Critical Reviews in Oncology / Hematology 144 (2019) 102821

Contents lists available at ScienceDirect

Critical Reviews in Oncology / Hematology

journal homepage: www.elsevier.com/locate/critrevonc

Evaluation of spin in oncology clinical trials

C. Wayant[a,*], D. Margalski[b], K. Vaughn[c], M. Vassar[d]

[a] Oklahoma State University Center for Health Sciences, Department of Biomedical Sciences, 1111 West 17th Street, Tulsa, OK, 74107, USA
[b] Wake Forest Baptist Medical Center, Department of Internal Medicine, 1 Medical Center Blvd, Winston-Salem, NC, 27157, USA
[c] University of Oklahoma School of Community Medicine, Department of Medicine/Pediatrics, 4502 E. 41st Street, Tulsa, OK, 74135, USA
[d] Oklahoma State University Center for Health Sciences, Department of Psychiatry, 1111 West 17th Street, Tulsa, OK, 74107, USA

ARTICLE INFO

Keywords:
Randomized controlled trial
Overall survival
Surrogate endpoints
Spin
Bias
Abstracts

ABSTRACT

Purpose: Spin, the misrepresentation of research findings, in clinical trial abstract has been shown to influence how oncologist rate a drug's efficacy.
Materials and methods: We searched PubMed for clinical trials published in ten key journals in 2017. Our primary objectives were to assess the frequency and manifestations of spin in the abstracts of those clinical trials that measured both overall survival and at least one surrogate efficacy endpoints.
Results: 124 trials were included for analysis. We found evidence of spin in 46 of 124 (37.1%, 95% CI 29.1%–45.9%) trial abstracts. Spin in the abstract results was most often due to authors emphasizing a statistically significant subgroup analysis (n = 6). Spin in the abstract conclusions was most often due to authors relying on a statistically significant surrogate endpoint to highlight the bioefficacy of the intervention (n = 17).
Conclusion: Spin is prevalent in the abstracts of oncology clinical trials that measure OS and a surrogate endpoint. The conclusion sections of abstracts were most prone to contain spin. When OS was the primary endpoint, spin was primarily used to distract from the nonsignificant OS data. To mitigate unintentional hype for cancer therapies, we recommend authors structure their conclusions around patient-important outcomes.

1. Introduction

When authors misrepresent, distort, or otherwise selectively feature specific research data they introduce "spin" to the literature. The prevalence of spin has been quantified in a recent systematic review, which found that a median of 56.8% of trials contain some form of spin (Chiu et al., 2017). The effect of spin has been demonstrated in a two-arm, parallel group randomized trial involving 300 oncologists who were asked to evaluate a trial abstract with a nonsignificant primary endpoint4. Half were assigned to read an abstract with an overly optimistic conclusion about the intervention, while the other half read an abstract that concluded no benefit of the intervention. Oncologists who received the abstract with the overly optimistic conclusion rated the intervention as more effective, the trial as less rigorous, and were more likely to read the full text of the trial.

In 2016, a trial was published that examined the effect of adjuvant sunitinib on advanced stage renal cell carcinoma post nephrectomy. This trial, which formed the basis for FDA approval, used improvements in a surrogate endpoint (disease-free survival) as evidence of drug efficacy, and relegated the collinear Kaplan-Meier curves for overall survival (OS) to the Supplement (Ravaud et al., 2016). Despite OS being the secondary endpoint, the eight year follow-up data and hazard ratio of 1.01 (95% CI, 0.72–1.44) was highly suggestive of no survival benefit. Indeed, a letter in reply to this trial was written to emphasize that the goal of cancer therapy is to extend survival and mitigate adverse events, neither of which was shown in this trial of sunitinib (Zhang, 2017). Some have described the FDA approval of sunitinib as "regulatory capture" (Gyawali and Goldstein, 2018), which occurs when parties with high-stakes interest in a policy decision overpower other parties with less interest to achieve an intended outcome. We have an additional concern: how certain research data can be highlighted or omitted in order to shift perceptions of a drug's efficacy.In this case, by not mentioning the nonsignificant OS data in the abstract and placing the Kaplan-Meier graph in the Supplement, the authors of the sunitinib study framed their printed study around a statistically significant surrogate endpoint and omitted visual evidence of the nonsignificant survival benefit.

