Bias in Reporting of Randomized Clinical Trials in Oncology

By

Francisco Emilio Vera-Badillo

Thesis
Submitted in conformity with the requirements
for the degree of Master in Sciences.

Institute of Medical Science
at the University of Toronto

© Copyright by Francisco Emilio Vera-Badillo (2014)
Bias in Reporting of Randomized Clinical Trials in Oncology

Francisco Emilio Vera-Badillo
Masters in Science 2014
Institute of Medical Science
University of Toronto

Abstract

Background:
Bias in reporting efficacy and toxicity in clinical trials can influence treatment decisions. We evaluated the quality of reporting the primary endpoint and of toxicity in Randomized Clinical Trials in medical oncology.

Methods:
Three independent studies were undertaken. The first and second studies, identify use of spin in the reporting of the primary endpoint in the conclusion statement of the abstract and underreporting of toxicity. The third study assessed the prevalence of honorary and ghost authorship.

Results:
All studies met their objectives showing that the prevalence of spin in the reporting of primary endpoints, underreporting of toxicity and the incidence of ghost and honorary authors are highly prevalent in oncology literature.

Conclusion:
Use of spin in reporting of outcomes is common for studies with a negative primary endpoint. Reporting of toxicity is limited, especially for studies with positive primary endpoints. Ghost and honorary authorship are prevalent.
Acknowledgements

First and foremost, I would like to thank my supervisor, Dr. Ian F. Tannock, for all his time, effort, expertise and guidance that he has given to this project. I also thank my co-supervisor Dr. Eitan Amir for his time and advice in all moments. I would like also thank to the members of my thesis committee: Dr. Monika Krzyzanowska, Dr. An-Wen Chan and Dr. Shabir Alibhai for their guidance and thoughtful reviews on my work. It has been a pleasure to work with all of you. I also thank Arnoud J. Templeton and Alberto Ocaña for helpful discussions and comments.

Thank you to my family for their support. I want to recognize especially to my wife and daughters that embarked on this amazing experience of moving to Canada and faced the challenges of a new life. Without their support it would not be possible.

This thesis was supported in part by a National Council for Science and Technology (CONACYT) – Mexico grant, and by an Institute for Innovation and Technology Transference of the state of Nuevo Leon grant.
Contributions

I want to thank Marc Napoleone, summer student assigned to Dr. Ian F. Tannock service from May to August 2013 for his invaluable efforts for data collection and for his contributions for improving our database.
Financial Support

I gratefully acknowledge the Consejo Nacional para la Ciencia y la Tecnología (México) and the Instituto de Innovación y Transferencia de Tecnología de Nuevo León (México) for grant funding to complete my Masters program.
Table of Contents
Abstract ........................................................................................................................................1
Acknowledgments .............................................................................................................................5
Contributions .......................................................................................................................................6
Financial support .................................................................................................................................7
Chapter 1: Introduction .......................................................................................................................8
   I. Thesis Objectives ..........................................................................................................................8
   II. Background ..................................................................................................................................9
   III. Overview of the Thesis ............................................................................................................13
   IV. Methodological Considerations ...............................................................................................14

Chapter 2: (Paper 1): Bias in Reporting of Endpoints of Efficacy and Toxicity in Randomized Clinical Trials for Women with Breast Cancer ........................................................................16
   I. Abstract .......................................................................................................................................16
   II. Introduction ..................................................................................................................................18
   III. Methods ....................................................................................................................................20
      i. Literature search and study selection .......................................................................................20
      ii. Data extraction and analysis ..................................................................................................20
      iii. Endpoints .................................................................................................................................21
      iv. Statistical analysis ..................................................................................................................23
      v. Results .....................................................................................................................................24
      vi. Bias in the reporting of the primary endpoint .......................................................................24
      vii. Bias in reporting of toxicity ................................................................................................25
      viii. Influence of funding on results ..........................................................................................26
      ix. Discussion ................................................................................................................................28
      x. Disclosures ..............................................................................................................................32
     xi. Figure 1 ....................................................................................................................................33
     xii. Figure 2 ......................................................................................................................................34
     xiii. Figure 3 ....................................................................................................................................36
     xiv. Table 1 .....................................................................................................................................37
     xv. Table 2 ......................................................................................................................................38
     xvi. References .............................................................................................................................39

Chapter 3: (Paper 2): Bias in Reporting of Randomized Clinical Trials in Oncology ..........................42
Chapter 5: Discussion

I. Contribution to literature ................................................................. 95
II. Strengths and limitations ................................................................. 99
   a. Strengths .................................................................................. 99
   b. Limitations ............................................................................... 101
III. Future Research ........................................................................... 103
IV. Conclusions .................................................................................. 104
V. References ..................................................................................... 106
Chapter 1: Introduction

1.1: Thesis Objectives

Evidence-based clinical medicine relies on publication of high-quality data to determine standards of patient care [1]. Accurate presentation of the results of a randomized controlled trial (RCT) is the cornerstone of the dissemination of the results and their implementation in clinical practice[2]. Scientific articles are not simply reports of facts, and authors have many opportunities to consciously or subconsciously shape the impression of their results for readers; bias in use of language reporting outcomes (i.e. spin) can distort the interpretation of results and mislead readers. The use of these techniques can result from ignorance of the scientific issues, unconscious bias, or willful intent to deceive[3]; and favorable results are often highlighted while unfavorable data are suppressed[4].

Appropriate authorship establishes accountability, responsibility, and credit for scientific information reported in biomedical publications. However, misappropriation of authorship undermines the integrity of the authorship system and can be associated with other types of bias (i.e. selection bias)[5].

The objectives of this thesis were therefore:

1. To develop tools to measure bias in reporting of efficacy and toxicity and use them in a study of RCTs for breast cancer;
2. To perform an analysis of all RCTs reported from July 2010 to December 2012 evaluating systemic therapy for cancer to assess bias in reporting outcomes and toxicity, and to evaluate the consistency of endpoints among protocols, clinical trial registries and final publications. Funding and conflicts of interest will be evaluated as potential predictors for biased reporting;

3. And finally, through the analysis of protocols and manuscripts to determine the prevalence of ghost and honorary authorship in the cohort of studies used for the second aim.

These studies draw on several methods relating to clinical epidemiology, including assessment of bias in reporting outcomes, evaluation of quality in reporting toxicity, and the use of different statistical methods to establish associations of bias and its potential predictors.

1.2. Background

Phase III randomized clinical trials (RCTs) are designed to detect or exclude clinically important differences between experimental and control groups in endpoints that reflect benefit to patients.[6] Such trials provide the gold standard to evaluate the efficacy and toxicity of new drugs before approval by regulatory authorities.[7, 8]

Appropriate design and objective reporting of RCTs in journals are essential to inform clinicians about the activity and safety of new medical interventions.
Outcomes should normally include at least one endpoint reflecting potential benefit and at least one reflecting potential harm (e.g. grade III-IV adverse events).

Several factors affect the quality and credibility of studies reported in the medical literature. Among those factors are how studies are reported (and especially the concluding statement of the abstract) and whether that is consistent with the statistical results [9, 10], if endpoints are changed during the course of a clinical trial (usually to allow reporting of a positive result) [11], if toxicity is clearly reported [12] and how funding (especially from the pharmaceutical industry) affects reporting of results.

Bias in the reporting of outcomes has been explored previously in RCTs with statistically nonsignificant results for primary outcomes. Boutron et al reported presence of spin in 58% and 50% of conclusions in the abstract and body of the manuscript, respectively. [2] Spin, a type of bias, is defined as use of reporting strategies to highlight that the experimental treatment is beneficial, despite a statistically non-significant difference for the primary outcome, or to distract the reader from statistically non-significant results [13]. It is important to recognize the presence of bias and spin in reports of clinical trials, and to evaluate their importance when placing an RCT in context and ascribing a level of credibility [14].

Reviews have shown that a substantial proportion of clinical trials have suboptimal reporting of harm [12]. Pitrou et al reported that adverse events were
described only in 88% of 133 RTCs analyzed, and were poorly reported in abstracts (71%); of relevance is that 27% of the included studies did not report the severity of adverse events, and only 16% of studies mentioned explicitly grading of severity[13]. Of great importance and under-recognized is that trials are usually underpowered to detect differences in harms between their arms, so that the commonly-used phrase “no significant differences were found” is misleading[15]. The lack of prominence given to side effects is such that a previous study by Seruga et al reported that 39% of potential serious adverse drug reactions were not described in the cohort of assessed studies[16].

Selection of endpoints or outcome measures is another concern, although in 2004 the ICMJE published guidelines for mandatory registration of clinical trials[17], and in 2007 the Surgical Journal Editors Group followed these recommendations[18]; consistency between a clinical trial registry and the final manuscript in the reporting of primary and secondary endpoints of surgical RCTs was reported recently to be poor: only 55% of the published papers showed no discrepancy while in 45% of manuscripts there was omission, addition, change in definition, downgrading or upgrading of outcomes[19]; another paper showed similar results, 49% discrepancies in the reporting of primary outcomes[20].

Authorship has been a difficult topic to address since it implies personal criticism of colleagues, but inappropriate authorship has substantial implications. In 2008 Ross et al analyzed the literature relating to rofecoxib, a non-steroidal anti-inflammatory drug removed from the market after showing a high risk of cardiovascular events, and showed that a substantial number of papers was
requested through contracts with medical publishing companies that recruited external, academically affiliated investigators be authors [21].

Financial ties might be expected to aggravate biased reporting and inappropriate authorship. Authors of studies funded by the pharmaceutical industry, and having a key role in the design, analysis, and interpretation or reporting of a trial may have conflicts of interest with the sponsor [22, 23]. Sponsorship, other sources of funding and conflicts of interest (COI) of authors can influence the reporting of results of trials relating to efficacy and safety. [24] Clinical trials are tools with potential to change the standard of care, and therefore have a substantial secondary economic impact [25] and can be used as a marketing strategy. Previously most research was linked to academic institutes; however in recent times this leadership has shifted to for-profit-organizations (the pharmaceutical industry), which sponsors >50% of published trials. [23, 26] Results that are unfavorable to the sponsor, for example negative studies, where the experimental drug was not superior to the standard or care, or when the new drug is significantly more toxic compared with other alternatives, can pose considerable financial risks to companies. Pressure to show that the drug causes a favorable outcome may results in biases in design, outcome, and reporting of industry-sponsored research. [27-30]

Most of the above studies relating to spin, bias and inappropriate authorship were related to studies in general internal medicine, and their incidence in the medical oncology literature is limited. Assessment of reporting in oncology is
important to establish guidelines that will impact on quality of information provided to readers.

1.3. Overview of the Thesis.

The first study is a cohort of RCTs in the breast cancer literature comprising published manuscripts from 1995-2011. This study was performed to develop tools to measure bias in reporting of outcomes and reporting toxicity.

The second study involves a cohort of studies in the field of medical oncology evaluating medical interventions in phase II and phase III RCTs, where possible original protocols were obtained. Consistency in the reporting of primary outcomes among protocols, clinical trial registries and final publications were evaluated by two investigators, including the MSc candidate through a previously design extraction form. Previously-developed tools were used to assess spin in reporting outcomes and under-reporting of toxicity in the selected studies.

For the first two studies funding was evaluated as a predictor of bias and for the second study financial ties of the first/corresponding author (when different) were also analyzed as a predictor.

The third and final study explores the incidence of ghost and honorary authorship in the second cohort of papers. My goal was that these studies should provide data to better inform clinicians about quality of reporting and inform editors about
how to improve the review process of the submitted literature before its publication.

1.4. Methodological Considerations.

For the first study data were presented descriptively as means or medians. Predictors of bias were assessed by the Chi-squared test and by univariable logistic regression (categorical variables) or univariable linear regression (continuous variables). Correlations between variables were tested using Spearman’s correlation and the magnitude of association was assessed as described by Burnand et al[31]. For the second and third studies data were presented descriptively as means with their standard deviations. Predictors of bias were assessed by univariable and multivariable logistic regression.

All statistical analyses were conducted using SPSS statistical software version 17 (IBM Corp, Arkmon, New York). All significance tests were two-sided using an alpha level of 0.05. No correction was applied for multiple statistical testing.

