

# Reporting of Safety Results in Published Reports of Randomized Controlled Trials

Isabelle Pitrou, MD, MSc; Isabelle Boutron, MD, PhD; Nizar Ahmad, MD, MSc; Philippe Ravaud, MD, PhD

**Background:** Reports of clinical trials usually emphasize efficacy results, especially when results are statistically significant. Poor safety reporting can lead to misinterpretation and inadequate conclusions about the interventions assessed. Our aim was to describe the reporting of harm-related results from randomized controlled trials (RCTs).

**Methods:** We searched the MEDLINE database for reports of RCTs published from January 1, 2006, through January 1, 2007, in 6 general medical journals with a high impact factor. Data were extracted by use of a standardized form to appraise the presentation of safety results in text and tables.

**Results:** Adverse events were mentioned in 88.7% of the 133 reports. No information on severe adverse events and

withdrawal of patients owing to an adverse event was given in 27.1% and 47.4% of articles, respectively. Restrictions in the reporting of harm-related data were noted in 43 articles (32.3%) with a description of the most common adverse events only (n=17), severe adverse events only (n=16), statistically significant events only (n=5), and a combination of restrictions (n=5). The population considered for safety analysis was clearly reported in 65.6% of articles.

**Conclusion:** Our review reveals important heterogeneity and variability in the reporting of harm-related results in publications of RCTs.

*Arch Intern Med.* 2009;169(19):1756-1761

**T**HE REPORTING OF HARM IS AS important as the reporting of efficacy in publications of clinical trials. Both are essential in estimating the ratio of benefit to harm of medical interventions. However, harm is frequently insufficiently reported<sup>1-4</sup>: reports usually emphasize benefits, especially when results for efficacy are statistically significant. A previous review of randomized controlled trials (RCTs) published from 1997

drawal owing to severe adverse events. These works did not focus on the potential influence of safety reporting on readers' interpretation of the trial results. As an example, the reporting of the safety of rofecoxib in the VIGOR study was highly criticized.<sup>8,9</sup> Krumholz et al<sup>10</sup> stated that the presentation of results obscured the cardiovascular risk associated with rofecoxib by reporting the hazard of myocardial infarction as if naproxen was the intervention group (relative risk 0.2; range, 0.1-0.7) and was protecting against myocardial infarction. The authors did not report the absolute number of cardiovascular events, even though all other results were presented appropriately with rofecoxib as the intervention group.

Several guidelines aim to improve the quality of the reporting of safety in RCTs. The extension of the Consolidated Standards of Reporting Trials (CONSORT) statement for harm provides specific guidelines for reporting harm-related results of clinical trials.<sup>11-13</sup> Particularly, the extension cautions against common mistakes and gives examples of inadequate reporting of harm-related data (eg, reporting

*For editorial comment  
see page 1737*

through 1998 by Ioannidis and Lau<sup>1</sup> showed that safety reporting varied widely in reports of trials and was in general largely inadequate. These results were confirmed by reviews of specific medical or surgical areas.<sup>5-7</sup> However, these works focused mainly on the reporting of indispensable parameters, such as reporting harm with numbers instead of generic statements, reporting severe adverse events per study harm, or reporting patient with-

**Author Affiliations:**  
Département d'Epidémiologie, Biostatistique, et Recherche Clinique, Groupe Hospitalier Bichat-Claude Bernard (AP-HP); Université Denis Diderot; INSERM U738, Paris, France.

qualitative instead of quantitative data, global reporting instead of reporting per study arm).

Our aim was to appraise the reporting of harm in reports of RCTs published in general medical journals with high impact factors and to focus particularly on the presentation of harm-related results. For this purpose, we systematically reviewed the reporting of safety data in a sample of recent results in reports of RCTs published in 6 core medical journals.

