Do Published Data in Trials Assessing Cancer Drugs Reflect the Real Picture of Efficacy and Safety?

Jia-Wei Lv, MD^{a,*}; Yu-Pei Chen, MD^{a,*}; Guan-Qun Zhou, MD^{a,*}; Xu Liu, MD^{a,*}; Ying Guo, MD^b; Yan-Ping Mao, MD^{a,c}; Jun Ma, MD^a; and Ying Sun, MD, PhD^a

Abstract

Background: The reporting quality of publications is of vital importance to ensure accurate evidence dissemination. This study aimed to compare the consistency of results reporting between the ClinicalTrials.gov results database and the respective matching publications. **Methods:** We identified 323 phase III/IV cancer drug trials with a randomized controlled design and searched PubMed for publications in a 50% random sample (n=160). Data were extracted independently from ClinicalTrials.gov and publications. A scoring system was applied to determine characteristics associated with reporting quality. **Results:** Of 117 reviewed trials with publications, result reporting was significantly more complete in ClinicalTrials.gov for efficacy measurement (92.3% vs 90.6%), serious adverse events (SAEs; 100% vs 43.6%), and other adverse events (OAEs; 100% vs 62.4%). For trials with both posted and published results for design information (n=117), efficacy measurements (n=98), SAEs (n=51), and OAEs (n=73), discrepancies were found in 16 (13.7%), 38 (38.8%), 26 (51.0%), and 54 (74.0%) trials, respectively. Overreporting of treatment effects (7 trials) and alteration of primary end points favoring statistically significant outcomes (11 trials) were the major discrepancies in efficacy reporting; incomplete (66 trials) and underreporting (20 trials) of SAEs were the predominant issues in benefit/risk reporting. Median quality score was 21 (range, 14–28). Trials that had parallel assignment, were phase IV, had primary funding by industry, were completed after 2009, and had earlier results posted possessed better reporting quality. **Conclusions:** Although most trials showed reasonable completeness and consistency, some discrepancies are prevalent and persistent, jeopardizing evidence-based decision-making. Our findings highlight the need to consult results systematically from both ClinicalTrials.gov and publications.

J Natl Compr Canc Netw 2017;15(11):1363–1371
doi: 10.6004/jnccn.2017.7002

Clinical trials are valuable in guiding evidence-based practice in medicine.¹ Oncologists' decision-making regarding therapeutic regimens generally depends on information from publications in peer-reviewed journals, the main channel through which trial results are publicly disclosed and communicated.² Thus, the reporting quality of publications is of vital importance to ensure accurate dissemination of evidence.^{3,4}

ClinicalTrials.gov is the largest publicly accessible trial registry, and the only one with a results database.^{5,6} In September 2007, the FDA Amendments Act (FDAAA; section 801) was passed mandating the timely reporting of results of applicable clinical trials to ClinicalTrials.gov,⁷ which greatly expanded the legal requirements for the public reporting of trial results and enhanced reporting transparency. In contrast to peer-

From the ^aDepartment of Radiation Oncology, and ^bClinical Trials Centre, Sun Yat-sen University Cancer Centre, State Key Laboratory of Oncology in South China, Collaborative Innovation Centre of Cancer Medicine, Guangzhou, People's Republic of China; and ^cDepartment of Radiation Oncology, University of Michigan, Ann Arbor, Michigan.

*These authors contributed equally.

Submitted April 20, 2017; accepted for publication June 30, 2017.

The authors have disclosed that they have no financial interests, arrangements, affiliations, or commercial interests with the manufacturers of any products discussed in this article or their competitors.

This work was supported by grants from the National Natural Science Foundation of China (No. 81372409), the Science and Technology Project

of Guangzhou City, China (No.132000507), the National Natural Science Foundation of China (No. 81402532), and the National Natural Science Foundation of China (No. 81572962).

