VOLUME 32 · NUMBER 36 · DECEMBER 20 2014

JOURNAL OF CLINICAL ONCOLOGY

ORIGINAL REPORT

Impact of Spin in the Abstracts of Articles Reporting Results of Randomized Controlled Trials in the Field of Cancer: The SPIIN Randomized Controlled Trial

Isabelle Boutron, Douglas G. Altman, Sally Hopewell, Francisco Vera-Badillo, Ian Tannock and Philippe Ravaud

Isabelle Boutron, Sally Hopewell, and Philippe Ravaud, Methods of Therapeutic Evaluation of Chronic Diseases Team, Epidemiology and Biostatistics Sorbonne Paris Cité Research Centre, Unite Mixte de Recherche 1153, L'Institut National de la Santé et de la Recherche Médicale; Isabelle Boutron and Philippe Ravaud, Assistance Publique des Hôpitaux de Paris, Hôpital Hôtel Dieu; Isabelle Boutron and Philippe Ravaud, Paris Descartes University, Sorbonne Paris Cité; Isabelle Boutron, Sally Hopewell, and Philippe Ravaud, French Cochrane Centre, Paris, France; Douglas G. Altman and Sally Hopewell, Centre for Statistics in Medicine, University of Oxford, Oxford, United Kingdom; Francisco Vera-Badillo and Ian Tannock, Princess Margaret Cancer Centre and University of Toronto, Toronto, Ontario, Canada; Philippe Ravaud, Columbia University Mailman School of Public Health, New York, NY.

Published online ahead of print at www.jco.org on November 17, 2014.

Supported by a grant from the French Ministry of Health, the National Institute for Cancer Research, and the Fondation Pour la Recherche Médicale (Equipe Espoir de la Recherche 2010). D.A. is funded by Cancer Research UK.

The funders had no influence on the study design; the collection, analysis, or interpretation of data; the writing of the article; or the decision to submit for publication.

Authors' disclosures of potential conflicts of interest are found in the article online at www.jco.org. Author contributions are found at the end of this article.

Clinical trial information: NCT01848704.

Corresponding author: Isabelle Boutron, MD, Centre d'Epidémiologie Clinique, Hôpital Hôtel Dieu, 1, Place du Parvis Notre Dame, 75181 Paris Cedex 4, France; e-mail: isabelle.boutron@htd.aphp.fr.

0732-183X/14/3236w-4120w/$20.00

DOI: 10.1200/JCO.2014.56.7503

ABSTRACT

Purpose
We aimed to assess the impact of spin (ie, reporting to convince readers that the beneficial effect of the experimental treatment is greater than shown by the results) on the interpretation of results of abstracts of randomized controlled trials (RCTs) in the field of cancer.

Methods
We performed a two-arm, parallel-group RCT. We selected a sample of published RCTs with statistically nonsignificant primary outcome and with spin in the abstract conclusion. Two versions of these abstracts were used—the original with spin and a rewritten version without spin. Participants were clinician corresponding authors of articles reporting RCTs, investigators of trials, and reviewers of French national grants. The primary outcome was clinicians' interpretation of the beneficial effect of the experimental treatment (0 to 10 scale). Participants were blinded to study hypothesis.

Results
Three hundred clinicians were randomly assigned using a Web-based system; 150 clinicians assessed an abstract with spin and 150 assessed an abstract without spin. For abstracts with spin, the experimental treatment was rated as being more beneficial (mean difference, 0.71; 95% CI, 0.07 to 1.35; $P = .030$), the trial was rated as being less rigorous (mean difference, −0.59; 95% CI, −1.13 to 0.05; $P = .034$), and clinicians were more interested in reading the full-text article (mean difference, 0.77; 95% CI, 0.08 to 1.47; $P = .029$). There was no statistically significant difference in the clinicians' rating of the importance of the study or the need to run another trial.

Conclusion
Spin in abstracts can have an impact on clinicians' interpretation of the trial results.

J Clin Oncol 32:4120-4126. © 2014 by American Society of Clinical Oncology

INTRODUCTION

Abstracts of articles presenting results of randomized controlled trials (RCTs) are an essential means of dissemination of research results because they allow for wide, fast, and free broadcast of these results.[1] When results of RCTs are not published, clinicians can decide only from the meeting abstract whether or not to use the treatment in clinical practice. Also, because half of scientific publications are behind a pay wall,[2] for many readers, the abstract is the only accessible part of a published article. Further, readers (eg, physicians, researchers) often decide from the abstract whether or not to obtain more information from the full-text article.[3] However, this mode of dissemination may have serious consequences for patients if abstracts do not present an accurate and unbiased reflection of the trial results.[4-11]

Spin or misrepresentation of study findings can be used to influence, positively, the interpretation of statistically nonsignificant RCTs, for example, by emphasizing the apparent benefit of a secondary outcome or findings from a subgroup of patients. Misrepresentation of study findings was found in approximately 60% of abstract conclusions of a representative sample of RCTs indexed in PubMed[12] with a statistically nonsignificant result for the primary outcome. In a review of RCTs for breast cancer, 59% of trials with a negative primary outcome used secondary end points to suggest benefit of experimental therapy.[13] However, there are no data on the possible impact of spin on the interpretation of trial results by the person reading the abstract.