---

## METHODS

---

### LITERATURE SEARCH

We searched MEDLINE via PubMed to identify all reports of RCTs published from January 1, 2006, through January 1, 2007, in 6 general medical journals with a high impact factor: *New England Journal of Medicine (NEJM)*, *Lancet*, *JAMA*, *BMJ*, *Annals of Internal Medicine (Ann Intern Med)*, and *PLoS Medicine (PLoS Med)*. We restricted our search to these journals because a high impact factor is usually considered a good predictor of quality in methodology and reporting of results.<sup>14</sup>

### STUDY SELECTION

Retrieved articles were assessed by 1 reviewer (I.P.) who screened all titles and abstracts to identify the relevant studies. Articles were included if the study was identified as an RCT with 2 parallel arms. We excluded observational studies (ie, cohort and case-control studies), case reports, editorials, letters, and RCTs with the following designs: factorial plan, crossover, cluster RCT, multiple-arm trial, and equivalence and noninferiority trial. Articles for which only an abstract was available were excluded. We also excluded reports of RCTs if safety was not a major issue (eg, public health intervention, assessment of diagnostic or prognostic tests and screening procedures, and medico-economic analysis of RCTs). We excluded RCTs if safety and efficacy outcomes were identical (eg, if the primary outcome was “death” in the assessment of stents or revascularization in cardiology).

### DATA EXTRACTION

A standardized data extraction form was developed on the basis of a review of the literature<sup>1</sup> and the CONSORT statement guidelines.<sup>13</sup> (The data extraction form is available from the authors on request.) Before data extraction, as a calibration exercise, the standardized form was tested on a separate set of 20 reports. A single reviewer (I.P.) extracted data from all articles on the basis of their full text. Supplementary material was retrieved if available. The reviewer was not blinded to the journal or the authors' names. In addition, a random sample of 30 articles was reviewed by one of us (N.A.) for quality assurance.

### CHARACTERISTICS OF THE REPORTS

Data were collected on general characteristics of reports, including journal, country of first author, funding source (public, private, or both), medical area of the trial (eg, cardiology, neurology), intervention classification (pharmacological or non-pharmacological treatment), number of centers, and sample size (ie, number of subjects randomized).

## HARM-RELATED DATA

For each report, we checked the title for the term “safety” or a term related to safety (eg, toxicity, or adverse effects or events). We determined whether the following indispensable safety parameters identified by Ioannidis and Lau<sup>1</sup> were reported: (1) harm with numerical data for each trial group instead of generic statements, (2) severity grades, (3) severe adverse events per trial group, and (4) patient withdrawal owing to adverse events per trial group with the description of the events.

In a second step, we focused on the results section in each report (text, figures, and tables). We determined whether restrictions were applied in reporting harm-related data, such as reporting only the adverse events observed at a certain frequency or rate (eg, >3% or >10% of participants), reporting adverse events that reached a *P* value threshold in the comparison of randomized groups (eg, *P* < .05) or reporting only the most severe events. If reporting was restricted to common events only, we collected the threshold values (eg, adverse events occurring in >20% patients). We also checked whether (1) safety data were described per event (eg, number of cases of unstable angina) or per patient (eg, number of patients with an episode of unstable angina), (2) data for different adverse events were combined per organ (eg, neurological, digestive events) into 1 composite outcome measure, (3) information on timing of adverse events was given, and (4) expected and unexpected adverse events were reported separately. We counted the total number of statistical tests reported for safety analysis and noted whether statistical comparisons were reported for each adverse event per arm, the totality of adverse events per arm, for patients with at least 1 adverse event, for organs or for severe adverse events.

In a third step, we focused on harm-related results reported in tables and figures. Tables and figures are a powerful means of conveying investigational results.<sup>15</sup> We counted the number of figures and tables allocated to safety reporting. We checked whether safety data in tables or figures were displayed per event or per patient, whether and how many statistical comparisons were displayed with a *P* value, and whether the safety analyses reported in the tables and figures were intention-to-treat analyses (ie, data for all patients randomized were analyzed for the allocated groups).

