

Evolution of the Randomized Controlled Trial in Oncology Over Three Decades

Christopher M. Booth, David W. Cescon, Lisa Wang, Ian F. Tannock, and Monika K. Krzyzanowska

A B S T R A C T

Purpose

The randomized controlled trial (RCT) is the gold standard for establishing new therapies in clinical oncology. Here we document changes with time in design, sponsorship, and outcomes of oncology RCTs.

Methods

Reports of RCTs evaluating systemic therapy for breast, colorectal (CRC), and non-small-cell lung cancer (NSCLC) published 1975 to 2004 in six major journals were reviewed. Two authors abstracted data regarding trial design, results, and conclusions. Conclusions of authors were graded using a 7-point Likert scale. For each study the effect size for the primary end point was converted to a summary measure.

Results

A total of 321 eligible RCTs were included (48% breast, 24% CRC, 28% NSCLC). Over time, the number and size of RCTs increased considerably. For-profit/mixed sponsorship increased substantially during the study period (4% to 57%; $P < .001$). There was increasing use of time-to-event measures (39% to 78%) and decreasing use of response rate (54% to 14%) as primary end point ($P < .001$). Effect size remained stable over the study period. Authors have become more likely to strongly endorse the experimental arm ($P = .017$). A significant P value for the primary end point and industry sponsorship were each independently associated with endorsement of the experimental agent (odds ratio [OR] = 19.6, 95% CI, 8.9 to 43.1, and OR = 3.5, 95% CI, 1.6 to 7.5, respectively).

Conclusion

RCTs in oncology have become larger and are more likely to be sponsored by industry. Authors of modern RCTs are more likely to strongly endorse novel therapies. For-profit sponsorship and statistically significant results are independently associated with endorsement of the experimental arm.

J Clin Oncol 26:5458-5464. © 2008 by American Society of Clinical Oncology

INTRODUCTION

The randomized controlled trial (RCT) has become the gold standard for developing new therapeutic agents in oncology. In translating the results of RCTs into practice, clinicians should consider the quality of study design and execution and generalizability of the trials' results. Previous reviews have reported that the numbers and sizes of oncology RCTs is increasing over time¹⁻⁴ and that their design has improved, but that important deficiencies remain.⁵⁻⁸

Since the 1980s, the funding of biomedical research has changed, with 70% of funding for clinical drug trials now coming from the pharmaceutical industry.⁹ With this shift in sponsorship, it is important to recognize the potential for bias and conflict of interest in reports of RCTs. Several studies have pre-

viously found an association between for-profit sponsorship and the reporting of positive results in RCTs.¹⁰⁻¹⁴

Trends in methodology, sponsorship, and outcomes of RCTs evaluating treatments for cancer have not been well described. Although major advances have occurred, it was our hypothesis that recent positive RCTs might be reporting smaller effect sizes than those in the past. Furthermore, author interpretation of this effect size may have evolved such that therapeutic benefit once considered clinically insignificant is now being considered practice changing. To address these issues, we designed the current study to provide a comprehensive review of published RCTs in breast, colorectal (CRC), and non-small-cell lung cancer (NSCLC) over a 30-year period (1975 through 2004). Our objectives were to describe trends in methodology and reporting of

From the National Cancer Institute of Canada Clinical Trials Group, Queen's University, Kingston; and Princess Margaret Hospital, University of Toronto, Toronto, Ontario, Canada.

Submitted February 1, 2008; accepted July 8, 2008; published online ahead of print at www.jco.org on October 27, 2008.

Presented in part at the 43rd Annual Meeting of the American Society of Clinical Oncology, Chicago, IL, June 1-5, 2007.

Authors' disclosures of potential conflicts of interest and author contributions are found at the end of this article.

Corresponding author: Monika Krzyzanowska, MD, MPH, Department of Medical Oncology and Hematology, Princess Margaret Hospital, 610 University Ave, Toronto, ON, M5G 2M9, Canada; e-mail: monika.krzyzanowska@uhn.on.ca.

