]; $P = .02$), and had larger sample sizes (median [IQR], 800 [353-2032] vs 302 [140-788]; $P < .001$). The median (IQR) duration of the follow-up was similar for commercially and noncommer-

Table 2. Summary of Outcomes Analyzed for the Trials

| Variable | No. (%) | | | P value |
|---|---|---|---|---|
| | Overall | RCT | | |
| | | Commercially sponsored | Noncommercially sponsored | |
| Favorable outcome | 123 (57.0) | 74 (64.3)[a] | 49 (48.5)[a] | .02 |
| Use of major clinical events as primary outcome | 115 (53.2) | 57 (49.6) | 58 (57.4) | .27 |
| Major discrepancy between published and registered primary outcome | 82 (38.0) | 41 (35.7) | 41 (40.6) | .13 |
| Total No. of trials evaluated for spin | 84 (100.0) | 36 (42.9) | 48 (57.1) | NA |
| Spin present | 55 (65.5) | 29 (80.6)[a] | 26 (54.2)[a] | .02 |
| Extent of spin | | | | |
| None | 29 (34.5) | 7 (19.4)[a] | 22 (45.8)[a] | .01 |
| In 1 section other than the conclusion | 14 (16.7) | 9 (25.0) | 5 (10.4) | .07 |
| In the conclusion section only | 7 (8.3) | 4 (11.1) | 3 (6.3) | .42 |
| In 2 sections | 14 (16.7) | 5 (13.9) | 9 (18.8) | .55 |
| In all sections | 0 | 0 | 0 | NA |
| Total No. of trials evaluated for Fragility Index | 62 | 34 (54.8) | 28 (45.2) | NA |
| Fragility Index, median (IQR) | 5.0 (2.8-12.5) | 5.0 (2.8-12.5) | 4.5 (2.3-14.0) | NA |

Abbreviations: IQR, interquartile range; RCT, randomized clinical trial.

[a] $P < .05$ for commercially vs noncommercially sponsored trials.

Table 3. Results of Multivariable Regression

| Variable[a] | Use of major clinical event as primary end point, Exp(β) (95% CI) | P value | Favorable outcome, Exp(β) (95% CI) | P value | Extent of spin, Exp(β) (95% CI) | P value | Discrepancy between registered and published outcomes, Exp(β) (95% CI) | P value | Fragility index, Exp(β) (95% CI) | P value |
|---|---|---|---|---|---|---|---|---|---|---|
| Commercial sponsorship | 0.31 (0.12-0.85) | .02[b] | 2.80 (1.09-7.18) | .03[b] | 4.64 (1.05-20.54) | .04[b] | 0.37 (0.14-1.01) | .05 | 1.29 (1.00-1.66) | .04[b] |
| Author conflict of interest | 1.59 (0.65-3.88) | .30 | 2.37 (0.91-6.15) | .07 | 2.14 (0.56-8.23) | .26 | 1.05 (0.41-2.69) | .92 | 0.76 (0.61-0.95) | .01[b] |
| Type of intervention in the experimental therapy (nonsurgical vs surgical) | 0.52 (0.24-1.14) | .10 | 2.94 (1.13-7.63) | .02[b] | 0.88 (0.52-1.49) | .62 | 0.63 (0.25-1.57) | .32 | 0.57 (0.47-0.69) | <0.001[b] |
| Power >80% | 0.77 (0.39-1.54) | .45 | 0.77 (0.31-1.90) | .56 | 1.33 (0.41-4.28) | .63 | 1.65 (0.65-4.22) | .29 | 0.84 (0.67-1.06) | .13 |
| Sample size >500 | 1.34 (0.21-8.62) | .75 | 1.44 (0.25-8.30) | .68 | 0.44 (0.03-7.53) | .57 | 1.81 (0.29-11.43) | .52 | 1.38 (0.95-2.02) | .09 |
| Year of trial publication[c] | 0.99 (0.90-1.08) | .78 | 1.04 (0.95-1.14) | .40 | 0.92 (0.81-1.05) | .20 | 0.94 (0.85-1.04) | .24 | 1.05 (1.02-1.08) | .001[b] |
| Multicenter trial | 1.05 (0.23-4.83) | .95 | 1.30 (0.32-5.26) | .71 | 3.74 (0.14-10.72) | .43 | 0.62 (0.14-2.82) | .53 | 2.02 (1.55-2.65) | <.001[b] |
| Composite end point | 2.42 (0.92-6.36) | .07 | 1.23 (0.48-3.16) | .66 | 0.39 (0.11-1.42) | .15 | 1.84 (0.66-5.08) | .24 | 0.83 (0.63-1.11) | .21 |

Abbreviation: Exp(β), exponent of regression coefficient (interpreted as odds ratio for logistic regression).

