Recommendations to improve adverse event reporting in clinical trial publications: a joint pharmaceutical industry/journal editor perspective

Neil Lineberry,[1,2] Jesse A Berlin,[3] Bernadette Mansi,[4] Susan Glasser,[5] Michael Berkwits,[6] Christian Klem,[7] Ananya Bhattacharya,[8] Leslie Citrome,[9] Robert Enck,[10] John Fletcher,[11] Daniel Haller,[12] Tai-Tsang Chen,[13] Christine Laine[14]

[1]Navigant Consulting, Boston, MA, USA
[2]Takeda Pharmaceuticals, Cambridge, MA, USA
[3]Johnson & Johnson, 1125 Trenton-Harbourton Road, PO Box 200, Mail Stop TE3-15, Titusville, NJ 08560, USA
[4]GlaxoSmithKline, Collegeville, PA, USA
[5]Janssen Research & Development, Raritan, NJ, USA
[6]*JAMA* and JAMA Network, Philadelphia, PA, USA
[7]AstraZeneca, Gaithersburg, MD, USA
[8]Bristol-Myers Squibb, Princeton, NJ, USA
[9]*International Journal of Clinical Practice*, Valhalla, NY, USA
[10]*American Journal of Hospice and Palliative Medicine*, Knoxville, TN, USA
[11]*Canadian Medical Association Journal*, Ottawa, ON, Canada
[12]*Gastrointestinal Cancer Research*, Philadelphia, PA, USA
[13]Bristol-Myers Squibb, Wallingford, CT, USA
[14]*Annals of Internal Medicine*, Philadelphia, PA, USA
Correspondence to: J A Berlin JBerlin@its.jnj.com

Additional material is published online only. To view please visit the journal online.

Cite this as: ***BMJ*** **2016;355:i5078**
http://dx.doi.org/10.1136/bmj.i5078

Accepted: 16 September 2016

Medical Publishing Insights & Practices (MPIP)—a partnership among pharmaceutical companies and the International Society for Medical Publication Professionals—aims to identify ways to improve transparency and credibility in publishing the results of industry sponsored research. This article provides guidance from MPIP on clinically relevant and more informative adverse event reporting, previously identified by journal editors as a significant unmet need to improve patient care and increase the credibility of industry sponsored publications. Our recommendations include highlighting adverse events of most relevance to practitioners and their patients, avoiding broad summary statements such as "generally safe" or "well tolerated," and including more detailed adverse event data (where appropriate) to offer additional clinically important insight. These recommendations complement the earlier recommendations in the Consolidated Standards of Reporting Trials (CONSORT) Harms Extension. Although developed for industry sponsored trials, the adoption of our recommendations would enhance adverse event reporting in clinical research publications regardless of the funding source and thereby facilitate clinical decision making.

SUMMARY POINTS

- Objective reporting of adverse event data within clinical trials publications could provide greater context and clarity for the application of trial results to daily clinical practice
- Conference and manuscript abstracts should include objective information on the incidence and type of clinically relevant adverse events instead of overly general statements such as "well tolerated"
- Clinically relevant adverse events should be identified and communicated with clarity around relevant clinical characteristics, such as severity, frequency, and timing, which could be more informative than incidence rates
- Adverse event reporting should include numerators and denominators for all events; formal statistical analyses should be used selectively, and post hoc analyses should be clearly identified

Balanced reporting of drug adverse events in medical publications provides important context for healthcare practitioners about the benefit-risk profile of drug interventions. In a recent publication, the authors express the relevant concern: "The extent of 'hidden' or 'missing' data prevents researchers, clinicians, and patients from gaining a full understanding of harm, and this may lead to incomplete or erroneous judgements on the perceived benefit to harm profile of an intervention."[1] In 2004, the CONSORT group (for Consolidated Standards of Reporting Trials) published minimum standards for improved harms reporting in response to variability and incomplete or uneven reporting of adverse events from clinical trials.[2] While adverse event reporting subsequently improved, the overall communication of adverse event data continues to be suboptimal.[3-9] For example, a 2013 review of 325 randomized clinical trials published between 2007 and 2011 described inadequate or uninformative reporting of adverse event collection and analysis methodology,[9] suggesting poor adherence to the CONSORT Harms Extension in that sizeable trial sample. Moreover, existing guidance in the CONSORT Harms Extension, while rigorous and broadly applicable across disease areas, can lack the level of specificity regarding clinically meaningful adverse events necessary for practical clinical application. Additional guidance could help authors better identify, communicate, and display clinically relevant adverse event information in ways that facilitate clinicians' benefit-risk assessments for shared treatment decisions with patients.

In 2010, Medical Publishing Insights & Practices (MPIP), a partnership among pharmaceutical companies and the International Society for Medical Publication Professionals (ISMPP), and journal editors held a roundtable meeting to identify ways to help close the credibility gap in industry sponsored clinical research.[10] Communicating drug adverse events in a more transparent and clinically meaningful manner

was highlighted as one of 10 key recommendations to improve industry sponsored publications. This article describes collaboration among MPIP co-sponsors, industry experts, and journal editors to develop specific recommendations to improve drug adverse event reporting in industry sponsored, clinical trial publications. We offer consensus recommendations (to complement the CONSORT recommendations for harms reporting) and examples of "best practices" from the published clinical trial literature to help authors and trial sponsors communicate drug adverse events in a more informative and clinically meaningful manner.

Methods

Definitions

We used the World Health Organization (WHO) definition of adverse event: "Any untoward medical occurrence in a patient or clinical investigation subject administered a pharmaceutical product and which does not necessarily have to have a causal relationship with this treatment."[11] We use the terms "healthcare practitioners" or "practitioners" to refer to any licensed individual who can prescribe medical treatment.

In-depth interviews

From January to March 2014, one of the authors (NL) conducted 28 in-depth phone interviews. Each interview lasted 60 minutes with three key groups involved with reporting drug adverse events in published manuscripts from clinical trials:

- Industry experts (n=18) who worked for a pharmaceutical or biotech company in a role that involves adverse event data collection, management, or reporting
- Journal editors (n=6) who worked in an editorial capacity for a journal that publishes clinical trials
- Clinical investigators (n=4) who were authors of clinical trial papers.

Interview topics discussed were drawn from a predefined script determined by MPIP co-sponsors and relevant follow-up questions determined by NL during the discussion:

- How are adverse event results currently reported in clinical trial publications?
- What policies, guidelines, or best practices exist and how are they followed?
- What are current challenges with adverse event data reporting in clinical trial publications?
- Where do existing guidelines have gaps related to these challenges?
- What are potential solutions or recommendations to address these challenges?

NL collected the responses. The MPIP Steering Committee then aggregated, summarized, and prioritized these research findings by areas of focus using a consensus process as the basis for a roundtable discussion with journal editors and industry experts. The Steering Committee comprised publication professionals (individuals employed by MPIP sponsoring companies to organize and disseminate scientific and clinical data through peer reviewed publications) with representation from ISMPP.

Recommendation development

Research results were discussed at a roundtable meeting hosted by MPIP (New York, USA, May 2014) with nine US based journal editors, 15 MPIP Steering Committee members, and two industry experts, some of whom had participated in the research. A detailed summary of the interview results was shared with participants along with draft recommendations to structure the conversation. The group engaged in open discussion in an all-day session that identified key areas for improvement in adverse event reporting and modified the proposed recommendations in a collaborative fashion. Decisions were not finalized until group wide consensus was reached. Following the roundtable meeting, the MPIP Steering Committee further refined these recommendations through multiple follow-up discussions with a subset of roundtable participants (six journal editors and two industry experts) who comprise the authors on this paper.

Results

"Best practice" recommendations (based on our research and consensus process) for reporting adverse events from industry sponsored trials in publications are summarized in table 1 with detailed examples. The CONSORT requirements for harms reporting[2] are included in table 1 to illustrate the complementary nature of our recommendations.

Recommendation 1—Identify and communicate the most clinically relevant drug adverse event data as part of a comprehensive safety profile

Clinical trial publications should include the most clinically relevant, representative drug adverse event data while providing balance and context in interpretation of those data to help practitioners assess the benefit-risk profile for the intervention. Adverse event measures that are always clinically relevant and that should always be reported are:

- Deaths
- Serious adverse events as defined by the US Food and Drug Administration (FDA; adverse events that—in the view of the investigator or trial sponsor—were fatal or life threatening, or resulted in inpatient hospital admission or prolongation of existing hospital stay, persistent or significant incapacity or substantial disruption in the patient's ability to perform normal life functions, or a congenital anomaly or birth defect)[12]
- Adverse events that led to discontinuation of trial agent.