Author contributions: Study concept and design: Lv, Chen, Sun. Data acquisition: Lv, Guo, Sun. Data analysis and interpretation: Lv, Chen, Zhou, Liu, Guo. Access to raw data: Lv, Sun. Creation of tables and figures: Lv, Chen, Zhou, Liu, Guo. Manuscript preparation: All authors. Critical revision: All authors.

Correspondence: Ying Sun, MD, PhD, Sun Yat-sen University Cancer Center, State Key Laboratory of Oncology in South China, Collaborative Innovation Center of Cancer Medicine, Department of Radiation Oncology, 651 Dongfeng Road East, Guangzhou 510060, People's Republic of China. E-mail: sunying@sysucc.org.cn

reviewed publications, which might be subject to the selective judgments of editors and reviewers, information posted on ClinicalTrials.gov goes through a quality assurance (QA) process when required information is missing or internally inconsistent.

Cancer is a major public health problem worldwide and is the leading and second-leading cause of death in China and the United States, respectively.^{8,9} The interpretation and accuracy of trial results are of particular concern in medical oncology, in which therapeutic regimens are often rapidly developed and prompt treatment decisions are important for saving lives. So how accurately does the published literature convey information to the oncologic community regarding the efficacy and safety of cancer drugs assessed in clinical trials?

Currently, only one study has investigated the reporting consistency between ClinicalTrials.gov result database and publications.¹⁰ However, trials included in the study were completed before January 1, 2009—2 years after the enactment of mandatory result reporting law—and only 5 trials (3%) related to oncology. The accuracy of published literature conveying information to the oncologic community on cancer drug trials remains unknown. Has reporting consistency improved in the 10 years since the mandatory reporting laws were enacted?

To address these questions, we included cancer drug trials with results posted on ClinicalTrials.gov and that were completed between 2004 and 2014. Our study had 2 objectives: to identify the degree of completeness and consistency of results reported between the ClinicalTrials.gov database and the subsequent publications, and to identify the trends of reporting quality and associated characteristics.

Methods

Data Source and Study Sample

Data were obtained through the Aggregate Analysis of ClinicalTrials.gov (AACT) database, reflecting data downloaded up until September 27, 2015. Among approximately 200,000 studies registered on ClinicalTrials.gov, we focused on clinical trials with results posted on the ClinicalTrials.gov results database (n=18,474). A total of 323 phase III/IV cancer drug trials with a randomized controlled design that posted results on ClinicalTrials.gov and were completed between January 1, 2004, and January 1, 2014,

were included in the final selection (Figure 1). The detailed selection process is shown in [supplemental eAppendix 1](#), available online with this article at JNCCN.org.

We then screened a 50% sample through random sampling by STATA, version 12 (StataCorp LP, College Station, TX) (n=160) and searched for matching publications. Trial characteristics were well-balanced between selected and unselected samples (Table 1); search strategies are detailed in [supplemental eAppendix 1](#). We chose the publication that first reported the primary outcome measures (POMs) for publication date records and reviewed it for completeness and consistency. Because no trials uploaded their results before the primary completion date, and the results posted on ClinicalTrials.gov were required to report on primary outcomes, we believed this method would largely help reduce bias. We found that some trials with publications only recruited a few participants; because small trials are less likely to influence clinical practice, we restricted our final comparison to trials with a minimum sample size of 60 participants (N=117; Figure 1).

Data Extraction and Criteria for Discrepancy

Information on the following 3 dimensions were extracted and compared: (1) basic design information, including study design, number of arms, and number of patients undergoing randomization; (2) efficacy measurements, including the number of POMs, descriptions, and measurements; timing of assessment; number of patients involved in efficacy analysis; and specific metrics; and (3) benefit/risk reports, including individuals affected by at least 1 serious adverse event (SAE) and other adverse event (OAE), risk difference (experimental arm vs control arm risk), and number of individuals at risk per group. For trials with multiple experimental arms, we selected the arm of primary interest stated in the registration. The criteria for discrepancy are stated in [supplemental eAppendix 1](#); specific items in each dimension with definitions and examples showing discrepant results are presented in [supplemental eTable 1](#).