© 2008 by American Society of Clinical Oncology

0732-183X/08/2633-5458/\$20.00

DOI: 10.1200/JCO.2008.16.5456

RCTs, in addition to sponsorship, outcomes, and authors' interpretation of results. From this overview, we expect to gain insight into how to improve the reporting and interpretation of contemporary clinical trials in oncology.

METHODS

Search Strategy

A search was undertaken for all RCTs of systemic therapy in breast, CRC, and NSCLC published during three decades (1975 through 2004) in the following journals: *Journal of Clinical Oncology*, *Journal of the National Cancer Institute*, *Cancer Treatment/Chemotherapy Reports*, *New England Journal of Medicine*, *Lancet*, and *Journal of the American Medical Association*. These journals were selected because they were felt to contain a high proportion of widely read and practice-changing clinical trials in oncology published over the past 30 years. Indexes and tables of contents of these journals were reviewed electronically and by hand to find relevant articles. The following were excluded: studies of a radiation and/or surgical intervention, studies of cancer screening and prevention, articles that presented data from multiple RCTs, studies comparing a single drug(s) given by different schedule, multiple reports of the same study (the first final report in a journal we reviewed was included), phase II or pilot studies, and studies that presented results only for a subgroup of the original study population.

Data Abstraction

A data abstraction form was designed to capture information regarding study methodology, sponsorship, results, and author conclusions. To guide the abstraction process, a data manual was created to ensure consistency between the two abstractors who reviewed eligible articles. The data abstraction form was piloted by two authors of the current study (C.M.B. and D.W.C.) on 30 RCTs; results were compared and the abstraction tool was subsequently modified.

All eligible articles were reviewed. Country of study origin was assigned based on the institutional affiliation of the first author. The primary end point of each study was identified; if there was no explicit statement, the end point implied to be of primary importance was recorded. We evaluated use of intention-to-treat (ITT) analysis based on raw data presented in the article. We recorded whether the analysis included all randomly assigned patients or only eligible patients.

Study conclusions were assigned a score from 1 to 7 based on a scale developed by Ridker and Torres:¹² 4 of 7 for a neutral statement, 7 of 7 for strong endorsement of experimental arm, and 1 of 7 for strong endorsement of the control arm (Table 1).

Two authors of the present study (C.M.B. and D.W.C.) each performed data abstraction independently on one half of the eligible articles. A score out of 7 for author conclusion was assigned based on the concluding section of the abstract (RCT author's score). This score was assigned before the reviewer reading any other part of the abstract or article and without knowing the

sponsorship status of the trial. For studies with no abstract the concluding paragraph of the article was used to assign the RCT author's score. To provide a comparison to the RCT author's conclusion, on reviewing the full article, the same reviewer assigned a score out of 7 (reviewer's score) based on their impression of the overall benefit (or lack thereof) and toxicity of the experimental arm compared with control.

Study sponsorship was determined based on explicit statements in the article and by the affiliation of study authors. Using definitions proposed by Ridker and others, studies were classified into one of four groups: those financed exclusively by for-profit pharmaceutical companies; those financed exclusively by government, foundation, or other not-for-profit agencies; those financed jointly by for-profit and not-for-profit sources; and those for which no source of funding was identified.^{12,15,16}

Statistical Analysis

Descriptive statistics were used to summarize trends over time. The study period was divided into three decades: 1975 to 1984, 1985 to 1994, and 1995 to 2004. A summary measure of effect size was calculated for the primary end point of each study (or if not stated explicitly, the most clinically relevant end point presented). Because of the considerable heterogeneity in primary end points, we were unable to calculate absolute effect size across all studies. For this reason we determined the relative effect size of the experimental compared with control arms, which allowed for comparison across studies, disease site, and time. For time-to-event end points (ie, overall survival, disease-free survival) a hazard ratio (HR) was calculated for each study from reported survival rates in experimental and control groups by assuming time to event was exponentially distributed. For studies in which the primary end point was response rate, the summary measure was calculated as a risk ratio (RR): the ratio of response rates between experimental and control arms. For studies with multiple treatment groups, the experimental arm with the best outcome was chosen for calculation of HR and RR relative to the control arm.