There are many forms of spin in the reporting of medical research findings, but at its core spin is an attempt to beautify or omit unfavorable results (Chiu et al., 2017). Abstracts may be most susceptible

* Corresponding author.
E-mail addresses: cole.wayant@okstate.edu (C. Wayant), daniel.margalski@okstate.edu (D. Margalski), kaleb.vaughn@okstate.edu (K. Vaughn), matt.vassar@okstate.edu (M. Vassar).

https://doi.org/10.1016/j.critrevonc.2019.102821
Received 19 February 2019; Received in revised form 8 October 2019; Accepted 14 October 2019
1040-8428/

to spin because of the limited word counts enforced by journals. Further, the consequences of spin in abstracts may be more severe. There is evidence that many physicians only read the abstract of most research articles (Barry et al., 2001; Marcelo et al., 2013; Saint et al., 2000). Moreover, institutions in low- or middle-income countries may not have the resources to access the full text of articles. Thus, abstracts must be accurate synopses of full manuscripts and avoid misleading conclusions about drug efficacy. In oncology, spin occurs in the abstracts as well as the full text of manuscripts (Altwairgi et al., 2012; Vera-Badillo et al., 2013). Trial authors may omit toxicity results or selectively report endpoints based on statistical significance. And while all areas of medicine including oncology are susceptible to spin, we argue that the oncology literature may be most susceptible due, in part, to the presence of surrogate endpoints that are designed to predict clinical benefits to patients.

Surrogate endpoints are often acceptable in oncology trials (Kemp and Prasad, 2017), even for FDA approval (Beaver et al., 2018). However, OS is considered the ideal survival endpoint in oncology trials owing to its objectivity and relevance to patients (Food And Drug, 2015). But, OS requires increased sample size and follow-up duration, which may delay the approval or development of new therapies (Cheema and Burkes, 2013). Owing to these factors, surrogate endpoints have increased in popularity as primary endpoints since they often require fewer patients and less time to measure (Kemp and Prasad, 2017). Despite their popularity, surrogate endpoints are often imperfect measures of OS and frequently have larger effect sizes (Tan et al., 2017). The fact that surrogate endpoints are often imperfect measures of treatment effectiveness, they may show discordant results and OS may be required to clarify treatment utility in practice. However, because OS is more often statistically nonsignificant, authors may be tempted to spin toward the surrogate endpoints that tend to have larger effect sizes.

Regardless of which endpoint is primary and which is secondary, selectively emphasizing a secondary endpoint or subgroup analysis means authors are emphasizing fragile, underpowered results. Since surrogate endpoints and OS are almost equally acceptable in clinical trials, oncologist-authors may feel comfortable focusing on whichever of these endpoints is statistically significant. Therefore, the primary objective of this novel investigation is to evaluate the frequency and manifestations of spin in oncology clinical trials that measured a surrogate endpoint and OS.

2. Methods

We searched PubMed on March 30, 2018 to identify clinical trials published in 2017 reporting at least one key surrogate endpoint and OS published in ten key journals. The exact search strategy is publicly available via the Open Science Framework (Wayant, 2018). The six key surrogate endpoints were progression-free survival, disease-free survival, objective response rate, complete response, time to progression, and time to treatment failure. These surrogate endpoints were selected based on the Food and Drug Administration's (FDA) "Guidance for Industry: Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics" document (U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER), 2007). The following journals were searched: *Journal of Clinical Oncology*, *The Lancet: Oncology*, *JAMA Oncology*, *Cancer*, *Annals of Oncology*, *Journal of the National Cancer Institute*, *British Journal of Cancer*, *European Journal of Cancer*, *New England Journal of Medicine*, and *The Lancet*. Search results were added to a PubMed collection and uploaded to Rayyan (Ouzzani et al., 2016).