Other adverse event measures might be of particular interest based on the disease(s) under investigation, comorbidities of the study population, intervention mechanism, trial duration, or other considerations. A general best practice is to specify adverse events of interest in the planning of a clinical trial (eg, in the clinical trial protocol) on the basis of the mechanism of

Table 1 | MPIP recommendations for adverse event reporting, in parallel to the CONSORT Extension for Harms

MPIP recommendations to improve adverse event reporting for industry sponsored clinical trial manuscripts	CONSORT Extension for Harms checklist[2] (numbers refer to item numbers in the standard CONSORT checklist)
Methods and results	
[Methods and results] 1. Identify and communicate the most clinically relevant* drug adverse event data as part of a comprehensive safety profile. We recommend authors develop a "clinical relevance" filter to identify adverse events of greatest clinical interest given these considerations, and clearly state the rationale for focusing on these in the publication's methods section with references to evidence when available.	[Methods] "(6) List addressed adverse events with definitions for each (with attention, when relevant, to grading, expected vs. unexpected events, reference to standardized and validated definitions, and description of new definitions). Clarify how harms-related information was collected (mode of data collection, timing, attribution methods, intensity of ascertainment, and harms-related monitoring and stopping rules, if pertinent)."
[Methods and results] 2. Report timing, frequency, duration, and other potentially relevant descriptors when clinically appropriate. Authors should also report whether adverse events were collected in a non-elicited (passively collected) or elicited (proactively collected) fashion, adding explicit detail about the data collection methodology for both types in the methods section.	[Methods] "(12) Describe plans for presenting and analyzing information on harms (including coding, handling of recurrent events, specification of timing issues, handling of continuous measures, and any statistical analyses)."
[Methods and results] 3. Use statistical analysis for clinically relevant adverse events (where appropriate). Numerators and denominators should be specified. Overall, we propose that formal statistical comparisons when reporting adverse events should be selectively used where there is clear justification. We recommend utilizing an approach that determines the appropriateness of statistical analyses via a tiered system that assumes it is important to report information about all events but only some events require formal hypothesis testing.	[Results] "(13) Describe for each arm the participant withdrawals that are due to harms and their experiences with the allocated treatment." [Results] "(16) Provide the denominators for analyses on harms." [Results] "(17-19) Present the absolute risk per arm and per adverse event type, grade, and seriousness, and present appropriate metrics for recurrent events, continuous variables, and scale variables, whenever pertinent." "Describe any subgroup analyses and exploratory analyses for harms."
Abstract and introduction	
[Abstract] 4. Avoid use of overly general text descriptions for adverse events, including in abstracts. Abstracts should include a phrase or sentence summarizing the most clinically relevant adverse event data with frequency percentages consistent with those presented in the main text of the publication.	[Abstract and introduction] "(1, 2) If the study collected data on harms and benefits, the title or abstract, and introduction, should so state."
Discussion	
[Discussion] 5. Discuss adverse events findings in the broader context of available evidence and maintain consistency of data across different public reports.	[Discussion] "(20-22) Provide a balanced discussion of benefits and harms with emphasis on study limitations, generalizability, and other sources of information on harms."

*Adverse events that are always clinically relevant and should always be reported include deaths; serious adverse events as defined by the US Food and Drug Administration—that is, adverse events that (in the view of the investigator or trial sponsor) were fatal or life threatening, or resulted in inpatient hospital admission or prolongation of existing hospital admission, persistent or significant incapacity or substantial disruption in the patient's ability to perform normal life functions, or a congenital anomaly/birth defect; and adverse events resulting in discontinuation of trial agent(s) (as specified in the clinical trial protocol).

action and prior clinical experience with the experimental agent, and to faithfully report all such harms with explanations for any deviations (eg, data not interpretable because of data collection issues). We strongly encourage submission of the clinical trial protocol (inclusive of information pertaining to definitions of adverse events of interest, and methods for monitoring and data collection) together with the publication, as is already mandated by several journals. Such clinical trial protocols should also be available publicly through the journal (via a weblink or on request) after the associated manuscript is published.

In addition, we recommend authors develop a "clinical relevance" filter to identify adverse events of greatest clinical interest based on their clinical experience (eg, those that are typically seen with that drug class [especially those that are serious], or affect patients' quality of life, or adversely affect adherence to prescribed treatment), and clearly state the rationale for focusing on these in the publication's methods section with references to evidence when available. For example, tardive dyskinesia would be an important outcome to assess in trials of psychoactive interventions, but it could be less relevant for trials of other agents or classes if there is little potential pharmacological or off-target receptor relation with the disease being studied. In addition, adverse events typically classified as "mild" or "nuisance" in some situations, such as nausea or vomiting, might be highly relevant if they lead to changes in dosing regimens that could affect outcomes (eg, oncology treatments).

Information about the absence of common adverse events typically seen for a class of drugs can be equally informative, especially for producing meta-analyses with other trials at a later date. Authors should also report in the methods section whether adverse events were collected in a non-elicited (passively collected) or elicited (proactively collected) fashion, adding explicit detail about the data collection methodology for both types (or refer the reader to the trial protocol if submitted together with the publication). It should be noted that our discussions did not result in any recommendations around the attribution of adverse events to the study intervention by clinical investigators. Given the inherent subjectivity in such attribution, it has limited value in the context of randomized, double blind clinical trials and was considered less important than the other adverse event reporting recommendations summarized here.

The intent of the "clinical relevance" recommendation is to broaden adverse event reporting beyond what is mandated by regulators and to leverage the clinical experience and expertise of physician investigators to judge which adverse events should be highlighted. Performing a systematic review of the published literature pertaining to an investigational agent can also help investigators to determine clinical relevance by taking into account previously reported adverse events with that agent. The "clinical relevance" filter should be applied against a background of comprehensive reporting of adverse events in the body, data tables, and supplemental section of the published paper. The tables can take many forms, such as highlighting both common adverse events using a traditional incidence threshold and events of particular clinical interest (table 2).[13] Adverse events can be grouped into clinically

Table 2 | Example of table presenting adverse event data (Recommendation 1).[13] This example includes both common adverse events along with adverse events of interest for the disease and intervention displayed within the same table. Numerators, denominators, and percentages are included for all measurements. Reproduced from reference 13, Copyright (2013), with permission from Elsevier

Adverse events	Placebo (n=204)	Daclizumab HYP 150 mg (n=208)	Daclizumab HYP 300 mg (n=209)
Summary of adverse events			
Any adverse event	161 (79%)	151 (73%)	159 (76%)
Any serious adverse event	53 (26%)	32 (15%)	36 (17%)
Any serious adverse event, including multiple sclerosis relapse	12 (6%)	15 (7%)	19 (9%)
Death	0	1 (<1%)	0
Common adverse events that took place in >5% of patients given daclizumab HYP			
MS relapse	77 (38%)	47 (23%)	42 (20%)
Nasopharyngitis	31 (15%)	30 (14%)	30 (14%)
Headache	21 (10%)	20 (10%)	20 (10%)
Upper respiratory-tract infection	14 (7%)	18 (9%)	22 (11%)
Pharyngitis	9 (4%)	13 (6%)	13 (6%)
Oral herpes	10 (5%)	10 (5%)	13 (6%)
Rash	6 (3%)	12 (6%)	11 (5%)
Adverse events of interest			
Infections	89 (44%)	104 (50%)	112 (54%)
Serious infections	0	6 (3%)	3 (1%)
Cutaneous events	27 (13%)	38 (18%)	45 (22%)
Serious cutaneous events	0	2 (<1%)	3 (<1%)
Other potential immune-mediated serious adverse events			
Autoimmune thyroiditis	0	0	1 (<1%)
Crohn's disease	0	0	1 (<1%)
Hypersensitivity	0	0	1 (<1%)
Lymphadenopathy	0	0	1 (<1%)
Incidence of ALT or AST abnormalities			
1-3×ULN	64 (31%)	54 (26%)	62 (30%)
3-5×ULN	6 (3%)	7 (3%)	6 (3%)
>5×ULN	1 (<1%)	9 (4%)	8 (4%)
Injection-site reactions, erythema	3 (1%)	4 (2%)	4 (2%)
Malignancy	1 (<1%)	1 (<1%)	2 (<1%)

Data are number (%). ALT=alanine aminotransferase; AST=aspartate transaminase; HYP=high yield process; ULN=upper limit of normal.

appropriate categories, such as by organ system, as long as the rationale is clearly defined.