The remainder of the thesis presents the results of each study in the form of three separate papers (Chapter 2-4), as summarized below. The final chapter (Chapter 5) provides discussion of the relevance of the findings to health care providers, limitations of the studies, and recommendations for future research.
Chapter 2: (Paper 1): Bias in Reporting of Endpoints of Efficacy and Toxicity in Randomized Clinical Trials for Women with Breast Cancer

This article was published as:


Annals of oncology : official journal of the European Society for Medical Oncology / ESMO. See Appendix for permission to re-produce.

2.1. Abstract

Introduction Phase III randomized controlled trials (RCTs) assess clinically important differences in endpoints that reflect benefit to patients. Here we evaluate the quality of reporting of the primary endpoint and of toxicity in RCTs for breast cancer.

Methods PUBMED was searched from 1995-2011 to identify RCTs for breast cancer. Bias in the reporting of the primary endpoint and of toxicity was assessed using pre-designed algorithms. Associations of bias with Journal Impact Factor (JIF), changes in primary endpoint compared to information in ClinicalTrials.gov and funding source were evaluated.

Results Of 164 included trials, 33% showed bias in reporting of the primary endpoint and 67% in the reporting of toxicity. The primary endpoint was more likely to be reported in the concluding statement of the abstract significant
differences favoring the experimental arm were shown; 59% of 92 trials with a negative primary endpoint used secondary endpoints to suggest benefit of experimental therapy. Only 32% of articles indicated the frequency of grade III-IV toxicities in the abstract. A positive primary endpoint was associated with under-reporting of toxicity.

**Conclusion** Bias in reporting of outcome is common for studies with negative primary endpoints. Reporting of toxicity is poor especially for studies with positive primary endpoints.
2.2. Introduction

Phase III randomized clinical trials (RCTs) are designed to detect or exclude clinically important differences between experimental and control groups in endpoints that reflect benefit to patients.\cite{6} Such trials provide the gold standard to evaluate the efficacy and toxicity of new drugs before approval by regulatory authorities.\cite{7, 8}

Appropriate design and objective reporting of RCTs in journals are essential to inform clinicians about the activity and safety of new medical interventions. It is good practice to design RCTs with no more than three outcomes for which hypothesis-testing is planned.\cite{32} Otherwise multiple significance testing may lead to apparently significant results that occur by chance. These outcomes should normally include at least one endpoint reflecting potential benefit and at least one reflecting potential harm (e.g. grade III-IV adverse events). Reviews have shown that a substantial proportion of clinical trials have suboptimal reporting of harm.\cite{15} Guidelines such as Consolidated Standards of Reporting Trials (CONSORT) can improve the quality of reporting of clinical trials.\cite{33}

Bias in reporting of clinical trials and selective publication can create false perceptions of drug efficacy and safety. There is evidence for selective reporting of favorable results and suppression of unfavorable data from publication, leading to inappropriate conclusions\cite{7, 34}. This may be influenced by publication bias - the association between positive results and acceptance of
reports for publication[32, 35]. Selection bias can affect not only the interpretation of the trial itself but also of subsequent systematic reviews or overviews, producing inaccurate summaries of research[7, 36] and misrepresentation of toxicity[37]. Reporting of harms may be viewed as discrediting the reporting of benefits.

Spin, a type of bias, is defined as use of reporting strategies to highlight that the experimental treatment is beneficial, despite a statistically non-significant difference for the primary outcome, or to distract the reader from statistically non-significant results[13]. It is important to recognize the presence of bias and spin in reports of clinical trials, and to evaluate their importance when placing a RCT in context and ascribing a level of credibility[14].

Here we review papers reporting RCTs for breast cancer to quantify the extent of biased reporting, and to guide readers in judging the credibility of their conclusions. Because busy clinicians often read only the abstracts of publications[38] we have emphasized accurate reporting of the primary endpoint and toxicity in the abstract. We hypothesized that despite the availability of guidelines to minimize bias in reporting, this remains prevalent.
2.3. Methods

2.3.1. Literature search and study selection

We performed an electronic search of MEDLINE (Host: PubMed) for publications from January 1995 to August 2011 using the following MeSH terms: Randomized Clinical Trial, Randomized Controlled Trial, Phase III AND Breast Neoplasms or Breast Cancer. Inclusion criteria were human studies published in English and including patients age 18 years or older. We excluded trials with sample size <200 patients as they were unlikely to be definitive studies and more likely to have higher levels of bias.[12] Furthermore the focus of this study was to assess reporting of clinical trials that potentially change clinical practice. Other exclusion criteria included trials where the primary endpoint was not a time to event endpoint, commentaries, review articles, observational studies, meta-analyses, ongoing studies, and articles for which only the abstract was available.

2.3.2. Data extraction and analysis

The following data were extracted independently from each RCT by two authors (FV and RS); setting of treatment (adjuvant vs metastatic), sponsorship (industry vs non-industry or not stated), year of publication, impact factor of the journal where the trial was published (as a continuous variable), the primary and secondary endpoints (overall survival [OS], progression-free survival [PFS], disease-free survival [DFS], response rate, toxicity, or quality of life), whether the
primary endpoint was defined in the abstract and/or in the paper and whether secondary endpoints were reported in the abstract and in the paper. If data on PFS or DFS were not reported, similar endpoints such as time to progression (TTP), time to treatment failure (TTF), or event free survival (EFS) were extracted instead. For papers with more than 2 arms, assessment of efficacy was not considered if at least one arm was positive (see figure 1); if all were negative it was included for analysis. For toxicity all trials were included regardless the number of study arms. Journal impact factor at time of the study was extracted from the Gerstein Science Information Center of the University of Toronto through the Web of Science (Host:BIOSIS) and was retrieved from the Journal Citation reports up to 2012.

For recent trials, we also explored whether the primary endpoint listed in the trial registry ClinicalTrials.gov was the same as that reported in abstracts or papers reporting the same trial. We initially searched articles for any reference to trial registration. For those not reporting such data, we manually searched the ClinicalTrials.gov database for trial-related information.

2.3.3. Endpoints

The primary analysis included assessment of the prevalence of spin in reporting the primary endpoint of the study, and in reporting toxicity; the secondary analysis evaluated predictors of bias and spin.
Bias was defined as inappropriate reporting of the primary endpoint and toxicity, with emphasis on reporting of these outcomes in the abstract. A decision tree was used to assess if the primary endpoint was reported with bias, and if a secondary endpoint was used to imply benefit of the experimental arm (Figure 1). Studies where multiple primary endpoints were reported and where at least one endpoint was positive were not considered for assessment of bias.

Bias in reporting of toxicity was assessed using a hierarchy scale from 1 (excellent) to 7 (very poor) to indicate whether reporting of grade 3 and 4 toxicities occurred in the concluding statement of the abstract, elsewhere in the abstract, in the results section of the paper, only in a table or not at all, with lower scores if they were also included in the discussion section of the paper (Figure 2). We defined reporting of grade 3 and 4 toxicities as poor if they were not mentioned in the abstract (scale of 5 to 7 in our hierarchy), and good (scale 1-2) if they were mentioned in the concluding statement of the abstract. Where there were no statistically significant differences in toxicity, a general statement in the abstract was deemed sufficient; where statistically significant differences were seen, it was expected that they would be reported in the abstract.

Spin was defined as the use of words in the concluding statement of the abstract to suggest that a trial with a negative primary endpoint was positive based on some apparent benefit shown in one or more secondary endpoints.

Predictors of bias included the impact of source of funding on reporting of the primary endpoint, the relationship between quality of reporting of the primary
endpoint and of toxicity and the frequency of a change in the primary endpoint from the original protocol (when this could be obtained from ClinicalTrials.gov) to a published paper. Also whether journal impact factor, setting of the trial (adjuvant versus metastatic), definitive versus surrogate endpoint (e.g. OS compared to PFS or DFS) and modification of the primary endpoint influenced the prevalence of bias or spin.

2.3.4. Statistical analysis

Data were presented descriptively as means or medians. Predictors of bias were assessed by the Chi-squared test and by univariable logistic regression (categorical variables) or univariable linear regression (continuous variables). Correlations between variables were tested using Spearman’s correlation and the magnitude of association was assessed as described by Burnand et al[31]. All statistical analyses were conducted using SPSS statistical software version 17 (IBM Corp, Armonk, New York). All significance tests were two-sided using an alpha level of 0.05. No correction was applied for multiple statistical testing.
2.4. Results

A total of 568 articles were identified initially and 164 RCTs (148 for systemic therapy, 11 for radiation therapy, 5 for surgical therapy) were eligible for analysis (Figure 3). Characteristics of the trials are reported in Table 1. Eighty-one trials (49.4%) were conducted in the adjuvant setting and 83 (50.6%) evaluated experimental therapy for women with metastatic breast cancer. Overall survival was the primary endpoint in only 27 trials (16.5%) and DFS or PFS was the primary endpoint in 137 studies (83.5%). Only 30 trials (18%) were identified as included in ClinicalTrials.gov. Among these studies, the primary endpoint was changed in the final report in 7 (23.3%) studies. Seventy two (43.9%) studies were positive with a significant p-value for the difference in primary endpoint favoring the experimental arm compared with 92 (56.1%) with a non-significant p-value. The majority, 150 trials (91.4%), were published in medium or high impact journals; the calculated median impact factor was 19. Bias in reporting efficacy (Spearman’s rho=0.27 and Chi-squared p=0.30) and toxicity (Spearman’s rho=0.46 and Chi-squared p=0.06) were not influenced by date of publication.

2.4.1. Bias in the reporting of the primary endpoint

Fifty-four trials (32.9% of the total sample) were reported as positive, based on a non-primary endpoint, despite not finding a statistically significant difference in the primary endpoint. These reports were biased and used spin in attempts to conceal that bias. When assessing only those reports with a non-significant
difference in the primary endpoint between the arms (N=92), the incidence of this bias increased to 59.0%. Compared to studies with a statistically significant difference between arms in the primary endpoint, studies with a non-significant difference showed a statistically significant association with not reporting the primary endpoint in the concluding statement of the abstract (27% vs 7%, OR=5.15, 95% CI=1.86-14.26, p=0.001. Compared to studies where the primary endpoint did not change, there was a trend for trials with a change of primary endpoint to report a statistically significant difference for (the new) primary endpoint (OR=2.29, 95% CI=0.37-14.32, p=0.47). There was no association between journal impact factor and bias (Spearman's rho= -0.10, p=0.20).

There were no apparent differences in the probability of bias in trials conducted in the adjuvant or metastatic settings (Chi-squared p=0.146). There was also no association between bias and the type of primary endpoint, (OS versus DFS or PFS, Chi-squared p=0.23).

2.4.2. Bias in reporting of toxicity

A total of 110 (67.1%) papers met our definition of biased reporting of toxicity. Distribution of bias according to the hierarchy scale is reported in table 2. There was a statistically significant association between biased reporting of toxicity and observation of a statistically significant difference in the arms for the primary endpoint (OR=2.00, 95% CI=1.02-3.94, p=0.044). There was no association between biased reporting of toxicity and biased reporting of efficacy (Chi-squared
p=0.43), or with change of primary endpoint (OR=0.58, 95% CI=0.1-3.2, p=0.17). The journal's impact factor was not associated with biased reporting of toxicity (Spearman’s rho=-0.153 Chi squared p=0.73). Bias in the reporting of toxicity was significantly associated with use of OS as the primary endpoint (OR= 3.30, 95% CI=1.1-10.1, p=0.028). Reporting of toxicity was not influenced by the setting of the trial (adjuvant versus metastatic, OR=1.68, 95% CI=0.9 to 3.3, p=0.12).

