Clinical Study Reports of randomized controlled trials – an exploratory review of previously confidential industry reports

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2012-002496
Article Type:	Research
Date Submitted by the Author:	14-Dec-2012
Complete List of Authors:	Doshi, Peter; Johns Hopkins University, Jefferson, Tom; Cochrane Vaccines Field
Primary Subject Heading:	Evidence based practice
Secondary Subject Heading:	Medical publishing and peer review, Ethics, Research methods
Keywords:	MEDICAL ETHICS, MEDICAL JOURNALISM, INTERNAL MEDICINE

SCHOLARONE™
Manuscripts

For peer review only

Clinical Study Reports of randomized controlled trials – an exploratory review of previously confidential industry reports

Authors:

Peter Doshi PhD
Johns Hopkins University School of Medicine, Baltimore, Maryland, USA

Tom Jefferson MD
The Cochrane Collaboration, Roma, Italy

Corresponding author: Peter Doshi <pnd@jhu.edu>

Word count: 2998

Tables: 3
Figures: 2
Appendices: 2
References: 41

For peer review only

Patient consent statement No consent was necessary as no patients were involved

Ethics approval statement No ethical approval was necessary as no patients were involved and all data were aggregate or anonymized and publicly available.

Role of the sponsor statement As the review had no extramural funding, there was no sponsor.

Author Contributions: Doshi had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. *Study concept and design*: Doshi and Jefferson. *Acquisition of data*: Doshi and Jefferson. *Analysis and interpretation of data*: Doshi and Jefferson. *Critical revision of the manuscript for important intellectual content*: Doshi and Jefferson. *Statistical analysis*: Doshi.

For peer review only

ARTICLE SUMMARY

Research questions or hypotheses addressed

What are Clinical Study Reports (CSRs)? What do they contain and how long are they?

Might CSRs help address reporting biases associated with the published literature, and improve the quality of evidence synthesis?

Key Messages (up to 3)

CSRs represent a hitherto hidden and untapped source of detailed RCT data (mean page length: 1,854 pages), increasingly becoming publicly available, and should form the basic unit for evidence synthesis to minimize the problem of reporting bias.

CSRs show that numerous individuals make important technical contributions to the design, conduct, and reporting of each trial, but journal publications often fail to record these details, resulting in a loss in individual responsibility for what is reported.

The E3 guideline to which most CSRs conform was published in 1995, and needs updating.

Strengths and Limitations

We cannot say whether our sample is representative and whether our conclusions are generalizable to an undefined and undefineable population of CSRs.

For peer review only

Abstract

Objective: To explore the structure and content of a non-random sample of clinical study reports (CSRs) to guide clinicians and systematic reviewers.

Search strategy: We searched public sources and lodged Freedom of Information requests for previously confidential CSRs primarily written by industry for regulators.

Selection criteria: CSR reporting sufficient information for extraction ("adequate")

Primary outcome measures: Presence and length of essential elements of trial design and reporting and compression factor (ratio of page length for CSR compared to its published counterpart in a scientific journal).

Data extraction: data were extracted on standard forms and cross-checked for accuracy

Results: We assembled a population of 84 CSRs (covering 90 RCTs; 144,610 pages total) dated 1991-2011 of 14 pharmaceuticals. 78 were adequate. Report synopses had a median length of 5 pages, efficacy evaluation 13.5 pages, safety evaluation 17 pages, attached tables 337 pages, trial protocol 62 pages, statistical analysis plan 15 pages, and individual efficacy and safety listings had a median length of 447 and 109.5 pages, respectively. While 16 (21%) of CSRs contained completed case report forms, these were accessible to us in only one case (765 pages representing 16 individuals). Compression factors ranged between 1 and 8805.

Conclusions: Clinical study reports represent a hitherto mostly hidden and untapped source of detailed and exhaustive data on each trial. They should be consulted by independent parties interested in a detailed record of a clinical trial, and should form the basic unit for evidence synthesis as their use is likely to minimize the problem of reporting bias. We cannot say whether our sample is representative and whether our conclusions are generalizable to an undefined and undefineable population of CSRs.

Word count: 275

Primary Funding Source: The review had no extramural funding.

For peer review only

Introduction

Systematic reviews are thought to provide one of the most robust ways to evaluate the effects of healthcare interventions. But the robustness of findings clearly rests upon reviewers' access to clinical trial information sufficient to critically evaluate and reproduce the original research. Research on reporting bias over the last decades has shown that trusting the published literature at face value, even peer-reviewed publications, can be fraught with difficulty—a problem that spans drug classes.[1–12]

Following the decision by the European regulator, European Medicines Agency (EMA) on 30 Nov 2010, to make available a broad spectrum of documents related to medicinal products for human and veterinary use,[13,14] attention is focusing on one particular type of regulatory document: clinical study reports (CSRs).[15–18] CSRs are usually written for regulators following guidelines developed by the industry-regulatory collaborative effort "International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use" (ICH). The ICH guidelines "Structure and Content of Clinical Study Reports"[19] (See Appendix 1) are known by the document code "E3". They were formalized in 1995 "to assist sponsors in the development of a report that is complete, free from ambiguity, well organised and easy [for regulators] to review."[19]. E3 has not been edited or changed since 1995.

CSRs are but one category of information that is transmitted from study sponsors to regulators (Figure 1), but are important as they contain substantially more information and detail on the intervention being tested than published versions of the same trial. The wealth of information may be sought with increasing frequency by researchers appraising single trials, entire trial programmes, or by those synthesizing evidence.[17,20] We are aware of two recent examples of systematic reviews carried out using CSRs and other regulatory material.[12,21] One group also concluded that journal publications insufficiently report clinical trials.[22]

Despite CSRs' potential importance very little is known about their structure and content outside of those individuals with direct involvement in regulatory processes. This knowledge gap may hinder development of methods for fair and reliable appraisal of CSRs and their use in evidence synthesis. We are not aware of any instruments specifically designed for appraising CSRs. Lack of visibility may also conceal the complexity of the organization and reporting of clinical trials.

We carried out an exploratory review to describe the structure and content of a non-random sample of clinical study reports. Our long-term intention is to improve the credibility of research synthesis by facilitating a move from the level of detail found in journal articles to the level of

For peer review only

detail found in regulatory documents, thus guiding clinicians and other decision makers at all levels.

Methods

We obtained CSRs from public sources, as follows:

1. Requesting from EMA, under its freedom of information (FOI) policy, CSRs for manufacturer sponsored trials of the 10 best-selling prescription-bound products in the United States in 2010.[23]
2. Reusing CSRs from our own previous research[12]
3. Downloading CSRs openly available on the Internet.
4. Corresponding with other researchers who have obtained CSRs through FOI requests
5. Requesting manufacturers fill any gaps in the completeness of reports that we believe are legally required to be publicly available.

To create as broad a database as possible, we did not apply restrictions in drug type or family or sponsor. We did not submit requests under the Freedom of Information Act to the Food and Drug Administration because such requests can take years to be fulfilled and—once fulfilled—may be heavily redacted.[24]

We did not draw a random sample of CSRs as there is no known sampling frame. No one knows how many reports have been written by intervention category as there is no central register of CSRs. Through familiarity with CSRs for oseltamivir and zanamivir, which were included in one of our Cochrane reviews,[12] we developed and piloted a data extraction sheet designed to capture the salient characteristics of CSRs . We created a list of around 40 potential sections we expected to find, generated from elements specified in E3. For each element in the list, we checked whether the obtained CSR included that section (confirmed either by direct identification of the section or an indication the section existed based on the CSR's table of contents), whether we had access to it, and its page length. Because of previous difficulties we had accessing CSR appendices, we also recorded whether sections were listed as appendices or not. Page length was calculated either by directly counting the pages or by estimating their size from the table of contents of each report, and was used as a crude proxy for the level of detail available. Page lengths were rounded up to the next integer, and were summarized by reporting medians and ranges. We also included questions relating to trial registration and authorship. Our (blank) data extraction sheet is in Appendix 2.

All variables from CSRs were first extracted in single. We subsequently audited each other's extractions, checking the accurateness of the information. We chose to present elements

For peer review only

analogous with those that typically appear in trials reported in scientific journals including the study Synopsis (a brief summary of the study), the study Protocol (written prospectively, describing the study methods), Efficacy and Safety Evaluations (a narrative summary of the efficacy and safety results of the study, including tables and figures), as well as attached tables. We also included elements rarely found in journal publications: sample (blank) and completed case report forms (CRFs are paper or electronic forms designed to capture pre-specified efficacy and safety related information for each study participant), the statistical analysis plan (a prospectively written narrative and/or statistical code indicating how trial data will be analyzed), and individual participant efficacy and safety listings. The corresponding E3 section numbers are listed in Table 2. Disagreements were resolved by discussion.

Our uncorrected (original) and corrected extraction sheets as well as audit records are available upon request from the corresponding author.

We calculated a compression factor for published trials: the ratio of CSR page length compared to the page length of the same trial as published in scientific journals. Trial publications were searched for in multiple sources: clinical trial registers, published systematic reviews, and correspondence with sponsors. Because in most cases we could not access all parts of all CSRs, we calculated both “conservative” and “realistic” compression factors. “Conservative” compression factors were calculated using the total number of pages of CSRs available to us divided by the length of journal reports, while “realistic” compression factors were based on the true total page length of the CSR, when known, even if inaccessible.