### STATISTICAL ANALYSIS

Statistical analysis involved use of SAS statistical software (version 9.1; SAS Institute Inc, Cary, North Carolina). Descriptive statistics (mean [SD], median and interquartile range [IQR], extreme values) were used for continuous variables. Categorical variables were described with frequencies and percentages. We determined the degree of agreement between reviewers by the  $\kappa$  coefficient for categorical variables or the intraclass correlation coefficient for continuous variables.

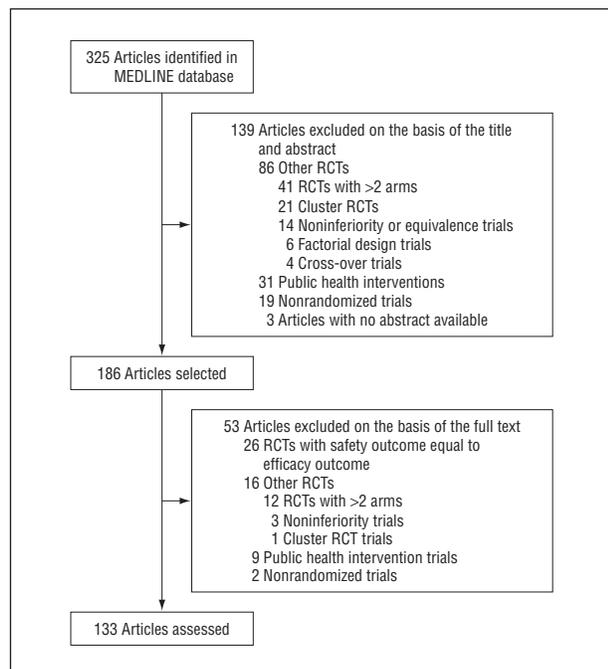
---

## RESULTS

---

### STUDY SELECTION AND CHARACTERISTICS OF SELECTED REPORTS

The electronic search identified 325 citations. From this list, we selected 186 articles after screening titles and abstracts (**Figure**). Finally, 133 articles were reviewed on the basis of their full text. (See the eAppendix [<http://www.archinternmed.com>] for the list of reports reviewed.)



**Figure.** Flowchart of screening reports of randomized controlled trials (RCTs) describing harm in medical journals with high impact factors.

**Table 1** describes the main characteristics of trials reviewed. Nearly half of the reports (n=59 [44.4%]) were published in *NEJM*, 26 (19.5%) in *Lancet*, 21 (15.8%) in *JAMA*, 19 (14.3%) in *BMJ*, and 8 (6.0%) in *Ann Intern Med*. No report reviewed was published in *PLoS Med*. Most reports reviewed assessed pharmacological interventions (n=110 [82.7%]). The trial was multicentric in 96 (72.2%) of the reports, and the median [IQR] sample size was 462 [185-1001]. The funding source was partially or completely for-profit funding in 73 reports (54.9%).

#### SAFETY REPORTING IN TEXTS

**Table 2** describes the main characteristics of the reporting of harm-related results. Interobserver reproducibility was good for the main items assessed, with  $\kappa$  coefficients of 0.57 to 0.90 (for 3 main items the  $\kappa$  coefficients was not calculated owing to the  $\kappa$  paradox) and interrater agreement of 0.80 to 0.97.

Adverse events were mentioned in the abstract of 95 articles (71.4%; range, 47.4% for *BMJ* to 84.8% for *NEJM*). Adverse events were reported in 118 articles (88.7%), and numerical data were provided in 112 (84.2%; range, 57.9% for *BMJ* to 91.5% for *NEJM*).

No information on severity of adverse events was given in 36 articles (27.1%), and 16 (12.0%) reported only generic statements. Severity grades were described in 21 reports (15.8%). The name of the scale used was given in 18 of these. The main scale used was the Common Terminology Criteria for Adverse Events toxicity scale (n=10). In 63 reports (47.4%), no data were given on withdrawals owing to adverse events. The description of adverse events leading to withdrawals was given in only 17 articles (12.8%).