When comprehensive reporting is not possible (eg, due to space constraints), the publication should emphasize adverse events of highest clinical relevance given the particular treatment or study population. Access to the complete adverse event dataset should be made alongside the publication through data sharing websites or portals that protect participant privacy without placing undue burdens on requestors.[14 15]

Recommendation 2—Report timing, frequency, duration, and other potentially relevant descriptors when clinically appropriate

Other adverse event measurements can help communicate a more clinically relevant safety profile (box 1).

Box 1: Other meaningful adverse event measurements to build a more complete adverse event profile (Recommendation 2)

Adverse event measurements that should be considered for clinical trial publications to help provide additional relevant clinical information based on the judgment of the authors:

- Timing, frequency, and duration of adverse events
- Exposure-time associations
- Dose modifications
- Adherence to study treatments
- Relevant laboratory measurements

Adverse events affecting treatment administration, adherence, and quality of life should be expressed not only in terms of severity and frequency but also in terms of duration. For example, those adverse events seen with newer classes of drugs causing chronic, low grade fatigue and other symptoms could be cumulatively disabling and, therefore, clinically relevant. The success or inadequacy of protocol specified strategies for mitigation of adverse events that are known to be intervention related (eg, interruption in dosing or dose reduction) should be reported whenever possible.

Information about timing and frequency of adverse events is also important for determining the clinical implications of an adverse event (table 3),[16] given that aggregated metrics often fail to state when an adverse event was observed, how many times it was reported in individual patients, or (where measurement is feasible) how long it lasted. Typically, clinical trial publications include a summary table of adverse event incidence, where the incidences often are reported singly irrespective of whether that patient had one or more given events. This focus allows broad coverage of the complete set of adverse events seen in that trial, including those of a more serious nature (such as myocardial infarction, stroke, or organ failure) where the fact that these occurred once is important for clinicians to know.

However, for other types of adverse events (especially those that are generally considered non-serious), there

Table 3 | Example of table presenting timing, frequency, and duration of adverse events (Recommendation 2).[16] This adverse events table highlights not only the total adverse events but also when they occurred and their duration until resolution. An additional measure to consider would be how often a patient experienced the same adverse event. Additional displays of duration, providing more information about the distribution of duration (not just mean and standard deviation (SD)), should also be considered. Reproduced from reference 16, Copyright 2002, with permission from Elsevier (original table title: "Numbers of subjects (n=22, *f*=12) reporting individual adverse reactions including time of onset (h) and resolution time (h)")

	Number of subjects affected			Time (h) of onset (mean ± SD)		Resolution time (h) (mean ± SD)	
Adverse event	Males	Females	Total (%)	Males	Females	Males	Females
Vertigo/dizziness	9 (6)*	12 (10)	96 (73)	19 ± 10	14 ± 11	171 ± 132	284 ± 170
Nausea/vomiting	7 (3)	11 (9)	82 (55)	28 ± 20	22 ± 19	65 ± 49	190 ± 152
Headache	5 (1)	11 (2)	73 (14)	19 ± 9	37 ± 29	53 ± 24	85 ± 53
Insomnia	5 (2)	8 (5)	59 (32)	37 ± 32	30 ± 16	182 ± 39	93 ± 47
Anxiety	0	6 (1)	27 (5)	0	28 ± 9	0	36 ± 12
Depression	2 (0)	3 (0)	23	84 ± 12	72 ± 52	48 ± 24	32 ± 11
Confusion	2 (0)	1 (1)	14 (5)	24 ± 0	24 ± 0	60 ± 12	120 ± 0
Hallucination	0	1 (0)	5	0	24 ± 0	0	72 ± 0
Diarrhea	3 (1)	6 (3)	41 (18)	24 ± 0	38 ± 9	40 ± 23	60 ± 36
Abdominal pain	2 (1)	7 (2)	41 (14)	24 ± 0	31 ± 21	48 ± 24	62 ± 34
Palpitations	1 (0)	6 (0)	32	72 ± 0	41 ± 17	120 ± 0	52 ± 32
Pruritus	0	0	0	0	0	0	0

*In brackets: grade 3 (severe) symptoms.

can be added value in reporting or summarizing their multiple occurrences and combined duration, because these might lead to treatment discontinuations because of tolerability problems. For example, chronic headaches over a week or more could have much different clinical implications than a single, transient headache, but both might only be communicated in a typical table as one patient experiencing headache. In addition, where feasible, it might be more informative to provide relative risks (with confidence intervals and appropriate statistical caveats—see Recommendation 3 for details) of relevant adverse events, rather than simply reporting incidence rates, as a means of better conveying the likelihood of a patient experiencing these events. Figures 1-3[17-19] and the supplementary appendix[20-23] provide examples of additional potentially relevant descriptors. These additional data could be included in the main adverse events table, as a separate figure, or as supplemental data depending on the clinical relevance of the results and space constraints. Visual depictions of adverse event data in non-traditional formats, such as graphical plots, might be desirable if they facilitate efficient visualization of event profiles or differences in incidence rates between treatment groups[18 24] (fig 3).[19] Some online publications allow readers to access the data underlying visual depictions that aid interpretation.

Recommendation 3—Use statistical analysis for clinically relevant adverse events (where appropriate)

While formal statistical analyses of adverse event data can be useful, most clinical trials (with the exception of those measuring a specific adverse event as a primary or co-primary endpoint) are not designed with sufficient power to make definitive conclusions about adverse events.[25-27] We offer four overarching principles to guide statistical analysis when adverse events were not included in the primary endpoints.

Firstly, specify numerators and denominators. When individual events are recurrent, numerators should include both the number of events (per person time) and the number of patients experiencing the event. Furthermore, reporting denominators for adverse event data is important, as specified in the CONSORT Harms Extension (table 1). Including denominators can become especially relevant in situations where the denominator for collecting adverse event data varies appreciably. For example, if individual events are captured separately in multiple time intervals within the trial (eg, during a titration phase or separately during a stable dose phase), there could be changes to the denominator over time as participants reach endpoints at earlier time points, withdraw from the study, or have missing data for other reasons. In such situations, differences in denominators should be reported explicitly to ensure adverse event reporting is not always limited to the overall treatment group denominator (table 2).[13]

Secondly, reporting confidence intervals around absolute risk differences should in most cases be considered the minimum standard if analysis of adverse event data is performed (table 4),[28] given that these provide measures of the range of values consistent with the observed difference.[29 30] However, for rare events or trials with small sample sizes, it might be more appropriate to list the number of occurrences or individual adverse events without confidence intervals.

Thirdly, as previously required by the CONSORT Harms Extension (table 1), if more formal statistical comparisons are reported (eg, P values) then the underlying rationale and methodology for any analyses should be prespecified in the statistical analysis plan and stated in the publication. Additionally, inferential analyses not prespecified in the protocol (eg, in response to peer review requests or to test unanticipated signals of differences in adverse events) should be clearly labelled as "post hoc" or "exploratory" in the publication.

Fourthly, all formal statistical tests should be reported, even if not statistically significant, to avoid selective reporting and because the results may still be clinically informative.