Logistic regression was used to identify factors associated with strong author endorsement (score of 6 or 7). Explanatory variables considered were decade (1975 to 1984, 1985 to 1994, or 1995 to 2004), disease site (breast, colorectal, lung), setting (palliative, adjuvant, neoadjuvant), type of control arm (active treatment or not), type of primary end point (time to event or response rate), statistical significance of primary end point, effect size, and sponsorship (for-profit, not-for-profit, mixed, or not known). Variables with *P* values less than .1 in univariate analysis were entered into the multivariate model. In multivariate analysis, stepwise selection techniques were used, and predictors were considered statistically significant if the *P* value was less than .05. All analyses were performed in SAS 9.1 (SAS Institute, Cary, NC).

RESULTS

The search strategy yielded 380 articles. Fifty-nine articles were subsequently excluded for the following reasons: multiple reports of the same study (*n* = 12); reports of subgroup analysis (*n* = 11); studies of different dosing schedules for the same drug (*n* = 9); other diseases (*n* = 6); radiation and/or surgical intervention (*n* = 5); studies that were not RCTs (*n* = 5); data pooled from multiple studies (*n* = 4); phase II/pilot studies (*n* = 3); prevention studies (*n* = 2); reports of preliminary safety data (*n* = 2). The remaining 321 eligible articles included 171,161 randomly assigned patients. As shown in Table 2, there was a substantial increase in the total number of reported RCTs over time. Although the proportion of RCTs for women with breast cancer remained constant over time, there was a considerable increase in the proportion of RCTs for colorectal cancer and a decrease in those for NSCLC.

Inter-observer agreement between the two data abstractors (based on independent data extraction from 10 studies) was found to be excellent: 96% agreement for all data points and κ statistic of 0.90 and 0.96 for reviewers and author's scores, respectively.

Table 1. Grading Scale for Authors' Conclusions

Score (/7)	Authors' Conclusion
1	Standard of care (SOC) was significantly better and/or highly preferred
2	The SOC was significantly better and/or moderately preferred
3	The SOC was nonsignificantly better
4	The SOC and newer treatment were equivalent
5	The newer treatment was nonsignificantly better
6	The newer treatment was significantly better and/or moderately preferred
7	The newer treatment was significantly better and/or highly preferred

Table 2. Trends in Oncology RCT Methodology

Characteristic	1975-1984 (n = 47)		1985-1994 (n = 107)		1995-2004 (n = 167)	
	No.	%	No.	%	No.	%
Disease site						
Breast	19	40	53	50	81	49
NSCLC	23	49	29	27	39	23
CRC	5	11	25	23	47	28
Setting						
Palliative	34	72	66	62	101	60
Adjuvant	13	28	40	37	61	37
Neoadjuvant	0	0	1	1	5	3
Organization						
Multicenter	26	60	87	85	158	95
International	12	26	30	28	86	52
Cooperative group	13	28	59	56	71	43
Study origin						
North America	28	60	69	64	61	37
Europe	17	36	35	33	100	60
Other	2	4	3	3	6	4
Study design						
Sample size						
Median		100		249		446
No. of studies for which these data were available	47		107		167	
Time for accrual						
Median, months		30		41		33
No. of studies for which these data were available	22		83		153	
Follow-up						
Median, months		30		55		47
No. of studies for which these data were available	6		38		105	
Interventions						
Any chemotherapy	36	77	70	66	124	75
Any placebo/observation	10	21	28	26	25	15
Any hormonal agent	4	9	14	13	18	11
Any targeted agent	1	2	3	3	7	4
ITT analysis						
Any	32	70	93	87	155	93
All randomized	15	33	24	22	91	54
Eligible only	17	37	69	64	64	38
Primary end point*						
Time-to-event end point	11	39	60	72	125	78
OS	6	21	51	61	82	51
DFS	2	7	5	6	16	10
RFS	3	11	4	5	8	5
TTP	0	0	0	0	19	12
RR	15	54	21	25	23	14
Other	2	7	2	2	13	8

Abbreviations: RCT, randomized controlled trial; NSCLC, non-small-cell lung cancer; CRC, colorectal cancer; ITT, intention to treat; OS, overall survival; DFS, disease-free survival; RFS, relapse-free survival; TTP, time to progression; RR, response rate.