Two investigators (J-W.L., X.L.) independently assessed the completeness and consistency of information between sources. The percentage of data that were discrepant for each item of comparison was calculated to show the disagreement between the 2

Discrepant Reporting in Cancer Trials

Figure 1. Flowchart showing the study selection process of clinical trials assessing cancer drugs.

investigators; percentages ranged from 0% to 6.8%, which was generally low. Any disagreement was solved by consensus, and a third reviewer (Y-P.C.) randomly rechecked a 50% sample for QA.

Scoring System and Statistical Analysis

We applied a scoring system to determine characteristics associated with reporting completeness and consistency (detailed in supplemental eAppendix 1). To compare the basic trial characteristics between selected and unselected samples, chi-square analysis was performed. The completeness of results posted on ClinicalTrials.gov and in publications was compared using McNemar’s test of equality of paired proportions.¹¹ To identify trial characteristics associated with reporting completeness and consistency, we used total score as the outcome variables using the linear regression model. The characteristics of each trial were obtained from the National Library of Medicine (definitions of trial characteristics

are presented in supplemental eAppendix 1). The multivariable model with backward elimination included every variable associated with $P \leq .10$ in univariate analysis. Variables significant at $P < .05$ in the final multivariate model were considered independent predictors. Collinearity among variables was identified by collinearity diagnostics. The significance level of the model was evaluated by F value. All analyses were performed using STATA, version 12. Two-sided $P < .05$ was considered significant.

Results

Trial Characteristics

Of the 50% random samples ($n=160$), 1 trial registered as phase IV was found to be phase II and another was found to be a noncancer trial. Thus, these trials were excluded ($n=2$). Of the remaining 158 trials, 121 (76.6%) had publications. Trial characteristics

Table 1. Comparison of the Basic Trial Characteristics			
Characteristic	Trials Unselected (N=163) n (%)	Selected Samples (N=160) n (%)	P Value ^a
Phase			.11
III	142 (87.1)	148 (92.5)	
IV	21 (12.9)	12 (7.5)	
Primary sponsor			.27
Industry	125 (76.7)	115 (71.9)	
NIH	7 (4.3)	4 (2.5)	
Other ^b	31 (19.0)	41 (25.6)	
Primary completion year			.62
Before 2007	19 (11.7)	18 (11.3)	
2008–2009	77 (47.2)	84 (52.4)	
2010–2013	67 (41.1)	58 (36.3)	
End point classification			.65
Safety & efficacy	103 (62.3)	103 (64.4)	
Efficacy only	58 (35.6)	53 (33.1)	
Safety only	2 (1.2)	4 (1.9)	
Interventional model			.35
Parallel assignment	158 (51.3)	150 (48.7)	
Crossover assignment	4 (30.8)	9 (69.2)	
Factorial assignment	1 (50.0)	1 (50.0)	
Masking			.49
Open-label	88 (54.0)	92 (57.5)	
Single-blind	0 (0)	1 (0.6)	
Double-blind	75 (46.0)	67 (41.9)	
Participant enrollment			.60
Median	330	317	
<150	44 (27.0)	47 (29.4)	
150–1,000	103 (63.2)	93 (58.1)	
>1,000	16 (9.8)	20 (12.5)	
Study location			.49
With US site	102 (62.6)	106 (66.3)	
Without US site	61 (37.4)	54 (33.7)	

^aP values were calculated using the chi-square or Fisher exact test if indicated.
^bOthers including government, academic institution, nonprofit, or nonfederal organizations.

Discrepant Reporting in Cancer Trials

and publication features are shown in [supplemental eTable 2](#).