Overall, we propose that when adverse events are being reported, formal statistical comparisons should be selectively used where there is clear justification. We recommend using an approach that determines the appropriateness of statistical analyses via a tiered system that assumes that it is important to report information about all events, but that only some events require formal hypothesis testing (box 2).[31]

Recommendation 4—Avoid use of overly general text descriptions for adverse events, including in abstracts

Clinical trial publications often rely on general statements such as "safe and well tolerated" or "no clinically significant or unexpected adverse events" to summarize the overall safety profile. Variations of the phrase "safe

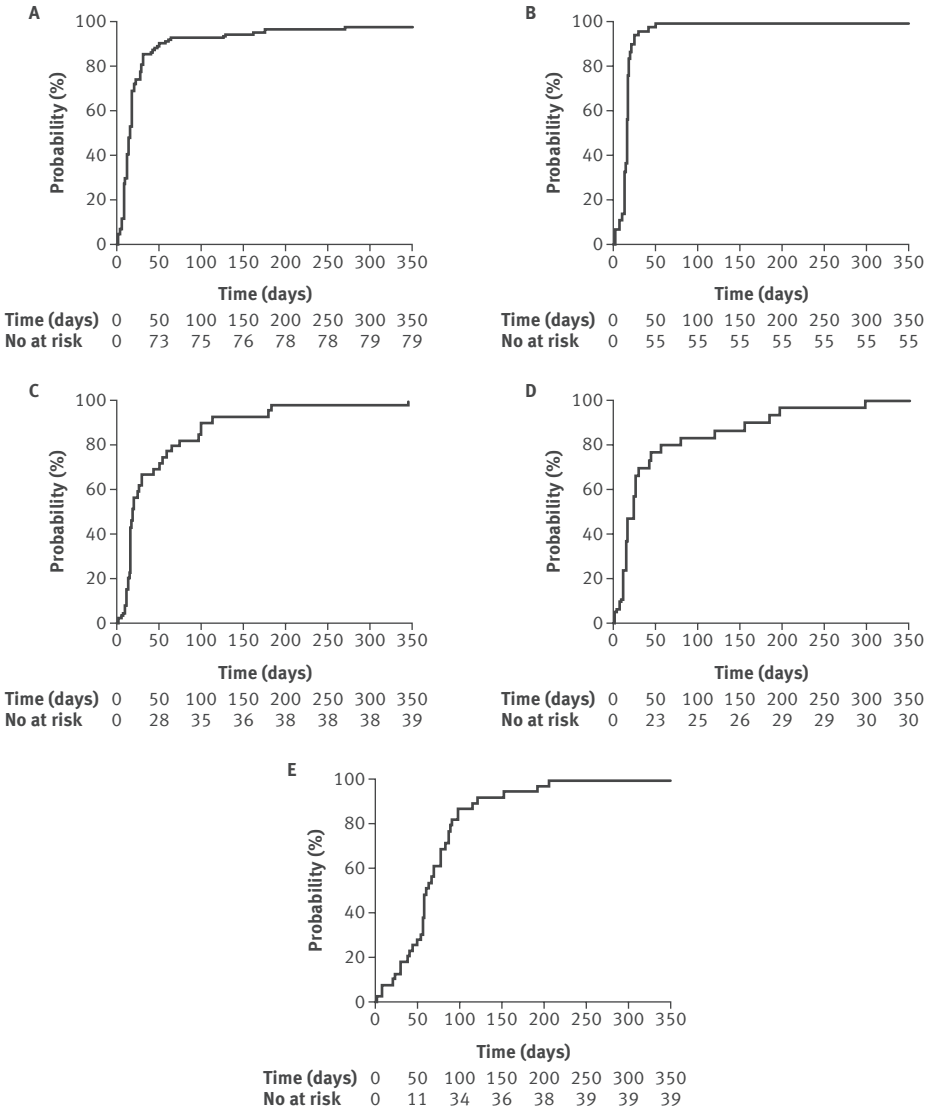


Fig 1 | Example of reporting exposure-time associations (Recommendation 2).[17] These figures show the cumulative chance of experiencing particular adverse events over time. When constrained for space, only adverse events (AEs) of clinical relevance should be prioritized. Reproduced from reference 17, with permission from Oxford University Press (original figure legend: "Cumulative incidences of major adverse events (AEs) after the start of everolimus treatment (A) stomatitis, (B) thrombocytopenia, (C) anemia, (D) hyperglycemia, and (E) pneumonitis")

and well tolerated" appeared in nearly 6000 published clinical trials from 1990 to 2015.[32] Authors instead should avoid broad generalizations, support summary statements about adverse event profiles with specific and informative data, and interpret those data in the context of the trial design and limitations (eg, small number of patients, short duration) and study population (eg, comorbidities).

The high visibility of abstracts makes the inclusion of relevant and informative adverse event data statements in this abbreviated section essential.[33 34] Both the CONSORT 2010 statement and the CONSORT Harms Extension recommend including harms data in the abstract. To provide more specificity, we recommend that abstracts include a phrase or sentence summarizing the most clinically relevant adverse event data and providing frequency percentages consistent with those presented in the main text of the publication. Examples of text providing this detail are shown in box 3.[35-40]

Recommendation 5—Discuss adverse events findings in the broader context of available evidence and maintain consistency of data across different public reports

Because individual trial results rarely provide a complete representation of a drug's likely adverse event profile, the publication should include context relevant to adverse events observed in previous studies of the same intervention and compared to the representative adverse event profile from similar agents in the same or a highly similar class if available. For example, where appropriate, trials of tumor necrosis factor (TNF) α inhibitors should measure and report data on infection and malignancy given the reported increased risk of

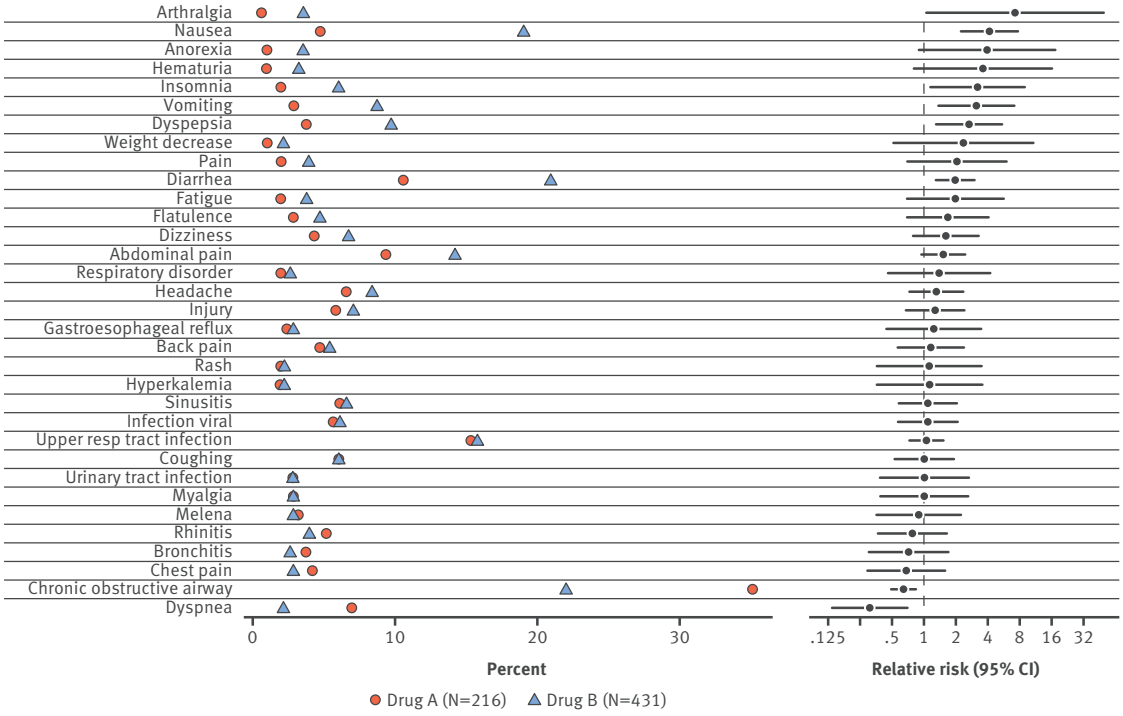
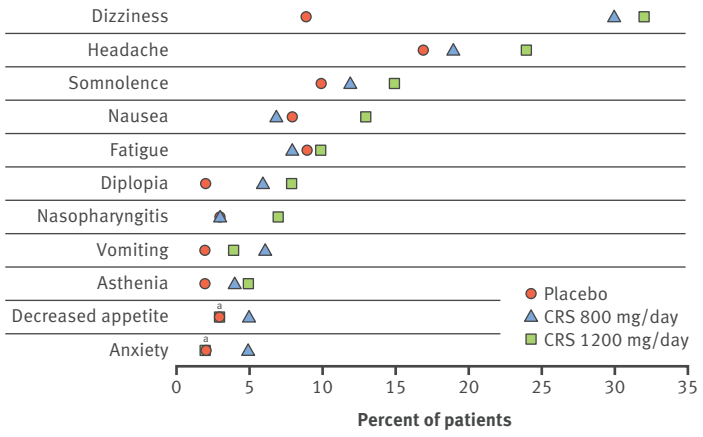


Fig 2 | Example of reporting relative risk differences (Recommendation 2).[18] Adverse events in this example are displayed by their frequency while also ranked by their relative risk using 95% confidence intervals between the investigational arm and placebo. Reproduced from reference 18, Copyright 2007 John Wiley & Sons, Ltd (original figure legend: "Most frequent on-therapy adverse events sorted by risk")



[a] Incidence for the placebo and carisbamate 1200 mg/day groups are the same

Fig 3 | Example of graphical representation of adverse event data (Recommendation 2).[19] This adverse event table displays typical adverse events in a visual fashion by plotting each treatment group along an axis by adverse event. This format allows the reader to more easily discriminate the difference in frequency of adverse events by treatment group. Reproduced from reference 19, *Epilepsia*, Copyright the International League Against Epilepsy (original figure legend: "Treatment-emergent adverse events in at least 5% of patients in any treatment group in the double-blind phase (safety analysis set)")

these findings with long term clinical exposure.[41-43] An exception might be that one might not expect to see malignancies arising during a short-term efficacy trial. These comparisons will allow healthcare practitioners to better understand whether observed adverse events reported in the publication are expected or unexpected and their relative frequency compared to previous reports.