2.4.3. Influence of funding on results

Funding from industry partners was reported in 103 (62.8%) studies, 32 (19.5%) studies were funded by academic or governmental grants and in 29 (17.7) studies the source of funding was not stated. Three studies reported this in the abstract[39-41], while all others reported this information in the body of the manuscript. Success in finding a significant difference between the arms for the primary endpoint and bias in the reporting of this endpoint were not influenced by source of funding (Chi-squared p=0.78 and p=0.71, respectively). Similarly, industry funding was not associated with biased reporting of toxicity (Chi-squared p=0.71). There was no effect of industry funding on odds of change in the primary endpoint (OR=3.20, CI=0.3-31.4, p=0.41). There was a significant, but weak association between funding from industry and higher journal impact factor of the published study (Spearman’s rho=0.39 and Chi-squared p=0.05).
2.5. Discussion

Several papers have evaluated the frequency and characteristics of bias in the reporting of efficacy[1, 2, 8, 32, 42, 43], but these reports have tended to focus on heterogeneous medical conditions and not on cancer clinical trials. Furthermore, there are limited data in the literature about bias in the reporting of toxicity[13]. Here we have explored the frequency of bias in reporting of efficacy and toxicity in randomized trials evaluating treatments for breast cancer. We focus our research on breast cancer given that it is the most common malignancy in women, has substantial mortality[44] and is a cancer site with a large number of trials.

The objective of a phase III RCT is to detect or exclude differences in endpoints that will reflect benefit to patients. The chosen endpoints should be measures of patient benefit (i.e. improved efficacy or better safety and tolerability). Overall survival is the gold standard for the assessment of benefit: it is unambiguous and is not subject to investigator interpretation[45]. PFS or DFS may be suitable endpoints if these measures are valid surrogates for OS[46], or possibly in trials where there is a high rate of cross-over to the experimental arm which confounds interpretation of OS. For women with breast cancer neither DFS nor PFS have been shown to be adequate surrogates for OS[47, 48] but 83.5% of our cohort of trials used these endpoints.

Bias in the reporting of the primary endpoint was prevalent especially when statistical significance of the difference in the primary endpoint between the arms
was not found. You et al[49] evaluated reports of RCTs published between 2005 and 2009, and found that there was misinterpretation of the PE in 21.6% of the trials; this included non-significance in a superiority trial interpreted as showing treatment equivalence, study conclusion based on endpoints other than PE, study considered positive despite a non-significant p-value, and study conclusions based only on one endpoint when there were co-primary endpoints. We found a higher incidence of inappropriate reporting of the PE in RCTs for breast cancer that increased dramatically when only the trials with a non-significant p-value were assessed. Consequently, spin was used frequently to influence, positively, the interpretation of negative trials, by emphasizing the apparent benefit of a secondary endpoint. We found bias in reporting efficacy and toxicity in 32.9% and 67.1% of trials, respectively, with spin and bias used to suggest efficacy in 59% of the trials that had no significant difference in their primary endpoint. These results are similar to those in other areas of medicine[8]. In contrast to those data where bias in the reporting of toxicity was less frequent when the primary endpoint was positive, we found that bias in the reporting of toxicity was higher when the trial had a significant p-value for the difference in the primary endpoint between experimental and control arms. A possible explanation for this finding may be that investigators and/or sponsors then focus on efficacy as the basis of registration and downplay toxicity to make the results more attractive.

To avoid selection for publication of positive trials, and/or publication of a subset of the original recorded outcomes on the basis of the results, registration of trials
is now mandatory. Due to our period of evaluation (1995-2011), only 18% of our trials were registered in ClinicalTrials.gov. In some of these trials the primary endpoint was changed between the time of registration and reporting of their results. Among these trials, there was a trend towards change of the primary endpoint being associated with positive results, suggesting that it may be a strategy to make a negative trial appear positive. This may be on the basis of a low likelihood of observing enough events for this endpoint to be statistically significant or even a lack of effect of the experimental therapy to modify the original primary endpoint (usually overall survival). Trial registration does not necessarily remove bias in reporting outcome, although it does make it easier to detect [42].

The pharmaceutical industry is increasingly influential in clinical trial sponsorship with data showing an increase in industry sponsorship of phase III RCTs from 24 to 72% over a 30 year period [2, 8, 43]. In our cohort of trials, 67% were industry sponsored, but we found no association between industry sponsorship and biased reporting of either efficacy or toxicity, and no association of for-profit sponsorship with change of the primary endpoint between that listed in trial registries and the final publication.

There are some limitations to our study. First, we searched only breast cancer trials and we cannot extrapolate our findings to published reports of trials for other types of cancer. Second, including studies <200 patients would likely to increase level of bias, but the clinical impact of such studies is low. Third, we utilized subjective measures for some of our outcome measures such as the
presence of spin. Fourth, our scales used to assess bias in reporting of efficacy and toxicity were based on our interpretation of the characteristics that a paper has to accomplish to be considered unbiased, but they have not been validated. Fifth, many of our included trials were not available at ClinicalTrials.gov. This database was established in 2002[50] and many trials initiated prior to this date were not included. Furthermore, many European trials were not initially included in the US-based ClinicalTrials.gov database and European Clinical Trials Registries do not have easily searchable databases[51]. Our analysis of change in the primary endpoint should therefore be interpreted with caution.

In conclusion, bias in the reporting of efficacy and toxicity remains prevalent. Clinicians, reviewers, journal editors and regulators should apply a critical eye to trial reports and be wary of the possibility of biased reporting. Guidelines are necessary to improve the reporting of both efficacy and toxicity.
Disclosures: authors declared not to have disclosures.
Figure 2.1. Decision Tree for Assessment of reporting of the Primary Endpoint in the Concluding Statement of the abstract.

PE= Primary endpoint, SE= Secondary endpoint.
Figure 2.2. Hierarchy scale for reporting of adverse events

Adverse Event

- Not in results table (NOT R)
- In results table (R)

- Not in abstract (NOT A)
- In Abstract (A)
  - In discussion (D)
  - Not in concluding statement (NOT C)
  - In concluding statement (C)

- Not in discussion (NOT D)
- In discussion (D)
  - R + (NOT A) + (NOT D)
  - R + A + (NOT C) + (NOT D)
  - R + A + C + (NOT D)
  - R + A + C + D
Not R= Not Reported in Results Table, Not D= Not Reported in Discussion, Not A= Not Reported in Abstract, Not C= Not Reported in Concluding Statement.

R= Reported in Results Table, D= Reported in Discussion, A= Reported in Abstract, C= Reported in Concluding Statement.
Figure 2.3. Flow diagram illustrating selection of articles for analysis.
Table 2.1. Characteristics of trials.

| Difference in Primary Endpoint between arms | Significant | 72 (43.90) |
|                                           | Non Significant | 92 (56.10) |
| Primary Endpoint                           | DFS or PFS     | 137 (83.5) |
|                                           | OS             | 27 (16.5)  |
| Setting                                    | Adjuvant       | 81 (49.4)  |
|                                           | Metastatic     | 83 (50.6)  |
| Funding                                    | Not Industry or Not Stated | 61 (37.20) |
|                                           | Industry       | 103 (62.80) |
| ClinicalTrials.gov                         | PE changed     | 7 (4.2)    |
|                                           | PE not changed | 23 (14)    |
|                                           | Protocol N/A   | 134 (81.8) |

PE= Primary Endpoint; N/A= Not Available; DFS= Disease Free Survival; PFS= Progression Free Survival; OS= Overall Survival.
Table 2.2. Distribution of Bias in Reporting Toxicity.

<table>
<thead>
<tr>
<th>Toxicity Hierarchy Scale</th>
<th>Number of Trials N=164</th>
<th>%</th>
<th>Positive PE (%)</th>
<th>Negative PE (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>17</td>
<td>10</td>
<td>7 (4)</td>
<td>10 (6)</td>
</tr>
<tr>
<td>2</td>
<td>7</td>
<td>4</td>
<td>2 (1)</td>
<td>5 (3)</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>12</td>
<td>4 (2)</td>
<td>16 (10)</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td>6</td>
<td>5 (3)</td>
<td>5 (3)</td>
</tr>
<tr>
<td>5</td>
<td>21</td>
<td>13</td>
<td>11 (7)</td>
<td>10 (6)</td>
</tr>
<tr>
<td>6</td>
<td>55</td>
<td>34</td>
<td>31 (19)</td>
<td>24 (15)</td>
</tr>
<tr>
<td>7</td>
<td>34</td>
<td>21</td>
<td>12 (7)</td>
<td>22 (13)</td>
</tr>
</tbody>
</table>

PE: Primary Endpoint.

Higher numbers and shaded area refers to poor reporting of toxicity.
References


Chapter 3: (Paper 2): Bias in Reporting of Randomized Clinical Trials in Oncology

This article will be submitted as:

3.1. Abstract

Background: Bias in reporting efficacy and toxicity in clinical trials can influence treatment decisions. Here, we describe quality of reporting of the primary endpoint and of toxicity in articles describing randomized controlled trials (RCTs) of cancer therapy, and how this is influenced by financial relationships of the first and corresponding authors with the sponsor.

Methods: We reviewed articles published from July 2010 to December 2012 in six high impact journals to identify phase II and phase III RCTs of systemic treatment for cancer. Bias in reporting of the primary endpoint and toxicity were assessed using pre-defined algorithms. Association of bias with funding source and financial ties of the first and corresponding author were evaluated.

Results: Two hundred articles were identified. For RCTs where there was no statistically significant difference in the primary endpoint, 47% of reports used biased reporting to imply benefit of the experimental treatment. Reporting of toxicity was biased in 18.5% of the studies and was associated with a positive primary
endpoint. Source of funding and financial ties were not associated with biased reporting.

**Conclusion:** Bias in reporting of outcomes is common for studies with a negative primary endpoint. Reporting of toxicity is limited, especially for studies with positive primary endpoints.
3.2. Introduction

Clinical trials are undertaken to evaluate efficacy and toxicity of new interventions. In oncology, phase II trials evaluate if a drug has biological activity in a given tumor site, and if results are encouraging, are followed by large phase III trials that determine if this new intervention is more effective or less toxic than the established standard of care. However, biased reporting can influence the interpretation of clinical trials and can lead to decisions that impact negatively on patient care. They can lead either to a decision to undertake a phase III randomized clinical trial (RCT) where hundreds of patients are exposed to a drug that has not shown appropriate activity or tolerance in a phase II trial, or to inappropriate clinical decisions based on biased reporting of a phase III RCT.

Several factors affect the quality and trustworthiness of studies reported in the medical literature. Among those factors are how studies are reported (and especially the concluding statement of the abstract) and whether that is consistent with the statistical results[9, 10], if endpoints are changed during the course of a clinical trial (usually to allow reporting of a positive result)[11], if toxicity is clearly reported[12] and how funding (especially from the pharmaceutical industry) affects reporting of results.

Scientific articles are not simply reports of facts. Authors have many opportunities to consciously or subconsciously shape the impression of their results for readers; that is, to add language bias to their scientific report.[2, 21, 52] Spin is defined as
use of reporting strategies to highlight that the experimental treatment is beneficial, despite a statistically non-significant difference in the primary outcome, or to distract the reader from statistically non-significant results.[13, 21] It is important to recognize the presence of spin in reports of clinical trials, and to evaluate its importance when placing a RCT in context and ascribing a level of credibility.[14]

Publication bias refers to selective reporting of trials with apparently beneficial results, and together with other strategies such as changing the primary endpoint from a negative one to a new positive endpoint, affects credibility of reported studies[42, 53]. These and other factors stimulated the development of registries of trials, which is publically available, as a key tool to reduce bias in reporting.[54, 55] In 2005 the International Committee of Medical Journal Editors (ICMJE) initiated a policy requiring investigators to deposit information about trial design into an accepted clinical trials registry before the onset of patient enrollment[56], thereby improving transparency. Registries need to meet minimum criteria and editors of most high impact journals have established this as a requirement for publication.[57] Information in clinical trial registries should reflect precisely the protocol used in the clinical trial but there are no reports confirming that data in the registry reflect accurately the protocol; discrepancies between endpoints reported in the registry and those finally reported in the manuscript have been reported in up to 49% of trials.[20] Clarity in registration is required to determine whether there was a deviation from the protocol.[58]
Since the 1990’s various authors have suggested that the original protocol should be submitted together with any manuscript that reports the results of a RCT; this would allow reviewers and editors to evaluate whether there is evidence of manipulation of the trial design or its statistical analysis to make the study appear “positive”. Interestingly, at least one editor-in-chief from the New England Journal of Medicine (NEJM) considered that to be unnecessary.[59] However, since July 2010 all papers published in the NEJM include a copy of the original protocol as supplementary data. Also, since 2009: “The Journal of Clinical Oncology (JCO) believes that, for the editors and reviewers to provide appropriate peer review, a redaction of the protocol or the entire protocol for all (randomized) phase II and III studies must be provided”. The goal of this measure is to increase standards for credibility and transparency of clinical trials reporting.[60]

Here we review manuscripts reporting RCTs evaluating systemic therapies for cancer to quantify the extent of biased reporting and impact of financial relationships on bias reporting. We also recorded potential conflicts of interest of lead authors. We hypothesized that despite the availability of guidelines to minimize bias, this remains prevalent and can be influenced by author’s financial ties.