Table 1. Example of an Abstract With Spin and an Abstract Rewritten Without Spin

Abstract With Spin	Rewritten Abstract Without Spin
Purpose: To evaluate the efficacy and tolerability of treatment A plus cyclophosphamide, doxorubicin, and fluorouracil (CAF) therapy in comparison with CAF alone, in patients with advanced or recurrent breast cancer.	**Purpose:** To evaluate the efficacy and tolerability of treatment A plus cyclophosphamide, doxorubicin, and fluorouracil (CAF) therapy in comparison with CAF + placebo, in patients with advanced or recurrent breast cancer.
Patients and Methods: In this randomized, double-blind, placebo-controlled trial, patients were treated with six cycles of CAF therapy: 28 days per cycle, with doxorubicin (25 mg/m^2) and fluorouracil (500 mg/m^2) administered on days 1 and 8 and cyclophosphamide (100 mg orally) administered on days 1 through 14. Primary end point was overall response rate (ORR; partial or complete response). In total, 221 patients were assessable.	**Patients and Methods:** In this randomized, double-blind, placebo-controlled trial, patients were treated with six cycles of CAF therapy: 28 d/cycle, with doxorubicin (25 mg/m^2) and fluorouracil (500 mg/m^2) administered on days 1 and 8 and cyclophosphamide (100 mg orally) administered on days 1 through 14. Primary end point was overall response rate (ORR; partial or complete response). In total, 221 patients were assessable.
Results: ORR was 42.6% for CAF compared with 53.1% for treatment A + CAF, a 24.6% relative improvement and 10.5% absolute increase (P = .077). There was a trend for prolonged progression-free survival (PFS; median, 241 days for CAF *v* 366 days for treatment A + CAF; P = .145). In retrospectively defined subgroups, significant improvement in PFS in favor of treatment A was observed in patients who were premenopausal, had no prior therapy, and were stage IV at diagnosis with an intact primary tumor. Except for neutropenia and leukopenia, there was no statistically significant excess of grade 3/4 adverse events compared with CAF. Treatment with treatment A did not affect the plasma concentration of doxorubicin.	**Results:** A total of 227 patients were recruited; the full analysis set consisted of 221 patients. ORR was 42.6% (95% CI, 33.1 to 52.5) for CAF + placebo compared with 53.1% (95% CI, 43.5 to 62.5) for treatment A + CAF (P = .077). The median progression-free survival (PFS) was 241 days for CAF + placebo *v* 366 days for treatment A + CAF (P = .145). Any adverse event occurred more frequently with CAF + treatment A (95.5% *v* 88.4%; P = .025). Neutropenia occurred in 91.2% with CAF + treatment A versus 77.7% with CAF + placebo, (P = .005). Treatment with treatment A did not affect the plasma concentration of doxorubicin.
Conclusion: Treatment A + CAF was well tolerated and is suggested to have efficacy in patients who had not received prior therapy.	**Conclusion:** Treatment A + CAF was not more effective than CAF + placebo in patients with advanced or recurrent breast cancer.

The objective of this study was to assess the impact of spin on the interpretation of results presented in abstracts of clinical trials in the field of cancer. Our primary hypothesis was that spin would increase the probability that readers would judge the experimental treatment to be beneficial, despite the absence of a statistically significant difference in the primary outcome.

METHODS

Design

We planned a two-arm RCT to compare the interpretation of results of RCTs when abstracts are reported with or without spin. For this purpose, we selected a sample of 30 abstracts of articles describing RCTs performed in the field of cancer with statistically nonsignificant differences in the primary outcome that were reported with spin in the abstract conclusion. We systematically rewrote the selected abstracts to obtain the following two formats of the same abstract: an abstract reported with spin and an abstract reported without spin. Clinicians were randomly assigned to evaluate one abstract that either did or did not contain spin. According to this design, one abstract would be read by several clinicians with a clustering between the readers of the same abstract.

The planning, implementation, analysis, and writing of this study followed the CONSORT guidelines.[14] The study obtained ethics approval from the Institutional Review Board of Paris Descartes University, Paris, France (registration No. 00001072), and the protocol is registered on ClinicalTrials.gov (NCT01848704).