For more common or more serious adverse events, authors should identify previously reported adverse event profiles and interpret current study findings in the context of what is known about the drug's safety. If the publication reports on a first-in-class medication or an early output from a longer clinical development plan, authors should highlight adverse events of special interest noted in human or animal studies for ongoing or future studies. As an example, the development of the first immune modulating agent in oncology identified "immune related" adverse events as associated with this new class of treatments.[44]

Results from an individual trial are now typically reported in multiple forms, including regulatory documents, clinical study reports (CSRs), clinical trial registries, clinical trial publications, medical meeting publications or presentations, patient level data portals, and other databases. Ensuring consistency among these various channels is helpful to ensure the interpretability and relevance of the data, whereas unexplained differences found through independent analyses could serve as grounds for skepticism about the data. For example, a recent analysis of adverse event reporting from clinical trials conducted with orlistat in the 1990s revealed that the clinical trial publications reported substantially lower incidences of adverse events than the CSRs submitted to the European Medicines Agency.[45] The major reasons for this discrepancy were counting multiple instances of the same adverse event only once (even though such methodology was not described in the CSRs), and the application of post hoc filters for adverse event reporting (which were not clarified in the relevant publications).[45]

Therefore, authors should make a reasonable attempt to compare adverse event data in the manuscript to those previously reported from the same trial in publications or conference presentations and other publicly available sources, including submissions to regulatory

Table 4 | Example of table presenting confidence intervals with adverse event data (Recommendation 3).[28] This table defines confidence intervals for the differences between investigational and control patient groups (right hand column). The number needed to harm and 95% confidence intervals can be derived from the same data by calculating 1/[difference] as a supplement to the information presented. However, the number needed to harm has potential drawbacks in terms of potential misunderstanding by clinicians and lack of communication about the uncertainty of the results that should be taken into account before use. Reproduced from reference 28, Copyright 2007 with permission from the American College of Physicians. All rights reserved (original table title: "Adverse events and symptoms that occurred in 5% or more patients who received at least 1 dose of study drug*")

Adverse event	Exenatide group (n=121), n (%)	Placebo group (n=112), n (%)	Difference (95% CI)†
Patients reporting ≥1 adverse event	92 (76.0)	73 (65.2)	10.9 (–1.7 to 23.4)
Patients reporting a serious adverse event	2 (1.7)	0	1.7 (–1.5 to 4.8)
Nausea	48 (39.7)	17 (15.2)	24.5 (12.7 to 36.3)
Nasopharyngitis	16 (13.2)	9 (8.0)	5.2 (–3.5 to 13.9)
Vomiting	16 (13.2)	1 (0.9)	12.3 (5.2 to 19.5)
Hypoglycemia	13 (10.7)	8 (7.1)	3.6 (–4.6 to 11.8)
Dyspepsia	9 (7.4)	1 (0.9)	6.5 (0.7 to 12.4)
Edema	7 (5.8)	9 (8.0)	–2.3 (–9.6 to 5.1)
Headache	7 (5.8)	5 (4.5)	1.3 (–5.2 to 7.8)
Diarrhea	7 (5.8)	3 (2.7)	3.1 (–2.9 to 9.1)
Influenza	6 (5.0)	5 (4.5)	0.5 (–5.8 to 6.8)

*Values refer to numbers and percentages of patients who reported particular symptoms that occurred ≥1 times during the trial.
†Difference (reported in percentage points) is calculated as exenatide minus placebo.

Box 2: Safety planning, evaluation, and reporting team (SPERT) recommendations (Recommendation 3)

The system categorizes events into one of three categories or tiers for analysis. The important distinction is between prespecified hypotheses (tier 1) and non-prespecified events, which are further subdivided based on whether the event is "common" (tier 2) or "uncommon" (tier 3):

- Tier 1 events should be reported with risk differences and confidence intervals and P values, potentially including adjustment for multiple testing over time.
- Tier 2 events should be reported with risk differences and confidence intervals, possibly including P values.
- Tier 3 events should be reported with descriptive statistics (numbers, percentages, and possibly rates per person time) but without P values or confidence intervals. Some Tier 3 events predefined in the protocol to ensure close monitoring might be expected to be so rare that statistical analysis will not be meaningful.

bodies. Important differences between current and previously disclosed data from the study should be clearly noted and an explanation provided. Examples include the evolution of adverse event signals (as the database of recruited patients matures or longer follow-up time accrues between publications for the same study cohort) as well as changes in adverse events definitions or groupings, data censoring conventions, and how subgroups of patient populations are reported (eg, age groups defined differently). Differences between these publications should be flagged proactively in the cover letter to the editors when submitting the manuscript and disclosed in the publication as appropriate and as agreed with the editor.

Discussion

Clinical trial publications provide a vital foundation for recommending treatments where the ultimate use of the product should be based on the product label and prescribing information. Thus, authors of these publications bear responsibility to present a balanced perspective of the findings so that practitioners can make well informed assessments of benefit-risk when deciding on patient treatment.[46] Journal editors previously highlighted that clinically relevant and more informative adverse event reporting is a significant unmet need to improve patient care and increase the credibility of industry sponsored publications.[10] To address this issue, we provide here consensus recommendations identified through sequential discussions with journal editors and industry representatives who participate in the generation and publishing of clinical trial manuscripts. This guidance is meant to supplement the guidance in the CONSORT Harms Extension and address issues that have emerged over the past decade.

Firstly, our discussions highlighted the suboptimal implementation of the recommendations in the CONSORT Harms Extension[9] (potentially because of the lack of specificity in adverse event reporting recommendations) as a key shortfall in clinical trial publications. We therefore provide additional complementary guidance with detailed examples to encourage more complete and clinically meaningful adverse event reporting. Moreover, our discussions emphasize that frequency of adverse events alone should not serve as a proxy for relevance. Instead, we propose that comprehensive adverse event reporting (including events identified by regulators as always being important) should be accompanied by more detailed reporting of those adverse events of greatest clinical significance for practitioners and patients. This more flexible approach depends heavily on the experience and judgment of authors to apply an appropriate filter and clearly communicate their rationale in the clinical trial publication.

Secondly, peer reviewed clinical trial publications should avoid the use of overly general summary phrases (which could convey subjectivity) and instead provide underlying detail for communicating the more informative and meaningful adverse event profile. In addition, adverse event reporting should be consistent throughout the body of the paper, abstract, results, and tables or figures. This approach is broadly applicable to clinical trial publications because no single journal, therapeutic area, author group, or sponsor performs consistently better than each other in this regard.[3-9]

Thirdly, authors should discuss how the individual trial results relate to the larger body of clinical evidence in order to help practitioners and future investigators better understand the evolving adverse event profile of a new intervention.

To facilitate widespread dissemination and adoption of these recommendations, all MPIP members are implementing these guidelines within their own research and medical divisions and are advocating broad industry adoption, including as part of a campaign to raise awareness of the importance of transparency in dissemination of clinical research. MPIP's representation from industry makes the group a good starting point, supported by the fact that industry is the largest single sponsor of clinical trials.[47 48] The next

Box 3: Recommendations for text descriptions (Recommendation 4)

Avoid vague phrases to describe adverse events

- "Safe and well tolerated"
- "Adverse events seen were minor and generally well tolerated"
- "No unexpected adverse events"

Recommendation for text: example 1

These examples from the results section[35] highlight the recommended use of confidence intervals (CIs) and communication of formal statistical analysis when looking at a particular adverse event of clinical relevance.