3.3. Methods

3.3.1. Literature search and study selection.

A comprehensive search of all papers published from July 2010 to December 2012 in NEJM, The Lancet, Journal of the American Medical Association (JAMA), Lancet
Oncology, JCO and the Journal of the National Cancer Institute (JNCI) was performed manually to extract papers reporting results of RCTs in cancer patients. The rationale for selection of the highest impact journals was the assumption of high-quality reports that are likely to impact clinical practice within these journals. Supplementary sections of articles were also accessed to obtain the trial protocol when available. When not available the editorial offices were contacted by email to request copies of the protocol; if this was not successful the corresponding author or the sponsor was contacted.

Inclusion criteria were all two-arm, parallel group, phase II and phase III superiority RCTs reporting results of experimental medical interventions in oncology. We excluded studies of biomarker analysis, non-primary publications, reports of non-randomized trials and phase I/II trials, radiation trials, surgical trials, trials comparing drug sequences, non inferiority trials, and multi-arm trials. These exclusion criteria were used to ensure reasonable homogeneity in the sample and focus on trial that are likely to change clinical practice.

3.3.2. Data extraction and analysis

Two authors independently extracted the following data from the primary manuscript describing each RCT by two authors (FV-B and MN): PubMed identifier number, year of publication, journal of publication, journal impact factor (JIF) at time of this study, country of origin of the first author, phase of the trial were the investigation was carried out, number of patients enrolled, disease site, origin of funding (pharma only, mixed, non-pharma), and population included in the study.
(i.e. adults, pediatric). For financial ties, we extracted information for the first author and for the corresponding author only when they were different; potential conflicts of Interest (COI) were those disclosed in the manuscript.

3.3.3. Endpoints

Bias was defined systematically. An article was considered biased if it met at least one of the following criteria; (1) Bias of efficacy was assessed using a decision tree whether the primary endpoint was reported with spin in the concluding statement of the abstract or the conclusion, and whether a secondary endpoint was used to imply benefit of the experimental arm as defined by Pitrou[13] and McGauran[21] and/or (2) underreporting of toxicity as described previously.[11] See Figure 1 in supplementary material. For studies with >1 primary endpoint, the endpoint analyzed was the one for which the study was powered.

Bias in reporting of toxicity was assessed using a hierarchical scale to indicate whether reporting of grade 3 and 4 toxicities occurred in the concluding statement of the abstract, elsewhere in the abstract, in the results section of the paper, only in a table or not at all, emphasizes was done in table presentation, given that as it was mentioned before, most of readers pay attention in summary of data.[38] We defined reporting of grade 3 and 4 toxicities as poor if they were not mentioned in the abstract or in a table and good if they were mentioned in the concluding statement of the abstract. When there were no statistically significant differences in toxicity, a general statement in the abstract was deemed to be sufficient; when
statistically significant differences were seen, it was expected that they would be reported in the abstract, see Figure 2 in supplementary material.

We analyzed source of funding and first-author financial ties with the sponsor as potential predictors of bias, when the corresponding author was different from first-author, assessment for this last was also included. Funding was evaluated in all studies as pharma versus non-pharma sponsored. Potential COI of first or corresponding author was analyzed under three headings: (1) no COI, (2) research funding, honoraria, consulting, expert testimony and/or other; and (3) employment and/or stock ownership.

Other predictors included for analysis were JIF, phase of the study, country of origin of the first author at the time of article publication, and the primary endpoint used (overall survival versus a surrogate).

3.3.4. Missing Data

Missing data were considered to be missing at random and no further analysis or correction was performed.

3.3.5. Statistical Analysis

Data are presented descriptively as means with their standard deviations. Predictors of bias were assessed using univariable logistic regression analysis and reported as odds ratio (OR) and their respective 95% confidence intervals (CI). Forward elimination multivariable logistic regression at the p<0.05 threshold was
planned; however only 1 variable met this criterion and therefore multivariable analysis was not conducted. All statistical analyses were conducted using SPSS statistical software version 17 (IBM Corp, Armonk, New York). All significance tests were two-sided using an alpha level of 0.05. No correction was applied for multiple statistical testing. Interreviewer agreement was assessed using Cohen κ statistic.

3.4. Results

A total of 403 articles were identified initially and 200 RCTs (48 phase II studies and 152 phase III studies) were eligible for analysis, see Figure 1. The characteristics of the trials are described in Table 1. Ninety-four protocols were available for comparison: 22 protocols for phase II and 72 protocols for phase III RCTs. In 33 studies (16.5%) corresponding authors were different from first authors. Cohen κ for inter-reviewer agreement was 0.88 (95% CI=0.80-0.96) for spin assessment and 0.88 (95% CI=0.81-0.96) for under-reporting of toxicity assessment.

3.4.1. Consistency and Spin in reporting outcomes

One hundred and ninety-three clinical trials were registered in a clinical trial registry (96.5%); for two studies authors declared that was not a requirement because studies were started in 2002 and for the remaining 5 studies information was not available despite our attempts to contact the authors. The primary endpoint was consistent among the protocol (n=94), clinical trial registry (n=193) and the final
Publication (n=193) in 99% of studies, but in two studies (1%) description of the primary endpoint in the protocol and clinical trial registry was vague and no inference about change of endpoint over time could be made.

In 107 RCTs the difference in the primary endpoint between arms of the trials was statistically non-significant, spin reporting outcomes was present in 50 (47.7%) concluding statements of the abstract and in 45 (42.1%) concluding statements of the manuscript. Ten studies (5%) used spin only in the abstract conclusion but not the manuscript conclusion; and no studies were biased in the conclusion of the manuscript, but not in the abstract. See supplementary material for examples.

In univariable analysis,Spin in reporting efficacy outcomes was associated with a high JIF (OR=0.89; 95%CI=0.82-0.97; p=0.004) of the journal where it was reported; but was not associated with the phase of the study (OR=0.75, 95%CI=0.37-1.56; p=0.45), with the country of origin of the first author (OR=0.87, 95%CI=0.45-1.69; p=0.68), with availability of the protocol or with having OS or a surrogate as primary endpoint (OR=0.91, 95%CI=0.58-1.43; p=0.69).

3.4.2. Under-reporting toxicity

A total of 37 (18.5%) papers met our definition of under-reporting of toxicity. Distribution of bias according to the hierarchial scale is reported in Table 3. There was a statistically significant association between under-reporting of toxicity and observation of a statistically significant difference in the arms for the primary
endpoint; studies with a positive primary endpoint were more likely to under-report toxicity (OR=0.21, 95% CI=0.087-0.503; p<0.001) compared to those studies with a negative primary endpoint. In univariable analysis, under-reporting of toxicity was not associated with JIF (OR=0.39; 95%CI=0.13-1.16; p=0.09), phase of the study (OR=1.44, 95%CI=0.59-3.53; p=0.43), country of origin of the first author (OR=0.72; 95%= 0.34-1.54; p=0.40), protocol availability or with having OS versus a surrogate as primary endpoint (OR=1.50, 95% CI=0.93-2.44; p=0.10). There was no association between bias in reporting outcomes of efficacy and bias in reporting toxicity (OR=1.35, 95% CI=0.61-2.97; p=0.46).

3.4.3. Funding and author’s financial ties.

The pharmaceutical industry funded 165 (82.5%) of the 200 included studies; of these 113 (56%) studies and 53 studies (26.5%) were funded totally or partially, respectively. Only 31 (15.5%) studies were funded by governmental research agencies or cooperative groups. Three studies did not report source of funding. Studies funded by the pharmaceutical industry were not associated with greater incidence of bias in reporting efficacy (OR=0.68; 95%CI=0.30-1.56; p=0.36) or toxicity (OR=0.52, 95%CI=0.22-1.25; p=0.14).

First authors and corresponding authors (when different) declared having financial ties with the sponsor in 141 studies (71.5%); of these, two authors (1%) were employees of the sponsor and the remaining 139 (69.5%) authors declared that they had received funding for research, consulting or expert testimony, honoraria, or other). Fifty-seven (28.5%) authors declared no conflicts of interest with the
sponsor. Financial ties were not associated with bias in reporting efficacy (OR=1.12, 95%CI= 0.56-2.22; p=0.76) but were associated with bias in reporting toxicity (OR=0.34, 95%CI=0.16-0.70; p=0.003).

3.5. Discussion

Although in 2004 the ICMJE published guidelines for mandatory registration of clinical trials[17]; consistency among a clinical trial registry and the final manuscript in the reporting of primary and secondary endpoints of surgical RCTs was reported recently to be low: only 55% of the published papers showed no discrepancy while in 45% of manuscripts there was omission, introduction, change in definition, downgrading or upgrading of outcomes.[19] Another surgery paper showed similar results, with 49% discrepancies in the reporting of primary outcomes.[20] We reported previously 4% inconsistency among the primary outcomes reported in breast cancer trials manuscripts, in all cases the primary endpoint in the clinical trial registry was OS and a surrogate was used at the time of publication. Here we present evidence that for 94 RCTs evaluating medical interventions for solid tumors reported since 2010, with available protocols, 99% of the studies did not change the original primary outcome. Only in 2 articles was the reporting vague. These results confirm a remarkable advance in consistency although we cannot comment on the 106 articles where the protocols were not available.
Access to protocols is a sign of trust from the investigators and sponsor to readers and allows an open assessment not only of the endpoints, but also of other important factors such as inclusion criteria, and assessment and management of toxicity; this is important information that is highly relevant when the results of clinical trials are applied in daily practice. Some journals are requiring submission of protocols as a requirement for peer review, but a substantial majority of protocols are not accessible to readers, even under request.[60] We were unable to obtain the protocols for half of the trials in our cohort despite contacting the author and sponsor.

We reported previously that bias in reporting outcomes is almost 60% in articles reporting studies with a negative primary endpoint in breast cancer RCTs[11]. Here we confirm biased reporting of efficacy in 47% of RCTs evaluating treatments for a variety of tumor sites, even when limiting our study to reports in journals with high impact factors, journals that are associated with changes in standards of clinical practice.

Under-reporting of toxicity to highlight a positive primary endpoint is a type of bias that has been reported previously by us and by other groups[11, 29, 61]. Reporting of toxicity or tolerance in a more positive way for the experimental arm has been associated with studies that have financial ties with for-profit sponsorship. However our analysis did not find an association of biased reporting of toxicity with either funding source or first author financial ties. Bias was more prevalent in reporting efficacy than toxicity.
Our study has limitations. First, protocols were available for approximately half of the published reports of RCTs; then consistency of the primary endpoint was assessed only among the clinical trial registry and the final publication. Second, we utilized subjective measures to determine some of our outcome measures such as the presence of spin; assessment of spin is linked to inherent bias of the reviewer, however high consistency among reviewers limits this risk. Third, classification of under-reporting of toxicity can be arbitrary for some manuscripts, but as it was mentioned, consistency among reviewers was strong, reducing this inherent bias. Fourth, we focused only on papers in journals with high JIF, which may represent higher quality trials, we consider this important because these publications lead the standard of care and are assumed that peer-review is very thoughtful. Fifth, for financial ties we explored the effect only of first and corresponding authors, although we recognize that other authors can influence decisions as to what is reported.

In conclusion, transparency in having primary endpoints available in clinical trial registries is high, although complete protocols are not often accessible; this should be a requirement to allow appropriate interpretation of studies. Spin is used frequently to distract the reader when the primary endpoint is negative, and editors should allow only simple concluding statements that apply only to the primary endpoint. Reporting of toxicity should also be improved. Funding and financial ties did not appear to have a significant influence to increase bias in reporting, and intrinsic bias of authors towards reporting positive studies may be as important as bias originating from financial motives.
Table 3.1. Characteristics of included studies.