*Data shown are for studies in which a primary end point was either explicitly identified or implied in the article. Time-to-event end points include OS, DFS, RFS, and TTP.

Study Methodology

Trends in study organization and methodology are shown in Table 2. Over time there was a substantial increase in multicenter and international trials: North American-led trials decreased from 60% to 37% (28 of 47 to 61 of 167 trials; $P < .0001$), whereas European-initiated studies increased from 36% to 60% (17 of 47 to 100 of 167 trials; $P < .0001$). Despite a considerable increase in median sample size, the duration of study accrual remained stable.

A primary end point was explicitly stated in 7% (three of 47 trials), 29% (31 of 107 trials), and 67% (112 of 167 trials) of RCTs in

the three decades, respectively ($P < .0001$). As shown in Table 2, there was a shift from response rate as primary end point (54% to 14% [15 of 28 trials to 23 of 161 trials]) to time-to-event end points (39% to 78% [11 of 28 trials to 125 of 161 trials]; $P < .001$). Ten percent of RCT reports explicitly stated that the trial was terminated prematurely, and this remained constant over time. ITT analysis was performed in 87% (280 of 321 trials) of RCTs, although many studies (47%, 150 of 321 trials) included only eligible patients in the ITT analysis. ITT analysis of all randomly assigned patients increased over the study period (33% to 54% [15 of 47 trials to 91 of 167 trials]; $P < .0001$).

Sponsorship, Effect Size, and Author Conclusions

There was a substantial increase in for-profit and mixed sponsorship between 1975 and 2004 (Table 3). Government-funded RCTs decreased from 60% to 31% (28 of 47 trials to 51 of 167 trials; $P < .0001$), whereas industry sponsorship has increased from 4% to 57% (two of 47 trials to 95 of 167 trials; $P < .0001$). Studies funded by for-profit organizations were more likely to be in the setting of metastatic disease (73% v 53% [88 of 120 and 87 of 165 trials]; $P < .001$) and to have larger median sample size (396 patients v 307 patients; $P = .015$) compared with studies sponsored by nonprofit groups.

As shown in Table 3, the relative benefit of the experimental arm compared with the control arm (ie, effect size) remained stable over time. There has been an increase over time in the proportion of trials with a significant P value for the primary end point (23% v 42% [11 of 47 and 70 of 167 trials]; $P = .007$). Although effect size was not found to vary with source of sponsorship, there was a trend toward a greater proportion of industry-funded trials having a significant P value as compared with nonprofit-funded studies (41% v 30% [49 of 120 and 50 of 165 trials]; $P = .07$).

The proportion of studies in which RCT authors strongly endorsed the experimental arm (defined as a score of 6 or 7) increased from 31% to 49% over the study period (11 of 37 trials to 82 of 167 trials; $P = .017$; Table 3). Scores assigned by the reviewers did not change significantly with time. Median RCT author score (5 v 4;

$P = .001$) and the proportion of authors who strongly endorsed the experimental arm (56% v 36% [66 of 118 and 55 of 153 trials]; $P = .001$) were greater for for-profit/mixed versus not-for-profit sponsored studies. As shown in Table 4, predictors of strong endorsement of the experimental arm (ie, RCT author score of 6 or 7) in univariate analysis were as follows: significant P value for the primary end point, control arm with no active treatment, time-to-event primary end point, for-profit/mixed sponsorship, adjuvant/neoadjuvant setting, and effect size. Significant P value, time-to-event end point, for-profit/mixed sponsorship, and effect size remained significant in multivariate analysis. The test for interaction between $P \leq .05$ for the primary end point and sponsorship was not significant ($P = .65$).