> "Investigators reported a total of 35 events of pancreatitis in 33 patients in the saxagliptin arm and 35 events in 30 patients in the placebo arm (hazard ratio [HR] 1.09 [95% CI 0.66–1.79]) (Table 1 and Supplementary Table 1). Pancreatitis was adjudicated in 24 patients (26 events) receiving saxagliptin and 21 (25 events) receiving placebo (HR 1.13 [95% CI 0.63–2.06], Fisher exact test P=0.77) (Table 1 and Supplementary Fig 1A).
>
> "Pancreatic cancer was reported in 5 patients in the saxagliptin arm and 12 in the placebo arm (HR 0.42 [95% CI 0.13–1.12], Fisher exact test P=0.09), including 1 patient with a neuroendocrine tumor in the placebo arm."

Material reproduced from reference 35, with permission from the American Diabetes Association. Copyright and all rights reserved.

Recommendation for text: example 2

This text description[36] illustrates a number of points—absolute numerators and denominators along with percentages, describing serious adverse events in addition to those of more relevance, and instances where adverse events were high in the investigational drug group. The supplementary appendix is used to capture other events that might be of interest to readers.

> "Adverse events occurred in 2060 of 2976 patients (69.2%) in the evolocumab group and in 965 of 1489 patients (64.8%) in the standard-therapy group (Table 3). Serious adverse events occurred in 222 patients (7.5%) in the evolocumab group and in 111 patients (7.5%) in the standard-therapy group. Elevations in aminotransferase or creatine kinase levels occurred at a similar rate in the two groups: 1.0% in the evolocumab group and 1.2% in the standard-therapy group for elevated amino transferase levels and 0.6% and 1.1%, respectively, for elevated creatine kinase levels. Although the rate of neurocognitive adverse events was low (<1%), such events were reported more frequently in the evolocumab group. Of note, the incidence of neurocognitive adverse events did not appear to be related to the LDL cholesterol level during treatment (Table S2 in the Supplementary Appendix). Other adverse events are listed in Table S3 in the Supplementary Appendix."

Material reproduced from reference 36 (*New England Journal of Medicine*); LDL=low density lipoprotein.

Recommendation for text: example 3

Despite similar incidence of overall adverse events between all three patient groups, the authors for this example[37] also include a detailed section focused on reporting adverse events of special interest. These are summarized in an adverse events table.

> "4.4.1. Adverse events of special interest. The proportion of patients reporting EPS-related TEAEs during the AC/CT phases was higher in the paliperidone ER-(n=207, 34%) than olanzapine (n=23, 16%) group; and during the maintenance phase was higher in the olanzapine group (n=8, 10%) than paliperidone ER (n=6, 4%) or placebo (n=4, 3%) groups. EPS-related TEAEs in the paliperidone ER group during the maintenance phase were dyskinesia, akathisia, hypokinesia, tremor (n=1, 1% each) and extrapyramidal disorder (n=2, 1%); the event of dyskinesia resulted in study discontinuation. Two patients in the olanzapine group were discontinued from the study (hypertonia [n=1], extrapyramidal disorder [n=1]) during the maintenance phase."

Material reproduced from reference 37, Copyright 2012, with permission from Elsevier; AC/CT=acute and continuation treatment phase; EPS=extrapyramidal side effect; ER=extended release; TEAE=treatment emergent adverse event.

Recommendation for abstract: example 1

This abstract from a phase 3 study[38] notes more serious adverse events and adds the number of total instances along with percentage seen. It then compares the adverse events from the investigational drug groups to those reported for the placebo control group.

> Results section: "The most common grade 3-4 adverse events were fatigue (72 [9%] of 791 patients in the abiraterone group vs 41 [10%] of 394 in the placebo group), anaemia (62 [8%] vs 32 [8%]), back pain (56 [7%] vs 40 [10%]), and bone pain (51 [6%] vs 31 [8%])."
>
> Conclusions section: "No new safety signals were identified with increased follow-up."

Reproduced from reference 38, Copyright 2012, with permission from Elsevier.

Recommendation for abstract: example 2

In addition to providing the number and frequency for serious adverse events, this abstract[39] highlights specifically those adverse events that led to discontinuation along with laboratory measurements of interest for these treatments for hepatitis C virus infection.

> "Serious adverse events occurred in 12 (6%) patients in the treatment-naive group; 11 (5%) non-responders, and 16 (7%) ineligible, intolerant, or ineligible and intolerant patients; adverse events leading to discontinuation (most commonly reversible increases in alanine or aspartate aminotransferase) occurred in six (3%), two (1%), and two (1%) patients, respectively, with no deaths recorded. Grade 3 or 4 laboratory abnormalities were uncommon, with low incidences of aminotransferase increases during the first 12 weeks with daclatasvir plus asunaprevir and placebo in treatment-naive patients (≤2% each)."

Reproduced from reference 39, Copyright 2014, with permission from Elsevier.

Recommendation for abstract: example 3

This abstract from a medical meeting poster[40] calls attention to the total adverse events in addition to those of greatest frequency, which also are of greatest relevance to infections associated with anti-tumor necrosis factor α agents. Both the number and percentage are highlighted for each data point.

> "Ninety-three AE were recorded (mean of 0,8 AE/patient/year): 73 with IFX and 20 with ADA. There was at least one AE in 21 of the patients (91%). The most frequently reported events were infections (n=50, in 16 patients), episodes of serum sickness-like disease (n=16, in 7 patients) and infusion reactions (n=14, in 4 patients)."

Reproduced from reference 40, with permission from the author; ADA=adalimumab; AE=adverse event(s); IFX=infliximab; TNF=tumor necrosis factor.

phase of driving adoption of these recommendations will focus on cooperative trial groups and academic sponsors because these recommendations have the potential for generalizability towards non-industry sponsored trials given the similar goal to improve adverse event reporting via published manuscripts.

The limitations of this report include the small number of participants, the lack of direct involvement of patients or patient groups and academic trialists, and that the findings represent statements that might not be universally accepted or comprehensive. However, we sought to involve participants from a cross section of functional areas with experience in the reporting of adverse events in clinical trial publications and to minimize the effects of not directly involving patients or practicing clinicians. For example, the recommendation to use a "clinical relevance" filter should address the concerns of patients as physician authors are generally expected to identify such clinically relevant adverse events based on their clinical experience and awareness of patients' willingness to accept a benefit-risk tradeoff. In addition, although our panel of interviewees did not separately include academic trial groups, the medical journal editors—as well as representatives from industry who participated in our panel—work closely with academic trialists and research organizations; as a result, the perspectives gained from such interactions are reflected in our recommendations.

Conclusions

There is room for improvement in the reporting of adverse event data from clinical trials. This paper provides recommendations intended to supplement existing guidelines and address key challenges when reporting adverse event data in clinical trial publications. Following these recommendations and "best practice" examples will help improve the transparency, clinical relevance, and credibility of adverse event reporting.

Contributors: The project was conceived and designed by the MPIP Steering Committee and executed by all the authors. NL was responsible for interviews and journal-industry roundtable meeting facilitation. All the authors were involved in discussions of the data and development of the recommendations. JAB, AB, SG, CK, NL, and BM wrote the first draft of the manuscript, which was critically reviewed, revised, and approved for submission by all the authors. JAB is guarantor.