<table>
<thead>
<tr>
<th>Characteristics of Included Studies</th>
<th>N = 200</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase III</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protocol available</td>
<td>152</td>
<td>76</td>
</tr>
<tr>
<td>No protocol</td>
<td>48</td>
<td>24</td>
</tr>
<tr>
<td>Phase II</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protocol available</td>
<td>72</td>
<td>36</td>
</tr>
<tr>
<td>No protocol</td>
<td>80</td>
<td>40</td>
</tr>
<tr>
<td>Protocol access</td>
<td></td>
<td></td>
</tr>
<tr>
<td>On-line</td>
<td>64</td>
<td>32</td>
</tr>
<tr>
<td>Not on-line but provided by:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Investigator</td>
<td>28</td>
<td>14</td>
</tr>
<tr>
<td>Sponsor</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Not provided</td>
<td>106</td>
<td>53</td>
</tr>
<tr>
<td>Year of Publication</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010</td>
<td>40</td>
<td>20</td>
</tr>
<tr>
<td>2011</td>
<td>61</td>
<td>30</td>
</tr>
<tr>
<td>2012</td>
<td>99</td>
<td>50</td>
</tr>
<tr>
<td>Journal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>New England Journal of Medicine</td>
<td>25</td>
<td>12</td>
</tr>
<tr>
<td>Lancet</td>
<td>12</td>
<td>6</td>
</tr>
<tr>
<td>JAMA</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>Lancet Oncology</td>
<td>35</td>
<td>18</td>
</tr>
<tr>
<td>Journal of Clinical Oncology</td>
<td>118</td>
<td>59</td>
</tr>
<tr>
<td>Journal of the National Cancer Institute</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Country</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-US</td>
<td>123</td>
<td>61</td>
</tr>
<tr>
<td>US</td>
<td>77</td>
<td>39</td>
</tr>
<tr>
<td>Disease Site</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast</td>
<td>41</td>
<td>20</td>
</tr>
<tr>
<td>Lung</td>
<td>46</td>
<td>23</td>
</tr>
<tr>
<td>Gastro-intestinal</td>
<td>49</td>
<td>25</td>
</tr>
<tr>
<td>Genito-urinary</td>
<td>25</td>
<td>13</td>
</tr>
<tr>
<td>Gynecology</td>
<td>13</td>
<td>6</td>
</tr>
<tr>
<td>Melanoma</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>Central Nervous System</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Sarcoma</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Head &amp; Neck</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Funding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not available</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Pharma Only</td>
<td>112</td>
<td>56</td>
</tr>
<tr>
<td>Pharma + Non-for-profit group</td>
<td>53</td>
<td>26</td>
</tr>
<tr>
<td>Non-for-profit group</td>
<td>31</td>
<td>16</td>
</tr>
<tr>
<td>---------------------</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td><strong>Primary Endpoint</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall survival</td>
<td>72</td>
<td>36</td>
</tr>
<tr>
<td>Disease Free Survival/Recurrence Free Survival/Progression Free Survival/Time to Tumor Progression</td>
<td>101</td>
<td>50</td>
</tr>
<tr>
<td>Response rate</td>
<td>23</td>
<td>12</td>
</tr>
<tr>
<td>Other</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td><strong>Primary Endpoint</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>93</td>
<td>46</td>
</tr>
<tr>
<td>Negative</td>
<td>107</td>
<td>54</td>
</tr>
<tr>
<td><strong>Reported in Clinical Trial Registry</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>Yes</td>
<td>193</td>
<td>96</td>
</tr>
<tr>
<td><strong>Population</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adults</td>
<td>197</td>
<td>98</td>
</tr>
<tr>
<td>Pediatric</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>
Table 3.2. Bias in reporting of efficacy and toxicity.

<table>
<thead>
<tr>
<th></th>
<th>Total (%)</th>
<th>Spin in Conclusion of the Abstract (%)</th>
<th>Spin in Conclusion of manuscript (%)</th>
<th>Under-reporting of toxicity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase III</td>
<td>152 (100)</td>
<td>36 (24)</td>
<td>31 (20)</td>
<td>30 (20)</td>
</tr>
<tr>
<td>Protocol available</td>
<td>72</td>
<td>14</td>
<td>14</td>
<td>16</td>
</tr>
<tr>
<td>No protocol</td>
<td>80</td>
<td>22</td>
<td>17</td>
<td>14</td>
</tr>
<tr>
<td>Phase II</td>
<td>48 (100)</td>
<td>14 (29)</td>
<td>11 (23)</td>
<td>7 (15)</td>
</tr>
<tr>
<td>Protocol available</td>
<td>22</td>
<td>4</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>No protocol</td>
<td>26</td>
<td>10</td>
<td>9</td>
<td>3</td>
</tr>
</tbody>
</table>
Table 3.3. Distribution of under-reporting of toxicity.

<table>
<thead>
<tr>
<th>Toxicity hierarchy scale</th>
<th>Number of trials N=200</th>
<th>%</th>
<th>Positive PE (%)</th>
<th>Negative PE (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>41</td>
<td>20.5</td>
<td>19 (9.5)</td>
<td>22 (11)</td>
</tr>
<tr>
<td>2</td>
<td>7</td>
<td>3.5</td>
<td>2 (1)</td>
<td>5 (2.5)</td>
</tr>
<tr>
<td>3</td>
<td>93</td>
<td>46.5</td>
<td>60 (30)</td>
<td>33 (16.5)</td>
</tr>
<tr>
<td>4</td>
<td>22</td>
<td>11</td>
<td>5 (2.5)</td>
<td>17 (8.5)</td>
</tr>
<tr>
<td>5</td>
<td>7</td>
<td>3.5</td>
<td>2 (1)</td>
<td>5 (2.5)</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>3.5</td>
<td>0 (0)</td>
<td>7 (3.5)</td>
</tr>
<tr>
<td>7</td>
<td>23</td>
<td>11.5</td>
<td>5 (2.5)</td>
<td>18 (9)</td>
</tr>
</tbody>
</table>
Figure 3.1. Study selection.

Excluded:
8 Biomarker analysis
5 Not cancer therapy.
42 Not primary publication.
95 Not RCT.
8 Phase I/II
4 Published before July 2010.
6 Radiation trials.
3 Surgical trials.
22 Sequence trials.
8 Non inferiority.
2 Multi-arm.
References


Supplementary Data

Figure 1. Decision Tree for Assessment of reporting of the Primary Endpoint in the Concluding Statement of the abstract.
Figure 2: Hierarchy scale for reporting of adverse events

Not R = Not Reported in Results Table, Not D = Not Reported in Discussion, Not A = Not Reported in Abstract, Not C = Not Reported in Concluding Statement.
R= Reported in Results Table, D= Reported in Discussion, A= Reported in Abstract, C= Reported in Concluding Statement. Shadowed area represents “under-reporting of toxicity”.
Examples of abstracts with bias in the reporting of efficacy.

Abstract #1.


**Bevacizumab in combination with chemotherapy as first-line therapy in advanced gastric cancer: a randomized, double-blind, placebo-controlled phase III study.**


**PURPOSE:**
The Avastin in Gastric Cancer (AVAGAST) trial was a multinational, randomized, placebo-controlled trial designed to evaluate the efficacy of adding bevacizumab to capecitabine-cisplatin in the first-line treatment of advanced gastric cancer.

**PATIENTS AND METHODS:**
Patients received bevacizumab 7.5 mg/kg or placebo followed by cisplatin 80 mg/m² on day 1 plus capecitabine 1,000 mg/m² twice daily for 14 days every 3 weeks. Fluorouracil was permitted in patients unable to take oral medications. Cisplatin was given for six cycles; capecitabine and bevacizumab were administered until disease progression or unacceptable toxicity. The primary end point was overall survival (OS). Log-rank test was used to test the OS difference.

**RESULTS:**
In all, 774 patients were enrolled; 387 were assigned to each treatment group (intention-to-treat population), and 517 deaths were observed. Median OS was 12.1 months with bevacizumab plus fluoropyrimidine-cisplatin and 10.1 months with placebo plus fluoropyrimidine-cisplatin (hazard ratio 0.87; 95% CI, 0.73 to 1.03; P = .1002). Both median progression-free survival (6.7 v 5.3 months; hazard ratio, 0.80; 95% CI, 0.68 to 0.93; P = .0037) and overall response rate (46.0% v 37.4%; P = .0315) were significantly improved with bevacizumab versus placebo. Preplanned subgroup analyses revealed regional differences in efficacy outcomes. The most common grade 3 to 5 adverse events were neutropenia (35%, bevacizumab plus fluoropyrimidine-cisplatin; 37%, placebo plus fluoropyrimidine-cisplatin), anemia (10% v 14%), and decreased appetite (8% v 11%). No new bevacizumab-related safety signals were identified.

**CONCLUSION:**
Although AVAGAST did not reach its primary objective, adding bevacizumab to chemotherapy was associated with significant increases in progression-free survival and overall response rate in the first-line treatment of advanced gastric cancer.
Abstract #2


**Carboplatin plus paclitaxel versus carboplatin plus pegylated liposomal doxorubicin as first-line treatment for patients with ovarian cancer: the MITO-2 randomized phase III trial.**


**PURPOSE:**
Carboplatin/paclitaxel is the standard first-line chemotherapy for patients with advanced ovarian cancer. Multicentre Italian Trials in Ovarian Cancer-2 (MITO-2), an academic multicenter phase III trial, tested whether carboplatin/pegylated liposomal doxorubicin (PLD) was more effective than standard chemotherapy.

**PATIENTS AND METHODS:**
Chemotherapy-naive patients with stage IC to IV ovarian cancer (age ≤ 75 years; Eastern Cooperative Oncology Group performance status ≤ 2) were randomly assigned to carboplatin area under the curve (AUC) 5 plus paclitaxel 175 mg/m² or to carboplatin AUC 5 plus PLD 30 mg/m², every 3 weeks for six cycles. Primary end point was progression-free survival (PFS). With 632 events in 820 enrolled patients, the study would have 80% power to detect a 0.80 hazard ratio (HR) of PFS.

**RESULTS:**
Eight hundred twenty patients were randomly assigned. Disease stages III and IV were prevalent. Occurrence of PFS events substantially slowed before obtaining the planned number. Therefore, in concert with the Independent Data Monitoring Committee, final analysis was performed with 556 events, after a median follow-up of 40 months. Median PFS times were 19.0 and 16.8 months with carboplatin/PLD and carboplatin/paclitaxel, respectively (HR, 0.95; 95% CI, 0.81 to 1.13; P = .58). Median overall survival times were 61.6 and 53.2 months with carboplatin/PLD and carboplatin/paclitaxel, respectively (HR, 0.89; 95% CI, 0.72 to 1.12; P = .32). Carboplatin/PLD produced a similar response rate but different toxicity (less neurotoxicity and alopecia but more hematologic adverse effects). There was no relevant difference in global quality of life after three and six cycles.

**CONCLUSION:**
Carboplatin/PLD was not superior to carboplatin/paclitaxel, which remains the standard first-line chemotherapy for advanced ovarian cancer. However, given the observed CIs and the different toxicity, carboplatin/PLD could be considered an alternative to standard therapy.
Chapter 4: (Paper 3): Honorary and Ghost-Authors of Reports of Randomized Clinical Trials in Oncology

This article will be submitted as:


4.1 Abstract

Background: The International Committee of Medical Journal Editors (ICMJE) developed guidelines for responsible and accountable authorship. Few data inform the frequency and nature of ghost and honorary authorship in oncology trials.

Methods: We conducted a systematic review of reports of randomized clinical trials (RCTs) evaluating systemic cancer therapy published July 2010 to December 2012 in six high-impact journals. Failure to include investigators and the statistician listed in protocols as authors in the paper, and/or use of non-author medical writers, were criteria used to define ghost authorship. The list of contributions for authors of published articles was recorded, and we defined an article as having an honorary author if any author did not meet all three criteria described by ICMJE.

Results: Two hundred publications were identified, of which 89 articles indicated use of a medical writer (45%). For 61 articles, protocols with listed investigators were available, and 40 (66%) met our definition of ghost authorship. Contributions of each author were provided in 193 articles and 63 (35%) met our definition for
honorary authorship. Funding source was not a predictor for either honorary or ghost authorship. Assistance of a medical writer was acknowledged only in sponsored trials. Journals with high Impact Factor were associated more commonly with honorary authorship.