One of us (CW) screened all articles for inclusion. To be included an article had to meet meet the following criteria: randomized clinical trial with a head to head comparison, measure both a surrogate endpoint from the FDA "Guidance to Industry" list and OS, and conduct a superiority analysis. We excluded articles that were not clinical trials, clinical trials with an incompatible design (e.g., cluster, crossover, single arm), pooled analyses, noninferiority analyses, and trials of non-oncologic interventions published in the included general medical journals. Three of us (CW, DM, and KV) extracted data for all included trials.

2.1. Definition of spin

Our definition of spin was based on Boutron et al. (2010) which states that spin is the "use of specific reporting strategies, from whatever motive, to highlight that the experimental treatment is beneficial, despite a statistically nonsignificant difference for the primary outcome, or to distract the reader from statistically nonsignificant results." We modified this definition to include all trials regardless of the statistical significance of the primary endpoint. Doing so allowed us to capture evidence of spin in abstract conclusions when a surrogate endpoint (primary endpoint) was statistically significant and OS (secondary endpoint) was nonsignificant.

2.2. Objectives

Our primary objective was to assess the frequency and manifestations of spin in oncology clinical trials that report both a surrogate endpoint and OS. Spin was assessed within a trial (i.e., emphasizing a subgroup analysis when the primary analysis is nonsignificant) and outside a trial (i.e., selective outcome reporting bias - adding, subtracting, or changing the order of trial endpoints compared to a trial registry before publication). In the former (within a trial), primary trial endpoints would match the registry, but they would be reported out of order or with different emphasis. In the latter (selective outcome reporting bias), trial endpoints would not match the registry. Our secondary objective is to compare trials with OS as the primary endpoint and with OS as the secondary endpoint.

2.3. Spin in the abstract title and results

We considered there to be evidence of spin if a study title suggested treatment effectiveness where none exists. For example, if a title began with "First line use of…", despite showing no significant benefit of the intervention, this may be spin and mislead readers about the study conclusions. We considered there to be spin in abstract results when a trial emphasized statistically significant results out of order (e.g. subgroup before overall analysis, secondary endpoint before primary endpoint), reported a per-protocol analysis when intention-to-treat was prespecified, or used "trend statements" in the description of statistical significance (e.g., "trend toward significance"). We did not consider it to be spin if a trial omitted a secondary endpoint (including OS) from the abstract results section, since this could be interpreted as the standard reporting of results.

2.4. Spin in the abstract conclusions

We considered there to be evidence of spin when a trial interpreted statistically nonsignificant results as showing treatment equivalence or comparable effectiveness, focused on a significant subgroup or within-group comparison, used unjustified, optimistic statements in the description of outcomes, emphasized subgroups or modified treatment populations, distracted from nonsignificant findings by stating that the nonsignificant results were due a trial design issue (e.g., underpowered), or claimed treatment benefit from a statistically significant surrogate endpoint when OS was nonsignificant. We considered using only a surrogate endpoint as evidence of treatment benefit as spin because surrogate endpoints have been shown to be poor predictors of OS (Prasad et al., 2015). Abstract conclusions frequently state that the primary, surrogate endpoint was met while maintaining a focus on