**Table 1. Description of Randomized Controlled Trials (RCTs) With Reports Published in Medical Journals With High Impact Factors**

Characteristics of Trials	All Reports (n=133)
Journal <sup>a</sup>	
<i>NEJM</i>	59 (44.4)
<i>Lancet</i>	26 (19.5)
<i>JAMA</i>	21 (15.8)
<i>BMJ</i>	19 (14.3)
<i>Ann Intern Med</i>	8 (6.0)
<i>PLoS Med</i>	0
Location of country of first author <sup>a</sup>	
In North America	64 (48.1)
In Europe	47 (35.3)
Other	22 (16.5)
Funding <sup>a</sup>	
Public	51 (38.3)
Private	40 (30.1)
Both	33 (24.8)
Not reported	9 (6.8)
Medical area of RCT <sup>a</sup>	
Cardiovascular diseases	21 (15.8)
Pediatrics	19 (14.3)
Oncology/hematology	16 (12.0)
Neurology/psychiatry	14 (10.5)
Infectious diseases	12 (9.0)
Nephrology/urology	9 (6.8)
Musculoskeletal diseases/orthopedics	8 (6.0)
Gynecology/obstetrics	8 (6.0)
Respiratory medicine	8 (6.0)
Emergency medicine	5 (3.8)
Endocrine disorders/nutrition	5 (3.8)
Gastroenterology/hepatology	5 (3.8)
Dermatology	2 (1.5)
Ear, nose, throat/ophthalmology/dentistry	1 (0.7)
Experimental arm <sup>a</sup>	
Pharmacological treatment	110 (82.7)
Nonpharmacological treatment	21 (15.8)
Both	2 (1.5)
Control arm <sup>a</sup>	
Use of placebo	71 (53.4)
Pharmacological treatment	35 (26.3)
Nonpharmacological treatment	14 (10.5)
Usual care	11 (8.3)
Other	2 (1.5)
Multicentric <sup>a</sup>	96 (72.2)
Centers, No. <sup>b</sup>	24 (7.0-65.0)
Patients randomized, No. <sup>b</sup>	462 (185-1001)

Abbreviations: *Ann Intern Med*, *Annals of Internal Medicine*; *NEJM*, *New England Journal of Medicine*; *PLoS Med*, *PLoS Medicine*.

<sup>a</sup>Data are expressed as number (%).

<sup>b</sup>Data are expressed as median (interquartile range).

In total, 43 articles (32.3%) exhibited some restrictions in the reporting of harm-related data. The adverse events reported were the most common adverse events only (n=17), severe adverse events only (n=16), statistically significant adverse events only (n=5), and a combination of restrictions (n=5). In reports that described only common adverse events, the median threshold rate was 5% (minimum=2%; maximum=20%). Safety results were reported at the level of the event (n=114 [85.7%]) or the patient (ie, patients with  $\geq 1$  adverse event; n=76 [57.1%]) or as a composite outcome combining adverse events per organ (n=32 [24.1%]).

**Table 2. Characteristics of Reporting of Harm-Related Data in Reports of Randomized Controlled Trials Published in Medical Journals With High Impact Factors<sup>a</sup>**