[a] For multivariable regression, each variable was adjusted for trial method–related covariates that were associated with the variable on univariate analyses.

[b] $P < .05$.

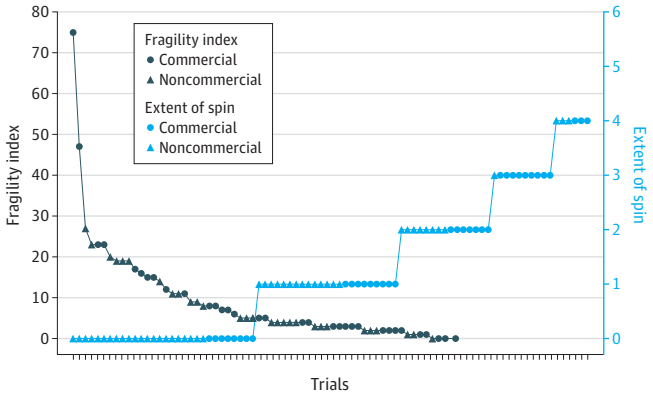
[c] Year of trial publication was associated with the fragility index only on univariate analysis; the results presented are for the univariate model.

cially sponsored trials (12 months for both; $P = .67$). The median (IQR) percentage of patients lost to follow-up was higher for commercially sponsored trials (1.0% [0.0%-2.5%] vs 0.1% [0.0%-2.4%]; $P = .01$). Commercially sponsored trials were published in journals with higher impact factors (median [IQR] impact factor, 27.1 [16.8-79.3] vs 16.8 [4.9-53.3]; $P = .001$) and had a greater number of citations (median [IQR] No. of citations, 192 [65-349] vs 100 [32-251]; $P = .001$) compared with noncommercially sponsored trials (Table 1).

Of the 115 trials that used major clinical events as primary outcomes, 57 (49.6%) were commercially sponsored and 58 (57.4%) were noncommercially sponsored. At multivariable regression analysis, commercial sponsorship was inversely associated with the use of major clinical events as a primary end point (Exp[β], 0.31; 95% CI, 0.12-0.85; $P = .02$) (**Table 3** and eTable 5 in the Supplement).

Overall, 123 trials (57.0%) reported results that favored the experimental intervention. After adjusting for differences in trial characteristics, commercial sponsorship was associated

Figure 2. The Fragility Index and Extent of Spin According to Funding Source




The Fragility Index is the number of patients needed to switch the significance of trial results. The extent of spin is the number of article sections that use specific reporting strategies to suggest that the experimental treatment was advantageous despite a nonsignificant difference in the primary outcome.

with results that favored the experimental therapy (Exp[β], 2.80; 95% CI, 1.09-7.18; *P* = .03) (Table 3 and eTable 5 in the Supplement).

Of the 84 trials that reported nonsignificant differences in the primary outcome, spin was identified in 55 (65.5%), including 29 (80.6%) of 36 commercially sponsored trials and 26 (54.2%) of 48 noncommercially sponsored trials (*P* = .02) (Table 2 and eTable 6 in the Supplement). The distribution of the extent of spin is shown in the eFigure in the Supplement. After adjusting for differences in trial characteristics, we found that commercial sponsorship was associated with significantly greater extent of spin (Exp[β], 4.64; 95% CI, 1.05-20.54; *P* = .04) (Table 3 and eTable 5 in the Supplement).

One hundred eighty-one trials (83.8%) had primary outcomes prospectively reported in a clinical registry. Commercially sponsored trials registered their protocol more often than noncommercially sponsored trials did (102 of 115 [88.7%] vs 79 of 101 [78.2%]; *P* = .04). Of the registered trials, 82 (38.0%) had at least 1 major discrepancy between the registered and published primary outcomes, and no significant differences were observed between commercially and noncommercially sponsored trials (41 of 115 [35.7%] vs 41 of 101 [40.6%]; *P* = .13), even after adjusting for trial characteristics (Table 2, Table 3, and eTable 5 in the Supplement).