Completeness of Reporting

After excluding trials without a minimum sample size of 60 participants, 117 of 121 trials entered into our final comparison. Table 2 compared the completeness of results reporting in the ClinicalTrials.gov results database and the matching publications. Reporting was significantly more complete on ClinicalTrials.gov than in publications for SAEs (100% vs 43.6%) and OAEs (100% vs 62.4%). No statistical significance was observed in basic design information (100% vs 100%) and efficacy measurements (92.3% vs 90.6%).

Consistency of Reporting

Table 3 summarized the major discrepancies among trials with both posted and published results for basic design information (n=117), efficacy measurements (n=98), SAEs (n=51), and OAEs (n=73). In basic design information, 16 of 117 trials (13.7%) indicated at least 1 discrepancy. Among these, 2 trials with discrepant study design were attributed to amended protocol during the trial processing; however, the crucial information was not updated in Clinical-

Trials.gov. Generally, published articles suggested broader study population for randomization compared with the information in the ClinicalTrials.gov results database. The median relative difference was 2.5% (range, 0.3%–22.8%).

For efficacy measurements, 86 of 98 (87.8%) trials reported at least 1 discrepancy. The most common discrepancy occurred with secondary outcome measurements (75/98; 76.5%). Of 18 trials with different numbers of POMs and measurement tools, 13 trials reported more POMs in ClinicalTrials.gov than in publications, 3 trials had the same number of POMs but referred to different measurement tools, and 2 trials reported more POMs in publications. A total of 18 trials differed in treatment effects (specific metrics), of which 2 trials were not comparable as they referred to different measurements. Among 16 comparable trials, 7 reported larger treatment effects in publications; 9 reported larger treatment effects in ClinicalTrials.gov (1 trial was noninferior). On an absolute scale, observed discrepancy did not change interpretation of results, except for 1 trial. We further investigated the alteration in POMs and the influence of selective reporting. Of the 18 accessible

Table 2. Completeness of Reporting in the ClinicalTrials.gov Results Database and the Matching Publications

Domains	Completeness Items	Complete Reporting in ClinicalTrials.gov, (N=117) n (%)	Complete Reporting in Publications, (N=117) n (%)	P Value
<i>Basic design information</i>				
	Total	117 (100)	117 (100)	
	Study design	117 (100)	117 (100)	1.00
	Study arms	117 (100)	117 (100)	1.00
	Randomized participants	117 (100)	117 (100)	1.00
<i>Efficacy measurements</i>				
	Total	108 (92.3)	106 (90.6)	
	POM numbers & measurements	117 (100)	117 (100)	1.00
	Timing of assessment	117 (100)	117 (100)	1.00
	Patients involved	116 (99.1)	117 (100)	1.00
	Specific metrics	117 (100.0)	114 (97.4)	.25
	SOM numbers & measurements	109 (93.2)	110 (94.0)	1.00
<i>SAE</i>				
	Total	117 (100)	51 (43.6)	
	Individuals affected by at least one SAE	117 (100)	86 (73.5)	<.01
	Risk difference ^a	117 (100)	51 (43.6)	<.01
	Individuals at risk per group	117 (100)	51 (43.6)	<.01
<i>OAE</i>				
	Total	117 (100)	73 (62.4)	
	Individuals affected by at least one OAE	117 (100)	111 (94.9)	.03
	Risk difference ^a	117 (100)	73 (62.4)	<.01
	Individuals at risk per group	117 (100)	74 (63.2)	<.01

Abbreviations: OAE, other adverse event; POM, primary outcome measure; SAE, serious adverse event; SOM, secondary outcome measure.

^aRefers to experimental arm/control arm risks.