*Conclusion:* Ghost and honorary authorship are prevalent in articles describing trials for systemic therapy of cancer. Guidelines should be established to improve transparency and accountability.
4.2. Introduction

Authorship establishes accountability, responsibility and credit for scientific information reported in biomedical publications.[5, 62] If inappropriate authorship is present, it can undermine the integrity of the research and can increase the risk of manipulation of the analysis and conclusions. This can, in turn, influence the interpretation by readers and lead to adoption of poor treatment strategies.

Concerns about integrity of authorship have been recognized for decades[62-65], and in 1985 the International Committee of Medical Journal Editors (ICMJE) developed the following guidelines for responsible and accountable authorship[66]: (1) Substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpretation of data for the work; AND (2) Drafting the work or revising it critically for important intellectual content; AND (3) Final approval of the version to be published. The 2013 update added: (4) Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved[66].

Inappropriate authorship can be described under two major headings: ghost-authorship and honorary-authorship. A ghost author is defined as someone who has contributed substantially to a paper but is not named as an author in the final publication[67]. An honorary author is a listed author who does not meet authorship criteria specified by the ICMJE.[68]
Medical writers are used commonly to assist in the preparation of manuscripts, especially those reporting randomized clinical trials (RCTs) sponsored by pharmaceutical companies.[69] Medical writers are usually either employed by or contracted to the sponsor, are rarely included as authors, and may serve as ghost authors whose participation is frequently not acknowledged. Editorial guidelines of many journals encourage acknowledgement of assistance in preparation of manuscripts, but this is almost certainly under-reported. In a previous study, participation of medical writers was reported in only 6% of the studies analyzed, and in only 10% of projects funded by the pharmaceutical industry.[70] There is no information as to the extent to which use of medical writers can lead to bias in reporting results and side effects of new therapies, but as direct or indirect employees of the sponsor, their contribution is unlikely to be completely objective and independent.

Statisticians may also be excluded as authors of RCTs, especially if they are employed by or are contracted to the sponsor. In a previous review, statisticians were listed as authors in only 7% of reports of clinical trials[71], although another study involving not only RCTs but several types of research found that statisticians were authors in about 65% of reports.[72] Statisticians play an essential role in the design analysis and interpretation of RCTs, and should be included as authors or at least acknowledged in papers reporting them.

All major medical journals require disclosure of contributions by the final authors of a paper, but that does not disclose the presence and contributions of unlisted ghost authors, while honorary authors might make inaccurate disclosures. Here we
review papers reporting results of RCTs evaluating systemic therapy for solid tumors in six high-impact journals, and the protocols on which these studies were based, in an attempt to quantify the extent of ghost and honorary authorship in reporting cancer trials. We hypothesized that there would be substantial frequency of ghost and honorary authorship, and these would be associated more frequently with pharmaceutical-sponsored studies.

4.3. Methods

4.3.1. Literature search and study selection

A comprehensive search of all articles published from July 2010 to December 2012 in the New England Journal of Medicine (NEJM), The Lancet, the Journal of the American Medical Association (JAMA), Lancet Oncology, the Journal of Clinical Oncology (JCO) and the Journal of the National Cancer Institute (JNCI) was performed manually to extract papers reporting results of phase II and phase III RCTs for solid tumor malignancies. The rationale for selection of the highest impact journals was the assumption that high-quality reports that impact clinical practice would be published in these journals. Supplementary sections of articles were also accessed to obtain the trial protocol when available. When not available the editorial offices were contacted by email to request copies of the protocol; if this was not successful the corresponding author or the sponsor was contacted.

Inclusion criteria were all phase II and phase III RCTs reporting results of experimental medical interventions. We excluded studies of biomarker analysis,
those not reporting trials of cancer therapy, non-primary publications, reports of non-randomized trials and phase I/II trials, trials or surgery or radiation, trials comparing drug sequences (rather than different drugs), non-inferiority trials, and multi-arm trials. These exclusion criteria were used to ensure reasonable homogeneity and high-quality evidence in the sample.

4.3.2. Data extraction and analysis

The following data were extracted independently from the primary manuscript describing each RCT by two authors (FV-B and MN): PubMed identifier number, year of publication, journal of publication, journal impact factor (JIF) at the present time, country of origin of the first author, phase of the trial, where the investigation was carried out, number of patients enrolled, disease site, origin of funding (pharmaceutical only, pharmaceutical partial, non-pharmaceutical), and population included in the study (i.e. adults, pediatric).

Authorship was extracted from published manuscripts, and the contribution of each author was evaluated to determine whether they satisfied the first three ICMJE criteria for authorship as described above. Statisticians were identified if they were listed in the protocol or the manuscript under that denomination.

Protocols were assessed for consistency of the listed investigators with listed authors of the published paper and vice versa. The role of the statistician was evaluated for designation in the protocol, authorship of the manuscript or both; the relationship of the statistician with the sponsor was also recorded. The sponsor’s contribution to statistical analysis and any declaration of use of a medical writer to
draft or assist with the manuscript were recorded (reporting is mandatory in all included journals). We understand statisticians can change over the time and perhaps some of them will be in charge of developing the protocol but will not be part at the time of analysis/final publication; to reduce the risk of bias for this assessment we have adjusted the analysis based on “delayed time of publication” that takes into account the difference of time between the date of final publication minus the date of the protocol available for analysis.

4.3.3. Objectives

The primary goal of the present study was to assess the frequency of ghost and honorary authorship.

Ghost authorship was determined to be present in any scenarios where (1) investigators listed in the protocol were not included as authors or were not acknowledged in the article describing the trial. We also determined (2) whether the individual who performed statistical analyses was listed as an author or was acknowledged in the publication and (3) whether assistance of a medical writer was acknowledged in preparing the manuscript for publication. Presence of at least one of these three criteria was used to define ghost authorship for those publications with a list of authors of protocol available. An exploratory analysis including medical writing assistance only was performed in the complete dataset.
Honorary authorship was assessed based on ICMJE criteria. We defined an article as having a honorary author if any of its authors did not meet all of the three authorship criteria described previously. For criterion #1 we separated contributions based on (a) conception and/or design of the study project and (b) patient recruitment and/or analysis of data. Each criterion for authorship is reported separately. We also included an exploratory analysis of consistency between each author’s declaration of contribution to the conception or design of the study and their listing as an investigator in the protocol, this analysis was carried out in the complete dataset.

We evaluated possible predictors for ghost and honorary authorship including source of funding, journal where the manuscript was published, country of origin of the first author and corresponding author (if it was different from the first author), phase of the study, and the time elapsed between the final version of the protocol available for analysis and the manuscript publication.

4.3.4. Missing Data

Missing data were considered to be missing at random and no further analysis or correction was performed.

4.3.5. Statistical Analysis
Data are presented descriptively as means with their standard deviations. The test of one proportion is utilized to assess the precision of the estimate for the proportion of studies with ghost authorship. For the purpose of this test, the null hypothesis was that the proportion of articles with ghost authorship should be 0%. *Post-hoc* power using an alpha of 0.05 and based on the above assumptions is also reported. Predictors for ghost and honorary authorship were assessed using univariable logistic regression analysis and reported as odds ratio (OR) and their respective 95% confidence intervals (CI). Forward elimination multivariable logistic regression at the p<0.05 threshold was planned; however only 1 variable met this criterion and therefore multivariable analysis was not conducted. All statistical analyses were conducted using SPSS statistical software version 17 (IBM Corp, Armonk, New York). All significance tests were two-sided using an alpha level of 0.05. No correction was applied for multiple statistical testing.

### 4.4. Results

A total of 403 articles were identified initially and 200 RCT’s (48 phase II studies and 152 phase III studies) were eligible for analysis (Figure 1). The characteristics of the trials are described in Table 1. Ninety-four protocols were available for analysis: 22 protocols for phase II and 72 protocols for phase III RCTs. Sixty-four protocols were available on-line, twenty-eight protocols were obtained from the principal investigator/corresponding author and two protocols were obtained from the sponsor. Sixty-one protocols included a list of investigators, and were included
in the analysis of ghost authorship according to criteria previously mentioned. A separate analysis of ghost authorship predictors was carried out in the 200 studies according to the basis of use of writing assistance.

A total of 193 studies were registered and seven studies were not registered (or information was not available) in a clinical trial registry; the protocol was not available for any of them. For 7 articles (3.5%) the corresponding author was different from the first author. For 77 articles (39%) the first author of the paper was based in the United States. Disease sites were distributed roughly according to the incidence of each disease in developed countries: 25% of studies were in gastrointestinal (GI) malignancies, 23% for lung cancer and 20% for breast cancer. Four studies did not disclose funding sources; 165 studies (82%) were partially or fully funded by for-profit industry. Overall survival was the primary endpoint for 36% of studies, while a surrogate endpoint was used in the remaining. Seventeen studies (8.5%) reported >1 primary endpoint; ten of these studies were not statistically powered for formal analysis of the second primary endpoint according to the statistical description of the study.

4.4.1. Assessment of Ghost Authorship

For the 61 protocols that listed investigators, all investigators listed in the protocol were also authors of or acknowledged in the manuscript in 26 studies (43%); of these, in 12 studies (20%) authors of protocol were also authors of the manuscript.
In the 49 (80%) protocols for which some of the listed investigators were not included as authors of the manuscript, they were acknowledged in 14 of these studies (23%). A statistician was identified in 44 studies (72%), was listed as an investigator in 40 protocols (66%), and was an author of the published manuscript in 36 publications (59%); in 4 publications (9%) the statistician named in the protocol was not an author or was not acknowledged in the manuscript. Based on our definition of ghost authorship 40 of 61 evaluable studies (66%, 95% CI 54%-78%) met criteria based on non-inclusion as authors, and lacking acknowledgement of investigators or statisticians listed in the protocol and writing assistance. This analysis provided 100% power to detect a difference between this finding and the null hypothesis. Among the 200 included studies, use of medical writing assistance was reported in 89 studies (44.5%), see Table 2.

4.4.2. Assessment of Honorary Authorship

Declarations of their role in the study were reviewed for all authors of the manuscript of the 200 papers, in 5 cases, they didn’t report any of the three criteria; therefore 195 studies, reporting at least one criteria, were included for analysis. All authors claimed that they met criterion #1 for authorship in 158 papers of 180 evaluable articles (158/180; 88%); in 116 papers authors declared participation in conception or design of the study, and in 42 studies, authors declared that they met these criteria by recruiting patients. Authorship criterion #2
refers to drafting the work or revising it critically for important intellectual content: all authors claimed to meet these criteria in 141 of 195 evaluable articles (72%). Authorship criterion #3 requires final approval of the version to be published: all authors claimed to meet this criterion in 184 of 188 evaluable articles (98%). Only 45 first-authors of 61 papers with available protocols that listed the investigators (74%) declared a role in conception or design of the study and also were listed as investigators in the protocol. Sixty-three manuscripts (35%) met our definition of honorary authorship. See Table 3.

4.4.3. Funding

The pharmaceutical industry provided funding for 165 studies (82.5%): full funding was reported in 112 studies (56%) and partial funding in 53 studies (26.5%); only 31 manuscripts (15.5%) reported funding from non-profit agencies or groups. All 89 studies (44.5%) that acknowledged assistance of a medical writer were funded by industry. In 133 studies the relationship of the statistician with the sponsor was declared, and in 57 studies (43%) the statistical analysis was funded by the sponsor. The probability that the statistician was an author of the manuscript was not influenced by funding (OR 1.3, 95% CI= 3.4-4.8, p=0.72).

4.4.4. Predictors

4.4.4.1. Ghost Authorship
An exploratory model of predictors of ghost authorship did not find association of ghost authorship with country of origin of the first author, number of patients enrolled in the study, phase of the study, disease site were the experiment was carried out, JIF of the publication, funding, and honorary authorship. Publication delay was assessed, and was not associated with ghost authorship, see Table 4.

Predictors for medical writing assistance only did not find an association of medical writers with country of origin of the first author, number of patients enrolled in the study, phase of the study, disease site and JIF of the publication. Medical writing assistance was used only in studies sponsored by pharmaceutical companies, and none of the studies funded by cooperative groups or with government sponsorship reported use of writing assistance. See table 4.