The Fragility Index of trial results was analyzed in 62 trials that reported a statistically significant difference in a dichotomous primary outcome (**Figure 2**, eFigure, and eTable 7 in the Supplement). The Fragility Index or the median (IQR) number of patients whose outcomes, if switched from a nonevent to an event, would render a statistically significant result not significant was 5.0 (2.8-12.5) for commercially sponsored trials and 4.5 (2.3-14.0) for noncommercially sponsored trials (Table 2). At multivariable regression analysis, commercial sponsorship was associated with a greater number of patients needed to change the significance of trial outcomes (Exp[β], 1.29; 95% CI, 1.00-1.66; *P* = .04) (Table 3). In 4 trials (6%) assessed for robustness, the change in condition of only 1 patient was needed to switch the statistical significance.

The sensitivity analysis confirmed the results of the main analysis (eTable 8 in the Supplement). Further details and the full reference list are provided in eAppendix, eTable 4, and eReferences in the Supplement.

## Discussion

We analyzed 216 RCTs of invasive cardiovascular interventions published from 2008 to 2019. Most of the trials were multicenter and used a superiority design and an intention-to-treat analysis. Trials were generally designed with limited power (80%) to detect relatively large treatment effects (30%). Overall, 59.3% of trials used composite end points as a primary outcome, whereas only 53.2% used major clinical events as their primary outcome. Trials enrolled 62.4% of the screened patients, and the median sample size was small (502 patients). The median follow-up was 12 months, with a small percentage of patients lost during follow-up (0.5%). Thirty-five trials (16.2%) were not prospectively registered. Most trials (123 trials [57.0%]) reported a favorable outcome, and a high percentage of trials that reported neutral or unfavorable outcomes (65.5%) interpreted and published findings that were inconsistent with the trial results.

More than half of the trials reported commercial support. Commercially sponsored trials had distinctive characteristics, including multicenter design, more frequent protocol registration, use of noninferiority analysis, larger sample size, inclusion of higher percentage of screened patients, higher number of patients lost to follow-up, and publication in journals with higher impact factors and greater number of citations after publication.

Commercial sponsorship was associated with statistically significant findings favoring the sponsored treatment. In trials without statistically significant primary outcomes, commercial support was associated with reporting strategies, suggesting that the sponsor treatment was advantageous.

In addition, in 82 trials (38.0%), at least 1 major discrepancy existed between the registered and published primary outcomes, regardless of trial sponsorship. The results of most contemporary trials of invasive cardiovascular interventions were not robust, with a median of 5 patients whose condition

needed to switch from a nonevent to an event to convert a significant difference in the primary outcome to a nonsignificant outcome. Commercially sponsored trials were more robust than noncommercially sponsored trials.

Commercial sponsorship in cardiovascular research was systematically evaluated in a survey of trials published between 2000 and 2005, in which a subgroup analysis of 38 device trials showed that 23 (76.7%) of 30 trials funded by for-profit entities reported favorable outcomes favoring the sponsored device compared with 4 (50%) of 8 device trials funded by nonprofit entities ($P$ = .07).[5] Ridker and Torres[5] restricted their analysis to trials published in *JAMA*, the *Lancet*, or *NEJM*, which potentially increased the risk of publication bias, and the subgroup analysis was likely underpowered. Because trials from a much broader range of publications are included in the development of cardiovascular consensus guideline recommendations, we reviewed all relevant published trials independent of the journal of publication. Our findings confirm that (1) more than half of the trials of invasive cardiovascular interventions published between 2008 and 2019 received some or full funding from commercial entities and (2) trials with commercial funding were significantly more likely to report a favorable outcome favoring the sponsored intervention.