Abstract Selection

We identified 30 abstracts of articles describing RCTs performed in the field of cancer with statistically nonsignificant differences in the primary outcome that were reported with spin in the abstract conclusion. Spin was defined as a way of reporting to convince readers that the experimental treatment was beneficial, despite a statistically nonsignificant primary outcome.[12]

To identify abstracts with spin, we relied on five existing collections of abstracts with spin. Details are reported in the Data Supplement. For all these

Table 2. Description of the Modification of the 30 Abstracts With Spin

Spin	Action	Abstracts (n = 30) No.	%
No identification of the primary outcome in the Methods section	Clearly report the primary outcome in the Methods section	11	37
Incomplete reporting of the results for the primary outcome	Report results for the primary outcome with numbers in both arms (if possible with some measure of variability) with no wording of judgment	21	70
Selective reporting of statistically significant SOs	Report results for all SOs or for no SO or for the most clinically important SO	14	47
No or incomplete reporting of safety data	Report safety data including reason for withdrawals; report treatment discontinuation, when applicable	11	37
Focus on statistically significant subgroup or secondary analysis	Delete subgroup or secondary analysis	15	50
Linguistic spin	Delete linguistic spin	6	20
Conclusions with spin	Standardize conclusion	30	100

Abbreviation: SO, secondary analysis.

collections, two independent reviewers systematically evaluated whether the presentation and interpretation of the results were distorted but did not evaluate the risk of bias of included studies. One reviewer (I.B.) screened the abstracts and reports of all these collections and selected abstracts reporting two-arm, parallel-group RCTs in the field of cancer that randomly assigned more than 100 participants and had a statistically nonsignificant difference for all primary outcomes. To obtain a more homogeneous sample, we excluded abstracts describing RCTs comparing the same treatment but with different dosage, different mode of administration, or different duration of treatment; abstracts with spin claiming equivalence or comparable effectiveness; and abstracts with unusual types of spin (eg, a focus on another study objective).

Of the 30 selected abstracts, 13 were published in a high–impact factor journal (ie, impact factor ≥ 10). Five were published before 2005, 12 between 2005 and 2009, and 13 after 2009. The median number of patients per trial was 361 (quartile 1 to quartile 3, 227 to 627 patients); 14 trials were completely or partially industry funded, 10 trials were funded by academic grants, and for six trials, the funding source was not clearly reported. The primary outcome was overall survival (n = 8); progression-free survival (n = 6); disease-free survival (n = 6); response rate (n = 3); outcome related to quality of life, fatigue, or sleep quality (n = 4); toxicity (n = 1); or other (n = 2). Most abstracts (n = 24) were structured. The main focus of the spin was subgroup analysis (n = 15), statistically significant secondary outcomes (n = 10), or other strategies (n = 5). Overall, 10 abstracts had a high level of spin in the conclusion, defined as a conclusion reported with no acknowledgment of the statistically nonsignificant primary outcome.[12]

Construction of Abstracts Without Spin

The 30 selected abstracts were rewritten to remove spin. One researcher rewrote each abstract according to specific guidelines described in Box 1.

The lengths of the two versions of each abstract were within 25 words. All of the modified abstracts were evaluated independently by another researcher, who checked whether the abstract was rewritten according to the guidelines to have no spin; any disagreement was discussed, and the abstract was modified according to the consensus reached. Thus, we generated the following two formats of each abstract: one with spin and one without spin. Table 1 provides an example of a rewritten abstract without spin. Table 2 summarizes the modifications.

Abstracts with and without spin were presented in their original format, but names of authors, references, journal name, registration number, trial name or acronym, and article title were deleted, and the description of treatment was masked systematically with generic terms (eg, treatment A and comparator B). There were no other modifications to the original abstracts. All selected abstracts with and without spin are shown in the Data Supplement.

Participants and Recruitment Strategy

Our target population consisted of clinicians in the field of cancer having some expertise in clinical research. To recruit such participants, we collected e-mail addresses of the following people: clinicians who were corresponding authors of published articles reporting clinical trials in the field of cancer (we systematically searched Medline via PubMed using the search term "cancer" to identify all publications indexed with the Publication Type term "Randomized Controlled Trial" or "Clinical Trial" and published from January 1, 2010, to June 30, 2013; we extracted the e-mail addresses of the corresponding authors of these papers; details of the search strategy are in the data supplement); expert clinicians who participate in reviewing grants in the field of cancer for the French National Institute for Cancer Research; and principal investigators in the field of cancer registered in ClinicalTrials.gov (identified using the search term "cancer" performed on April 17, 2013).