DISCUSSION

We have observed several important trends in this review of 321 RCTs involving more than 170,000 patients with cancer conducted between 1975 and 2004. Contemporary RCTs are larger and more likely to be multicenter and international. Time-to-event end points have largely replaced response rate as the primary end point. This encouraging move toward more clinically relevant end points is tempered by the observation that one third of RCTs published between 1995 and 2004 failed to explicitly identify the primary end point. Authors of modern

Table 3. Sponsorship, Results, and Author Conclusions for Oncology RCTs

Characteristic	1975-1984 (n = 47)		1985-1994 (n = 107)		1995-2004 (n = 167)	
	No.	%	No.	%	No.	%
Sponsorship classification						
For profit	1	2	12	11	64	38
Nonprofit	33	72	72	67	60	36
Mixed	1	2	11	10	31	19
Not known	12	26	12	16	12	7
Source of RCT funding						
Government	28	60	66	62	51	31
Industry	2	4	23	23	95	57
Foundation	7	15	14	13	24	14
Cooperative group	12	26	59	55	75	45
Other	1	2	0	0	2	1
Effect size						
HR						
Median	1.4		1.2		1.2	
95% CI	1.0 to 2.3		1.0 to 1.4		1.1 to 1.3	
RR						
Median	0.9		1.1		1.3	
95% CI	0.6 to 1.5		0.9 to 1.3		1.1 to 1.4	
$P \leq .05$ for primary end point	11	23	32	30	70	42
Study conclusion*						
RCT author score						
Median	4		4		5	
RCT author score 6 or 7†	11	31	40	39	82	49
Reviewer score						
Median	5		5		5	
Reviewer score 6 or 7	17	37	42	39	78	47

Abbreviations: RCT, randomized controlled trial; HR, hazard ratio; RR, response rate.

*For score definitions, refer to Table 1.

†Data shown for author score are for studies in which there was an assessable concluding statement (1975-1984, n = 36; 1985-1994, n = 103; 1994-2004, n = 167).

Table 4. Results of Univariate and Multivariate Analysis for Variables Associated With RCT Author Strong Endorsement of the Experimental Arm

Variable	Univariate Analysis			Multivariate Analysis		
	Odds Ratio	95% CI	P	Odds Ratio	95% CI	P
Decade						
1975-1984	Reference			—	—	—
1985-1994	1.4	0.6 to 3.3				
1995-2004	2.2	1.0 to 4.8	.06			
Disease site						
Breast	Reference			—	—	—
Colorectal	1.0	0.6 to 1.8				
NSCLC	0.7	0.4 to 1.3	.465			
Setting						
Palliative	Reference			—	—	—
Adjuvant neoadjuvant	2.3	1.4 to 3.6	.0007			
Control arm						
Active agent	Reference			—	—	—
No active agent	4.7	2.0 to 10.8	.0003			
Primary end point						
Response rate	Reference			Reference		
Time to event	2.9	1.6 to 5.2	.0003	4.1	1.6 to 10.3	.003
Effect size, unit = 0.1	1.3	1.2 to 1.4	< .0001	1.1	1.1 to 1.2	.0004
P for primary end point						
> .05	Reference			Reference		
≤ .05	25.2	13.6 to 46.9	< .0001	19.6	8.9 to 43.1	< .0001
Not known	1.5	0.3 to 7.4		0.7	0.02 to 26.3	
Sponsorship						
Nonprofit	Reference			Reference		
For-profit/mixed	2.3	1.4 to 3.7	.003	3.5	1.6 to 7.5	.004
Not known	1.0	0.4 to 2.1		1.0	0.3 to 3.5	

NOTE. Strong endorsement defined as RCT author score of 6 or 7.
Abbreviations: RCT, randomized controlled trial; NSCLC, non-small-cell lung cancer.

RCTs are more likely to strongly endorse the experimental arm. Consistent with the recent report by Djulbegovic et al, we found that effect size in new cancer therapies has remained stable over time.¹⁷

There has been a dramatic shift in sponsorship of RCTs in oncology from government to for-profit organizations. Although sponsorship status is not associated with increased effect size, industry-funded RCTs are more likely to strongly endorse novel treatments. Multivariate analyses suggest that independent predictors of endorsing the experimental therapy are significant *P* value for primary end point, time-to-event primary end point, for-profit/mixed sponsorship status, and effect size.