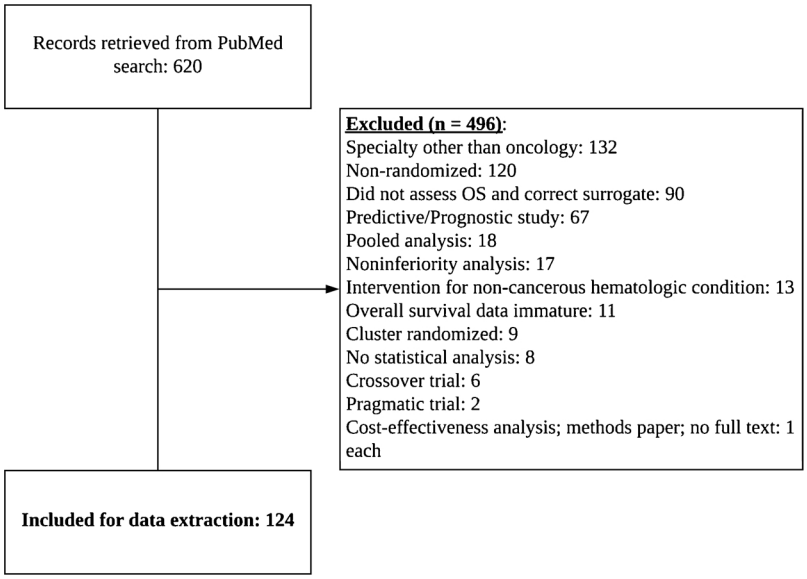

Fig. 1. Flow diagram of included and excluded studies.

patient-important endpoints, such as OS. A clear statement that OS data was nonsignificant was not required. In phase 2 trials or interim analyses of phase 3 trials, we did not consider there to be spin if authors stated that further investigation was necessary to confirm the present findings.

2.5. Statistical analysis

Summary statistics (frequencies and proportions) were calculated using Google Sheets. We used Stata 15.1 (Stata Corp, LLC; College Station, TX) and Fisher's exact test to compare differences in categorical endpoints.

3. Results

Of the 620 articles retrieved, 124 were included. Articles were excluded mostly for not being oncology trials published in general medical journals, being nonrandomized (including single-arm trials), or for not measuring both a surrogate from the FDA guidance document and OS. Fig. 1 itemizes all exclusions. Table 1 itemizes characteristics of the included trials and describes the proportion of trials with spin associated with each characteristic. Overall, in the 124 trials there were 126 primary endpoints: 71 were surrogate endpoints and 55 were OS. In five trials, OS and a surrogate endpoint were co-primary endpoints, while in nine trials the surrogate endpoint and OS were co-secondary endpoints. The most common surrogate endpoints measured were progression-free survival (n = 46), followed by disease-free survival (n = 17), and overall response rate (n = 3). The primary endpoint was clearly described in 86.3% (107/124) of abstracts. In the majority of cases, the primary endpoint was not statistically significant (79/126; 62.7%, 95%CI 54.0%–70.7%).

We found evidence of spin in 46 of 124 (37.1%, 95%CI 29.1%–45.9%) of trial abstracts (Table 2). There was no evidence of spin in trial titles. Spin was present in 19 (15.3%, 95%CI 9.9%–22.4%) abstract results and 40 (32.3%, 95%CI 24.7%–40.9%) abstract conclusions. Of the 118 trials that reported a trial registration number, 10 (8.5%, 95%CI 4.7%–14.9%) selectively reported their endpoints. Sixteen (12.9%, 95%CI 8.1%–19.9%) RCTs had spin in both the abstract results and conclusions.

When OS was a primary endpoint there was evidence of spin 29.1% (16/55; 95%CI 18.8%–42.1%) of the time. OS was statistically significant in 3 of these trials; however, evidence of spin in each trial was due to selective reporting bias, which indicates that authors deviated from the trial protocol to report the statistically significant OS data as the primary endpoint. Thus, when OS was the primary endpoint, spin was primarily used to distract from the nonsignificant OS data. When OS was a secondary endpoint there was evidence of spin 43.5% (30/69; 95%CI 32.4%–55.2%) of the time. OS was statistically significant in 2 of these trials, meaning authors were most likely to frame their conclusions around statistically significant surrogate endpoint data, rather than patient important outcomes. There was no significant difference in the rates of spin when OS was a primary or secondary endpoint (p = .13), indicating that evidence of spin was used for different reasons depending on which endpoint is primary and which is statistically significant.