An invitation e-mail was sent to all the e-mail addresses collected. If respondents agreed to participate, an Internet link included in the invitation e-mail gave them access to a screening question asking them whether they were a clinician; nonclinicians were excluded from the study. If they answered yes, respondents were randomly assigned to receive either an abstract with spin or an abstract without spin, which they were asked to assess. Invitation e-mails were sent in waves until the planned number of clinicians had logged on and completed the assessment. A maximum of two reminders were sent to participants. The e-mail invitation and details of the study are described in the Data

Guidelines for writing an abstract without spin

1. In the Context section:
   a. Delete all information that could distort the understanding of the aim of the trial.
      i. The aim is to evaluate the treatment effect on a secondary outcome.
      ii. The aim is to evaluate the treatment effect for a subgroup.
      iii. The aim is to evaluate overall improvement.
2. In the Methods section:
   a. Clearly report the primary outcome.
   b. According to space constraints, report all secondary outcomes evaluated in the Methods section or report no secondary outcome evaluated in the Methods section to avoid specifically highlighting statistically significant secondary outcomes.
   c. Delete information that could distort the understanding of the aim (eg, within-group comparison, modified population analysis, subgroup analysis).
3. In the Results section:
   a. Delete subgroup analyses that were not prespecified, based on the primary outcome, and interpreted in light of the totality of prespecified subgroup analyses undertaken.
   b. Delete within-group comparisons.
   c. Delete linguistic spin.
   d. Report the results for the primary outcome with numbers in both arms (if possible with some measure of variability) with no wording of judgment.
   e. Report results for all secondary outcomes, for no secondary outcome, or for the most clinically important secondary outcome.
   f. Report safety data including reason for withdrawals; report treatment discontinuation when applicable.
4. In the Conclusions section:
   a. Delete the author conclusion, and only add the following standardized conclusion: "the treatment A was not more effective than comparator B in patients with...."
   b. Specify the primary outcome in the conclusion when some secondary outcomes were statistically significant: "the treatment A was not more effective on overall survival than the comparator B in patients with...."

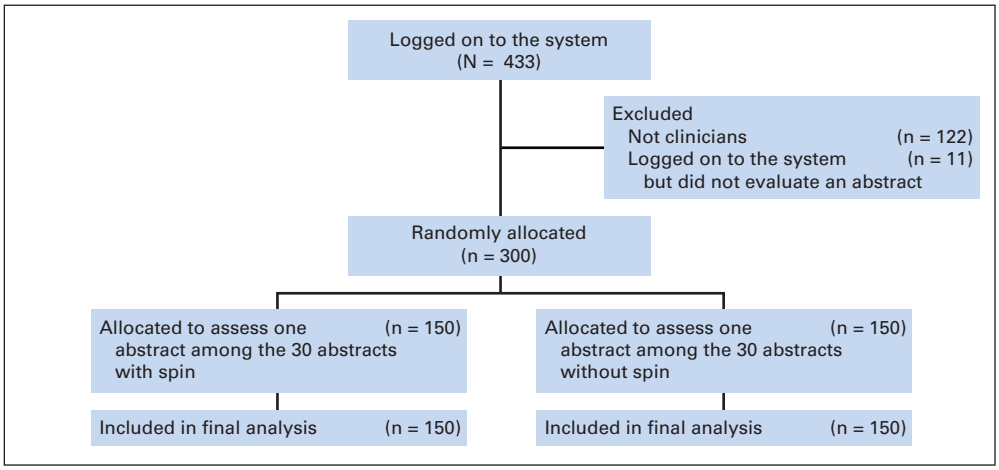

Fig 1. Flow of participants.

Supplement. As an incentive, participants were informed that after completion of the study, they would be entered into a drawing, along with all other participants, for a chance to win an iPad. This strategy led to a large number of potential participants, but we cannot determine how many participants received the invitation e-mail because the e-mails may have been directed to junk mail folders or filtered as spam.

Random Assignment

The random assignment sequence was computer generated by a statistician using blocks of 60 (ie, the number of abstracts selected times two). The list and the block size were not disclosed to the investigators. Allocation concealment was assured by use of a computerized random assignment system. If a clinician logged on to the system but did not evaluate an abstract, the clinician was excluded and the abstract was automatically allocated to another clinician.

Outcomes

Our primary outcome was the clinicians' interpretation of the beneficial effect of the experimental treatment, measured using a scale from 0 to 10. The secondary outcomes were the clinicians' assessment of the rigor of the study methodology and the study importance, their interest in reading the full text, and their interest in running another trial evaluating this treatment. The questions posed were as follows:

- Based on this abstract, do you think treatment A would be beneficial to patients? (Scale, 0 [very unlikely] to 10 [very likely]; primary outcome).
- Rate the overall rigor of the study methodology. (Scale, 0 [not at all rigorous] to 10 [very rigorous]).[15]
- Rate the importance of the study. (Scale, 0 [not at all important] to 10 [very important]).[15]
- Are you interested in reading the full-text article for the study described in this abstract? (Scale, 0 [not at all interested] to 10 [very interested]).[15]
- Do you think it would be interesting to run another trial evaluating this treatment? (Scale, 0 [not interesting at all] to 10 [very interesting]).

The face validity of these outcomes was evaluated with a sample of 23 clinicians. To minimize bias, clinicians were blinded to the study hypothesis.[16] All clinicians were informed that they were participating in a survey about the interpretation of abstracts of articles describing clinical trials. They were not informed of the objectives and assumptions of the study.