The conclusions drawn from an RCT can be influenced substantially by the quality of methodology and reporting. Several reports in the oncology literature have described some improvement in quality of trial design and reporting over time, but consistent deficiencies persist, including failure to perform ITT analysis on all randomly assigned patients.^{2,3,5-8,18} In a review of abstracts describing RCTs presented at the American Society of Clinical Oncology annual meeting (1989 through 1998), Krzyzanowska et al¹⁹ found that only 22% of studies identified the primary end point explicitly and 74% of abstracts reported multiple end points. Here we found multiple studies without explicit definition of the primary end point in full publications. Although inadequate reporting does not necessarily imply deficient methodology,²⁰ it is essential that the primary end point of a trial be described a priori, because otherwise, statistical analysis may be mis-

leading. In light of these factors, medical journals have made significant strides toward improving the quality of clinical trial design and reporting through the development and adoption of the CONSORT statement and mandatory trial registration.^{21,22}

Previous studies have shown that study outcome may be correlated with sponsorship. A meta-analysis by Bekelman et al²³ that pooled data from eight (nononcology) studies showed a significant association between industry sponsorship and pro-industry conclusions (odds ratio, 3.6; 95% CI, 2.6 to 4.9). Within the oncology literature, two reviews of pharmacoeconomic studies have found that studies sponsored by for-profit organizations are more likely to draw favorable conclusions about novel anticancer agents.^{24,25} In RCTs evaluating treatments for multiple myeloma, Djulbegovic et al²⁶ found that equipoise was maintained in studies funded by nonprofit organizations, with 47% of RCTs favoring the experimental arm. However, in RCTs supported by industry, 74% of studies favored the experimental treatment. Peppercorn et al²⁷ reviewed published clinical trials for breast cancer from 10 journals published in 1993, 1998, and 2003 and found evidence of increasing pharmaceutical industry involvement over time. Although these reports have suggested an association between outcome and source of sponsorship, they were each limited to a single disease site and did not compare the authors' conclusion with any differences in effect size. Our study provides information about study sponsorship, relative outcome in control and experimental arms, and author interpretation of this benefit.

Within the existing literature and clinical practice, there is considerable heterogeneity regarding what constitutes a so-called positive trial. At least three domains contribute to an RCT being classified as positive or negative: the benefit of one arm compared with the other, the statistical significance of this difference, and the interpretation of results as presented by study authors. In this study, we have evaluated the association between each of these three variables. Our results suggest that the strongest predictor of RCT authors endorsing new therapies is a statistically significant difference in outcome between study arms. We have also found that industry sponsorship is an independent predictor of studies being reported as positive. The causal mechanism of the latter observation is not clear. Given the stability of effect size across studies, it seems unlikely that industry-funded RCTs are finding a greater magnitude of benefit between treatment arms. Although our multivariate analysis found that industry sponsorship and significant *P* value were each associated with strong author endorsement, the lack of significant interaction between these two variables may simply represent an underpowered analysis. Accordingly, it remains plausible that industry-sponsored trials are more likely to be positive because they are larger and therefore have more power to detect a smaller difference. Our observation that modern authors are more likely to endorse the experimental arm may simply reflect changes in sponsorship, study design, and/or journal review process and reporting criteria that also occurred during the study period. Furthermore, it is possible that authors are endorsing novel agents that have less toxicity or are more convenient to deliver.

By reviewing RCTs for three disease sites over a period of three decades, our data provide information that may be of use in the reporting and interpretation of contemporary cancer clinical trials. A potential weakness of our study is that by including only RCTs of systemic therapy in breast cancer, NSCLC, and CRC, our findings may not be generalizable to other disease sites. Also, by limiting our search to six journals, we did not capture every RCT published during the study period. However, we were most interested in methodology, sponsorship, and outcomes of practice-changing RCTs; a high proportion of which are published in the journals we included. Publication bias has been well described,²⁸ and we recognize that our cohort of trials does not represent the entire body of RCTs in oncology. Although we were unable to evaluate whether absolute effect size has decreased over time, given the stability of relative benefit and the marked increase in sample size (and likely power) during the study period, our initial hypothesis of reduced absolute benefit over time likely remains valid, although not directly proven. Finally, perception

of what constitutes a positive trial is a complex and multifactorial process and may involve other variables that we have not measured in this study.