Participants and their affiliations at the time of the Roundtable meeting were as follows: Patricia Baskin (executive editor, *Neurology*); Michael Berkwits (electronic editor and deputy editor, *JAMA*); Jesse A Berlin (vice president, global epidemiology, Johnson & Johnson); Ananya Bhattacharya (director, publication policy and education, Bristol-Myers Squibb); Angela Bickford, (publications director, GlaxoSmithKline); Matthew Cahill (associate vice president, global scientific and medical publications, Merck & Co); Edward Campion (senior deputy editor and online editor, *New England Journal of Medicine*); Tai-Tsang Chen (executive director, head of global biometric sciences, Bristol-Myers Squibb); Leslie Citrome (editor in chief, *International Journal of Clinical Practice*); Juli Clark (executive director, global medical writing, Amgen); Robert Enck (editor in chief, *American Journal of Hospice and Palliative Medicine*); John Fletcher (editor in chief, *Canadian Medical Association Journal*); Maureen Garrity (director, publication, Astellas); Thomas Gesell (Envision Pharma, on behalf of ISMPP); Jodie Gillon (director, publications management, Pfizer); Susan Glasser (senior director, scientific and medical publications, Janssen Research & Development); Samantha Gothelf (executive director, global medical publications, Bristol-Myers Squibb); Daniel Haller (editor in chief, *Gastrointestinal Cancer Research*); Carolyn Hustad (executive director, global scientific and medical publications, Merck & Co); Christian Klem (senior director, global publications and scientific communications, CV/Met/CKD, AstraZeneca); Christine Laine (editor in chief, *Annals of Internal Medicine*); Mary-Margaret Lannon (senior director, global publications, Takeda); Neil Lineberry (associate director, Navigant Consulting); Bernadette Mansi (director, publications & disclosure practices, GlaxoSmithKline); Doug Martin (managing director, Navigant Consulting); LaVerne Mooney (director, publications management, Pfizer); Teresa Peña (global director, publications, AstraZeneca); Karen Pinette (director, medical publications, Biogen); Lindor Qunaj (analyst, Navigant Consulting); and Douglas Robertson (associate editor, *Gastroenterology*).

Patricia Baskin (executive editor, *Neurology*); Kim Tran (associate director, publications, Bristol-Myers Squibb); and Brian Scheckner (director, publication policy and education, Bristol-Myers Squibb) assisted with development of the manuscript.

Participants in the research and their affiliations at the time of interview include: Sabina Alam (editor in chief, *BMC Medicine*); Harold Bays (medical director, Louisville Metabolic and Atherosclerosis Research Center); Jesse A Berlin (vice president, global epidemiology, Johnson & Johnson); Peter Calverley (pulmonologist, University of Liverpool); Tai-Tsang Chen (executive director, head of global biometric sciences, Bristol-Myers Squibb); Nancy Chescheir (editor in chief, *Obstetrics and Gynecology*); Jessica Chinault-Jalboot (safety management business process lead, Merck & Co); Kathryn Connor (clinical lead, Merck & Co); Anh Tuan Dinh-Xuan (editor in chief, *European Respiratory Journal*); Mark Dransfield (pulmonary, allergy and critical care medicine, University of Alabama at Birmingham); Susan Duke (biostatistics, GlaxoSmithKline); Ashwin Gollerkeri (medical monitoring, Pfizer); Stuart Kupfer (clinical science, Takeda); Elizabeth Loder (editor, *BMJ*); Louis Mini (vice president, medical affairs, Takeda Pharmaceuticals); Karen Outten (clinical trial safety reporting, Merck & Co); Frank Rockhold (safety and pharmacovigilance, GlaxoSmithKline); Anthony Rodgers (biostatistics, Merck & Co); Robert Spanheimer (global outcomes and epidemiology, Takeda); Stuart Spencer (executive editor, *The Lancet*); Murray Stewart (head of metabolic pathways and cardiovascular, GlaxoSmithKline); Bartholomew Tortella (medical affairs, Pfizer); Bonnie Vlahos (clinical lead, Pfizer); Michele Wible (biostatistics, Pfizer); Robert Williams (clinical science, Pfizer); Shawn Yu (outcomes research, Takeda); and two interview participants who could not be contacted subsequently and therefore did not provide permission for acknowledgment.

Medical editorial assistance was provided by Shalini Murthy (ProEd Communications) and funded by Medical Publishing Insights & Practices.

Funding: Medical Publishing Insights & Practices (MPIP) (www.mpip-initiative.org/) is a collaboration of pharmaceutical companies and the International Society for Medical Publication Professionals (ISMPP) and is sponsored by the member companies. NL was a paid consultant to MPIP in the conduct of research, data collection and analysis, development of the manuscript, and the design and management of the adverse events roundtable meeting. NL was also an employee of Navigant at the time of the roundtable and initiation of the manuscript. NL continued to be a paid consultant on the development of the manuscript and management of author meetings after leaving Navigant. NL is currently employed by Takeda, Cambridge, MA.

Competing interests: We have read and understood the BMJ policy on declaration of interests and declare the following: JAB, AB, TC, SG, CK, NL, and BM are employees of companies sponsoring MPIP, as shown by their individual affiliations. The remaining authors declare no competing interests.

Provenance and peer review: Encouraged, but not commissioned; externally peer reviewed.

1 Golder S, Loke YK, Wright K, Norman G. Reporting of adverse events in published and unpublished studies of health care interventions: a systematic review. *PLoS Med* 2016;13:e1002127. doi:10.1371/journal.pmed.1002127.
2 Ioannidis JP, Evans SJ, Gøtzsche PC, et al. CONSORT Group. Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Ann Intern Med* 2004;141:781-8. doi:10.7326/0003-4819-141-10-200411160-00009.
3 Ioannidis JP. Adverse events in randomized trials: neglected, restricted, distorted, and silenced. *Arch Intern Med* 2009;169:1737-9. doi:10.1001/archinternmed.2009.313.
4 Haidich AB, Birtsou C, Dardavessis T, Tirodimos I, Arvanitidou M. The quality of safety reporting in trials is still suboptimal: survey of major general medical journals. *J Clin Epidemiol* 2011;64:124-35. doi:10.1016/j.jclinepi.2010.03.005.
5 Smith SM, Chang RD, Pereira A, et al. Adherence to CONSORT harms-reporting recommendations in publications of recent analgesic clinical trials: an ACTTION systematic review. *Pain* 2012;153:2415-21. doi:10.1016/j.pain.2012.08.009.