Spin in the abstract results was most often due to authors emphasizing a statistically significant subgroup analysis (n = 6). Five trials used rhetoric to spin their data, 4 emphasized a statistically significant secondary endpoint, and 2 reported only a ratio of events that favored the intervention and omitted the confidence interval or p-value that would have shown that the intervention effect was not statistically significant.

Spin in the abstract conclusions was most often due to authors using a statistically significant surrogate endpoint to highlight the efficacy of their intervention, without caution because of nonsignificant OS data (n = 17). All of the included trials had mature OS data. Five trials emphasized a statistically significant subgroup analysis, 5 wrongly interpreted a nonsignificant p-value as showing equivalence between the experimental and control groups, and 4 distracted from nonsignificant findings by critiquing their trial design. Nine trials were classified as having "Other" evidence of spin, and in these cases the authors either claimed that the intervention was beneficial, despite reporting no significant endpoints, or claimed to accomplish another objective that was not established *a priori* (ex., they conclude that administration of the drug is feasible when they were only assessing survival).

4. Discussion

Our results show that spin is prevalent in the abstracts of oncology clinical trials that measure OS and a surrogate endpoint. The conclusion sections of abstracts were most prone to contain spin. OS was more often a secondary endpoint. As a secondary endpoint, OS was statistically significant only twice; therefore, authors frequently concluded a treatment was effective based on significant surrogate endpoint data

Table 1
Characteristics of all included trials (n = 124) and those with evidence of spin (n = 46).

Characteristic	Total No. (%, 95%CI)	No. With Spin (%)[a]
Journal		
Journal of Clinical Oncology	31 (25.0; 95%CI 18.2%–33%)	16 (34.8; 95%CI 22.7%–49.2%)
Lancet Oncology	30 (24.2; 95%CI 17.5%–32.4%)	5 (10.9; 95%CI 4.7%–23.0%)
Annals of Oncology	19 (15.3; 95%CI 10.0%–22.7%)	10 (21.7; 95%CI 12.3%–35.6%)
New England Journal of Medicine	16 (12.9; 95%CI 8.1%–19.9%)	4 (8.7; 95%CI 3.4%–20.3%)
JAMA Oncology	9 (7.3; 95%CI 3.9%-13.2%)	3 (6.5; 95%CI 2.2%-17.5%)
British Journal of Cancer	8 (6.5; 95%CI 3.3%–12.2%)	5 (10.9; 95%CI 4.7%–23.0%)
Cancer	6 (4.8; 95%CI 2.2%–10.2%)	3 (6.5; 95%CI 2.2%–17.5%)
Lancet	4 (3.2; 95%CI 1.3%–8.0%)	0 (0.0)
Journal of the National Cancer Institute	1 (0.8; 95%CI 0.1%–4.4%)	0 (0.0)
Control arm		
Active drug only	85 (68.5; 95%CI 59.9%–76.1%)	31 (67.4; 95%CI 53.0%–79.1%)
Active + Placebo	22 (17.7; 95%CI 12.0%–25.4%)	4 (8.7; 95%CI 3.4%–20.3%)
Placebo	12 (9.7; 95%CI 5.6%–16.2%)	9 (19.6; 95%CI 10.7%–33.2%)
Other	4 (3.2; 95%CI 1.3%–8.0%)	1 (2.2; 95%CI 0.4%–11.3%)
Surgery	1 (0.8; 95%CI 0.1%–4.4%)	1 (2.2; 95%CI 0.4%–11.3%)
Funding source		
Industry	68 (54.8; 95%CI 46.1%–63.3%)	22 (47.8; 95%CI 34.1%–61.9%)
Mixed (with Industry)	25 (20.2; 95%CI 14.0%–28.1%)	13 (28.3; 95%CI 17.3%–42.6%)
Public (e.g., government)	24 (19.4; 95%CI 13.4%–27.2%)	6 (13.0; 95%CI 6.1%–25.7%)
Private (e.g., foundation)	4 (3.2; 95%CI 1.3%–8.0%)	2 (4.3; 95%CI 1.2%–15.5%)
Hospital	1 (0.8; 95%CI 0.1%–4.4%)	1 (2.2; 95%CI 0.4%–11.3%)
Not mentioned	1 (0.8; 95%CI 0.1%–4.4%)	1 (2.2; 95%CI 0.4%–11.3%)
Other	1 (0.8; 95%CI 0.1%–4.4%)	1 (2.2; 95%CI 0.4%–11.3%)