4.4.4.2. Honorary authorship

Predictors described above for ghost authorship were also assessed for honorary authorship. In a univariable analysis honorary authorship was associated with the JIF, phase and number of patients in the study. In a multivariable analysis, only JIF remained significant, however the odds were minimally increased (OR=1.03, 95% CI=1.004-1.065; p=0.03); see Table 4.

4.5. Discussion
Three previous studies have analyzed the prevalence of ghost and/or honorary authorship in the medical literature.[5, 68, 71] Here, we report that the frequency of both honorary and ghost authorship is high in reports of RCTs evaluating systemic therapy for cancer.

Ghost authorship was analyzed from different perspectives: investigators listed in the protocol not considered for authorship or acknowledged in the manuscript, statisticians not listed as authors of the final publication or the use of medical writers. Some investigators have proposed that statistical analysis is not enough to deserve authorship,[73] but statisticians meet criteria 1 and 2 of authorship guidelines by analyzing data and reviewing accuracy of the data reported. Failure to list the statistician as an author is a source of ghost authorship more frequently associated with pharmaceutical-industry sponsored trials[71], here we confirmed that assertion and report and new finding were academic sponsorship is associated with the statistician always included in the list of authors of the final publication; this is very important in order to meet the new 4th criterion of accountability. Finally, we found that many investigators listed in protocols were not included as authors of the manuscript, despite protocol development being the most critical part of any research project.

Ghost authorship has been reported previously to be a source of important bias, making of knowledge dissemination a tool for marketing of specific drugs, therefore affecting sometimes, as in the case of robecoxib[21] health of consumers. Incidence of ghost authors associated with industry has been an area of concern for years, as we discussed here, of all trials analyzed, the presence of medical
writers is a unique characteristic of trials supported by industry. Although, the non-significant correlation among our definitions of ghost and honorary ghost authorship and funding, editors have raised concern about how senior academics can be contributing for authorship as a marketing strategy.

While some have argued that the use of medical writers does not constitute ghost authorship because they may be involved only in providing grammatical assistance, and that don’t meet any criteria for authorship, their potential to increase bias is substantial.[67] Others have proposed to apply the GATE principles, that means assessment for guarantee (are the authors guarantors of the article?), advice (was the professional writer “advised” by the authors before starting the assignment?), transparency (were the writers and researchers identified appropriately on the authorship line or in the Acknowledgments section?) and expertise (does the professional writer have sufficient knowledge in a specific field to critically comment, interpret and analyse the literature?).[69] Establishing the former criteria, professional writers can have their role and limitations clearly defined.

Honorary authorship can occur if a sponsor deems it advantageous to include among the authors an individual, usually a well-known “thought leader”, even if that individual has had minimal involvement in the study and does not meet ICMJE criteria for authorship. Honorary authorship is also prevalent, and is associated with JIF. We found that authors, who did not meet required authorship criteria were included more frequently in the three general medical journals with the highest impact factor, compared with the journals focused only on oncology research. In
2% of the studies all authors did not meet the last criterion of authorship, approval of the manuscript. When analyzing the new added criterion, which refers to accountability of all aspects of the manuscript, we can infer that at least 2% of the authors did not meet this criterion since they did not approve the final version of the manuscript. It is also apparent that many of the lead authors of the publication were not involved in the conception and planning of a study; even those who declared themselves to be part of developing a protocol were often not listed as authors of it.

When comparing these results to those published previously, honorary authorship was more prevalent, being present in 35% of the trials analyzed in our data set, compared with 9% and 19% reported previously[5, 68]. Previous estimates of ghost authorship have been more variable. In our study, investigators of the protocol were not all listed as authors of the manuscript in 80% of the studies, whereas previously this applied to only 13% of the studies analyzed.[71] In previous reports, statisticians were not named as authors of published studies in a variable proportion ranging from 9-65%,[71, 72] however in our study, this was only present in 9% of the studies. The use of a medical writer is prevalent and was acknowledged in 44.5% of the present articles; it was reported in only 6% in a previous study[70], and this recognition may have increased because all journals included for analysis now require that information at the time of submission. The acknowledgement of medical writing assistance reflects more transparency. Previous experience suggests that the use of medical writers and ghost authors
was associated with substantial bias in reporting efficacy of drugs and could provide misleading results of trials.[21]

Funding was not associated with honorary or ghost authorship, but it was associated with greater use of medical writers. This has not been reported previously.

This study has limitations. The evaluation of protocols is one of the most objective measures for ghost authorship but for 34% of the studies the protocol did not provide the list of investigators, limiting the analysis of ghost authorship. The declaration of authorship varies among publications, but we were able to extract these data and information was entered in a uniform format according to the ICMJE guidelines.

In conclusion, while honorary and ghost authorship remains highly prevalent, reporting of medical writing assistance has improved. More judicious evaluation of authorship from information available in original protocols and self-declaration should be required by journals reporting clinical trials.
Table 4.1. Characteristics of included studies

<table>
<thead>
<tr>
<th>Characteristics of Included Studies</th>
<th>N = 200</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase III</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protocol available.</td>
<td>152</td>
<td>76</td>
</tr>
<tr>
<td>Not available</td>
<td>72</td>
<td>36</td>
</tr>
<tr>
<td>Phase II</td>
<td>80</td>
<td>40</td>
</tr>
<tr>
<td>Protocol available</td>
<td>48</td>
<td>24</td>
</tr>
<tr>
<td>Not available</td>
<td>22</td>
<td>11</td>
</tr>
<tr>
<td>Protocol access</td>
<td>26</td>
<td>13</td>
</tr>
<tr>
<td>On-line</td>
<td>64</td>
<td>32</td>
</tr>
<tr>
<td>Not on-line but provided by:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Investigator</td>
<td>28</td>
<td>14</td>
</tr>
<tr>
<td>Sponsor</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Not provided</td>
<td>106</td>
<td>53</td>
</tr>
<tr>
<td>Year of publication</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010</td>
<td>40</td>
<td>20</td>
</tr>
<tr>
<td>2011</td>
<td>61</td>
<td>30</td>
</tr>
<tr>
<td>2012</td>
<td>99</td>
<td>50</td>
</tr>
<tr>
<td>Journal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>New England Journal of Medicine</td>
<td>25</td>
<td>12</td>
</tr>
<tr>
<td>Lancet</td>
<td>12</td>
<td>6</td>
</tr>
<tr>
<td>JAMA</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>Lancet Oncology</td>
<td>35</td>
<td>18</td>
</tr>
<tr>
<td>Journal of Clinical Oncology</td>
<td>118</td>
<td>59</td>
</tr>
<tr>
<td>Journal of the National Cancer Institute</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Country of first author</td>
<td></td>
<td></td>
</tr>
<tr>
<td>US</td>
<td>77</td>
<td>39</td>
</tr>
<tr>
<td>Non-US</td>
<td>123</td>
<td>61</td>
</tr>
<tr>
<td>Disease site</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast</td>
<td>41</td>
<td>20</td>
</tr>
<tr>
<td>Lung</td>
<td>46</td>
<td>23</td>
</tr>
<tr>
<td>Gastro-intestinal</td>
<td>49</td>
<td>25</td>
</tr>
<tr>
<td>Genito-urinary</td>
<td>25</td>
<td>13</td>
</tr>
<tr>
<td>Gynecology</td>
<td>13</td>
<td>6</td>
</tr>
<tr>
<td>Other</td>
<td>26</td>
<td>12</td>
</tr>
<tr>
<td>Funding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not available</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Pharmaceutical industry only</td>
<td>112</td>
<td>56</td>
</tr>
<tr>
<td>Pharmaceutical + Non-profit group</td>
<td>53</td>
<td>26</td>
</tr>
<tr>
<td>Non-profit group</td>
<td>31</td>
<td>16</td>
</tr>
<tr>
<td>Primary endpoint</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall survival</td>
<td>72</td>
<td>36</td>
</tr>
<tr>
<td>Disease free survival/recurrence free survival/progression free survival/time to tumor</td>
<td>101</td>
<td>50</td>
</tr>
<tr>
<td>Category</td>
<td>Count (Yes)</td>
<td>Count (No)</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>-------------</td>
<td>------------</td>
</tr>
<tr>
<td>Progression</td>
<td>23</td>
<td>12</td>
</tr>
<tr>
<td>Response rate</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primer Endpoint</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>93</td>
<td>46</td>
</tr>
<tr>
<td>Negative</td>
<td>107</td>
<td>54</td>
</tr>
<tr>
<td>Reported in Clinical Trial Registry</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>Yes</td>
<td>193</td>
<td>96</td>
</tr>
<tr>
<td>Population</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adults</td>
<td>197</td>
<td>98</td>
</tr>
<tr>
<td>Pediatric</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

*From July 2011.*
Table 4.2. Incidence of Ghost Authorship

<table>
<thead>
<tr>
<th>Eligible for Full Assessment of Ghost Authorship</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Investigators of protocols also authors of manuscript</td>
<td>12</td>
<td>20</td>
</tr>
<tr>
<td>All Investigators of protocols either authors or listed in acknowledgment</td>
<td>14</td>
<td>23</td>
</tr>
<tr>
<td>Some Investigators of protocols not listed in manuscript</td>
<td>35</td>
<td>57</td>
</tr>
</tbody>
</table>

| Medical writer assistance | 21  | 34  |
| Statistician author of manuscript | 36  | 51  |
| Met ghost authorship definition | 40  | 66  |

<table>
<thead>
<tr>
<th>Eligible for Assessment of Medical Writing Assistance Only</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reported medical writer assistance</td>
<td>89</td>
<td>45</td>
</tr>
</tbody>
</table>
Table 4.3. Incidence of Honorary Authorship

<table>
<thead>
<tr>
<th>Honorary Authorship</th>
<th>Total*</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criterion #1</td>
<td>180</td>
<td>158</td>
<td>88</td>
</tr>
<tr>
<td>Study Design</td>
<td>116</td>
<td>42</td>
<td>65</td>
</tr>
<tr>
<td>Patient accrual</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Criterion #2</td>
<td>195</td>
<td>141</td>
<td>72</td>
</tr>
<tr>
<td>Criterion #3</td>
<td>188</td>
<td>184</td>
<td>98</td>
</tr>
<tr>
<td>Met honorary authorship criteria</td>
<td>193</td>
<td>63</td>
<td>33</td>
</tr>
<tr>
<td>First authors declared conception/design and listed as investigators of protocol</td>
<td>61</td>
<td>45</td>
<td>74</td>
</tr>
</tbody>
</table>
### Table 4.4 Univariable analysis.

<table>
<thead>
<tr>
<th></th>
<th>Ghost Authorship (61 studies)</th>
<th>Honorary Authorship (193 studies)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR</td>
<td>95% CI</td>
</tr>
<tr>
<td>Country of origin</td>
<td>2.22</td>
<td>0.63-7.86</td>
</tr>
<tr>
<td># Patients</td>
<td>1.00</td>
<td>1.00-1.01</td>
</tr>
<tr>
<td>Study Phase</td>
<td>0.69</td>
<td>0.13-3.73</td>
</tr>
<tr>
<td>JIF</td>
<td>1.03</td>
<td>0.99-1.08</td>
</tr>
<tr>
<td>Funding</td>
<td>2.47</td>
<td>0.36-16.84</td>
</tr>
<tr>
<td>Publication Delay</td>
<td>0.99</td>
<td>0.84-1.15</td>
</tr>
<tr>
<td>Honorary Authorship</td>
<td>0.26</td>
<td>0.06-1.15</td>
</tr>
</tbody>
</table>
Figure 4.1.

403 papers identified

200 meet criteria

94 protocol available

Phase II: 22
Phase III: 72

106 protocol not available

Phase II: 26
Phase III: 80

Excluded:
8 Biomarker analysis
5 Not cancer therapy.
42 Not primary publication.
95 Not RCT.
8 Phase I/II
4 Published before July 2010.
6 Radiation trials.
3 Surgical trials.
22 Sequence trials.
8 Non inferiority.
2 Multi-arm.