Characteristic	All Reports (n=133)	NEJM (n=59)	Lancet (n=26)	JAMA (n=21)	BMJ (n=19)	Ann Intern Med (n=8)
Indispensable safety parameters						
Reporting of adverse events <sup>b</sup>	118 (88.7)	55 (93.2)	24 (92.3)	20 (95.2)	12 (63.2)	7 (87.5)
Reporting of adverse events with numerical data <sup>b</sup>	112 (84.2)	54 (91.5)	21 (80.8)	19 (90.5)	11 (57.9)	7 (87.5)
No. of events reported (n=112) <sup>c</sup>	8.0 [5-16]	10.5 [5-18]	7.0 [4-10]	7.0 [5-13]	8.0 [2-18]	13.0 [6-27]
Reporting of severity (n=97) <sup>b</sup>	97 (72.9)	52 (88.1)	17 (65.4)	13 (61.9)	10 (52.6)	5 (62.5)
Numerical data <sup>d</sup>	81 (83.5)	45 (86.5)	14 (82.4)	10 (76.9)	8 (80.0)	4 (80.0)
Reporting per arm <sup>d</sup>	76 (78.4)	43 (82.7)	13 (76.5)	11 (84.6)	5 (50.0)	4 (80.0)
Use of severity grades <sup>d</sup>	21 (21.6)	13 (25.0)	3 (17.7)	2 (15.4)	1 (10.0)	2 (40.0)
Reporting of withdrawals due to adverse events <sup>b</sup>	70 (52.6)	37 (62.7)	13 (50.0)	8 (38.1)	6 (31.6)	6 (75.0)
Restrictions in reporting of safety results (n=43) <sup>b</sup>	43 (32.3)	26 (44.1)	4 (15.4)	6 (28.6)	4 (21.0)	3 (37.5)
To common events	17 (39.5)	11 (42.3)	1 (25.0)	2 (33.3)	0	3 (100)
To severe events	16 (37.2)	9 (34.6)	2 (50.0)	3 (50.0)	2 (50.0)	0
To statistically significant results	5 (11.6)	2 (7.7)	0	1 (16.7)	2 (50.0)	0
To severe events and statistically significant results	2 (4.7)	2 (7.7)	0	0	0	0
To common events and statistically significant results	3 (7.0)	2 (7.7)	1 (25.0)	0	0	0
Reporting of safety results <sup>b</sup>						
Reporting of safety data displayed per event	114 (85.7)	54 (91.5)	23 (88.5)	19 (90.5)	11 (57.9)	7 (87.5)
Reporting of patients with ≥ 1 adverse event	76 (57.1)	38 (64.4)	15 (57.7)	11 (52.4)	7 (36.8)	5 (62.5)
Combining different adverse events per organ into 1 composite outcome	32 (24.1)	19 (32.2)	5 (19.2)	5 (23.8)	2 (10.5)	1 (12.5)
Reporting of timing information	42 (31.6)	19 (33.2)	6 (23.1)	12 (57.1)	4 (21.0)	1 (12.5)
Reporting of expected and unexpected events separately	4 (3.0)	0	1 (3.6)	1 (4.8)	1 (5.3)	1 (12.5)
Reporting of any statistical tests in safety analysis	63 (47.4)	35 (59.3)	8 (30.8)	11 (52.4)	5 (26.3)	4 (50.0)
Reporting of statistical tests for each event per arm	37 (27.8)	23 (39.0)	5 (19.2)	8 (38.1)	1 (5.3)	0
Reporting of global statistical tests (n=51)	51 (38.3)	26 (44.1)	7 (26.9)	11 (52.4)	5 (26.3)	2 (25.0)
Test for the totality of adverse events	5 (9.8)	4 (15.4)	0	1 (9.0)	0	0
Test for patients with ≥1 adverse event	8 (15.7)	2 (7.7)	1 (14.3)	3 (27.3)	2 (40.0)	0
Test for severe adverse events	12 (23.5)	4 (15.4)	3 (42.8)	3 (27.3)	2 (40.0)	0
Test for organs	5 (9.8)	1 (3.8)	2 (28.6)	1 (9.0)	1 (20.0)	0
Combinations of different tests	21 (41.2)	15 (57.7)	1 (14.3)	3 (27.3)	0	2 (100)

Abbreviations: *Ann Intern Med*, *Annals of Internal Medicine*; *NEJM*, *New England Journal of Medicine*.

<sup>a</sup> *PLoS Medicine* was not included in this table because no report reviewed was published in this journal.

<sup>b</sup> Data are expressed as number (%).

<sup>c</sup> Data are expressed as median [interquartile range].

<sup>d</sup> Among those reporting severity of adverse events.

Sixty-three articles (47.4%) described the use of at least 1 statistical test to compare safety data among groups. The median number of statistical tests reported for comparison of safety data was 5 (IQR, 2-12; range, 1-75). Thirty-seven articles (27.8%) gave the statistical tests for each type of event described. Fifty-one articles (38.3%) described global statistical comparisons, for the total number of adverse events per group (n=5 [9.8%]), the number of patients with at least 1 adverse event per group (n=8 [15.7%]), the total number of severe adverse events per group (n=12 [23.5%]), and the number of adverse events combined by organ per group (n=5 [9.8%]).