Sample Size

Each clinician read either an abstract with spin or an abstract without spin. We sought to detect a mean difference of 1.0 for the primary outcome with a within-group standard deviation of 2.5 (based on a pilot study). A sample of 266 assessments of abstracts was needed to detect as significant an effect size of 0.4 with a power of 90% and an α risk equal to 5%. In order that each abstract would be read the same number of times, we increased the sample to 300 clinicians (150 in each arm). Therefore, each abstract was assessed five times in each group.

Statistical Analysis

The statistical analysis was undertaken using R v2.15.1 (R Foundation for Statistical Computing, Vienna, Austria) by a statistician at the Center for Clinical Epidemiology, in Paris, France. All outcomes were quantitative, and the number of participants and abstracts in each group was balanced. The difference between groups was analyzed using a linear mixed model with fixed group effect and random abstract and abstract-group interaction effects. Random effects allow us to take into account the following two levels of clustering: within-group clustering as a result of the abstract (each abstract was assessed five times in each group) and between-group clustering (pairing between the abstracts used in the two arms of the trial). Inference was based on the restricted maximum likelihood. This model compared the mean difference between the two groups. For the primary and secondary outcomes, we estimated the difference between means (with 95% CIs). $P < .05$ was considered statistically significant. We performed a prespecified subgroup analysis by level of spin[12] in the abstract conclusions for the primary outcome using a test of interaction with the following two fixed factors: level of spin and an interaction term between group and level of spin.

RESULTS

Participants

Between May 27, 2013, and June 26, 2013, we sent e-mails of invitation to 5,618 e-mail addresses in waves until we reached the required sample size of 300 clinicians. Of the first 433 participants who logged on to the system, 122 were excluded because they were not clinicians and 11 logged on but did not evaluate the abstract (Fig 1). Thus, 300 clinicians were randomly assigned and assessed an abstract with spin (n = 150) or without spin (n = 150). The characteristics of the clinicians are listed in Table 3.

Clinicians were based mainly in Europe (n = 169; 56%), the United States (n = 77; 26%), and Canada (n = 20; 7%), they practiced mainly in tertiary care centers (n = 216; 72%), and most had more than 15 years of clinical practice (n = 165; 55%). About half (n = 142; 47%) had received some training in epidemiology, and 78% (n = 235) had received training in the methods of RCTs. In addition, 49% (n = 147) were involved as a principal investigator or co-investigator in more than 10 RCTs, and 67% (n = 200) had peer reviewed more than five articles in the last year. About half of the clinicians (n =

Table 3. Baseline Demographic and Other Characteristics of Clinicians

Characteristic	Abstract With Spin (n = 150)		Abstract Without Spin (n = 150)		Total (N = 300)	
	No. of Clinicians	%	No. of Clinicians	%	No. of Clinicians	%
Age, years						
Mean	47.0		48.3		47.7	
SD	9.7		9.9		9.8	
Male	124	82.7	111	74.0	235	78.3
Degree						
MD	73	48.7	82	54.7	155	51.7
MD/PhD	71	47.3	62	41.3	133	44.3
Other	6	4.0	6	4.0	12	4.0
Location						
Europe	84	56.0	85	56.7	169	56.3
United States	34	22.7	43	28.7	77	25.7
Canada	12	8.0	8	5.3	20	6.7
South America	6	4.0	0	0.0	6	2.0
Asia	6	4.0	10	6.7	16	5.3
Oceania	8	5.3	4	2.7	12	4.0
Duration of practice, years						
< 5	6	4.0	8	5.3	14	4.7
5 to 15	63	42.0	58	38.7	121	40.3
> 15	81	54.0	84	56.0	165	55.0
Current practice						
Primary care center	25	16.7	35	23.3	60	20.0
Secondary care center	15	10.0	9	6.0	24	8.0
Tertiary care center	110	73.3	106	70.7	216	72.0
No. of randomized trials involved in						
< 3	35	23.3	36	24.0	71	23.7
4 to 10	39	26.0	43	28.7	82	27.3
> 10	76	50.7	71	47.3	147	49.0
No. of randomized trials published as a corresponding author						
< 3	96	64.0	89	59.3	185	61.7
4 to 10	37	24.7	39	26.0	76	25.3
> 10	17	11.3	22	14.7	39	13.0
No. of abstracts of randomized trials read in the last month						
< 5	22	14.7	24	16.0	46	15.3
5 to 10	51	34.0	48	32.0	99	33.0
10 to 15	46	30.7	47	31.3	93	31.0
> 15	32	21.3	32	21.3	64	21.3
No. of articles peer reviewed in the last year						
< 5	47	31.3	53	35.3	100	33.3
5 to 10	48	32.0	50	33.3	98	32.7
> 10	55	36.7	47	31.3	102	34.0
No. of grant proposals peer reviewed in the last year						
< 5	110	73.3	112	74.7	222	74.0
5 to 10	20	13.3	21	14.0	41	13.7
> 10	20	13.3	17	11.3	37	12.3
Some training in epidemiology	73	48.7	69	46.0	142	47.3
Some training in the methods of randomized trials	114	76.0	121	80.7	235	78.3

Abbreviation: SD, standard deviation.