Table 3. Reporting Discrepancy Between the ClinicalTrials.gov Results Database and the Matching Publications

Domain	Item of Interest	Number of Trials With Discrepant Reporting	Percentage (%)
<i>Basic design information</i>			
	Total	16/117	13.7
	Study design	4/117	3.4
	Study arms	1/117	0.8
	Randomized participants	13/117	11.1
<i>Efficacy measurements</i>			
	Total	86/98	87.8
	POM numbers & measurements	18/98	18.4
	Timing of assessment	2/98	2.0
	Patients involved	11/98	11.2
	Specific metrics	18/98	18.4
	SOM numbers & measurements	75/98	76.5
<i>SAE</i>			
	Total	26/51	51.0
	Individuals affected by at least one SAE	2/51	3.9
	Risk difference ^a	7/51	13.7
	Individuals at risk per group	23/51	45.1
<i>OAE</i>			
	Total	54/73	74.0
	Individuals affected by at least one OAE	1/73	1.4
	Risk difference ^a	10/73	13.7
	Individuals at risk per group	54/73	74.0

Abbreviations: OAE, other adverse event; POM, primary outcome measure; SAE, serious adverse event; SOM, secondary outcome measure.
^aRefers to experimental arm/control arm risks.

trials with altered POMs, statistically significant results were favored in 11 trials (61.1%).

In SAE reporting, 26 of 51 trials (51.0%) reported at least 1 discrepancy. Among these, 20 of 23 trials reported fewer numbers of individuals at risk in publications (median relative difference, 7.1%; range, 0.4%–229.0%); 3 of 23 trials reported fewer numbers of individuals at risk on ClinicalTrials.gov (median relative difference, 45.2%; range, 25.0%–108.0%). It is worth noting that although relative/absolute differences were minor in some trials, due to the discrepancy between risk groups (experimental arm vs control arm) in 7 trials, the discrepant or even opposite interpretation of SAEs may occur.

In OAE reporting, 54 of 73 trials (74.0%) reported at least 1 discrepancy. Among these, 14 reported fewer numbers of individuals at risk in publications (median relative difference, 4.4%; range, 0.4%–220.9%) and 40 reported fewer numbers of individuals at risk on ClinicalTrials.gov (median relative difference, 5.1%; range, 0.8%–100.0%). Of note, different interpretations of OAEs may have occurred occur due to the discrepancy between risk groups (experimental arm vs control arm) in 10 tri-

als, although the relative/absolute differences were minor.

Multivariable Factors Associated With Reporting Completeness and Consistency

We then moved one step further to investigate the characteristics associated with reporting quality. The ClinicalTrials.gov results database was chosen as reference because of its greater completeness; trials with incomplete reporting on ClinicalTrials.gov were excluded (n=9). A total of 14 (see supplemental eTable 1) items compared in the previous discussion entered the quality scoring. For the 108 trials eligible for quality scoring, the median score was 21 (range, 14–28). Linear regression identified parallel assignment, phase IV trials, primary funding by industry (vs other funding, but not NIH funding), primary completion after 2009, and earlier results posted after primary completion as independent factors associated with greater completeness and consistency (Table 4). The significance level was indicated by *F* value, which equals 5.28 (*P*<.001). Collinearity diagnostics showed no evidence of collinearity among the variables (Tolerance_{all}>0.9). Greater completeness and consistency did not favor statistically significant primary outcomes (*P*=.21).

Discrepant Reporting in Cancer Trials

Table 4. Trial Characteristics Associated With Completeness and Consistency of Reporting (N=108)^a