Results

We identified 84 documents believed to be CSRs for 14 compounds. These covered therapeutic and biological interventions including antipsychotics, antidepressants, antivirals, natural antiarthritics, anti-inflammatory agents, pandemic influenza vaccines, statins, erythropoietins, and anti-platelet compounds. We included English-language summaries of two Japanese oseltamivir studies (JV15823, JV15824) as they had been presented to EMA in this form. We excluded CSRs which were too fragmentary to evaluate (olanzapine F1D-LC-HGAV, F1D-MC-HGAJ and F1D-MC-HGAO) and documents which were not in fact CSRs (reboxetine 14, 22 and 37). This left 78 CSRs (144,610 pages) (Figure 2). The median pages obtained per CSR was 644 (range 9 to 15,440). Only 4 of 78 CSRs (reboxetine 8, 16, 17, and 91) were written prior to November 30 1995 when ICH E3 was approved. Table 1 summarizes the pharmaceutical, manufacturer, date and provenance of the CSRs in our review. EMA reported not holding studies for esomepazole magnesium (Nexium), Advair diskus, quetiapine fumarate (Seroquel), montelukast sodium (Singulair), epoetin alfa (Epogen), and simvastatin.

For peer review only

All of the 78 included CSRs comprised of a synopsis (median page length 5 pages). The efficacy evaluation was identifiable and directly accessible in 76 (97%, median length 13.5 pages) and safety in 77 (99%; median length 17 pages). Attached tables were likewise present in 63 (81%) of CSRs, and were a median of 337 pages long (range: 1 to 3665). Seventy-three CSRs (94%) reported including the study protocol. In the 40 we could access, the median page length was 62. We found blank CRFs included in 68 (87%) of CSRs. Of the 33 we could directly access, the median length was 133 pages (range 14 to 981). For completed CRFs, 16 (21%) reports made direct mention of a section on completed CRFs, but we had access to completed CRFs in only 1 case (Arthronat; length 765 pages).

Fifty-five (71%) of 78 included CSRs included a statistical analysis plan in some form. Of those for which we could directly access the content (n=37), the median page length was 15 (range 3 to 85). Individual efficacy and safety listings were included in 53 (69%) and 62 (81%) CSRs respectively. The median page length was 447 (range 15 to 21,698) for efficacy and 109.5 (range 2 to 10,954) for safety.

A summary is presented in Table 2.

All trial reports in our review were sponsored by industry.

Median conservative compression factors ranged between 1 and 1221. The realistic compression factors, calculated for the Arthronat, paroxetine, and clopidogrel CAPRIE trials, were 379, 1021, and 8805, respectively. (Table 3)

Discussion

We collected and described a sizeable number of CSRs written in the last two decades. All CSRs contained a table of contents (as specified in E3 section 3); this, together with optical character recognition (to enable searching the full text of the scanned documents) and the occasional need to combine multiple files to create a single document, substantially improved the ease of navigating CSRs.

The future basic currency of research synthesis?

The median length of 644 pages for reports in this study confirms that CSRs are the most detailed and complete, integrated form of reporting of the design, conduct, and results of clinical trials. They far surpass the level of detail available in journal publications, and as such they are prime candidates for the next basic currency of evidence synthesis and appraisal of a trial. Given the EMA's new policy making such documents publicly available, access to these

For peer review only

documents is now relatively straightforward. However including CSRs in reviews is labor-intensive, given their size and complexity.[12]

Accessing complete CSRs

While CSRs may trump other forms of trial reporting in the public domain (such as conference abstracts or journal publications), serious limitations remain. Despite obtaining 144,610 pages for 78 CSRs, in almost all instances, we lacked full access to the CSRs' numerous appendices. Even for the sole complete CSR we obtained (Arthronat MA-CT-10-002) case report forms were provided for only 20% of participants. The text does not provide a reason for this omission, but it reflects the vagueness of the relevant section of the E3 guidance (16.3.2) which does not define "Other CRF's submitted." Also, we could only access the original trial protocol in 40 (51%) of 78 CSRs obtained. This is important because trial protocols, written prior to patient enrollment in a trial, are an important way to guard against reporting biases.[25,26]

We could obtain individual patient listings in only a minority of cases despite confirming their inclusion in the majority of CSRs (Table 2). This may be a significant limitation, as the E3 specifies that "the report with its appendices should also provide enough individual patient data, including the demographic and baseline data, and details of analytical methods, to allow replication of the critical analyses…"[19] Unavailability was possibly due to the fact that EMA allows manufacturers to submit CSRs omitting a number of appendices including individual patient data and case report forms (which EMA states should be available within 48 hours if requested).[27] In the case of oseltamivir, the subject of a Cochrane review we conducted,[12] the manufacturer refused to share with us report appendices not submitted to EMA,[28] and EMA declined requesting them on our behalf.[8] Although FDA likely possesses more complete CSRs and patient level data, it historically has treated such data as trade secret and/or confidential.[29–31] EMA is therefore at present the only reliable source of obtaining CSRs. As such, despite European regulators' progressive stance—announcing that "clinical trial data should not be considered commercial confidential information"[32]—the completeness gap is unlikely to be filled any time soon.

Individual participant listings

Individual participant listings—which identify participants by a unique ID—were accessible in 29 of the 78 CSRs we reviewed. But these data are difficult to analyze because they are presented as database printouts rather than in electronic form. This is understandable considering that CSRs are a written/archival format, but because EMA does not accept SAS datasets,[33,34] the industry standard, third-party access to databases of patient-level data remains elusive. We see no compelling reason why all regulators should not request these from sponsors and make

For peer review only

them publicly available. Whether availability of individual listings and CRFs, with its attendant laborious analysis, would increase our understanding of the trial and its results is unclear. But there is at least one case where the re-analysis of CRFs added invaluable knowledge to that already available in CSRs.[35]

Despite the apparent size of our non-random sample, we are not sure our conclusions are broadly generalisable to all other CSRs because we have extremely limited knowledge about the total population of CSRs in regulators' and sponsors' possession. Nevertheless, we found that the structure of CSRs was, within different house styles of presentation, strikingly similar across medical products and sponsors, probably thanks to ICH's E3.[36] This suggests that the structure and content of other CSRs is likely to be similar.

The public-private debate

One manufacturer has claimed that the non-release of case report forms is motivated by concerns over protecting participant confidentiality.[37] Nothing we have seen so far corroborates this claim. The EMA has deemed case report forms and individual patient listings to be, in principle, releasable in their entirety (after a preliminary review).[38] Furthermore, individual patient listings are intended to duplicate information contained in filled case report forms. The release of case report forms would ensure the accuracy of individual patient listings with little additional risk to patient confidentiality. Moreover, extra checks such as registration of protocols by bona fide research groups could deter any inappropriate use. We also believe that the sheer bulk of the forms acts as a deterrent against malice.

Size matters

Our range of compression factors show the scale of selection and synthesis which must (consciously or unconsciously) occur in the process of transforming CSRs into journal-length articles. We found a strong resemblance in detail, page length, structure, and purpose between the short Synopsis section of CSRs and reports of trials as published in scientific journals. . In some cases essential items of information such as the trial protocol and its subsequent amendments are simply not included in journal articles or are replaced by methods written post facto. In other cases of items essential for the interpretations of the trial results (such as the statistical analysis plan), tens of pages are reduced to a paragraph on sample size calculation in the journal report, underscoring the lack of detail (and its attendant problems) common to public forms of trial reporting. This is true even in databases not restricted by length, such as ClinicalTrials.gov.[39]

For peer review only

Our study raises the question of why the medical community has accepted the low (summary, aggregate) level of detail found in most peer-reviewed journal publications compared to the depth of detail available in CSRs. European regulators recently noted: "Documents that provide critical information on a study, such as the protocol (16.1.1), statistical methods (16.1.9), list of investigators and study sites and sample case report forms, would always be needed by reviewers assessing a study"[40] Why have those outside of the regulatory world tolerated journal publications lacking such details?

One possibility may be that while the clinical trial enterprise has changed dramatically in the last half century, the scientific journal publication model has not. Since the 1950s, there have been considerable transformations in the political economy of clinical trials driven by the increasingly commercialized and global nature of the pharmaceutical industry, the rise in academic-industry "partnerships" in medicine, and increased communication among regulators. It is now common to find trials with study centers scattered around the globe. This increasing complexity and the need to provide an audit record is reflected in the comprehensive tomes documenting the trials—CSRs—but trial reporting in scientific journals remains limited to summary and aggregate details.

Authorship or Contributorship?

Examination of CSRs revealed scores of important technical contributions to the design, conduct, and reporting of each trial. These included contributions from database programmers, records officers, and CSR writers, often invisible in the published journal article. In some cases, we found no mention in CSRs of individuals who figured as authors of subsequent published trial reports while individuals named as CSR authors went unacknowledged in journal publications. Current ICMJE guidelines on authorship and contributorship are largely focused on ensuring those placed on by-lines deserve to be authors. But the guidelines also suggest that "all contributors who do not meet the criteria for authorship should be listed in an acknowledgments section."[41] Given the complexity of clinical trials, the ICMJE should call for itemized contributorship: the names of all contributors to be specified along with their role in the design, conduct, analysis, or reporting of the trial. If the contribution of most people goes unrecorded, so does their individual responsibility for what is produced. Itemized contributorship records, to all phases of a trial, could be piloted in trial registers.

E3 guidance

The E3 guideline set an excellent standard, but it needs formal updating and further development. For example, there should be a self-standing set of definitions for terms such as "case report forms" and "Other CRF's submitted," (section 16.3.2) and a description of how a

For peer review only

particular trial fits within a sponsor's trial programme of pharmaceutical development. Apparently forgotten items such as certificates of analysis (describing the appearance and content of the interventions being tested) and post-1995 details such as trial registration numbers should be mentioned.

We hope our review has given CSRs what they have lacked so far: visibility. CSRs represent a largely untapped source of detailed data that we believe can serve as a means of addressing the ravages of reporting bias in all its forms, leading to a more accurate understanding of the effects of medicines.