61 Listed Investigators in the protocol
References

2. Shapiro DW, Wenger NS, Shapiro MF: The contributions of authors to multi-authored biomedical research papers. JAMA 271:438-42, 1994
Chapter 5: Discussion

This chapter will summarize the results of the thesis in terms of its contribution to the literature and implications of bias in reporting outcomes and toxicity in clinical trials, and the effect of ghost and honorary authorship and funding and conflicts of interest as predictor of bias. The limitations and strengths of the study will be discussed and suggestions for key areas of future research will be presented.

5.1. Contribution to the Literature.

Several papers have evaluated the frequency and characteristics of bias in the reporting of efficacy [1, 2, 8, 32, 42, 43], but these reports have tended to focus on heterogeneous medical conditions and not on cancer clinical trials. Furthermore, there are limited data in the literature about bias in the reporting of toxicity.[13] Here we have explored the frequency of bias in reporting of efficacy and toxicity in randomized trials evaluating treatments for breast cancer in the first manuscript, and after developing tools for assessment, in reports of RCTs evaluating systemic therapy for all cancer types. We focused our first project on research for breast cancer given that it is the most common malignancy in women, has substantial mortality [44] and is a cancer site with a large number of trials.

Bias in the reporting of the primary endpoint was prevalent among studies where there was a no statistically significant difference in the primary endpoint between
the arms. You et al [49] evaluated reports of RCTs published between 2005 and 2009, and found that there was misinterpretation of results relating to the primary endpoint in 21.6% of the trials; this included non-significance in a superiority trial interpreted as showing treatment equivalence, study conclusion based on endpoints other than the PE, study considered positive despite a non-significant p-value, and study conclusions based only on one endpoint when there were co-primary endpoints. We found a higher incidence of inappropriate reporting of the PE in RCTs for breast cancer that increased dramatically when only the trials with a non-significant p-value were assessed. Spin was used frequently to positively influence, the interpretation of negative trials, by emphasizing the apparent benefit of a secondary endpoint. We found bias in reporting efficacy and toxicity in 33% and 67% of trials, respectively, with spin and bias used to suggest efficacy in 59% of the trials that had no significant difference in their primary endpoint. These results are similar to those in other areas of medicine [8]. We found that bias in the reporting of toxicity was higher when the trial had a significant p-value for the difference in the primary endpoint between experimental and control arms. A possible explanation for this finding may be that investigators and/or sponsors then focus on efficacy as the basis of registration and downplay toxicity to make the results more attractive.

In the second study, we assessed papers reporting outcomes using systemic therapies for all types of cancer. Although in 2004 the ICMJE published guidelines for mandatory registration of clinical trials [17], and in 2007 the Surgical Journal Editors Group
followed these recommendations[18]; consistency between a clinical trial registry and the final manuscript in the reporting of primary and secondary endpoints of surgical RCTs was reported recently to be low: only 55% of the published papers showed no discrepancy while in 45% of manuscripts there was omission, introduction, change in definition, downgrading or upgrading of outcomes.[19] Another paper showed similar results, with 49% discrepancies in the reporting of primary outcomes.[20] In our initial study we identified only 30 trials included in ClinicalTrials.gov. Among these studies, the primary endpoint was changed in the final report in seven (23.3%) studies; in all cases the primary endpoint in the clinical trial registry was OS and a surrogate was used at the time of publication. Here we present evidence that for 94 RCTs evaluating medical interventions for cancer reported since 2010, with available protocols, 99% of the studies did not change the original primary outcome. Only in 2 articles was the reporting vague. These results confirm an advance in transparency although we cannot comment on whether findings are similar in the 106 articles (53%) where the protocols were not available.

We reported previously that bias in reporting outcomes is almost 60% in articles reporting studies with a negative primary endpoint in breast cancer RCTs [11]. Here we confirm biased reporting of 47% of RCTs evaluating treatments for a variety of tumor sites, even when limiting our study to reports in journals with high impact factors - journals that are associated with changes in standards of clinical practice.
Underreporting of toxicity to highlight a positive primary endpoint is a type of bias that has been reported previously by us and by other groups [11, 29, 61]. Reporting of toxicity or tolerance in a more positive way for the experimental arm has been associated with studies that have financial ties with for-profit sponsorship.[15] However our analysis did not find an association of biased reporting of toxicity with either funding source or with first author financial ties. Bias was more prevalent in reporting efficacy than toxicity. However not only does financial COI exist, there are also intrinsic COIs, that relate to an investigator’s perception of the need to engage in and publish research to achieve career advancement, to receive accolades from peers and professional societies, and to be competitive for grant funding[74, 75]. In the perception of patients, intrinsic COI are as relevant as financial COIs and should be acknowledge by the investigator at the time of discussion [76]; intrinsic COIs can probably be linked to bias in reporting efficacy and toxicity to make the study look positive, however our dataset was not powered to detect this type of COIs. Further exploration of this issue will need to take place in future research projects.

Our third study explored ghost and honorary authorship. Ghost authorship was analyzed from different perspectives: investigators listed in the protocol not considered for authorship or acknowledged in the manuscript, statisticians not listed as authors of the final publication or the use of medical writers. While some have argued that the use of medical writers does not constitute ghost authorship because they may be involved only in providing grammatical assistance [67], their potential to increase bias is substantial. We found that many investigators listed in
protocols were not included as authors of the manuscript, despite protocol development being the most critical part of any research project. Sixty percent of the analyzed studies met our definition of ghost authorship.

Honorary authorship can occur if a sponsor deems it advantageous to include an individual among the authors, usually a well-known “thought leader”, even if that individual has had minimal involvement in the study and does not meet ICMJE criteria for authorship. We found that authors, who did not meet required authorship criteria were included more frequently in the three general medical journals with the highest impact factor, compared with the journals focused only on oncology research. In 2% of the studies all authors did not meet the last criterion of authorship, approval of the manuscript. When analyzing the new added criterion, which refers to accountability of all aspects of the manuscript, we can infer that at least 2% of the authors did not meet this criterion since they did not approve the manuscript. Thirty-three percent of the studies analyzed met our criteria of honorary authorship.

5.2. Strengths and limitations.

5.2.1. Strengths

All three studies contributing to this thesis involved the analysis of retrospectively collected data that were evaluated rigorously and objectively. The study described in Chapter 2 has been cited on 18 occasions since its release in January 2013, and
also received wide media coverage that raised concerns about the quality of reporting of clinical trials. The objective of the second study reported in chapter 3 was to evaluate consistency in the reporting of the primary outcome among protocols, clinical trial registries and manuscripts; these assessments could not be undertaken in the first study, mainly because data were not sufficient given the earlier time frame of publication of the papers analyzed. Unlike other studies evaluating RCTs in oncology, we had access to a substantial number of protocols. We were able to describe some of the challenges to obtain this information, which adds value to our project and may facilitate future investigations. Most of the RCTs evaluated in this dataset were listed in a clinical trial registry. We were able to evaluate the impact of COI of the first/corresponding author in the reporting of bias, which is relevant since the leading author is the one who usually directs the message of the final study. In the third study described in chapter 4, we were able to estimate the incidence of ghost and honorary authorship through declarations in the manuscript and from direct information in protocols, and we are unaware of previous studies of this type in oncology. Although it is difficult to directly relate ghost and honorary authorship to author-bias their prevalence suggests substantial potential for such bias, as documented in the two earlier chapters.

5.2.2.Limitations

The first study was designed partly to develop and test tools to evaluate spin and reporting bias. We limited our investigation to RCTs with a sample size of at least
200; including studies with <200 patients would likely increase the level of bias, but the clinical impact of such studies is low. Second, we utilized subjective measures for some of our outcome measures such as the presence of spin. Third, our scales used to assess bias in reporting of efficacy and toxicity were based on our interpretation of the characteristics that a paper has to accomplish to be considered unbiased, but they have not been validated. Fourth, for chapter 2 many of our included trials were not available at ClinicalTrials.gov. This database was established in 2002 [50] and many trials initiated prior to this date were not included. Furthermore, many European trials were not included initially in the US-based ClinicalTrials.gov database and European Clinical Trials Registries do not have easily searchable databases [51], this condition changed radically for assessment in chapter 3, where all but seven studies were identified in a clinical trials registry. Our analysis of change in the primary endpoint was based only on 30 studies and should therefore be interpreted with caution.

Major limitations for the second study were first that protocols were available for under half of the published reports of RCTs. Second, we utilized subjective measures to determine some of our outcome measures such as the presence of bias in language. Third, we focused only on papers in journals with high impact factor, although these are where practice changing studies are most often published. Fourth, we focused only on papers reporting outcomes for systemic therapies, and excluded those reporting outcomes in surgical and radiation oncology, so that our conclusions cannot be generalized to the field of oncology. Fifth, for financial ties we explored the effect only of first and corresponding
authors, although we recognize that other authors can influence decisions as to what is reported.

In the third study evaluation of protocols is one of the most objective measures for ghost authorship but for 34% of the studies the protocol did not provide the list of investigators, thereby limiting the analysis of ghost authorship. The declaration of authorship varies among publications, but we were able to extract these data and information was entered in a uniform format according to the ICMJE guidelines.

5.3. Future Research

As a product of the initial study, we have participated in the project entitled: Impact of spin in the abstract of articles reporting results of Randomised Controlled Trials in the field of cancer, the SPIIN Randomised Controlled Trial. This study evaluates in a randomized design how readers interpret results of abstracts written with and without spin. The manuscript has been submitted to JCO for consideration of publication.

As mentioned before, COI of interest are not only financial, intrinsic COI of interest describe the need of succeed in the research world and researches have pressure to publish, these can cloud investigator´s judgments, leading to research activities that inappropriately increase risks to prospective participants[77, 78], and theoretically when wrongly reported results can bias the perception of peers about efficacy and safety of a new treatment. An important step would be to develop a
questionnaire about the perception of intrinsic conflict of interest from the cohort of studies analyzed in this thesis, and associate those responses with bias in publication of results, if it becomes significant, in comparison with the non-significant association with financial COI, there will be evidence to be more critical about motivations in academic careers; subjectivity of this topic makes evaluation difficult, but tools to measure these will need to be developed.

Our results can serve as a guide to editors to improve the mechanisms of peer review and increase transparency in how research is presented. As we presented in Chapter 2 even after the establishment of guidelines, bias in reporting outcomes remained common and was not influenced by time; in Chapter 3 this assessment could not be done given the inclusion of short time frame of 2.5 years. However our results suggested a similar trend. To evaluate bias in the future this can be done by establishing two transversal cohorts with a time difference of 5 years from the publication of the current thesis in a peer review journal, to assess if these publications and other that have been reported in the last 2 years have impacted the practices of editorial offices.

Although authors can have COI that can lead to bias about how research is presented[79], it is also important to consider the effect of COI that reviewers have at the time of accepting or rejecting a submitted manuscript. It has been reported previously that a reviewer shared adverse information of a meta-analysis of Avandia’s safety and toxicity[80] with the pharmaceutical who owned the product ahead of publication in the New England Journal of Medicine, and two weeks later the RECORD study, funded by GlaxoSmithKline, published in the same journal an
interim analysis arguing that the data from such a study was insufficient to make an statement about toxicity of the drug. This type of study will place greater scrutiny on the expectations of reviewers and if it confirms the independence of the peer review process it will strength the confidence on the research we read in medical literature.

5.4. Conclusions

Bias in reporting efficacy and toxicity of systemic therapy in oncology is prevalent. Measures to improve transparency in the reporting of outcomes through clinical trial registries have met expectations; however access to original protocols is limited. Access to protocols not only improves reporting, but also guides the reader through characteristics of the study that usually are not reported in the final manuscript because of space limitations, availability of protocols would improve the ability of readers to make clinical judgements. Transparency in reporting of authorship is an area of opportunity where editors can expand guidelines in order to reduce ghost and honorary authorship.

6. References.

[22] Bariani GM, de Celis Ferrari AC, Hoff PM, Krzyzanowska MK, Riechelmann RP. Self-reported conflicts of interest of authors, trial sponsorship, and the interpretation of editorials and related
[38] Barry HC, Ebell MH, Shaughnessy AF, Slawson DC, Nietzke F. Family physicians' use of medical abstracts to guide decision making: style or substance? The Journal of the American Board of Family Practice / American Board of Family Practice. 2001;14:437-42.