157) had read more than 10 abstracts of articles describing RCTs in the last month.

Clinicians' Interpretation of Abstracts

All outcomes were evaluated on a scale of 0 to 10 (0 = not at all likely, 10 = very likely). For the primary outcome (Table 4), the experimental treatment was rated as being more beneficial (mean score, 3.6; standard deviation, 2.5) in abstracts with spin than those abstracts without spin (mean score, 2.9; standard deviation, 2.6; Fig 2). The mean difference in scores between abstracts with and without spin was 0.71 (95% CI, 0.07 to 1.35; $P = .030$), which corresponds to an effect size of 0.25.

There was no evidence of heterogeneity of intervention effect in the subgroups defined by level of spin ($P = .65$). Evaluation of secondary outcomes showed that the trial was rated as less rigorous (mean difference, −0.59; 95% CI, −1.13 to −0.05; $P = .034$) and clinicians

Table 4. Comparison of the Interpretation of Abstracts of Randomized Controlled Trials Reported With and Without Spin

Question	Abstract With Spin (n = 150)		Abstract Without Spin (n = 150)		Difference		P	Effect Size
	Mean Score	Standard Deviation	Mean Score	Standard Deviation	Mean	95% CI		
Do you think treatment A would be beneficial to patients (score, 0 to 10)?	3.6	2.5	2.9	2.6	0.71	0.07 to 1.35	.030	0.25
Rate the overall rigor of the study methodology (score, 0 to 10)	4.5	2.4	5.1	2.5	−0.59	−1.13 to −0.05	.034	0.25
Rate the importance of the study (score, 0 to 10)	4.6	2.4	4.9	2.4	−0.38	−0.95 to 0.19	.18	0.15
Are you interested in reading the full-text article for the study described in this abstract (score, 0 to 10)?	5.1	3.2	4.3	3.0	0.77	0.08 to 1.47	.029	0.35
Do you think it would be interesting to run another trial evaluating this treatment (score, 0 to 10)?	4.8	2.9	4.2	2.9	0.64	−0.03 to 1.31	.061	0.22

were more interested in reading the full-text article (mean difference, 0.77; 95% CI, 0.08 to 1.47; $P = .029$) when abstracts were reported with spin. There was no statistically significant difference in the clinicians' interpretation on the importance of the study (mean difference, −0.38; 95% CI, −0.95 to 0.19; $P = .18$) or the need to run another trial (mean difference, 0.64; 95% CI, −0.03 to 1.31; $P = .061$).

DISCUSSION

Our results show that spin in the abstract of an RCT in the field of cancer can impact on clinicians interpretation of the trial results, such that they are more likely to rate a treatment as beneficial despite the primary outcome being statistically nonsignificant. These results are important because they demonstrate that a slight modification to the focus of the abstract can impact on clinicians' interpretation of a trial's result.

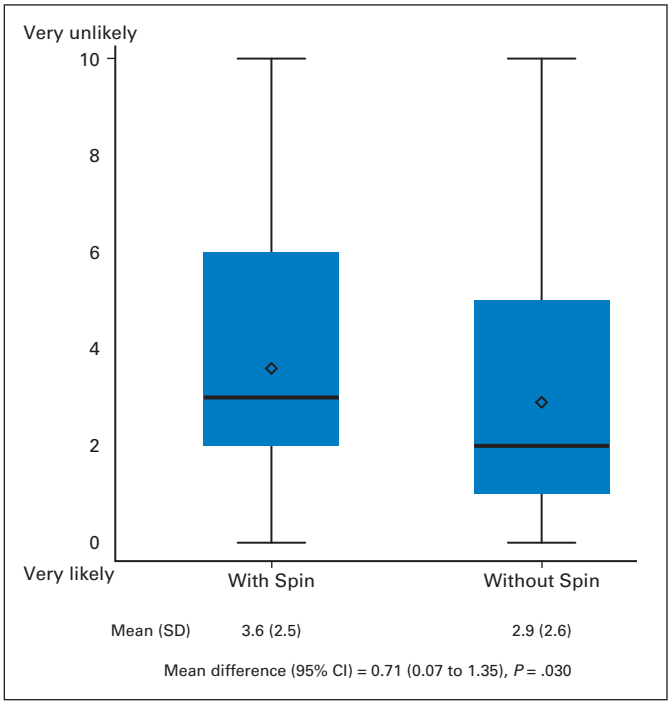

Fig 2. Readers' assessment of the beneficial effect of the treatments when assessed from abstracts of randomized controlled trials reported with or without spin. Scores are based on a numerical rating scale, ranging from 0 (very unlikely) to 10 (very likely). Boxes represent median observations (horizontal rule) with 25th and 75th percentiles of observed data (top and bottom of box). The diamonds represent the mean. The error bars represent the minimum and maximum values. SD, standard deviation.