Publication Characteristic	Univariate Analyses		Multivariable Analyses	
	Estimate (95% CI) ^b	P Value	Estimate (95% CI) ^b	P Value
Interventional model				
Crossover assignment	Ref		Ref	
Parallel assignment	3.68 (0.85 to 6.50)	.01	4.63 (1.95 to 7.32)	<.01
Primary purpose			Not included	
Treatment	Ref			
Prevention	1.14 (−0.95 to 3.12)	.27		
Others ^c	−1.70 (−3.91 to 0.66)	.16		
Masking			Not included	
Double-blind	Ref			
Open-label	−0.24 (−1.58 to 1.10)	.73		
Phase				
III	Ref		Ref	
IV	2.70 (0.43 to 5.36)	.05	2.99 (0.48 to 5.52)	.02
Sponsor				
Industry	Ref		Ref	
NIH	−0.44 (−4.46 to 3.58)	.83	0.08 (−3.67 to 3.83)	.97
Others ^d	−1.83 (−3.64 to −0.02)	.04	−1.92 (−3.65 to −0.19)	.03
Sample size			Not included	
<150	Ref			
150–1,000	0.40 (−1.62 to 2.43)	.69		
>1,000	0.71 (−1.77 to 3.19)	.57		
Journal impact factor ^e			Not included	
<10	Ref			
>10	0.19 (−1.19 to 1.58)	.78		
Primary outcomes			Not included	
Negative	Ref			
Positive	0.78 (−0.56 to 2.13)	.25		
Others ^f	2.67 (−1.43 to 6.76)	.20		
Time to results posted (mo)	−0.04 (−0.06 to −0.01)	.01	−0.03 (−0.06 to −0.01)	.03

^aTrials with incomplete reporting in the ClinicalTrials.gov results database were excluded (n=9).

^bThe estimates indicate the benefit observed compared with the reference. Any positive value indicates incremental benefit compared with the reference, whereas any negative value indicates detriment compared with the reference.

^cOthers including diagnostic, supportive care, and health services research.

^dOthers including government, academic institution, nonprofit, or nonfederal organizations.

^eAccording to Web of Science 2015 (formerly the Institute for Scientific Information Web of Knowledge).

^fOthers include primary outcomes could not be defined or primary outcomes with ≥1 positive and others negative.

Discussion

Discrepancy in reporting clinical trials has triggered widespread concerns. Previous literature has endeavored to explore the reporting discrepancy between publications and other relevant sources, including regulatory documents, clinical study reports, and registrations.^{12–16} However, these documents are often less available. Resorting to freely available descriptions of trial results, such as the ClinicalTrials.gov results database, is therefore more applicable and convenient for the public. With the FDAAA requiring mandatory posting of results within 1 year after the primary completion date and standardized reporting of results,⁷ ClinicalTrials.gov has become an interesting source of trial results. To date, only one study by Hartung

et al¹⁰ has investigated the reporting bias between the ClinicalTrials.gov results database and publications. However, trials included in that study were completed before January 1, 2009, 2 years after the enactment of mandatory result reporting law. The accuracy and trends of modern publications conveying information to the oncology community after long-term enactment of mandatory reporting law remain unknown. By including cancer drug trials with results posted on ClinicalTrials.gov and completed between 2004 and 2014, we found that the median score of reporting completeness and consistency was 21, indicating generally reasonable reporting quality. However, certain discrepancies are prevalent and persistent, and need to be addressed.

Overestimation and Selective Reporting in Efficacy Measurements

In POMs, we identified 18.4% trials with inconsistent reporting. This estimate was lower than other studies that explored inconsistencies between trial protocols and published results (62%),¹⁷ or trial registrations and publications (31%)¹⁸; however, it was much the same as the report by Hartung et al.¹⁰ It seems that the past years did not witness great improvements in consistent reporting of primary outcomes after the implementation of FDAAA. Additional efforts and tailored policy alongside with FDAAA section 801 should be made to improve reporting quality.

Overestimation and alteration of primary end points were the predominant discrepancy, which increase the prevalence of spurious results and give a false impression about cancer drugs. The purpose of reporting on primary outcomes is to define the most clinically relevant outcomes and protect against selective reporting.¹⁹ Additionally, primary outcomes are generally used for the calculation of sample size. However, if primary outcomes are subsequently omitted or altered, their protective mechanism may no longer function.