Conflicts of interest statement

All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare that:

Both authors are co-recipients of a UK National Institute for Health Research grant to carry out a Cochrane review of neuraminidase inhibitors (http://www.hta.ac.uk/2352).

Tom Jefferson was an ad hoc consultant for F. Hoffman-La Roche Ltd in 1998-1999. He receives royalties from his books published by Blackwells and Il Pensiero Scientifico Editore, none of which are on clinical study reports. He is occasionally interviewed by market research companies for anonymous interviews about Phase 1 or 2 products unrelated to products in this review. In 2011-12 he has acted as an expert witness in a litigation case related to one of the compounds in the review (oseltamivir). He is on a legal retainer for expert advice on litigation for influenza vaccines in health care workers.

Peter Doshi received €1500 from the European Respiratory Society in support of his travel to the society's September 2012 annual congress where he gave an invited talk on oseltamivir.

Both authors' spouses and children have no financial relationships that may be relevant to the submitted work.

Data sharing statement

The original extraction forms and audit record are available on request from the corresponding author.

For peer review only

Acknowledgements

We thank Drs Vallance and Kraus of GlaxoSmithKline for making public selected report appendices from the 9 paroxetine trials. We also thank Daniel Coyne for sharing the CSR that FDA sent him in response to his Freedom of Information request, and Iain Chalmers for guidance.

Peter Doshi is funded by an institutional training grant from the Agency for Healthcare Research and Quality #T32HS019488.

Figure Legends:

Figure 1. Types of clinical trial data typically held within and transferred between three realms: trial sponsor, regulatory, and public.

Figure 2. Study flow

For peer review only

References

1. Rising K, Bacchetti P, Bero L. Reporting Bias in Drug Trials Submitted to the Food and Drug Administration: Review of Publication and Presentation. Plos Med. 2008;5(11):e217.

2. Doshi P, Jones M, Jefferson T. Rethinking credible evidence synthesis. BMJ. 2012 Jan 17;344(jan17 2):d7898.

3. Chan A-W, Hróbjartsson A, Haahr MT, et al. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. JAMA. 2004 May 26;291(20):2457–2465.

4. Gøtzsche PC. Why we need easy access to all data from all clinical trials and how to accomplish it. Trials. 2011;12(1):249.

5. Turner EH, Matthews AM, Linardatos E, et al. Selective publication of antidepressant trials and its influence on apparent efficacy. N. Engl. J. Med. 2008 Jan 17;358(3):252–260.

6. Dickersin K, Min YI. Publication bias: the problem that won't go away. Ann. N. Y. Acad. Sci. 1993 Dec 31;703:135–146; discussion 146–148.

7. Hopewell S, Loudon K, Clarke MJ, et al. Publication bias in clinical trials due to statistical significance or direction of trial results. Cochrane Database Syst Rev. 2009;(1):MR000006.

8. Melander H. Evidence b(i)ased medicine--selective reporting from studies sponsored by pharmaceutical industry: review of studies in new drug applications. BMJ. 2003 May 29;326:1171–1173.

9. McGauran N, Wieseler B, Kreis J, et al. Reporting bias in medical research - a narrative review. Trials. 2010;11(1):37.

10. Sismondo S, Doucet M. Publication ethics and the ghost management of medical publication. Bioethics. 2009 Feb;24(6):273–283.

11. Vedula SS, Bero L, Scherer RW, et al. Outcome reporting in industry-sponsored trials of gabapentin for off-label use. N. Engl. J. Med. 2009 Nov 12;361(20):1963–1971.

12. Jefferson T, Jones MA, Doshi P, et al. Neuraminidase inhibitors for preventing and treating influenza in healthy adults and children. Cochrane Database of Systematic Reviews. 2012;(1):CD008965.

13. European Medicines Agency. European Medicines Agency policy on access to documents (related to medicinal products for human and veterinary use) POLICY/0043 [Internet]. 2010 Nov 30 [cited 2012 May 14];Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Other/2010/11/WC500099473.pdf

14. European Medicines Agency. Output of the European Medicines Agency policy on access to documents related to medicinal products for human and veterinary use [Internet]. 2010 Nov 30 [cited 2011 Sep 25];Available from:

For peer review only

http://www.ema.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2010/11/WC500099472.pdf

15. Gøtzsche PC, Jørgensen AW. Opening up data at the European Medicines Agency. BMJ. 2011;342:d2686.

16. Jefferson T, Doshi P, Thompson M, et al. Ensuring safe and effective drugs: who can do what it takes? BMJ. 2011;342:c7258.

17. Chan A-W. Out of sight but not out of mind: how to search for unpublished clinical trial evidence. BMJ. 2012 Jan 3;344(jan03 2):d8013–d8013.

18. Rodwin MA, Abramson JD. Clinical trial data as a public good. JAMA. 2012 Sep 5;308(9):871–872.

19. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. Structure and Content of Clinical Study Reports: E3 [Internet]. 1995 Nov 30 [cited 2012 Jul 8];Available from: http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E3/E3_Guideline.pdf

20. Grens K. Data Diving. The Scientist [Internet]. 2012 Apr [cited 2012 May 3];Available from: http://the-scientist.com/2012/05/01/data-diving/

21. Eyding D, Lelgemann M, Grouven U, et al. Reboxetine for acute treatment of major depression: systematic review and meta-analysis of published and unpublished placebo and selective serotonin reuptake inhibitor controlled trials. BMJ. 2010 Oct 12;341(oct12 1):c4737–c4737.

22. Wieseler B, Kerekes MF, Vervoelgyi V, et al. Impact of document type on reporting quality of clinical drug trials: a comparison of registry reports, clinical study reports, and journal publications. BMJ. 2012 Jan 3;344(jan03 1):d8141–d8141.

23. IMS Institute for Healthcare Informatics. The Use of Medicines in the United States: Review of 2010 [Internet]. 2011 Apr [cited 2012 Jul 30];Available from: http://www.imshealth.com/imshealth/Global/Content/IMS%20Institute/Documents/IHII_UseOfMed_report%20.pdf

24. Coyne DW. The health-related quality of life was not improved by targeting higher hemoglobin in the Normal Hematocrit Trial. Kidney International [Internet]. 2012 Mar 21 [cited 2012 Apr 13];Available from: http://www.nature.com/ki/journal/vaop/ncurrent/full/ki201276a.html

25. Chan A-W. Bias, Spin, and Misreporting: Time for Full Access to Trial Protocols and Results. Plos Med. 2008;5(11):e230.

26. Miller JD. Registering clinical trial results: the next step. JAMA. 2010 Feb 24;303(8):773–774.

27. European Medicines Agency. Note for guidance on the inclusion of appendices to clinical study reports in marketing authorisation applications [Internet]. 2004 Jun 23 [cited 2012 Jul

For peer review only

30];Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003638.pdf

28. Doshi P, Jefferson T, Del Mar C. The Imperative to Share Clinical Study Reports: Recommendations from the Tamiflu Experience. PLoS Med. 2012 Apr 10;9(4):e1001201.

29. Unlu M. It is time: why the FDA should start disclosing drug trial data. Mich. Telecomm. Tech. L. Rev. 2010;16:511–511.

30. Kesselheim AS, Mello MM. Confidentiality laws and secrecy in medical research: improving public access to data on drug safety. Health Aff (Millwood). 2007 Apr;26(2):483–491.

31. Halperin RM. FDA Disclosure of Safety and Effectiveness Data: A Legal and Policy Analysis. Duke Law Journal. 1979 Feb 1;1979(1):286–326.

32. Eichler H-G, Abadie E, Breckenridge A, Leufkens H, Rasi G. Open Clinical Trial Data for All? A View from Regulators. PLoS Med. 2012 Apr 10;9(4):e1001202.

33. European Medicines Agency. EMEA implementation of electronic-only submission and eCTD submission: Practical guidelines relating to non-eCTD electronic submissions [Internet]. 2008 Dec [cited 2012 Jul 31];Available from: http://www.emea.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2009/10/WC500004100.pdf

34. European Medicines Agency. EMEA implementation of electronic-only submission and eCTD submission: Questions and answers relating to practical and technical aspects of the implementation [Internet]. 2008 Dec [cited 2012 Jul 31];Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2009/10/WC500004095.pdf

35. Psaty BM, Prentice RL. Minimizing bias in randomized trials: the importance of blinding. JAMA. 2010 Aug 18;304(7):793–794.

36. Molzon JA, Giaquinto A, Lindstrom L, et al. The value and benefits of the International Conference on Harmonisation to drug regulatory authorities: advancing harmonization for better public health. Clin. Pharmacol. Ther. 2011 Apr;89(4):503–512.

37. GlaxoSmithKline. Paroxetine and pediatric and adolescent patients [Internet]. 2012 Jul 25 [cited 2012 Jul 30];Available from: http://www.gsk.com/media/paroxetine.htm

38. HMA/EMA Working Group on Transparency. HMA/EMA guidance document on the identification of commercially confidential information and personal data within the structure of the marketing authorisation (MA) application - release of information after the granting of a marketing authorisation [Internet]. 2012 Mar 14 [cited 2012 Jul 30];Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Other/2012/03/WC500124536.pdf

39. Zarin DA, Tse T, Williams RJ, et al. The ClinicalTrials.gov Results Database — Update and Key Issues. New England Journal of Medicine. 2011;364(9):852–860.

For peer review only

40. European Medicines Agency. ICH guideline E3 - questions and answers (R1) [Internet]. 2012 Jul [cited 2012 Jul 30];Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2012/07/WC500130356.pdf

41. International Committee of Medical Journal Editors. Uniform Requirements for Manuscripts Submitted to Biomedical Journals: Ethical Considerations in the Conduct and Reporting of Research: Authorship and Contributorship [Internet]. 2009 [cited 2012 Jul 31];Available from: http://www.icmje.org/ethical_1author.html

For peer review only

Table 1. Pharmaceutical, trials, producers, dates and sources of CSRs in the review.