[a] Percentage is based on the individual sample, rather than the entire population of trials (n = 124).

alone. And while it is not spin to discuss statistically significant surrogate endpoints when they are the primary endpoint of a trial, we considered it to be spin to ignore nonsignificant OS data and conclude that a treatment is definitively effective based on a surrogate endpoint alone. These results are in line with previous investigations of spin both in oncology (Altwairgi et al., 2012; Djulbegovic et al., 2011; You et al., 2012) and the overall medical literature (Boutron et al., 2010; Boutron and Ravaud, 2018; Chiu et al., 2017), indicating that spin continues to affect the accurate interpretation of trial results by physicians. The implication of this finding is that misrepresented, distorted research findings are being purported as true to oncologists.

Misrepresented or distorted research findings affect oncologist beliefs about drug efficacy. The SPIIN randomized trial demonstrated that oncologists believe drugs are more effective if the clinical trial abstract has spin in the conclusions (Boutron et al., 2014). Furthermore, oncologists are also more likely to read the full text of a clinical trial that has spin in the abstract. The reading habits of oncologists indicate that the effects of spin may be compounded since investigations of spin in the full text of trials have demonstrated that spin occurs at a similar rate (Boutron et al., 2010). It is known that internists often read only study abstracts, either to quickly learn or to screen out uninteresting results (Saint et al., 2000). If these findings hold true for oncologists, trial authors may be incentivized to spin nonsignificant results, since spin leads to more interest among readers, and may improve the chances of favorable peer reviews and publication.

The tendency for trial authors to emphasize statistically significant surrogate endpoints when OS is nonsignificant is not surprising. Surrogate endpoints are increasingly important to the field of oncology and often are the basis for new drugs approvals under the FDA Accelerated Approval pathway (Beaver et al., 2018). Clinical trial authors may believe that the intervention drug is truly effective, even without OS data. Optimism bias toward new oncology interventions has been described previously (Djulbegovic et al., 2011). However, when OS data is mature, available, and nonsignificant it may be difficult for authors to conclude that their drug is beneficial. Even if OS is a secondary endpoint, concluding that a treatment is beneficial may be

Table 2
Frequency and manifestations of spin among included trials (n = 124).

Location, type, and frequency of spin	No. (%)
Abstracts with evidence of spin	46 (37.1; 95%CI 29.1%–45.9%)
Abstracts with spin in the title	0 (0.0)
Abstracts with spin in the results[a]	19 (15.3; 95%CI 10.0%–22.7%)
Focus on statistically significant subgroup analysis	6
Use of suggestive language (e.g., trend toward significance)	5
Omit statistically nonsignificant OS primary endpoint data	4
Focus on hazard ratio, omit confidence interval and p-value	2
Focus on statistical significance, ignoring small effect size	1
Other	5
Abstracts with spin in the conclusions[a]	40 (32.3; 95%CI 24.7%–40.9%)
Recommend use of drug based on surrogate endpoint alone	17
Emphasis on statistically significant subgroup analysis	5
Interpreting a nonsignificant *P* value as showing noninferiority	5
Focus on flaws in trial design rather than nonsignificant results	4
Use of suggestive language (e.g., trend toward significance)	3
Focus on statistical significance, ignoring small effect size	1
Other	9

[a] Some abstract results and conclusions had more than one type of spin present, simultataneously.

difficult when only significant surrogate endpoint data are available at the time.