Incompleteness and Underestimation of Benefit/Risk Reports

Our study found that all trials posted complete AE reporting on ClinicalTrials.gov, whereas approximately 50% of the trials published complete AE reports in the literature. Similar to our findings, Riveros et al²⁰ pointed out that AE reporting was significantly more complete on ClinicalTrials.gov than in publications (73% vs 45%). The completeness rate for ClinicalTrials.gov in our study was higher. Possible reasons were that Riveros et al only included 4% of cancer drug trials, and the search was completed by 2012. The higher rates also reflected the inspiring work performed by the oncology community to the results posted on ClinicalTrials.gov. However, little improvement was seen in the publications. Although this was not surprising in light of word count limits imposed by journal editors,²⁰ it might also be attributed to the poorly measured benefit/harm events or purposeful concealment of unfavorable data.^{4,21} The findings underlined the need to consult ClinicalTrials.gov for more information on benefits/risks reported in trials.

Among trials with both posted and published AE reports, discrepant reporting was prevalent, with

most trials reporting fewer SAEs in publications. Underreporting of SAEs was of particular concern, because even if some differences did not alter the interpretation of safety issues, it may distort how oncologists balance the benefits and harms of cancer drugs. Moreover, these distortions may be amplified in systematic reviews and meta-analyses.^{22,23}

Improving the Reliability of Trial Results Reporting

We were pleased to see that trials with earlier results posted on ClinicalTrials.gov tended to have better reporting completeness and consistency. In addition, compared with trials completed before 2009, those completed in the subsequent 5 year period (2009–2014) showed much improvement in reporting quality. The improved reporting quality of later publications reflected the positive feedback for timely posting of results to ClinicalTrials.gov, and the supervisory function of the mandatory reporting law. Disappointedly, publishing in high-impact journals (impact factor >10) and trials with larger sample sizes (>1,000) did not guarantee better reporting conditions. Special attention should be paid to this phenomenon, because these trials are more likely to influence evidence-based clinical oncology practice.

To improve the reliability of results reporting in clinical trials, active participation of the various stakeholders is needed. First, our findings highlight a growing sense that both clinicians and peer reviewers should access trial results systematically from both ClinicalTrials.gov and the published literatures, when available. Consulting participant-level “raw data” for reference and resolving discrepancies between the 2 sources are the most crucial procedures to improve transparency and disclosure in clinical trials. Second, the mandatory result reporting law alone is still not enough, although a positive correlation was identified between timely result posting to ClinicalTrials.gov and better reporting consistency. Tailored policy is also needed, such as regarding QA and reporting timeline, alongside FDAAA section 801, which guarantees the quality of results posting to ClinicalTrials.gov and reduces discrepancy between the 2 sources.

Limitations

We note a number of limitations to our study. First, the ClinicalTrials.gov result database was used as

Discrepant Reporting in Cancer Trials

a reference for comparison and scoring. Although result reporting might be more complete and objective in ClinicalTrials.gov, the information remains suboptimal and invalid. Some discrepancies may be entry mistakes on ClinicalTrials.gov due to the urgency to report results or inexperience with the submission requirements. Second, we did not evaluate the changes of outcome measurements archived over time; only those reported in the results database were evaluated. Modifications of registered trial protocols, specifically secondary outcome measure additions or deletions, are common prior to publication. Third, the discrepancy in OAEs could be exaggerated possibly due to the different timelines for safety follow-up between primary publication and ClinicalTrials.gov. If there is a longer timeline for safety evaluation in the ClinicalTrials.gov entry, then it could explain why patients on ClinicalTrials.gov would be more likely to experience OAEs. Special attention should be paid to interpreting these findings.

Conclusions

Although results reporting of clinical trials assessing cancer drugs showed generally reasonable completeness and consistency, some discrepancies are prevalent and persistent, which jeopardizes evidence-based clinical decision-making. Making results publicly accessible on ClinicalTrials.gov may provide participant-level “raw data” as reference and help ameliorate reporting bias.

Acknowledgments

We would like to thank the staff members of the National Library of Medicine and NIH, and their colleagues across the United States, who have been involved with the development and maintenance of ClinicalTrials.gov. We thank the Clinical Trials Center, Sun Yat-sen University Cancer Center, for assistance in data interpretation.