Pharmaceutical and number (n) of assessed trial documents	Trial IDs	Manufacturer	Date of CSRs	Provenance in our study
Aripiprazole (Abilify) n=1	CN1368135	Bristol-Myers Squibb	2007	Freedom of Information request to EMA
Arthronat n=1	MA-CT-10-002	Rowtasha	2011	Manufacturer website http://arthronat.com/clinical-study.php
Atorvastatin (Lipitor) n=1	981-080	Pfizer	1999	Freedom of Information request to EMA
Clopidogrel (Plavix) n=5	CURE, CLARITY, COMMIT-CCS2, CAPRIE, PICOLO	Bristol-Myers Squibb	1997-2007	Freedom of Information request to EMA
Epoetin alfa (Epogen) n=1	930107	Amgen	1996	Freedom of Information request to FDA
H5N1 influenza vaccine n=1	H5N1-008, H5N1-011 EXT 008	GSK	2006	Freedom of Information request to EMA
H5N1 influenza vaccines n=2	V87P1, V87P6	Novartis	2008-2009	Freedom of Information request to EMA
Olanzapine (Zyprexa) n=3	F1D-LC-HGAV*, F1D-MC-HGAO*, F1D-MC-HGAJ*	Eli Lilly	1995†	Litigation http://zyprexalitigationdocuments.com/unsealed.php http://www.furiousseasons.com/zyprexadocs.html
Oseltamivir (Tamiflu) n=19	JV15823, JV15824, M76001, NP15757, NV16871, WP16263, WV15670, WV15671, WV15673 WV15697, WV15707, WV15708, WV15730, WV15758, WV15759 WV15871, WV15799, WV15812 WV15872, WV15819 WV15876 WV15978, WV15825, WV16193	Roche	1999-2004	Documents obtained as part of previous Cochrane review[12]
Paroxetine (Paxil, Aropax, Pexeva, Seroxat, Sereupin) n=9	329, 377, 453, 511, 676, 701, 704, 715, 716	GSK	1998-2002	Litigation (2004 legal settlement mandated release of clinical study reports on manufacturer's website of 9 studies on

For peer review only

				pediatric and adolescent patients) http://www.gsk.com/media/paroxetine.htm
Quetiapine (Seroquel) n=7	015, 041, 049, 125, 126, 127, 135	AstraZeneca	1996-2007	Litigaton http://psychrights.org/research/Digest/NLPs/Seroquel/UnsealedSeroquelStudies/
Reboxetine (Edronax, Norebox, Prolift, Solvex, Davedax, Vestra) n=24	8, 9, 13, 14*, 15, 16, 17, 22*, 32, 32a, 34, 35, 37*, 43, 45, 46, 47, 49, 50, 52, 71, 83, 91, 96	Pfizer	1991-2009	Health Technology Assesesment website (The German IQWiG obtained CSRs as part of its health technology assessment work) https://www.iqwig.de/information-on-studies-of-reboxetine.980.en.html
Rofecoxib (Vioxx) n=1	78	Merck	2003	Litigation http://dida.library.ucsf.edu/
Zanamivir (Relenza) n=9	NAI30009, NAI300010, NAIA2005, NAIA3002, NAIA3005, NAIB2005, NAIB2007, NAIB3001, NAIB3002	GSK	1998-1999	Documents obtained as part of previous Cochrane review[12]

483 * Subsequently excluded because of insufficient documentation

484 † H1D-MC-HGAO clinical study report date unknown

485 EMA = European Medicines Agency

486 FDA = Food and Drug Administration

487

For peer review only

Table 2. Key characteristics of the CSRs in the review

Section of CSR (corresponding section of E3)	Presence	Length	
	CSRs including section, n	CSRs with section length available, n	Median length (range), pages
Synopsis (E3 section 2)	78 (100%)	78	5 (1 - 15)
Efficacy evaluation (E3 sec. 11)	76 (97%)	77	13.5 (2 - 132)
Safety evaluation (E3 sec. 12)	77 (99%)	58	17 (2 - 188)
Attached tables not in report text (E3 sec. 14)	63 (81%)	76	337 (1 - 3665)
Protocol (E3 sec 16.1.1)	73 (94%)	41	62 (21 - 139)
Blank Case Report Form (CRF) (E3 sec. 16.1.2)	68 (87%)	33	133 (14 - 981)
Statistical Analysis Plan (E3 sec. 16.1.9)	55 (71%)	37	15 (3 - 85)
Individual participant efficacy listings (E3 sec. 16.2.6)	53 (69%)	19	447 (15 - 21698)
Individual participant safety listings (E3 sec. 16.2.7)	62 (81%)	26	109.5 (2 - 10954)
Completed CRFs (E3 sec. 16.3.2)	16 (21%)	1	765

For peer review only

Table 3. Conservative and realistic compression factors. A ratio of CSR page length to corresponding journal publication page length.

Pharmaceutical	Studies published in journals, n	Mean compression factor (range)
Conservative compression factors		
Aripiprazole	1	672
Clopidogrel	5	11 (4 - 19)
Epoetin Alfa	1	41
Fluad	2	488 (367 - 609)
GSK H5N1 vaccine	1	19
Oseltamivir	12	195 (1 - 1221)
Quetiapine	2	578 (352 - 803)
Reboxetine	5	88 (9 - 245)
Zanamivir	8	54 (28 - 92)
Realistic compression factors		
Arthronat*	1	379
Clopidogrel	1	8805
Paroxetine	9	1021 (50 - 5473)

* The Arthronat trial has not yet been published. Compression factor calculation is based on the page length of a draft manuscript "to be published soon," according to Arthronat.com.

For peer review only

Appendix 1. Elements specified ICH E3 "Structure and Content of Clinical Study Reports" (1995)[19]

1. TITLE PAGE
2. SYNOPSIS
3. TABLE OF CONTENTS FOR THE INDIVIDUAL CLINICAL STUDY REPORT
4. LIST OF ABBREVIATIONS AND DEFINITION OF TERMS
5. Ethics
    - 5.1. Independent Ethics Committee (IEC) or Institutional Review Board (IRB)
    - 5.2. Ethical conduct of the study
    - 5.3. Patient information and consent
6. INVESTIGATORS AND STUDY ADMINISTRATIVE STRUCTURE
7. INTRODUCTION
8. STUDY OBJECTIVES
9. INVESTIGATIONAL PLAN
    - 9.1. Overall study design and plan – description
    - 9.2. Discussion of study design, including the choice of control groups
    - 9.3. Selection of study population
        - 9.3.1. Inclusion criteria
        - 9.3.2. Exclusion criteria
        - 9.3.3. Removal of Patients from Therapy or Assessment
    - 9.4. Treatments
        - 9.4.1. Treatments Administered
        - 9.4.2. Identity of Investigational Product(s)
        - 9.4.3. Method of Assigning Patients to Treatment Groups
        - 9.4.4. Selection of Doses in the Study
        - 9.4.5. Blinding
        - 9.4.6. Prior and Concomitant Therapy
        - 9.4.7. Treatment Compliance
    - 9.5. Efficacy and safety variables
        - 9.5.1. Efficacy and Safety Measurements Assessed and Flow Chart
        - 9.5.2. Appropriateness of Measurements
        - 9.5.3. Primary Efficacy Variable(s)
        - 9.5.4. Drug Concentration Measurements
    - 9.6. Data quality assurance
    - 9.7. Statistical methods planned in the protocol and determination of sample size

For peer review only

9.7.1. Statistical and Analytical Plans
9.7.2. Determination of Sample Size
9.8. Changes in the conduct of the study or planned analyses
10. STUDY PATIENTS
10.1. Disposition of patients
10.2. Protocol deviations
11. EFFICACY EVALUATION
11.1. Data sets analyzed
11.2. Demographic and other baseline characteristics
11.3. Measurements of treatment compliance
11.4. Efficacy results and tabulations of individual patient data
11.4.1. Analysis of efficacy
11.4.2. Statistical/analytical issues
11.4.2.1. Adjustments for covariates
11.4.2.2. Handling of Dropouts or Missing Data
11.4.2.3. Interim Analyses and Data Monitoring
11.4.2.4. Multicentre Studies
11.4.2.5. Multiple Comparison/Multiplicity
11.4.2.6. Use of an "Efficacy Subset" of Patients
11.4.2.7. Active-Control Studies Intended to Show Equivalence
11.4.2.8. Examination of Subgroups
11.4.3. Tabulation of Individual Response Data
11.4.4. Drug Dose, Drug Concentration, and Relationships to Response
11.4.5. Drug-Drug and Drug-Disease Interactions
11.4.6. Drug Dose, Drug Concentration, and Relationships to Response
11.4.7. By-Patient Displays
12. SAFETY EVALUATION
12.1. Extent of exposure
12.2. Adverse events (AES)
12.2.1. Brief Summary of Adverse Events
12.2.2. Display of Adverse Events
12.2.3. Analysis of Adverse Events
12.2.4. Listing of Adverse Events by Patient
12.3. Deaths, other Serious Adverse Events and Other Significant Adverse Events