### SAFETY REPORTING IN TABLES AND FIGURES

**Table 3** describes the main characteristics of tables and figures related to the reporting of harm-related results. In total, 43 articles (32.3%) contained no table or figure describing safety. Numerical data in tables were described at the level of events in 25 of 90 reports (27.8%), the patient in 31 of 90 (34.4%), and both in 8 (8.9%). No distinction between events and patients was discernable in 26 reports (28.9%). Among the 90 articles with at least 1 table or figure dedicated to safety, statistical tests

were given in 56 (62.2%). The population considered for safety analysis was clearly reported in tables of 59 of 90 reports (65.6%), in which 32 described the analysis as intention-to-treat.

### COMMENT

We assessed the reporting of harm-related results in reports of RCTs published in 6 core medical journals in 2006. These results highlight the finding that despite the publication of a CONSORT statement<sup>13</sup> extension for harm-related data, the reporting of harm remains inadequate: descriptions of adverse events with numerical data in each trial arm were missing in 18% of the reports, and information related to the severity of adverse events and the withdrawal of patients owing to adverse events was lacking in 27.1% and 47.4% reports, respectively. We also found important heterogeneity and variability in the reporting of safety results. About one-third of the reports restricted the description of adverse events to the most common events or to severe or statistically significant results. This reporting of safety data goes against the recommendations of the CONSORT statement<sup>13</sup> extension checklist for reporting harm. The restrictions in report-

**Table 3. Reporting of Harm-Related Results in Tables and Figures in Reports of Randomized Controlled Trials Published in Medical Journals With High Impact Factors<sup>a</sup>**

Characteristics of Tables and Figures	Total Reports (n=133)	NEJM (n=59)	Lancet (n=26)	JAMA (n=21)	BMJ (n=19)	Ann Intern Med (n=8)
≥1 Table or figure reporting safety data <sup>b</sup>	90 (67.7)	50 (84.7)	13 (50.0)	14 (66.7)	8 (42.1)	5 (62.5)
No. of tables and figures reporting safety data <sup>c</sup>	1.0 [1-2]	1.0 [1-2]	1.0 [1-1]	1.0 [1-4]	1.5 [1-2]	1.0 [1-2]
No. of tables and figures in report (total) <sup>c</sup>	6.0 [5-7]	6.0 [5-6]	6.0 [5-8]	7.0 [6-8]	6.0 [5-7]	6.0 [5-8]
Expression of data in tables (n=90)						
Frequency of events <sup>b,d</sup>	25 (27.8)	12 (24.0)	5 (38.5)	2 (14.3)	4 (50.0)	2 (40.0)
Frequency of patients <sup>b,d</sup>	31 (34.4)	22 (44.0)	1 (7.7)	4 (28.6)	3 (37.5)	1 (20.0)
Both <sup>b,d</sup>	8 (8.9)	6 (12.0)	1 (7.7)	0	1 (12.5)	0
Unclear <sup>b,d</sup>	26 (28.9)	10 (20.0)	6 (46.1)	8 (57.1)	0	2 (40.0)
Reporting of statistical tests (n=90) <sup>b,d</sup>	56 (62.2)	33 (66.0)	8 (61.5)	8 (57.1)	4 (50.0)	3 (60.0)
Population of safety analysis clearly identified (n=90) <sup>b,d</sup>	59 (65.6)	32 (64.0)	8 (61.5)	12 (85.7)	4 (50.0)	3 (60.0)
Intent-to-treat analysis (n=59) <sup>b</sup>	32 (54.2)	14 (43.8)	3 (37.5)	9 (75.0)	4 (100.0)	2 (66.7)

Abbreviations: *Ann Intern Med*, *Annals of Internal Medicine*; *NEJM*, *New England Journal of Medicine*.

<sup>a</sup> *PLoS Medicine* was not included in this table because no report reviewed was published in this journal.