6 Hodkinson A, Kirkham JJ, Tudur-Smith C, Gamble C. Reporting of harms data in RCTs: a systematic review of empirical assessments against the CONSORT harms extension. *BMJ Open* 2013;3:e003436. doi:10.1136/bmjopen-2013-003436.
7 Shukralla AA, Tudur-Smith C, Powell GA, Williamson PR, Marson AG. Reporting of adverse events in randomised controlled trials of antiepileptic drugs using the CONSORT criteria for reporting harms. *Epilepsy Res* 2011;97:20-9. doi:10.1016/j.eplepsyres.2011.06.015.
8 Bagul NB, Kirkham JJ. The reporting of harms in randomized controlled trials of hypertension using the CONSORT criteria for harm reporting. *Clin Exp Hypertens* 2012;34:548-54. doi:10.3109/10641963.2012.681724.
9 Péron J, Maillet D, Gan HK, Chen EX, You B. Adherence to CONSORT adverse event reporting guidelines in randomized clinical trials evaluating systemic cancer therapy: a systematic review. *J Clin Oncol* 2013;31:3957-63. doi:10.1200/JCO.2013.49.3981.
10 Mansi BA, Clark J, David FS, et al. Ten recommendations for closing the credibility gap in reporting industry-sponsored clinical research: a joint journal and pharmaceutical industry perspective. *Mayo Clin Proc* 2012;87:424-9. doi:10.1016/j.mayocp.2012.02.009.
11 International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. Clinical safety data management: definitions and standards for expedited reporting. 1994. www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E2A/Step4/E2A_Guideline.pdf.
12 United States Food and Drug Administration. What is a serious adverse event? 2014. www.fda.gov/safety/medwatch/howtoreport/ucm053087.htm.
13 Gold R, Giovannoni G, Selmaj K, et al. SELECT study investigators. Daclizumab high-yield process in relapsing-remitting multiple sclerosis (SELECT): a randomised, double-blind, placebo-controlled trial. *Lancet* 2013;381:2167-75. doi:10.1016/S0140-6736(12)62190-4.
14 Battisti WP, Wager E, Baltzer L, et al. International Society for Medical Publication Professionals. Good publication practice for communicating company-sponsored medical research: GPP3. *Ann Intern Med* 2015;163:461-4. doi:10.7326/M15-0288.
15 Committee on Strategies for Responsible Sharing of Clinical Trial Data, Board on Health Sciences Policy, Institute of Medicine. Sharing clinical trial data: maximizing benefits, minimizing risk. National Academies Press (US), April 2015.
16 Rendi-Wagner P, Noedl H, Wernsdorfer WH, Wiedermann G, Mikolasek A, Kollaritsch H. Unexpected frequency, duration and spectrum of adverse events after therapeutic dose of mefloquine in healthy adults. *Acta Trop* 2002;81:167-73. doi:10.1016/S0001-706X(01)00210-8.
17 Nozawa M, Nonomura N, Ueda T, et al. Adverse event profile and dose modification of everolimus for advanced renal cell carcinoma in real-world Japanese clinical practice. *Jpn J Clin Oncol* 2013;43:1132-8. doi:10.1093/jjco/hyt121.
18 Amit O, Heiberger RM, Lane PW. Graphical approaches to the analysis of safety data from clinical trials. *Pharm Stat* 2008;7:20-35. doi:10.1002/pst.254.
19 Halford JJ, Ben-Menachem E, Kwan P, et al. A randomized, double-blind, placebo-controlled study of the efficacy, safety, and tolerability of adjunctive carisbamate treatment in patients with partial-onset seizures. *Epilepsia* 2011;52:816-25. doi:10.1111/j.1528-1167.2010.02960.x.
20 McGovern PC, Wible M, El-Tahtawy A, Biswas P, Meyer RD. All-cause mortality imbalance in the tigecycline phase 3 and 4 clinical trials. *Int J Antimicrob Agents* 2013;41:463-7. doi:10.1016/j.ijantimicag.2013.01.020.
21 Prosperi MC, Fabbiani M, Fanti I, et al. Predictors of first-line antiretroviral therapy discontinuation due to drug-related adverse events in HIV-infected patients: a retrospective cohort study. *BMC Infect Dis* 2012;12:296. doi:10.1186/1471-2334-12-296.
22 Leonard R, O'Shaughnessy J, Vukelja S, et al. Detailed analysis of a randomized phase III trial: can the tolerability of capecitabine plus docetaxel be improved without compromising its survival advantage?*Ann Oncol* 2006;17:1379-85. doi:10.1093/annonc/mdl134.
23 Miller AE, Wolinsky JS, Kappos L, et al. TOPIC Study Group. Oral teriflunomide for patients with a first clinical episode suggestive of multiple sclerosis (TOPIC): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet Neurol* 2014;13:977-86. doi:10.1016/S1474-4422(14)70191-7.
24 Duke SP, Bancken F, Crowe B, Soukup M, Botsis T, Forshee R. Seeing is believing: good graphic design principles for medical research. *Stat Med* 2015;34:3040-59. doi:10.1002/sim.6549.
25 O'Neill RT. Statistical concepts in the planning and evaluation of drug safety from clinical trials in drug development: issues of international harmonization. *Stat Med* 1995;14:117-27. doi:10.1002/sim.4780140932.
26 Chuang-Stein C. Safety analysis in controlled clinical trials. *Drug Inf J* 1998;32:1363S-72Sdoi:10.1177/00928615980320S132.
27 Singh S, Loke YK. Drug safety assessment in clinical trials: methodological challenges and opportunities. *Trials* 2012;13:138. doi:10.1186/1745-6215-13-138.
28 Zinman B, Hoogwerf BJ, Durán García S, et al. The effect of adding exenatide to a thiazolidinedione in suboptimally controlled type 2 diabetes: a randomized trial. *Ann Intern Med* 2007;146:477-85. doi:10.7326/0003-4819-146-7-200704030-00003.
29 Borenstein M. The case for confidence intervals in controlled clinical trials. *Control Clin Trials* 1994;15:411-28. doi:10.1016/0197-2456(94)90036-1.
30 Coulson M, Healey M, Fidler F, Cumming G. Confidence intervals permit, but do not guarantee, better inference than statistical significance testing. *Front Psychol* 2010;1:26.doi:10.3389/fpsyg.2010.00026.
31 Crowe BJ, Xia HA, Berlin JA, et al. Recommendations for safety planning, data collection, evaluation and reporting during drug, biologic and vaccine development: a report of the safety planning, evaluation, and reporting team. *Clin Trials* 2009;6:430-40. doi:10.1177/1740774509344101.
32 United States National Library of Medicine. PubMed. 2015. www.ncbi.nlm.nih.gov/pubmed.
33 Naylor CD, Chen E, Strauss B. Measured enthusiasm: does the method of reporting trial results alter perceptions of therapeutic effectiveness?*Ann Intern Med* 1992;117:916-21. doi:10.7326/0003-4819-117-11-916.
34 Bobbio M, Demichelis B, Giustetto G. Completeness of reporting trial results: effect on physicians' willingness to prescribe. *Lancet* 1994;343:1209-11. doi:10.1016/S0140-6736(94)92407-4.
35 Raz I, Bhatt DL, Hirshberg B, et al. Incidence of pancreatitis and pancreatic cancer in a randomized controlled multicenter trial (SAVOR-TIMI 53) of the dipeptidyl peptidase-4 inhibitor saxagliptin. *Diabetes Care* 2014;37:2435-41. doi:10.2337/dc13-2546.
36 Sabatine MS, Giugliano RP, Wiviott SD, et al. Open-Label Study of Long-Term Evaluation against LDL Cholesterol (OSLER) Investigators. Efficacy and safety of evolocumab in reducing lipids and cardiovascular events. *N Engl J Med* 2015;372:1500-9. doi:10.1056/NEJMoa1500858.
37 Berwaerts J, Melkote R, Nuamah I, Lim P. A randomized, placebo- and active-controlled study of paliperidone extended-release as maintenance treatment in patients with bipolar I disorder after an acute manic or mixed episode. *J Affect Disord* 2012;138:247-58. doi:10.1016/j.jad.2012.01.047.
38 Fizazi K, Scher HI, Molina A, et al. COU-AA-301 Investigators. Abiraterone acetate for treatment of metastatic castration-resistant prostate cancer: final overall survival analysis of the COU-AA-301 randomised, double-blind, placebo-controlled phase 3 study. *Lancet Oncol* 2012;13:983-92. doi:10.1016/S1470-2045(12)70379-0.
39 Manns M, Pol S, Jacobson IM, et al. HALLMARK-DUAL Study Team. All-oral daclatasvir plus asunaprevir for hepatitis C virus genotype 1b: a multinational, phase 3, multicohort study. *Lancet* 2014;384:1597-605. doi:10.1016/S0140-6736(14)61059-X.
40 Bernardes C, Carvalho D, Russo P, Saiote J, Ramos J. Anti-TNF alpha therapy in inflammatory bowel disease—safety profile in elderly patients [poster]. Presented at: 10th Annual Congress of European Crohn's and Colitis Organization. February 18-21, 2015; Barcelona: poster P637. https://www.ecco-ibd.eu/index.php/publications/congress-abstract-s/abstracts-2015/item/p637-anti-tnf-alpha-therapy-in-inflammatory-bowel-disease-safety-profile-in-elderly-patients.html.
41 Burmester GR, Panaccione R, Gordon KB, McIlraith MJ, Lacerda AP. Adalimumab: long-term safety in 23 458 patients from global clinical trials in rheumatoid arthritis, juvenile idiopathic arthritis, ankylosing spondylitis, psoriatic arthritis, psoriasis and Crohn's disease. *Ann Rheum Dis* 2013;72:517-24. doi:10.1136/annrheumdis-2011-201244.
42 Smolen JS, Emery P. Infliximab: 12 years of experience. *Arthritis Res Ther* 2011;13(suppl 1):S2.doi:10.1186/1478-6354-13-S1-S2.
43 Weinblatt ME, Bathon JM, Kremer JM, et al. Safety and efficacy of etanercept beyond 10 years of therapy in North American patients with early and longstanding rheumatoid arthritis. *Arthritis Care Res (Hoboken)* 2011;63:373-82.doi:10.1002/acr.20372.
44 Weber J. Review: anti-CTLA-4 antibody ipilimumab: case studies of clinical response and immune-related adverse events. *Oncologist* 2007;12:864-72. doi:10.1634/theoncologist.12-7-864.
45 Schroll JB, Penninga EI, Gøtzsche PC. Assessment of adverse events in protocols, clinical study reports, and published papers of trials of orlistat: a document analysis. *PLoS Med* 2016;13:e1002101. doi:10.1371/journal.pmed.1002101.
46 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;62:609-16. doi:10.1016/j.jclinepi.2008.08.005.
47 Ehrhardt S, Appel LJ, Meinert CL. Trends in National Institutes of Health funding for clinical trials registered in ClinicalTrials.gov. *JAMA* 2015;314:2566-7. doi:10.1001/jama.2015.12206.
48 Anderson ML, Chiswell K, Peterson ED, Tasneem A, Topping J, Califf RM. Compliance with results reporting at ClinicalTrials.gov. *N Engl J Med* 2015;372:1031-9. doi:10.1056/NEJMsa1409364.

Supplementary appendix: Additional materials