Explanations for similar findings in other fields have focused on bias and differential quality in commercially sponsored trial design and reporting.[3-7,10-12,14,16-18] In the present study, we found no difference in estimated treatment effect, length of follow-up, use of composite or clinically significant outcomes, or outcome modification compared with the published protocol between commercially and noncommercially sponsored trials. The more frequent use of noninferiority design, especially when coupled with the higher number of patients lost to follow-up, or the more subtle differences in trial characteristics that were not captured in this cross-sectional study may explain the association between commercial sponsorship and favorable findings. Alternatively, commercially sponsored trials are larger, more inclusive, and more often multicenter than noncommercially sponsored studies, which may make their results more generalizable. Commercially sponsored trials are more often preregistered, more robust, and more authoritative in the research world, as evidenced by publication in journals with higher impact factor and by more frequent citations.

In more than 65% of the studies with no statistically significant difference in primary outcomes, we found evidence of interpretation bias. Spin was significantly associated with commercial sponsorship.

As mentioned, at least 1 major discrepancy existed between the registered and published outcomes in 38.0% of the trials. This percentage is higher than the 31% to 33% reported in analyses of a broad range of medical conditions and interventions.[22,24] Changes in primary outcomes have been associated with larger intervention effect[22] and may suggest selective reporting based on statistical significance.

The RCTs that we analyzed were not particularly robust. In 4 trials, the change in condition of only 1 patient was needed to switch the statistical significance. This finding is concerning given the substantial role that RCT results play in federal device approvals, payer criteria, and clinical consensus guidelines. Power and sample size are not the sole determinants of a trial's robustness; the event rate is also a factor. Consequently, in trials with thousands of patients but low event rates, statistical significance has been reported to hinge on the outcome of only 1 patient.[11] Given that it may not be feasible to consistently fund and design invasive interventional trials with more robust results, a case can be made for specifying the number of patients needed to change the statistical significance of the outcomes in the primary reports of these trials as well as in the meta-analyses and consensus guidelines that cite them.

### Limitations

This study has several limitations. First, this cross-sectional study did not capture unpublished trials and is, therefore, subject to potential publication bias whereby studies reporting unfavorable outcomes are less likely to be published. Second, this study only shows the association of funding sources with reporting outcomes and does not demonstrate causality, nor is it explanatory. Third, the focus was on trials of invasive cardiovascular interventions and thus may not be generalizable to other trials, populations, or specialties. In addition, the sample size may be insufficient to detect significant differences in outcomes, particularly for subgroup analyses.

## Conclusions

Current RCTs of invasive cardiovascular interventions appeared to be relatively small and have short follow-up. Most of the trials had limited power to detect relatively large treatment effects. A substantial proportion of these trials was not prospectively registered and used minor clinical events as the primary outcome. In addition, major inconsistency between the registered and the published outcomes was found in a concerning number of trials, and only 5 patients were needed to change the statistical significance of the primary outcome for most of the trials. Commercial sponsorship supported more than half of the trials and was associated with differences in study design, results, and reporting. We believe these findings should inform efforts to improve RCT design, reporting, and consensus guidelines.

ARTICLE INFORMATION

**Accepted for Publication:** April 14, 2020.

**Published Online:** June 1, 2020. doi:10.1001/jamainternmed.2020.1670

**Author Affiliations:** Department of Cardiothoracic Surgery, Weill Cornell Medicine, New York, New York (Gaudino, Hameed, Rahouma, Khan, Girardi); Schulich Heart Centre, Sunnybrook Health Sciences Centre, University of Toronto, Toronto, Ontario, Canada (Tam, Fremes); Department of Medico-Surgical Sciences and Biotechnologies, Sapienza University of Rome, Rome, Italy (Biondi-Zoccai); IRCCS Neuromed, Pozzilli, Italy (Biondi-Zoccai); Samuel J. Wood Library and C. V. Starr Biomedical Information Centre, Weill Cornell Medicine, New York, New York (Demetres); Division

of General Internal Medicine, Weill Cornell Medical College, New York, New York (Charlson); Division of Cardiac Surgery, University of Ottawa Heart Institute, Ottawa, Ontario, Canada (Ruel); Fondazione Policlinico Universitario A. Gemelli IRCCS, Rome, Italy (Crea); Department of Cardiovascular and Thoracic Sciences, Università Cattolica de Sacro Cuore, Roma, Italy (Crea); Department of Cardiovascular Surgery, Charite, Berlin, Germany (Falk); Department of Cardiothoracic and Vascular Surgery, German Heart Center Berlin, Berlin, Germany (Falk); German Center for Cardiovascular Research, Berlin, Germany (Falk); Department of Health Science and Technology, Swiss Federal Institute of Technology, Zurich, Switzerland (Falk); Department of Cardiac Surgery, Smidt Heart Institute, Cedars-Sinai Medical Center, Los Angeles, California (Chikwe).