For peer review only

12.3.1. Listing of Deaths, other Serious Adverse Events and Other Significant Adverse Events
12.3.1.1. Deaths
12.3.1.2. Other Serious Adverse Events
12.3.1.3. Other Significant Adverse Events
12.3.2. Narratives of Deaths, Other Serious Adverse Events and Certain Other Significant Adverse Events
12.3.3. Analysis and Discussion of Deaths, Other Serious Adverse Events and Other Significant Adverse Events
12.4. Clinical laboratory evaluation
12.4.1. Listing of Individual Laboratory Measurements by Patient (16.2.8) and Each Abnormal Laboratory Value (14.3.4)
12.4.2. Evaluation of Each Laboratory Parameter
12.4.2.1. Laboratory Values Over Time
12.4.2.2. Individual Patient Changes
12.4.2.3. Individual Clinically Significant Abnormalities
12.5. Vital signs, physical findings and other observations related to safety
12.6. Safety conclusions
13. DISCUSSION AND OVERALL CONCLUSIONS
14. TABLES, FIGURES, AND GRAPHS REFERRED TO BUT NOT INCLUDED IN THE TEXT
14.1. Demographic data
14.2. Efficacy data
14.3. Safety data
14.3.1. Displays of Adverse Events
14.3.2. Listings of Deaths, Other Serious and Significant Adverse Events
14.3.3. Narratives of Deaths, Other Serious and Certain Other Significant Adverse Events
14.3.4. Abnormal Laboratory Value Listing (Each Patient)
15. REFERENCE LIST
16. APPENDICES
16.1. Study Information
16.1.1. Protocol and protocol amendments
16.1.2. Sample case report form (unique pages only)

For peer review only

16.1.3. List of IECs or IRBs (plus the name of the committee Chair if required by the regulatory authority) - Representative written information for patient and sample consent forms

16.1.4. List and description of investigators and other important participants in the study, including brief (1 page) CVs or equivalent summaries of training and experience relevant to the performance of the clinical study

16.1.5. Signatures of principal or coordinating investigator(s) or sponsor's responsible medical officer, depending on the regulatory authority's requirement

16.1.6. Listing of patients receiving test drug(s)/investigational product(s) from specific batches, where more than one batch was used

16.1.7. Randomisation scheme and codes (patient identification and treatment assigned)

16.1.8. Audit certificates (if available)

16.1.9. Documentation of statistical methods

16.1.10. Documentation of inter-laboratory standardisation methods and quality assurance procedures if used

16.1.11. Publications based on the study

16.1.12. Important publications referenced in the report

16.2. Patient Data Listings

16.2.1. Discontinued patients

16.2.2. Protocol deviations

16.2.3. Patients excluded from the efficacy analysis

16.2.4. Demographic data

16.2.5. Compliance and/or drug concentration data (if available)

16.2.6. Individual efficacy response data

16.2.7. Adverse event listings (each patient)

16.2.8. Listing of individual laboratory measurements by patient, when required by regulatory authorities

16.3. Case Report Forms

16.3.1. CRFs for deaths, other serious adverse events and withdrawals for AE

16.3.2. Other CRFs submitted

16.4. Individual Patient Data Listings (US Archival Listings)

For Peer review only

Basic Extraction Information

Questions	Answer	Notes
1. Drug common name:		
2. Trial ID:		
➔ Now, fill in the drug and trial ID in the bottom-right corner the page.	E.g. "Tamiflu, WV15670"	
➔ Now, save this file under a new filename	Use the naming convention "*Drugname Trial ID - Extractor's initials - YYYYMMDD.docx*", e.g. "Seroquel 015 - TJ - 20120311.docx"	
3. Report/CSR ID (if different from Trial ID):		
4. Extractor's name (Initials)		
5. Date of extraction		

Notes to extractor:

- Page numbers should be referred to by the format p.(page # as printed)/PDFp.(PDF page number, possibly indicating volume), e.g.
  - p.V-235/PDFp.945 = page "V-235", on PDF page 945
  - p.234/PDF(3)p.18 = page "234", on the 3rd PDF for this CSR, PDF page 18
- Most questions can be answered with a Y or N (indicating Yes or No) or a number (e.g. the number of PDF pages.
- Where specified as "Free form answer", the extractor may answer in his/her own words based on the extractor's reading of the CSR.

Item	Content	Notes
Overview questions 6. Does the CSR list a ISRCTN/NCT or equivalent registration number for this trial?		
7. List CSR number of authors		
8. List CSR authors & trialists (Copy names if available; "redacted" if redacted; "not listed" if not listed)		
9. Total length of CSR obtained, in PDF pages		
10. List CSR completion date		
11. Is the trial published?		
12. If Y give publication citation		
13. If Y give publication size (in pages)		
14. Who appears to be responsible for CSR? (Free form answer)		
Trial programme questions 15. How many trials appear to be in the trial programme?		
16. Does CSR indicate where this trial fits in the trial programme? (Free form answer)		
17. Does CSR say how much of the trial programme is published?		
18. How many trials are in possession of a ISRCTN/NCT or equivalent registration number?		
Basic elements of the Clinical Study Report		

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

19.	Does the CSR contain a **table of contents**?		
20.	If Y, is the **table of contents** listed as an Appendix?		
21.	If Y, is the **table of contents** accessible to us?		
22.	If Y, how long is the **table of contents** (in pages)?		
23.	Does the table of contents list a **title page**?		
24.	If Y, is the **title page** listed as an Appendix?		
25.	If Y, is the **title page** accessible to us?		
26.	If Y, how long is the **title page** (in pages)?		
27.	Does the table of contents list a **synopsis**?		
28.	If Y, is the **synopsis** listed as an Appendix?		
29.	If Y, is the **synopsis** accessible to us?		
30.	If Y, how long is the **synopsis** (in pages)?		
31.	Does the CSR contain a **list of abbreviations and definitions**?		
32.	If Y, is the **list of abbreviations and definitions** listed as an Appendix?		
33.	If Y, is the **list of abbreviations and definitions** accessible to us?		
34.	If Y, how long is the **list of abbreviations and definitions** (in pages)?		
35.	Does the CSR contain an **ethics section**?		
36.	If Y, is the **ethics section** listed as an Appendix?		
37.	If Y, is the **ethics section** accessible to us?		
38.	If Y, how long is the **ethics section** (in pages)?		
39.	Does the CSR contain a **investigators and study administrative structure**?		
40.	If Y, is the **investigators and study administrative structure** listed as an Appendix?		
41.	If Y, is the **investigators and study administrative structure** accessible to us?		
42.	If Y, how long is the **investigators and study administrative structure** (in pages)?		
43.	Does the CSR contain an **introduction**?		
44.	If Y, is the **introduction** listed as an Appendix?		
45.	If Y, is the **introduction** accessible to us?		
46.	If Y, how long is the **introduction** (in pages)?		
47.	Does the CSR contain a section on study objectives?		
48.	If Y, is the **section on study objectives** listed as an Appendix?		
49.	If Y, is the **section on study objectives** accessible to us?		
50.	If Y, how long is the **section on study objectives** (in pages)?		
51.	Does the CSR contain an **investigational plan** (from IHR 1995 E3, PDF p.13)?		
52.	If Y, is the **investigational plan** listed as an Appendix?		
53.	If Y, is the **investigational plan** accessible to us?		
54.	If Y, how long is the **investigational plan** (in pages)?		
55.	Does the CSR contain a section on **study patients**?		
56.	If Y, is the **study patients** listed as an Appendix?		
57.	If Y, is the **study patients** accessible to us?		
58.	If Y, how long is the **study patients** (in pages)?		

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

59.	If Y, does it include a list of protocol deviations?		
60.	Does the CSR contain a section on **efficacy evaluation**?		
61.	If Y, is the **efficacy evaluation** listed as an Appendix?		
62.	If Y, is the **efficacy evaluation** accessible to us?		
63.	If Y, how long is the **efficacy evaluation** (in pages)?		
64.	Does the CSR contain a section on **safety evaluation**?		
65.	If Y, is the **safety evaluation** listed as an Appendix?		
66.	If Y, is the **safety evaluation** accessible to us?		
67.	If Y, how long is the **safety evaluation** (in pages)?		
68.	Does the CSR contain a **discussion and overall conclusions** section?		
69.	If Y, is the **discussion and overall conclusions** listed as an Appendix?		
70.	If Y, is the discussion and overall conclusions accessible to us?		
71.	If Y, how long is the **discussion and overall conclusions** (in pages)?		
72.	Does the CSR contain a section on **tables, figures and graphs referred to but not included in the text**?		
73.	If Y, is the **tables, figures and graphs referred to but not included in the text** listed as an Appendix?		
74.	If Y, is the **tables, figures and graphs referred to but not included in the text** accessible to us?		
75.	If Y, how long is the **tables, figures and graphs referred to but not included in the text** (in pages)?		
76.	Does the CSR contain a **references** section?		
77.	If Y, is the **references** listed as an Appendix?		
78.	If Y, is the **references** accessible to us?		
79.	If Y, how long is the **references** (in pages)?		
Appendices related questions			
80.	Does the **table of contents** indicate that the CSR contains **appendices**?		
81.	If Y, does the **table of contents** list the **titles of the appendices**?		
82.	Does the CSR include the **study Protocol**?		
83.	If Y, is the **study Protocol** accessible to us?		
84.	If Y, how long is the **study Protocol** (in pages)?		
85.	Does the CSR contain a section on **Protocol amendments**?		
86.	If Y, is the section on **Protocol amendments** accessible to us?		
87.	If Y, how long is the section on **Protocol amendments** (in pages)?		
88.	Does the CSR contain a section on **Sample case report form (unique pages only)**?		
89.	If Y, is the section on **Sample case report form (unique pages only)** accessible to us?		
90.	If Y, how long is the section on **Sample case report form (unique pages only)** (in pages)?		
91.	Does the CSR contain a section on **List of IECs or IRBs (plus the name of the committee Chair if required by the**		