In summary, we have found that modern RCTs in breast cancer, NSCLC, and CRC are substantially larger and more international in scope than those of earlier decades. Although methodology and quality of reporting seems to be improving over time, serious deficiencies persist, particularly in the identification of the primary end point and by not including all randomly assigned patients in ITT analyses. There has been a substantial shift toward industry sponsorship of oncology RCTs. Over the past 30 years, authors' endorsement of novel therapies has increased while relative effect size has remained stable. A significant *P* value for the primary end point and industry sponsorship are independently associated with strong endorsement of the experimental therapy. Investigators and medical journals should continue to strive toward publication of high-quality studies and recognize the importance of adequate and unbiased reporting of study methodology and results. Finally, clinicians, investigators, and policy makers should maintain and refine perspective on what constitutes a meaningful benefit to patients beyond the *P* value associated with the result. Further research is needed to determine whether newly adopted therapies are truly worthwhile to patients.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The author(s) indicated no potential conflicts of interest.

AUTHOR CONTRIBUTIONS

Conception and design: Christopher M. Booth, David W. Cescon, Lisa Wang, Ian F. Tannock, Monika K. Krzyzanowska

Administrative support: Christopher M. Booth, David W. Cescon, Ian F. Tannock, Monika K. Krzyzanowska

Provision of study materials or patients: Christopher M. Booth, David W. Cescon, Ian F. Tannock, Monika K. Krzyzanowska

Collection and assembly of data: Christopher M. Booth, David W. Cescon, Lisa Wang, Ian F. Tannock, Monika K. Krzyzanowska

Data analysis and interpretation: Christopher M. Booth, David W. Cescon, Lisa Wang, Ian F. Tannock, Monika K. Krzyzanowska

Manuscript writing: Christopher M. Booth, David W. Cescon, Lisa Wang, Ian F. Tannock, Monika K. Krzyzanowska

Final approval of manuscript: Christopher M. Booth, David W. Cescon, Lisa Wang, Ian F. Tannock, Monika K. Krzyzanowska

REFERENCES

- Vale C, Stewart L, Tierney J: Trends in UK cancer trials: Results from the UK Coordinating Committee for Cancer Res National Register of Cancer Trials. *Br J Cancer* 92:811-814, 2005
- Breathnach OS, Freidlin B, Conley B, et al: Twenty-two years of phase III trials for patients with advanced non-small-cell lung cancer: Sobering results. *J Clin Oncol* 19:1734-1742, 2001
- Ioannidis JP, Polycarpou A, Ntais C, et al: Randomised trials comparing chemotherapy regimens for advanced non-small cell lung cancer: Biases and evolution over time. *Eur J Cancer* 39:2278-2287, 2003
- Reizenstein P, Delgado M, Gastiburu J, et al: Efficacy of and errors in randomized multicenter

trials: A review of 230 clinical trials. *Biomed Pharmacother* 37:14-24, 1983

5. Machin D, Stenning SP, Parmar MK, et al: Thirty years of Medical Research Council randomized trials in solid tumours. *Clin Oncol (R Coll Radiol)* 9:100-114, 1997

6. Nicolucci A, Grilli R, Alexanian AA, et al: Quality, evolution, and clinical implications of randomized, controlled trials on the treatment of lung cancer: A lost opportunity for meta-analysis. *JAMA* 262:2101-2107, 1989

7. Liberati A, Himmel HN, Chalmers TC: A quality assessment of randomized control trials of primary treatment of breast cancer. *J Clin Oncol* 4:942-951, 1986

8. Lai R, Chu R, Fraumeni M, et al: Quality of randomized controlled trials reporting in the primary

treatment of brain tumors. *J Clin Oncol* 24:1136-1144, 2006

9. Bodenheimer T: Uneasy alliance: Clinical investigators and the pharmaceutical industry. *N Engl J Med* 342:1539-1544, 2000