When authors deemphasize available OS data, there may be consequences for patients. Patients receiving adjuvant sunitinib may, like the oncologist-authors and the FDA reviewers, believe the drug is more effective than it truly is. Surrogate endpoints are useful when they predict OS early and accurately, but surrogate endpoints are incapable of completely replacing OS. Caution may be warranted in a trial that has statistically significant surrogate endpoint data and nonsignificant OS data, regardless of which endpoint is primary or secondary, since OS is what the surrogate endpoint is trying to predict.

To conclude, this investigation of spin in the abstracts of oncology clinical trials measuring OS and a surrogate endpoint shows that spin is common. Further, as a secondary endpoint, OS was statistically significant twice, raising questions about trial design and the utility of OS as a secondary endpoint. Nevertheless, authors frequently conclude a treatment is effective based on only statistically significant surrogate endpoint data. Spin was most common in the conclusion sections of abstracts, where authors interpret their results. The consequences of spin may include confusion about the true efficacy of a drug for patients and the dissemination of distorted conclusions to oncologists.

This study is limited by the 1-year cross section that was chosen for analysis. It is possible that our results do not reflect the reporting of oncology trials outside the chosen time frame, including clinical trials that were published in 2018 and later. Readers should account for this limitation when interpreting our study results.

Key message

Spin, the misrepresentation of research findings, may affect the interpretation of trial results. In oncology, surrogate endpoints are used as an indirect measure of the patient-important endpoint, overall survival. We found that authors of trials commonly use spin to emphasize surrogate endpoints when overall survival was not statistically significant.

Declaration of competing interest

The authors have no conflicts of interest.

Disclosures

None.

Funding

None.

CRediT authorship contribution statement

C. Wayant: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Validation, Writing - original draft, Writing - review & editing. **D. Margalski:** Data curation, Formal analysis, Investigation, Validation, Writing - original draft, Writing - review & editing. **K. Vaughn:** Data curation, Formal analysis, Investigation, Validation, Writing - original draft, Writing - review & editing. **M. Vassar:** Conceptualization, Investigation, Methodology, Project administration, Supervision, Validation, Writing - original draft, Writing - review & editing.

Acknowledgement

None.

References

Altwairgi, A.K., Booth, C.M., Hopman, W.M., Baetz, T.D., 2012. 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.

Barry, H.C., Ebell, M.H., Shaughnessy, A.F., Slawson, D.C., Nietzke, F., 2001. Family physicians' use of medical abstracts to guide decision making: style or substance? J. Am. Board Fam. Pract. 14, 437–442.

Beaver, J.A., Howie, L.J., Pelosof, L., Kim, T., Liu, J., Goldberg, K.B., Sridhara, R., Blumenthal, G.M., Farrell, A.T., Keegan, P., Pazdur, R., Kluetz, P.G., 2018. A 25-year experience of US food and drug administration accelerated approval of malignant hematology and oncology drugs and biologics: a review. JAMA Oncol. 4 (6), 849–856. https://doi.org/10.1001/jamaoncol.2017.5618.

Boutron, I., Altman, D.G., Hopewell, S., Vera-Badillo, F., Tannock, I., Ravaud, P., 2014. Impact of spin in the abstracts of articles reporting results of randomized controlled trials in the field of cancer: the SPIIN randomized controlled trial. J. Clin. Oncol. 32, 4120–4126.

Boutron, I., Dutton, S., Ravaud, P., Altman, D.G., 2010. Reporting and interpretation of randomized controlled trials with statistically nonsignificant results for primary outcomes. JAMA 303, 2058–2064.

Boutron, I., Ravaud, P., 2018. Misrepresentation and distortion of research in biomedical literature. Proc. Natl. Acad. Sci. U. S. A. 115, 2613–2619.

Cheema, P.K., Burkes, R.L., 2013. Overall survival should be the primary endpoint in clinical trials for advanced non-small-cell lung cancer. Curr. Oncol. 20, e150–e160.