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

	regulatory authority) - Representative written information for patient and sample consent forms?		
92.	If Y, is the section on **List of IECs or IRBs (plus the name of the committee Chair if required by the regulatory authority) - Representative written information for patient and sample consent forms** accessible to us?		
93.	If Y, how long is the section on **List of IECs or IRBs (plus the name of the committee Chair if required by the regulatory authority) - Representative written information for patient and sample consent forms** (in pages)?		
94.	Does the CSR contain a section on **List and description of investigators and other important participants in the study, including brief (1 page) CVs or equivalent summaries of training and experience relevant to the performance of the clinical study**?		
95.	If Y, is the section on **List and description of investigators and other important participants in the study, including brief (1 page) CVs or equivalent summaries of training and experience relevant to the performance of the clinical study** accessible to us?		
96.	If Y, how long is the section on **List and description of investigators and other important participants in the study, including brief (1 page) CVs or equivalent summaries of training and experience relevant to the performance of the clinical study** (in pages)?		
97.	Does the CSR contain a section on **Signatures of principal or coordinating investigator(s) or sponsor's responsible medical officer, depending on the regulatory authority's requirement**?		
98.	If Y, is the section on **Signatures of principal or coordinating investigator(s) or sponsor's responsible medical officer, depending on the regulatory authority's requirement** accessible to us?		
99.	If Y, how long is the section on **Signatures of principal or coordinating investigator(s) or sponsor's responsible medical officer, depending on the regulatory authority's requirement** (in pages)?		
100.	Does the CSR contain a section on **Listing of patients receiving test drug(s)/investigational product(s) from specific batches, where more than one batch was used**?		
101.	If Y, is the section on **Listing of patients receiving test drug(s)/investigational product(s) from specific batches, where more than one batch was used** accessible to us?		
102.	If Y, how long is the section on **Listing of patients receiving test drug(s)/investigational product(s) from specific batches, where more than one batch was used** (in pages)?		
103.	Does the CSR contain a section on **Randomisation scheme and codes (patient identification and treatment assigned)**?		
104.	If Y, is the section on **Randomisation scheme and codes**		

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

	(patient identification and treatment assigned) accessible to us?		
105.	If Y, how long is the section on **Randomisation scheme and codes (patient identification and treatment assigned)** (in pages)?		
106.	Does the CSR contain a section on **Audit certificates (if available) (see Annex IVa and IVb of the guideline)**?		
107.	If Y, is the section on **Audit certificates (if available) (see Annex IVa and IVb of the guideline)** accessible to us?		
108.	If Y, how long is the section on **Audit certificates (if available) (see Annex IVa and IVb of the guideline)** (in pages)?		
109.	Does the CSR contain a section on **Documentation of statistical methods**?		
110.	If Y, is the section on **Documentation of statistical methods** accessible to us?		
111.	If Y, how long is the section on **Documentation of statistical methods** (in pages)?		
112.	If Y, is the **Documentation of statistical methods** dated?		
113.	If Y, what is the date of the **Documentation of statistical methods**?		
114.	Does the CSR contain a section on **Documentation of inter-laboratory standardisation methods and quality assurance procedures if used**?		
115.	If Y, is the section on **Documentation of inter-laboratory standardisation methods and quality assurance procedures if used** accessible to us?		
116.	If Y, how long is the section on **Documentation of inter-laboratory standardisation methods and quality assurance procedures if used** (in pages)?		
117.	Does the CSR contain a section on **Publications based on the study**?		
118.	If Y, is the section on **Publications based on the study** accessible to us?		
119.	If Y, how long is the section on **Publications based on the study** (in pages)?		
120.	Does the CSR contain a section on **Important publications referenced in the report**?		
121.	If Y, is the section on **Important publications referenced in the report** accessible to us?		
122.	If Y, how long is the section on **Important publications referenced in the report** (in pages)? Edfgyh+		
123.	Does the CSR contain a section on **Discontinued patients**?		
124.	If Y, is the section on **Discontinued patients** accessible to us?		
125.	If Y, how long is the section on **Discontinued patients** (in pages)?		
126.	Does the CSR contain a section on **Protocol deviations**?		
127.	If Y, is the section on **Protocol deviations** accessible to us?		
128.	If Y, how long is the section on **Protocol deviations** (in		

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

	pages)?		
129.	Does the CSR contain a section on **Patients excluded from the efficacy analysis**?		
130.	If Y, is the section on **Patients excluded from the efficacy analysis** accessible to us?		
131.	If Y, how long is the section on **Patients excluded from the efficacy analysis** (in pages)?		
132.	Does the CSR contain a section on **Demographic data**?		
133.	If Y, is the section on **Demographic data** accessible to us?		
134.	If Y, how long is the section on **Demographic data** (in pages)?		
135.	Does the CSR contain a section on **Compliance and/or drug concentration data (if available)**?		
136.	If Y, is the section on **Compliance and/or drug concentration data (if available)** accessible to us?		
137.	If Y, how long is the section on **Compliance and/or drug concentration data (if available)** (in pages)?		
138.	Does the CSR contain a section on **Individual efficacy response data**?		
139.	If Y, is the section on **Individual efficacy response data** accessible to us?		
140.	If Y, how long is the section on **Individual efficacy response data** (in pages)?		
141.	Does the CSR contain a section on Adverse event listings (each patient)?		
142.	If Y, is the section on **Adverse event listings (each patient)** accessible to us?		
143.	If Y, how long is the section on **Adverse event listings (each patient)** (in pages)?		
144.	Does the CSR contain a section on **Listing of individual laboratory measurements by patient, when required by regulatory authorities**?		
145.	If Y, is the section on **Listing of individual laboratory measurements by patient, when required by regulatory authorities** accessible to us?		
146.	If Y, how long is the section on **Listing of individual laboratory measurements by patient, when required by regulatory authorities** (in pages)?		
147.	Does the CSR contain a section on **Case Report Forms for deaths, other serious adverse events and withdrawals for AE**?		
148.	If Y, is the section on **Case Report Forms for deaths, other serious adverse events and withdrawals for AE** accessible to us?		
149.	If Y, how long is the section on **Case Report Forms for deaths, other serious adverse events and withdrawals for AE** (in pages)?		
150.	Does the CSR contain a section on **Other Case Report Forms** submitted?		
151.	If Y, is the section on **Other Case Report Forms submitted** accessible to us?		
152.	If Y, how long is the section on **Other Case Report Forms**		

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

	submitted (in pages)?		
153.	Does the CSR contain a section on **Individual patient data listings**?		
154.	If Y, is the section on **Individual patient data listings** accessible to us?		
155.	If Y, how long is the section on **Individual patient data listings** (in pages)?		

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Mllner, Marcus

General comments

Is this new? “Yes but, no but, yes but…” (Vicky Pollard in most likely every show of Little Brittain).

Everybody who ever saw a clinical study report (CRS) – several meters of printed paper or many mega-bites of data, usually too large to be sent via regular email – just knows what this review shows. And regulators frequently work with clinicians and academics, provided they have no conflicts of interest. Likewise, everybody who works with an ethics committee reviews industry sponsored study protocols covering up to hundreds of pages. And it seems that nobody ever was greatly surprised that these pages are then compressed to a half-page maximum (which the majority doesn’t even want to read). To a certain extent this was hidden in plain sight. This paper makes a blind spot visible – well done!

I particularly like the discussion and my feeling is that very important points are not reflected in the “what this study adds” section. In particular I am referring to the point on authors and contributors (and their ghosts) but also to a necessary overhaul of E3 – maybe this is an opportunity where the academic community can realign with regulators? At the level of ICH the European Commission is very active to reduce the role and influence of industry.

Finally I propose to modify the last sentence of the abstract’s conclusion. From a scientifically purist perspective this is certainly correct but you can safely assume that there is always much, much, much more information submitted to a regulatory authority than to a journal for publishing (take also my example above concerning protocol submitted to ethics committees – this is just the front end of the same stick). If so, this should be addressed in slightly more detail in the discussion.

In case the BMJ will publish that paper, which I would greatly support, I assume that some of the contents will go on the BMJ’s website, such as the appendix, table 1 and figure 1? Actually I think figure one could be generally improved or even omitted – it is very well described in the text.

In case the BMJ will not publish this paper I dare to recommend editorialising these findings in some form. I really think that this paper touches a few related important things which went largely unquestioned for a long time.

Minor comments (in chronological order)

Page 4, line 72 (abstract): “78 were adequate”. Later in the text you say that four in fact were not CSRs. In this case isn’t it rather that four were just incorrectly classified? This is also an issue with figure 2 (page 23) where instead of 6 only 2 had inadequate data and 4 were no CRS (but maybe I misunderstand).

Page 5, line 112: I am sure there are more systematic reviews using regulatory data, actually I happen to

For peer review only

be an author of one (http://www.ncbi.nlm.nih.gov/pubmed/17907590?dopt=Citation) and there are FDA reviews as well. I believe, however, that there are no systematic reviews using such data on the efficacy and safety of medicinal products. The authors might narrow down their statement to such systematic reviews.

For peer review only

Ross, Joseph

In this manuscript, Doshi and Jefferson search public data sources and file Freedom of Information Act requests to obtain clinical study reports (CSRs) which they then descriptively explore in an effort to guide clinicians and systematic reviewers and inform evidence based medicine. While I am strongly supportive of better understanding the use of additional data sources such as CSRs to ensure better systematic reviews and summary analysis of clinical trial research, I do not think this research project achieved its maximum potential impact.

Originality and Importance

The investigators descriptively analyze 78 CSRs of 14 pharmaceuticals, providing information on page length and presence of key sections of information. While such a description has not been done previously to my knowledge, it does not provide sufficient insights to advance the field. This past January, Wieseler and colleagues published their findings in BMJ (2012;344:d8141) that demonstrated that CSRs reported higher quality information for clinical trials when compared with publications or results reporting systems. Their study was limited by the inaccessibility of CSRs for many of the comparisons they conducted. I had hoped that this study would advance the field further by making a comparison of this sort for a complete sample of study article-CSR pairs.