<sup>b</sup> Data are expressed as number (percentage).

<sup>c</sup> Data are expressed as median [interquartile range].

<sup>d</sup> Among reports providing safety data in tables and figures.

ing harm-related results in published reports may obscure some important rare and potentially severe adverse events.

For statistical comparisons, a median of 5 statistical tests was reported for safety analyses. We observed important variability, ranging from 1 to 75, in the number of statistical tests reported. About one-third of the reports described at least 10 statistical tests. Randomized controlled trials are well known to have insufficient statistical power to assess safety outcomes. Tsang et al<sup>16</sup> showed that in a sample of RCTs, the power to detect a statistically significant difference in serious adverse events yielded values ranging from 0.07 to 0.37. Consequently, as previously highlighted by Ioannidis et al<sup>17(p299)</sup>: “We must no longer accept confusing lists of noncomparable percentages of adverse events for clinical or scientific purposes. These lists can needlessly alarm patients and physicians or invite dismissal of real medication hazards.” Jonville-Béra et al<sup>18</sup> also emphasized that statements about “no differences” in adverse events between groups on the basis of nonsignificant *P* values is almost always inappropriate. Composite safety outcomes may be a method to deal with this inadequate statistical power<sup>19</sup>; however, only 32 reports described adverse events according to the organs involved, and 5 gave statistical comparisons per organ.

Although when reporting safety, the type of study population is of great importance, and any exclusion of treated patients could influence the validity of harm-related results, the population considered for safety analyses was not clearly reported in one-third of the reports. To our knowledge, no previous review has described with this precision the reporting of harm-related results from clinical trials. We identified few reviews that assessed the general reporting of harm-related results in clinical trials. Ioannidis and Lau<sup>1</sup> assessed the completeness of safety reporting in RCTs published from 1997 through 1998. This review was restricted to 7 areas of drug therapy. Loke and Derry<sup>4</sup> assessed the reporting of adverse drug reactions in reports of RCTs in 7 core medical journals pub-

lished in 1997 before the publication of the CONSORT statement extension for harm-related data, whereas our review assessed reports published in 2006 after the publication of the extension. Ethgen et al<sup>7</sup> assessed the reporting of harm-related results but compared pharmacological and nonpharmacological interventions in terms of harm.

Our study has some limitations. We considered reports of RCTs published in 6 core medical journals with a high impact factor. We did not consider specialized medical journals or journals with lower impact factors. Consequently, the generalizability of our results is limited. We excluded specific designs, such as factorial plans, crossover designs, cluster RCTs, and multiple-arm trials, because results of such trials are difficult and complicated to interpret and report. Consequently our results cannot be generalized to studies of such designs. Finally, a single reviewer extracted all the data; however, the quality assurance procedure gave good reproducibility.

In conclusion, our results highlight the need for improving the reporting of harm-related results from clinical trials. Despite the CONSORT statement extension for harm-related data, efforts should still be made to describe safety results with accuracy in reports of RCTs and to standardize practices for reporting.

**Accepted for Publication:** February 2, 2009.

**Correspondence:** Philippe Ravaud, MD, PhD, Département d’Épidémiologie, Biostatistique, et Recherche Clinique, Groupe Hospitalier Bichat-Claude Bernard, 46 rue Henri Huchard, 75877 Paris, CEDEX 18, France (philippe.ravaud@bch.aphp.fr).

**Author Contributions:** All authors had full access to all of the data (including statistical reports and tables) in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. *Study concept and design:* Pitrou and Ravaud. *Acquisition of data:* Pitrou and Ahmad. *Analysis and interpretation of data:* Pitrou, Boutron, Ahmad, and Ravaud. *Drafting of the manuscript:* Pitrou, Ahmad, and Ravaud. *Critical revision of the manu-*

script for important intellectual content: Boutron and Ravaud. *Obtained funding: Ravaud. Administrative, technical, and material support: Ravaud. Study supervision: Ravaud.*

**Financial Disclosure:** None reported.

**Funding/Support:** Dr Pitrou was funded by a grant from the Ministry of Higher Education and Research, France. The researchers were independent from the funding source.