Chiu, K., Grundy, Q., Bero, L., 2017. "Spin" in published biomedical literature: a methodological systematic review. PLoS Biol. 15, e2002173.

Djulbegovic, B., Kumar, A., Magazin, A., Schroen, A.T., Soares, H., Hozo, I., Clarke, M., Sargent, D., Schell, M.J., 2011. Optimism bias leads to inconclusive results-an empirical study. J. Clin. Epidemiol. 64, 583–593.

Food And Drug, 2015. Clinical Trial Endpoints for the Approval of Non Small Cell Lung Cancer Drugs and Biologics: Guidance for Industry [WWW Document]. URL: https://www.fda.gov/downloads/drugs/guidances/ucm259421.pdf. (Accessed 23 January 2018). .

Gyawali, B., Goldstein, D.A., 2018. The US Food and Drug Administration's approval of adjuvant sunitinib for renal cell cancer: a case of regulatory capture? JAMA Oncol. 4, 623–624.

Kemp, R., Prasad, V., 2017. Surrogate endpoints in oncology: when are they acceptable for regulatory and clinical decisions, and are they currently overused? BMC Med. 15, 134.

Marcelo, A., Gavino, A., Isip-Tan, I.T., Apostol-Nicodemus, L., Mesa-Gaerlan, F.J., Firaza, P.N., Faustorilla Jr, J.F., Callaghan, F.M., Fontelo, P., 2013. A comparison of the accuracy of clinical decisions based on full-text articles and on journal abstracts alone: a study among residents in a tertiary care hospital. Evid. Med. 18, 48–53.

Ouzzani, M., Hammady, H., Fedorowicz, Z., Elmagarmid, A., 2016. Rayyan—a web and mobile app for systematic reviews. Syst. Rev. 5, 210.

Prasad, V., Kim, C., Burotto, M., Vandross, A., 2015. The strength of association between surrogate end points and survival in oncology: a systematic review of trial-level meta-analyses. JAMA Intern. Med. 175, 1389–1398.

Ravaud, A., Motzer, R.J., Pandha, H.S., George, D.J., Pantuck, A.J., Patel, A., Chang, Y.-H., Escudier, B., Donskov, F., Magheli, A., Carteni, G., Laguerre, B., Tomczak, P., Breza, J., Gerletti, P., Lechuga, M., Lin, X., Martini, J.-F., Ramaswamy, K., Casey, M., Staehler, M., Patard, J.-J., S-TRAC Investigators, 2016. Adjuvant sunitinib in high-risk renal-cell carcinoma after nephrectomy. N. Engl. J. Med. 375, 2246–2254.

Saint, S., Christakis, D.A., Saha, S., Elmore, J.G., Welsh, D.E., Baker, P., Koepsell, T.D., 2000. Journal reading habits of internists. J. Gen. Intern. Med. 15, 881–884.

Tan, A., Porcher, R., Crequit, P., Ravaud, P., Dechartres, A., 2017. Differences in treatment effect size between overall survival and progression-free survival in immunotherapy trials: a meta-epidemiologic study of trials with results posted at ClinicalTrials.gOv. J. Clin. Oncol. 35, 1686–1694.

U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER), 2007. Guidance for Industry Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics [WWW Document]. URL: https://www.fda.gov/media/71195/download. (Accessed 6 February 2018). .

Vera-Badillo, F.E., Shapiro, R., Ocana, A., Amir, E., Tannock, I.F., 2013. Bias in reporting of end points of efficacy and toxicity in randomized, clinical trials for women with breast cancer. Ann. Oncol. 24, 1238–1244.

Wayant, C., 2018. Search Strategy [WWW Document]. Open Science Framework.

You, B., Gan, H.K., Pond, G., Chen, E.X., 2012. Consistency in the analysis and reporting of primary end points in oncology randomized controlled trials from registration to publication: a systematic review. J. Clin. Oncol. 30, 210–216.

Zhang, S., 2017. Adjuvant sunitinib in Renal-Cell Carcinoma. N. Engl. J. Med.