Instead, the investigators predominantly focus their analysis on descriptive information and imply the significance of missing information for systematic reviewers and summary analysis, without proving the impact of the absence. I strongly agree that the information missing is likely to be consequential, but as a research project, the purpose is to generate evidence that proves or disproves the hypothesis.

Moreover, some of the investigators conclusions are focused on what is missing from CSRs. But it is unclear what the implication of missing that information is for the field.

Scientific Reliability

The investigators explain that they did an exploratory review with a long-term intention of improving the credibility of research synthesis. I think the research question could be more clearly defined. It is not clear what the purpose of exploring the structure and content of CSRs, how new insights would be gained from this research, and so forth.

I am also not clear how this research is “exploratory”. That term is usually reserved for qualitative research that seeks to generate hypotheses, rather than test hypotheses. Although the investigators do not state an explicit hypothesis, neither are they using qualitative methods to develop one.

Overall Design of Study

For peer review only

I am concerned about the sample of trials used for analysis. By using a non-random sample, the findings are not generalizable. Moreover, more than a third were obtained from the investigators Tamiflu work, which they have already discussed in great detail in previous articles (PLoS Med 9(4): e1001201).

A stronger study would include a larger number of CSRs, ideally all from more recent time periods after ICH E3 approval (why include the 4 written prior?). 78 CSRs is a very small number. Moreover, given the number of products for which documents have been produced as part of litigation, it is likely that more CSRs are available in the UCSF DIDA web-base or in other places.

Methods

More methodological information would be useful. For instance, in the 5 steps for obtaining CSRs, I had a number of questions. How were CSRs identified for downloading on the internet? How were additional investigators identified for correspondence about CSRs obtained via Freedom of Information Act requests? What CSRs were manufacturers approached about?

I am not sure that I agree with the investigators contention that there is no known sampling frame to obtain CSRs. I would expect that a CSR would have been generated for every trial conducted as part of an application to a pharmaceutical regulator like the FDA or EMA. From regulator documents, all phase III trials could be identified and CSRs could have been requested.

Why did the investigators not simply abstract the information requested by ICH E3? Or maybe they did, but the text suggests to me that they developed their own abstraction form.

The compression factor objective was not established in the Introduction and the Methods are unclear, particularly the generation of “conservative” versus “realistic” compression factors. How many were inaccessible?

Results

The results are predominantly focused on page length and presence of content; a deeper analysis is necessary to provide new insight for the field. The new knowledge that is generated by the study is not convincing that key information is being lost when reporting clinical trial results in a CSR format as opposed to a journal article.

Given the narrow focus of the results, perhaps this article would be better structured as a research letter.

Interpretation and Conclusions

I thought the interpretation and conclusions, of the manuscript text and the abstract, went well beyond

For peer review only

the data presented. The investigators engaged in a substantial amount of editorializing, which detracts from the objectivity of their research. I would suggest a full re-write of these sections that were focused on summarizing their findings and clarifying the implications for the field.

For instance, the 2nd paragraph of the results states "[CSRs] far surpass the level of detail available in journal publications …" Any reader would assume this to be true based on a general understanding of the field, so this statement could be appropriately made in a commentary. However, the purpose of this article was to examine this question – and no measurable comparison to journal article content was made (to assess the level of detail), just journal article length. So this statement, in the context of this article, is unproven.

Abstract

The abstract should only make reference to the 78 CSRs that were the sample for the analysis.

For peer review only

Scherer, Roberta

Review for Doshi and Jefferson "Clinical Study Reports of randomized controlled trials – an exploratory review of previously confidential industry reports"

The authors of this report were able to obtain 84 Clinical Study Reports (CSR) from a variety of sources, and have reviewed the contents of 78, assessing whether specific sections as recommended by International Committee on Harmonisation were included in the report and noting the length of the various sections when included in the report.

Given that the decision by the European Medicines Agency to allow public access to these reports was made late in 2010, there is little current work in the literature on describing the contents. Other than sporadic articles on individual reports obtained through litigation, I am aware of only one report in the literature (Weiseler et al BMJ 2012 which the authors cite). It is likely that a great deal of effort was required to obtain the number of reports presented in this paper. This work is original and so adds to the current state of understanding in this field.

The authors argue that access to CSR will allow systematic reviewers to obtain trial information in more detail than can presently be obtained through journal articles. They state that little is known about the structure and content and aimed to describe what was included is a report. In this sense this information is of general use and of special interest to those persons performing and using systematic reviews of drug interventions. By its nature, however, this study cannot address the inadequacies of reporting other types of trials nor will it likely be of specific interest to practitioners.

It would be helpful if the author explicitly describe their definition of "adequate" for purposes of inclusion of a CSR in this report. They do say "too fragmentary" but that is somewhat vague. Using a detailed extraction form, the authors scored the presence of each of the sections recommended by the ICH, either by direct observation or by noting the table of contents, and then recorded the page length of each section. The study is straightforward and the authors have appropriately audited each others' extraction as a check for bias. I found that the authors make a fairly large assumption, however, in that they equate the length of a section (in number of pages) with the amount of detail that is provided by the report. While this assumption may be true, there is no data to support it. For example, although the number of pages in a typical journal trial report may be equivalent from article to article, the trial elements reported may vary widely. While page length might well be a reasonable surrogate for "amount of detail" I would have liked to have seen at least one or two direct comparisons to support this claim. Possibly, the information from Weiseler would support this assumption, but the authors do not describe it.
Because of this, I did not find the "compression factor" (a measure of the ratio of number of pages in a journal report to that in the CSR) to be a particularly useful measure and I wasn't sure how to interpret

For peer review only

it, especially the “conservative” vs the “realistic” factors. Further, the authors are over-interpreting the data when they say “The median length of 644 pages for reports in this study confirms that CSRs are the most detailed and complete, integrated form of reporting of the design, conduct and results of clinical trials” [line 218-220] when all they have shown is the number of pages in the report. This conclusion is based completely on the equation that page length is proportional to amount of detail and the authors provide no evidence in the paper or in the cited literature to support this assumption.

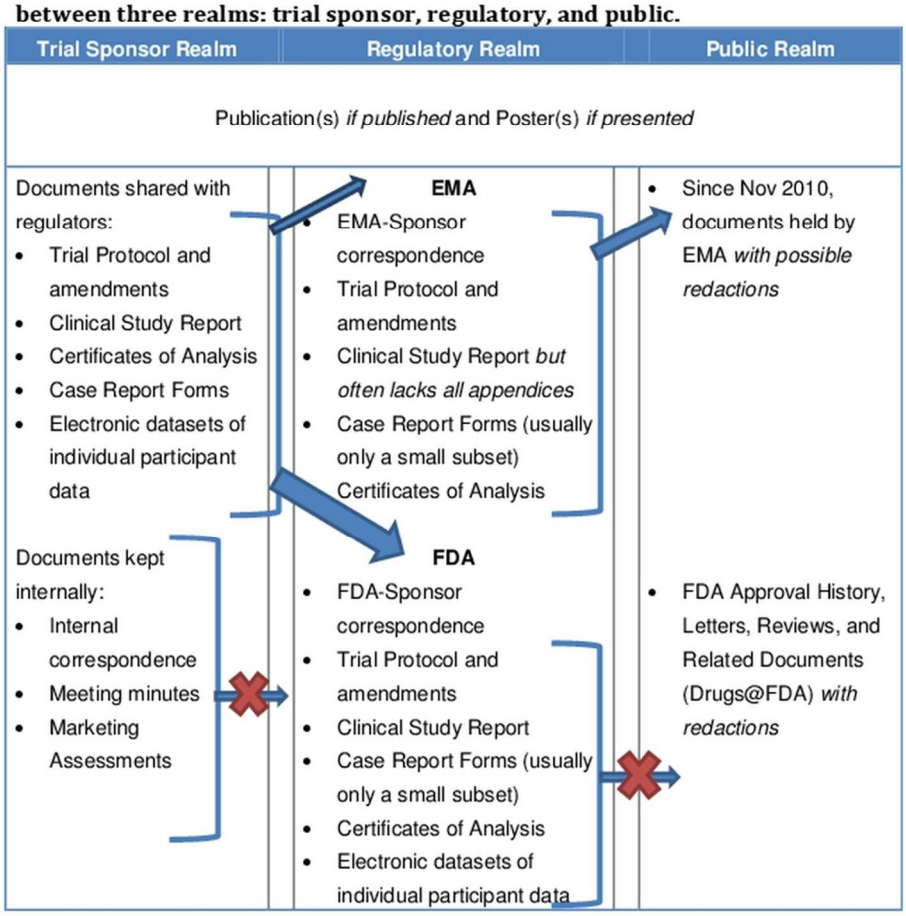
The authors also note the presence of individual case report forms available in one of the CSRs. Although the authors perceive the presence of individual patient data to be a good thing, it would be important to consider safeguards in place to protect patient confidentiality. For example, is there any assurance that the data have been correctly de-identified beyond simply changing the study ID.