**Additional Information:** An eAppendix is available at <http://www.archinternmed.com>.

## REFERENCES

1. Ioannidis JP, Lau J. Completeness of safety reporting in randomized trials: an evaluation of 7 medical areas. *JAMA*. 2001;285(4):437-443.
2. Ioannidis JP, Contopoulos-Ioannidis DG. Reporting of safety data from randomized trials. *Lancet*. 1998;352(9142):1752-1753.
3. Ioannidis JP, Lau J. Improving safety reporting from randomised trials. *Drug Saf*. 2002;25(2):77-84.
4. Loke YK, Derry S. Reporting of adverse drug reactions in randomised controlled trials: a systematic survey. *BMC Clin Pharmacol*. 2001;1:3.
5. Papanikolaou PN, Churchill R, Wahlbeck K, Ioannidis JP. Safety reporting in randomized trials of mental health interventions. *Am J Psychiatry*. 2004;161(9):1692-1697.
6. Bryant J, Loveman E, Cave C, Chase D, Milne R. Endocrinology trial design: adverse event reporting in randomised controlled trials of recombinant human GH in GH-deficient adults. *J Endocrinol*. 2002;175(2):545-552.
7. Ethgen M, Boutron I, Baron G, Giraudeau B, Sibilia J, Ravaud P. Reporting of harm in randomized, controlled trials of nonpharmacologic treatment for rheumatic disease. *Ann Intern Med*. 2005;143(1):20-25.
8. Curfman GD, Morrissey S, Drazen JM. Expression of concern: Bombardier et al, "Comparison of upper gastrointestinal toxicity of rofecoxib and naproxen in patients with rheumatoid arthritis," *N Engl J Med* 2000;343:1520-8. *N Engl J Med*. 2005;353(26):2813-2814.
9. Curfman GD, Morrissey S, Drazen JM. Expression of concern reaffirmed. *N Engl J Med*. 2006;354(11):1193.
10. Krumholz HM, Ross JS, Presler AH, Egilman DS. What have we learnt from Vioxx? *BMJ*. 2007;334(7585):120-123.
11. Altman DG, Schulz KF, Moher D, et al; CONSORT GROUP (Consolidated Standards of Reporting Trials). The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med*. 2001;134(8):663-694.
12. Moher D, Schulz KF, Altman DG; CONSORT Group (Consolidated Standards of Reporting Trials). The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *Ann Intern Med*. 2001;134(8):657-662.
13. Ioannidis JP, Evans SJ, Gotzsche PC, et al; CONSORT Group. Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Ann Intern Med*. 2004;141(10):781-788.
14. Lee KP, Schotland M, Bacchetti P, Bero LA. Association of journal quality indicators with methodological quality of clinical research articles. *JAMA*. 2002;287(21):2805-2808.
15. Schriger DL, Sinha R, Schroter S, Liu PY, Altman DG. From submission to publication: a retrospective review of the tables and figures in a cohort of randomized controlled trials submitted to the *British Medical Journal*. *Ann Emerg Med*. 2006;48(6):750-756.
16. Tsang R, Colley L, Lynd LD. Inadequate statistical power to detect clinically significant differences in adverse event rates in randomized controlled trials. *J Clin Epidemiol*. 2009;62(6):609-616.
17. Ioannidis JP, Mulrow CD, Goodman SN. Adverse events: the more you search, the more you find. *Ann Intern Med*. 2006;144(4):298-300.
18. Jonville-Béra AP, Giraudeau B, Autret-Leca E. Reporting of drug tolerance in randomized clinical trials: when data conflict with authors' conclusions. *Ann Intern Med*. 2006;144(4):306-307.
19. Tugwell P, Judd MG, Fries JF, Singh G, Wells GA. Powering our way to the elusive side effect: a composite outcome "basket" of predefined designated endpoints in each organ system should be included in all controlled trials. *J Clin Epidemiol*. 2005;58(8):785-790.