**Author Contributions:** Dr Gaudino had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
*Concept and design:* Gaudino, Hameed, Rahouma, Biondi-Zoccai, Crea, Girardi, Fremes, Chikwe.
*Acquisition, analysis, or interpretation of data:* Gaudino, Hameed, Rahouma, Khan, Tam, Biondi-Zoccai, Demetres, Charlson, Ruel, Falk, Chikwe.
*Drafting of the manuscript:* Gaudino, Hameed, Rahouma, Khan, Demetres, Ruel, Chikwe.
*Critical revision of the manuscript for important intellectual content:* Gaudino, Hameed, Rahouma, Tam, Biondi-Zoccai, Charlson, Ruel, Crea, Falk, Girardi, Fremes, Chikwe.
*Statistical analysis:* Gaudino, Hameed, Rahouma, Khan, Tam, Biondi-Zoccai, Falk.
*Obtained funding:* Girardi.
*Administrative, technical, or material support:* Hameed, Rahouma, Demetres, Girardi.
*Supervision:* Gaudino, Hameed, Rahouma, Charlson, Falk, Girardi, Fremes.

**Conflict of Interest Disclosures:** Dr Tam reported receiving Canadian Institutes of Health Research (CIHR) Fellowship. Dr Falk reported receiving grants from Abbott, Medtronic, Edwards Lifescience, and Boston Scientific during the conduct of the study; grants from Abbott, Medtronic, Edwards Lifescience, Boston Scientific, and the ETH outside the submitted work; and other from Biotronic, Berlin Heart, and Norvartis Pharma; receiving patents issued by the US (Patent No. 6,468,265, 6,398,726, and 08,0265), a patent issued by the German Utility (Patent No. 299 23 269.7) and Deutsche Gebrauchsmuster Amt. (AkZ: 20 2015 009 444), and patents issued by the European Union (Patent No. EP 12175533.4, EP16158181.4-165, and EP15 17 7575.6); and serving on the advisory boards of Medtronic, Berlin Heart, Novartis, and Boston Scientific. Dr Fremes reported being supported in part the Bernard S Goldman Chair in Cardiovascular Surgery. Dr Chikwe reported receiving institutional support for speaker activities from Edwards Lifesciences and Medtronic. No other disclosures were reported.

## REFERENCES

1. Neumann F-J, Sousa-Uva M, Ahlsson A, et al; ESC Scientific Document Group. 2018 ESC/EACTS Guidelines on myocardial revascularization. *Eur Heart J*. 2019;40(2):87-165. doi:10.1093/eurheartj/ehy394

2. Docherty KF, Campbell RT, Jhund PS, Petrie MC, McMurray JJV. How robust are clinical trials in heart failure? *Eur Heart J*. 2017;38(5):338-345. doi:10.1093/eurheartj/ehw427

3. Boutron I, Dutton S, Ravaud P, Altman DG. Reporting and interpretation of randomized controlled trials with statistically nonsignificant results for primary outcomes. *JAMA*. 2010;303(20):2058-2064. doi:10.1001/jama.2010.651

4. Hopewell S, Dutton S, Yu L-M, Chan A-W, Altman DG. The quality of reports of randomised trials in 2000 and 2006: comparative study of articles indexed in PubMed. *BMJ*. 2010;340:c723. doi:10.1136/bmj.c723

5. Ridker PM, Torres J. Reported outcomes in major cardiovascular clinical trials funded by for-profit and not-for-profit organizations: 2000-2005. *JAMA*. 2006;295(19):2270-2274. doi:10.1001/jama.295.19.2270

6. Flacco ME, Manzoli L, Boccia S, et al. Head-to-head randomized trials are mostly industry sponsored and almost always favor the industry sponsor. *J Clin Epidemiol*. 2015;68(7):811-820. doi:10.1016/j.jclinepi.2014.12.016