References

1. Reith C, Landray M, Devereaux PJ, et al. Randomized clinical trials—removing unnecessary obstacles. *N Engl J Med* 2013;369:1061–1065.
2. Craig JC, Irwig LM, Stockler MR. Evidence-based medicine: useful tools for decision making. *Med J Aust* 2001;174:248–253.
3. Simes RJ. Publication bias: the case for an international registry of clinical trials. *J Clin Oncol* 1986;4:1529–1541.
4. Chan AW, Altman DG. Identifying outcome reporting bias in randomised trials on PubMed: review of publications and survey of authors. *BMJ* 2005;330:753.
5. Zarin DA, Ide NC, Tse T, et al. Issues in the registration of clinical trials. *JAMA* 2007;297:2112–2120.
6. Edge S, Byrd DR, Compton CC, et al, eds. *AJCC Cancer Staging Manual*, 7th ed. New York, NY: Springer; 2010.
7. Food and Drug Administration Amendments Act of 2007. US Public Law. 2007;110–185. Washington, DC: Food and Drug Administration.
8. Chen W, Zheng R, Baade PD, et al. Cancer statistics in China, 2015. *CA Cancer J Clin* 2016;66:115–132.
9. Chen W, Zheng R, Zeng H, Zhang S. The updated incidences and mortalities of major cancers in China, 2011. *Chin J Cancer* 2015;34:502–507.
10. Hartung DM, Zarin DA, Guise JM, et al. Reporting discrepancies between the ClinicalTrials.gov results database and peer-reviewed publications. *Ann Intern Med* 2014;160:477–483.
11. Gonen M. Sample size and power for McNemar's test with clustered data. *Stat Med* 2004;23:2283–2294.
12. Patterson R, Nuttall JR. An evaluation of the risk of biopsy in squamous carcinoma: a clinical experiment. *Am J Cancer* 1939;37:64–68.
13. Common Terminology Criteria for Adverse Events (CTCAE), version 4.0. Available at: https://evs.nci.nih.gov/ftp1/CTCAE/CTCAE_4.03_2010-06-14_QuickReference_8.5x11.pdf. Accessed March 1, 2016.
14. National Cancer Institute. Summary Staging Guide, SEER Program. Bethesda, MD: National Institute of Health; 1981. NIH Publication No. 81.
15. Turner EH, Matthews AM, Linardatos E, et al. Selective publication of antidepressant trials and its influence on apparent efficacy. *N Engl J Med* 2008;358:252–260.
16. Wieseler B, Kerekes MF, Vervoelgyi V, et al. Impact of document type on reporting quality of clinical drug trials: a comparison of registry reports, clinical study reports, and journal publications. *BMJ* 2012;344:d8141.
17. Chan AW, Hrobjartsson A, Haahr MT, et al. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA* 2004;291:2457–2465.
18. Mathieu S, Boutron I, Moher D, et al. Comparison of registered and published primary outcomes in randomized controlled trials. *JAMA* 2009;302:977–984.
19. Mills JL. Data torturing. *N Engl J Med* 1993;329:1196–1199.
20. Riveros C, Dechartres A, Perronneau E, et al. Timing and completeness of trial results posted at ClinicalTrials.gov and published in journals. *PLoS Med* 2013;10:e1001566; discussion e1001566.
21. Ioannidis JP, Lau J. Completeness of safety reporting in randomized trials: an evaluation of 7 medical areas. *JAMA* 2001;285:437–443.
22. Fu R, Selph S, McDonagh M, et al. Effectiveness and harms of recombinant human bone morphogenetic protein-2 in spine fusion: a systematic review and meta-analysis. *Ann Intern Med* 2013;158:890–902.
23. Carragee EJ, Hurwitz EL, Weiner BK. A critical review of recombinant human bone morphogenetic protein-2 trials in spinal surgery: emerging safety concerns and lessons learned. *Spine J* 2011;11:471–491.

See JNCCN.org for supplemental online content.