The effect size related to the presence of spin in an abstract is nevertheless small. However, we should not determine whether the impact of an intervention is meaningful solely on the effect size; it also depends on the frequency of use of spin. In the context of cancer RCTs, abstracts with spin are frequent.[17] Consequently, spin in abstracts of RCTs may be responsible for a gap between the perceived and actual beneficial effect of the treatments studied.[11] However, to our knowledge, our study is the first RCT evaluating the impact of spin, and the results need to be replicated.

Spin is a classical concept in fields such as public relations, business, marketing, politics, and journalism, where it is defined as a form of propaganda to influence public opinion.[18,19] Interest in spin in the field of medicine and clinical research is relatively recent. Some studies have looked at the prevalence of spin in published articles of RCTs[12,17,20,21] or in studies of diagnostic test accuracy,[22,23] nonrandomized studies,[24-26] or epidemiologic studies.[27] These studies have shown that spin is common, particularly in the abstract of the article. The types of spin used by authors are diverse. In RCTs, spin could consist of a particular focus on statistically significant results or an inappropriate interpretation of non–statistically significant differences as demonstrating equivalence in treatment effectiveness. Interpretation could also be distorted through the use of linguistic spin or the lack of consideration of the implication of harms.

Our study has several strengths. First, it is a proof of concept study that used a strong design, an RCT, to evaluate the impact of spin on readers' interpretation. To our knowledge, few RCTs have been conducted in the field of research interpretation,[15,28-30] with none on spin. Second, we used a panel of 30 real-life abstracts of RCTs indexed in PubMed, which is a primary information source for clinicians and researchers. Further, we used a strategy of recruitment that allowed the recruitment of an international sample, with participation of clinicians from Europe, the United States, Canada, South America, Asia, and Oceania.

Our study has some limitations. First, we only evaluated the impact of spin in abstracts, and the clinicians who participated in this study did not have access to the full-text article to fully appraise the

study results. However, because of time constraints and the difficulties of accessing the full-text article, clinicians often read only the abstract of a journal article. Second, the clinicians included in this RCT had some experience in the conduct and review of clinical trials in the field of cancer, and this could underestimate the impact of spin. Third, we selected abstracts with various levels of spin, but our sample was too small to assess the impact of different types of spin. Finally, this study was performed in the field of cancer where the type of spin consists mainly of a focus on statistically significant secondary outcomes and subgroup analysis.[13]

To try to minimize the impact of spin, and thus biased dissemination of research results, authors should be educated on how to interpret research results. Peer reviewers and journal editors also play an important role; they should systematically check whether the abstract conclusions are consistent with the study results and whether the results reported in the abstract are free from bias.

In conclusion, spin in abstracts of articles describing RCTs in the field of cancer can impact on clinicians' interpretation of the beneficial effect of the treatment evaluated in these trials. It is important to raise the awareness of this problem within the scientific community to limit potential harm.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Disclosures provided by the authors are available with this article at www.jco.org.

AUTHOR CONTRIBUTIONS

Conception and design: All authors
Collection and assembly of data: Isabelle Boutron
Data analysis and interpretation: All authors
Manuscript writing: All authors
Final approval of manuscript: All authors