7. Lundh A, Lexchin J, Mintzes B, Schroll JB, Bero L. Industry sponsorship and research outcome. *Cochrane Database Syst Rev*. 2017;2:MR000033. doi:10.1002/14651858.MR000033.pub3

8. Ruel M, Falk V, Farkouh ME, et al. Myocardial revascularization trials. *Circulation*. 2018;138(25):2943-2951. doi:10.1161/CIRCULATIONAHA.118.035970

9. Atianzar K, Zhang M, Newhart Z, Gafoor S. Why did COAPT win while MITRA-FR failed? Defining the appropriate patient population for MitraClip. *Interv Cardiol*. 2019;14(1):45-47. doi:10.15420/icr.2018.40.1

10. Feinstein AR. The unit fragility index: an additional appraisal of "statistical significance" for a contrast of two proportions. *J Clin Epidemiol*. 1990;43(2):201-209. doi:10.1016/0895-4356(90)90186-S

11. Walsh M, Srinathan SK, McAuley DF, et al. The statistical significance of randomized controlled trial results is frequently fragile: a case for a fragility index. *J Clin Epidemiol*. 2014;67(6):622-628. doi:10.1016/j.jclinepi.2013.10.019

12. Fletcher RH, Black B. "Spin" in scientific writing: scientific mischief and legal jeopardy. *Med Law*. 2007;26(3):511-525.

13. Hernandez AV, Pasupuleti V, Deshpande A, Thota P, Collins JA, Vidal JE. Deficient reporting and interpretation of non-inferiority randomized clinical trials in HIV patients: a systematic review. *PLoS One*. 2013;8(5):e63272. doi:10.1371/journal.pone.0063272

14. Centers for Disease Control and Prevention. Vital Signs. Preventable deaths from heart disease & stroke. Published September 3, 2013. Accessed August 8, 2019. https://www.cdc.gov/vitalsigns/heartdisease-stroke/index.html

15. Web of Science. Clarivate. Accessed June 20, 2019. https://clarivate.com/products/web-of-science/

16. Djulbegovic B, Lacevic M, Cantor A, et al. The uncertainty principle and industry-sponsored research. *Lancet*. 2000;356(9230):635-638. doi:10.1016/S0140-6736(00)02605-2

17. Kjaergard LL, Als-Nielsen B. Association between competing interests and authors' conclusions: epidemiological study of randomised clinical trials published in the BMJ. *BMJ*. 2002;325(7358):249. doi:10.1136/bmj.325.7358.249

18. Als-Nielsen B, Chen W, Gluud C, Kjaergard LL. Association of funding and conclusions in randomized drug trials: a reflection of treatment effect or adverse events? *JAMA*. 2003;290(7):921-928. doi:10.1001/jama.290.7.921

19. Hopewell S, Loudon K, Clarke MJ, Oxman AD, Dickersin K. Publication bias in clinical trials due to statistical significance or direction of trial results. *Cochrane Database Syst Rev*. 2009;(1):MR000006. doi:10.1002/14651858.MR000006.pub3

20. Ioannidis JP. Effect of the statistical significance of results on the time to completion and publication of randomized efficacy trials. *JAMA*. 1998;279(4):281-286. doi:10.1001/jama.279.4.281

21. Chan A-W, Hróbjartsson A, Haahr MT, Gøtzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA*. 2004;291(20):2457-2465. doi:10.1001/jama.291.20.2457

22. Chen T, Li C, Qin R, et al. Comparison of clinical trial changes in primary outcome and reported intervention effect size between trial registration and publication. *JAMA Netw Open*. 2019;2(7):e197242. doi:10.1001/jamanetworkopen.2019.7242

23. Gaudino M, Hameed I, Biondi-Zoccai G, et al. Systematic evaluation of the robustness of the evidence supporting current guidelines on myocardial revascularization using the fragility index. *Circ Cardiovasc Qual Outcomes*. 2019;12(12):e006017. doi:10.1161/CIRCOUTCOMES.119.006017

24. Jones CW, Keil LG, Holland WC, Caughey MC, Platts-Mills TF. Comparison of registered and published outcomes in randomized controlled trials: a systematic review. *BMC Med*. 2015;13:282. doi:10.1186/s12916-015-0520-3