10. Bhandari M, Busse JW, Jackowski D, et al: Association between industry funding and statistically significant pro-industry findings in medical and surgical randomized trials. *CMAJ* 170:477-480, 2004

11. Yaphe J, Edman R, Knishkowsky B, et al: The association between funding by commercial interests and study outcome in randomized controlled drug trials. *Fam Pract* 18:565-568, 2001

12. Ridker PM, Torres J: Reported outcomes in major cardiovascular clinical trials funded by for-profit and not-for-profit organizations: 2000-2005. *JAMA* 295:2270-2274, 2006

13. Als-Nielsen B, Chen W, Gluud C, et al: Association of funding and conclusions in randomized drug trials: A reflection of treatment effect or adverse events? *JAMA* 290:921-928, 2003
14. Kjaergard LL, Als-Nielsen B: Association between competing interests and authors' conclusions: Epidemiological study of randomised clinical trials published in the *BMJ*. *BMJ* 325:249, 2002
15. Clifford TJ, Barrowman NJ, Moher D: Funding source, trial outcome and reporting quality: Are they related? Results of a pilot study. *BMC Health Serv Res* 2:18, 2002
16. Rochon PA, Gurwitz JH, Simms RW, et al: A study of manufacturer-supported trials of nonsteroidal anti-inflammatory drugs in the treatment of arthritis. *Arch Intern Med* 154:157-163, 1994
17. Djulbegovic B, Kumar A, Soares HP, et al: Treatment success in cancer: New cancer treatment successes identified in phase 3 randomized controlled trials conducted by the national cancer institute-sponsored cooperative oncology groups, 1955 to 2006. *Arch Intern Med* 168:632-642, 2008
18. Djulbegovic B, Adams JR, Lyman GH, et al: Evaluation and appraisal of randomized controlled trials in myeloma. *Ann Oncol* 12:1611-1617, 2001
19. Krzyzanowska MK, Pintilie M, Brezden-Masley C, et al: Quality of abstracts describing randomized trials in the proceedings of American Society of Clinical Oncology meetings: Guidelines for improved reporting. *J Clin Oncol* 22:1993-1999, 2004
20. Soares HP, Daniels S, Kumar A, et al: Bad reporting does not mean bad methods for randomised trials: Observational study of randomised controlled trials performed by the Radiation Therapy Oncology Group. *BMJ* 328:22-24, 2004
21. Begg C, Cho M, Eastwood S, et al: Improving the quality of reporting of randomized controlled trials: The CONSORT statement. *JAMA* 276:637-639, 1996
22. De Angelis C, Drazen JM, Frizelle FA, et al: Clinical trial registration: A statement from the International Committee of Medical Journal Editors. *N Engl J Med* 351:1250-1251, 2004
23. Bekelman JE, Li Y, Gross CP: Scope and impact of financial conflicts of interest in biomedical research: A systematic review. *JAMA* 289:454-465, 2003
24. Friedberg M, Saffran B, Stinson TJ, et al: Evaluation of conflict of interest in economic analyses of new drugs used in oncology. *JAMA* 282:1453-1457, 1999
25. Hartmann M, Knoth H, Schulz D, et al: Industry-sponsored economic studies in oncology vs studies sponsored by nonprofit organisations. *Br J Cancer* 89:1405-1408, 2003
26. Djulbegovic B, Lacey M, Cantor A, et al: The uncertainty principle and industry-sponsored research. *Lancet* 356:635-638, 2000
27. Peppercorn J, Blood E, Winer E, et al: Association between pharmaceutical involvement and outcomes in breast cancer clinical trials. *Cancer* 109:1239-1246, 2007
28. Krzyzanowska MK, Pintilie M, Tannock IF: Factors associated with failure to publish large randomized trials presented at an oncology meeting. *JAMA* 290:495-501, 2003

Acknowledgment

We thank Allan Detsky, MD, PhD, and Ralph Meyer, MD, for the thoughtful comments provided on earlier drafts of this manuscript, and Aoife O'Carroll and Shannon Godin for assistance provided in the conduct of this study.