REFERENCES

1. Hopewell S, Clarke M, Moher D, et al: CONSORT for reporting randomised trials in journal and conference abstracts. Lancet 371:281-283, 2008
2. Kurata K, Morioka T, Yokoi K, et al: Remarkable growth of open access in the biomedical field: Analysis of PubMed articles from 2006 to 2010. PLoS ONE 8:e60925, 2013
3. Barry HC, Ebell MH, Shaughnessy AF, et al: Family physicians' use of medical abstracts to guide decision making: Style or substance? J Am Board Fam Pract 14:437-442, 2001
4. Pop GH, Fesperman SF, Ball DA, et al: Duplicate presentations on prostate cancer at American Urological Association and European Association of Urology annual meetings. J Urol 182:674-678, 2009
5. De Bellefeuille C, Morrison CA, Tannock IF: The fate of abstracts submitted to a cancer meeting: Factors which influence presentation and subsequent publication. Ann Oncol 3:187-191, 1992
6. Takeda A, Loveman E, Harris P, et al: Time to full publication of studies of anti-cancer medicines for breast cancer and the potential for publication bias: A short systematic review. Health Technol Assess 12:1-46, 2008
7. Toma M, McAlister FA, Bialy L, et al: Transition from meeting abstract to full-length journal article for randomized controlled trials. JAMA 295:1281-1287, 2006
8. Booth CM, Le Maître A, Ding K, et al: Presentation of nonfinal results of randomized controlled trials at major oncology meetings. J Clin Oncol 27:3938-3944, 2009
9. Krzyzanowska MK, Pintilie M, Brezden-Masley C, et al: Quality of abstracts describing randomized trials in the proceedings of American Society of Clinical Oncology meetings: Guidelines for improved reporting. J Clin Oncol 22:1993-1999, 2004
10. Pitkin RM, Branagan MA, Burmeister LF: Accuracy of data in abstracts of published research articles. JAMA 281:1110-1111, 1999
11. Altwairgi AK, Booth CM, Hopman WM, et al: Discordance between conclusions stated in the abstract and conclusions in the article: Analysis of published randomized controlled trials of systemic therapy in lung cancer. J Clin Oncol 30:3552-3557, 2012
12. Boutron I, Dutton S, Ravaud P, et al: Reporting and interpretation of randomized controlled trials with statistically nonsignificant results for primary outcomes. JAMA 303:2058-2064, 2010
13. Vera-Badillo FE, Shapiro R, Ocana A, et al: Bias in reporting of end points of efficacy and toxicity in randomized, clinical trials for women with breast cancer. Ann Oncol 24:1238-1244, 2013
14. Moher D, Hopewell S, Schulz KF, et al: CONSORT 2010 explanation and elaboration: Updated guidelines for reporting parallel group randomised trials. BMJ 340:c869, 2010
15. Kesselheim AS, Robertson CT, Myers JA, et al: A randomized study of how physicians interpret research funding disclosures. N Engl J Med 367:1119-1127, 2012
16. Wood L, Egger M, Gluud LL, et al: Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: Meta-epidemiological study. BMJ 336:601-605, 2008
17. Sacher AG, Le LW, Leighl NB: Shifting patterns in the interpretation of phase III clinical trial outcomes in advanced non-small-cell lung cancer: The bar is dropping. J Clin Oncol 32:1407-1411, 2014
18. Safire W: The spinner spun. The New York Times. http://www.nytimes.com/1996/12/22/magazine/the-spinner-spun.html
19. Politics, spin, and science. Lancet 364:1994, 2004
20. Latronico N, Metelli M, Turin M, et al: Quality of reporting of randomized controlled trials published in intensive care medicine from 2001 to 2010. Intensive Care Med 39:1386-1395, 2013
21. Wilson JR: Rhetorical strategies used in the reporting of implantable defibrillator primary prevention trials. Am J Cardiol 107:1806-1811, 2011
22. Ochodo EA, de Haan MC, Reitsma JB, et al: Overinterpretation and misreporting of diagnostic accuracy studies: Evidence of "spin." Radiology 267:581-588, 2013
23. Lumbreras B, Parker LA, Porta M, et al: Overinterpretation of clinical applicability in molecular diagnostic research. Clin Chem 55:786-794, 2009
24. Linden A: Identifying spin in health management evaluations. J Eval Clin Pract 17:1223-1230, 2011
25. Brown AW, Bohan Brown MM, Allison DB: Belief beyond the evidence: Using the proposed effect of breakfast on obesity to show 2 practices that distort scientific evidence. Am J Clin Nutr 98:1298-1308, 2013
26. Prasad V, Jorgenson J, Ioannidis JP, et al: Observational studies often make clinical practice recommendations: An empirical evaluation of authors' attitudes. J Clin Epidemiol 66:361-366, 2013
27. Tzoulaki I, Liberopoulos G, Ioannidis JP: Assessment of claims of improved prediction beyond the Framingham risk score. JAMA 302:2345-2352, 2009
28. Woloshin S, Schwartz LM: Communicating data about the benefits and harms of treatment: A randomized trial. Ann Intern Med 155:87-96, 2011
29. Emerson GB, Warme WJ, Wolf FM, et al: Testing for the presence of positive-outcome bias in peer review: A randomized controlled trial. Arch Intern Med 170:1934-1939, 2010
30. Wegwarth O, Schwartz LM, Woloshin S, et al: Do physicians understand cancer screening statistics? A national survey of primary care physicians in the United States. Ann Intern Med 156:340-349, 2012

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Impact of Spin in the Abstracts of Articles Reporting Results of Randomized Controlled Trials in the Field of Cancer: The SPIIN Randomized Controlled Trial

The following represents disclosure information provided by authors of this manuscript. All relationships are considered compensated. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. For a detailed description of the disclosure categories, or for more information about ASCO's conflict of interest policy, please refer to the Author Disclosure Declaration and the Disclosures of Potential Conflicts of Interest section in Information for Contributors.

Isabelle Boutron
No relationship to disclose

Douglas G. Altman
No relationship to disclose

Sally Hopewell
No relationship to disclose

Francisco Vera-Badillo
No relationship to disclose

Ian Tannock
Research Funding: Sanofi

Philippe Ravaud
No relationship to disclose

Acknowledgment

We thank Gabriel Baron (statistician in the Center of Clinical Epidemiology, Paris, France), who performed the statistical analysis.