In the discussion, lines 302-304, the authors should note that the presence of open access journals increase the possibility of more detailed reporting in journal articles so that trial reports may no longer be “limited to summary and aggregate details”

Some minor issues:
Line 119: Not quite a mixed metaphor says that “lack of visibility may also conceal”
Line 191: sentence is unclear (are there words missing?). Also in that paragraph, there are some “of’’s” that should not be present (e.g. line 194)
Line 281 – too many periods

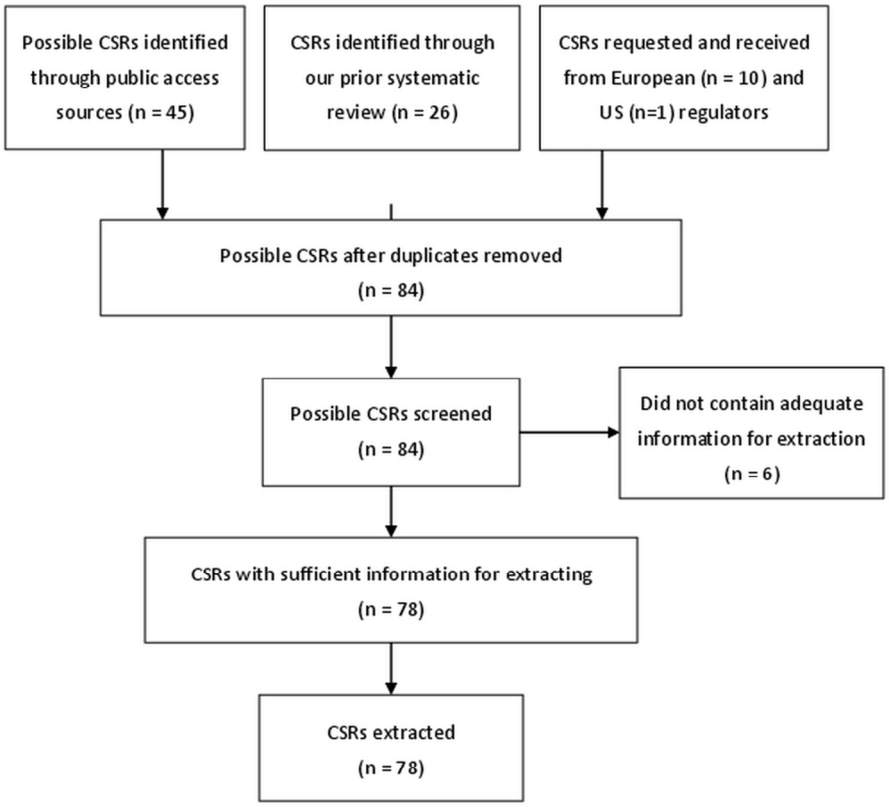
For peer review only

Figure 1. Types of clinical trial data typically held within and transferred between three realms: trial sponsor, regulatory, and public.



90x92mm (300 x 300 DPI)

Figure 2. Study flow



92x90mm (300 x 300 DPI)

Clinical Study Reports of randomized controlled trials – an exploratory review of previously confidential industry reports

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2012-002496.R1
Article Type:	Research
Date Submitted by the Author:	23-Jan-2013
Complete List of Authors:	Doshi, Peter; Johns Hopkins University, Jefferson, Tom; Cochrane Vaccines Field
Primary Subject Heading:	Evidence based practice
Secondary Subject Heading:	Medical publishing and peer review, Ethics, Research methods, Pharmacology and therapeutics
Keywords:	MEDICAL ETHICS, MEDICAL JOURNALISM, INTERNAL MEDICINE

SCHOLARONE™
Manuscripts

For peer review only

Clinical Study Reports of randomized controlled trials – an exploratory review of previously confidential industry reports

Authors:

Peter Doshi PhD
Johns Hopkins University School of Medicine, Baltimore, Maryland, USA

Tom Jefferson MD
The Cochrane Collaboration, Roma, Italy

Corresponding author: Peter Doshi <pnd@jhu.edu>

Word count: 3464

Tables: 3
Figures: 2
Appendices: 2
References: 42

For peer review only

Patient consent statement No consent was necessary as no patients were involved

Ethics approval statement No ethical approval was necessary as no patients were involved and all data were aggregate or anonymized and publicly available.

Role of the sponsor statement As the review had no extramural funding, there was no sponsor.

Author Contributions: Doshi had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. *Study concept and design*: Doshi and Jefferson. *Acquisition of data*: Doshi and Jefferson. *Analysis and interpretation of data*: Doshi and Jefferson. *Critical revision of the manuscript for important intellectual content*: Doshi and Jefferson. *Statistical analysis*: Doshi.

For peer review only

Research questions or hypotheses addressed

What are Clinical Study Reports (CSRs)? What do they contain and how long are they?

Might CSRs help address reporting biases associated with the published literature, and improve the quality of evidence synthesis?

Key Messages (up to 3)

CSRs represent a hitherto hidden and untapped source of detailed RCT data (mean page length: 1,854 pages), increasingly becoming publicly available, and should form the basic unit for evidence synthesis to minimize the problem of reporting bias.

CSRs show that numerous individuals make important technical contributions to the design, conduct, and reporting of each trial, but journal publications often fail to record these details, resulting in a loss in individual responsibility for what is reported.

The E3 guideline to which most CSRs conform was published in 1995, and needs updating.

Strengths and Limitations

We cannot say whether our sample is representative and whether our conclusions are generalizable to an undefined and undefineable population of CSRs.

For peer review only

Abstract

Objective: To explore the structure and content of a non-random sample of clinical study reports (CSRs) to guide clinicians and systematic reviewers.

Search strategy: We searched public sources and lodged Freedom of Information requests for previously confidential CSRs primarily written by industry for regulators.

Selection criteria: CSR reporting sufficient information for extraction ("adequate")

Primary outcome measures: Presence and length of essential elements of trial design and reporting and compression factor (ratio of page length for CSR compared to its published counterpart in a scientific journal).

Data extraction: data were extracted on standard forms and cross-checked for accuracy

Results: We assembled a population of 78 CSRs (covering 90 RCTs; 144,610 pages total) dated 1991-2011 of 14 pharmaceuticals. Report synopses had a median length of 5 pages, efficacy evaluation 13.5 pages, safety evaluation 17 pages, attached tables 337 pages, trial protocol 62 pages, statistical analysis plan 15 pages, and individual efficacy and safety listings had a median length of 447 and 109.5 pages, respectively. While 16 (21%) of CSRs contained completed case report forms, these were accessible to us in only one case (765 pages representing 16 individuals). Compression factors ranged between 1 and 8805.

Conclusions: Clinical study reports represent a hitherto mostly hidden and untapped source of detailed and exhaustive data on each trial. They should be consulted by independent parties interested in a detailed record of a clinical trial, and should form the basic unit for evidence synthesis as their use is likely to minimize the problem of reporting bias. We cannot say whether our sample is representative and whether our conclusions are generalizable to an undefined and undefineable population of CSRs.

Word count: 272

Primary Funding Source: The review had no extramural funding.

For peer review only

Introduction

Systematic reviews are thought to provide one of the most robust ways to evaluate the effects of healthcare interventions. But the robustness of findings clearly rests upon reviewers' access to clinical trial information sufficient to critically evaluate and reproduce the original research. Research on reporting bias over the last decades has shown that trusting the published literature at face value, even peer-reviewed publications, can be fraught with difficulty—a problem that spans drug classes.[1–12]

Following the decision by the European regulator, European Medicines Agency (EMA) on 30 Nov 2010, to make available a broad spectrum of documents related to medicinal products for human and veterinary use,[13,14] attention is focusing on one particular type of regulatory document: clinical study reports (CSRs).[15–18] CSRs are usually written for regulators following guidelines developed by the industry-regulatory collaborative effort "International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use" (ICH). The ICH guidelines "Structure and Content of Clinical Study Reports"[19] (See Appendix 1) are known by the document code "E3". They were formalized in 1995 "to assist sponsors in the development of a report that is complete, free from ambiguity, well organised and easy [for regulators] to review."[19]. E3 has not been edited or changed since 1995.

CSRs are but one category of information that is transmitted from study sponsors to regulators (Figure 1), but are important as they contain substantially more information and detail on the intervention being tested than published versions of the same trial. The wealth of information may be sought with increasing frequency by researchers appraising single trials, entire trial programmes, or by those synthesizing evidence.[17,20] We are aware of two recent examples of systematic reviews of the effects of pharmaceuticals carried out using CSRs and other regulatory material.[12,21] One group also concluded that journal publications insufficiently report clinical trials.[22]

Despite CSRs' potential importance very little is known about their structure and content outside of those individuals with direct involvement in regulatory processes. This knowledge gap may hinder development of methods for fair and reliable appraisal of CSRs and their use in evidence synthesis. We are not aware of any instruments specifically designed for appraising CSRs. Lack of visibility may also hinder understanding of the complexity of the organization and reporting of clinical trials.

We carried out an exploratory review to describe the structure and content of a non-random sample of clinical study reports. By describing the contents of CSRs, this research seeks to

For peer review only

transform CSRs from an obscure document only known to regulators and industry into a more widely known and accessible document. Our long-term intention is to improve the credibility of research synthesis by facilitating a move from the level of detail found in journal articles to the level of detail found in regulatory documents, thus guiding clinicians and other decision makers at all levels.

Methods

We obtained CSRs from public sources, as follows:

1. Requesting from EMA, under its freedom of information (FOI) policy, CSRs for manufacturer sponsored trials of the 10 best-selling prescription-bound products in the United States in 2010.[23]
2. Reusing CSRs from our own previous research (oseltamivir, zanamivir) [12]
3. Downloading CSRs openly available on the Internet. Search terms were not predefined, but sites searched included Google (http://www.google.com), the Drug Industry Document Archive (http://dida.library.ucsf.edu/), and IQWIG's library of reboxetine studies (https://www.iqwig.de/information-on-studies-of-reboxetine.980.en.html)
4. Corresponding with one researcher who obtained CSRs through a FOI request to FDA (epoetin alfa)
5. Requesting manufacturers fill any gaps in the completeness of reports that we believe are legally required to be publicly available (paroxetine).

To create as broad a database as possible, we did not apply restrictions in drug type or family or sponsor. We did not submit requests under the Freedom of Information Act to the Food and Drug Administration because such requests can take years to be fulfilled and—if fulfilled—may be heavily redacted.[24]

We did not draw a random sample of CSRs as there is no known sampling frame. No one knows how many reports have been written by intervention category as there is no central register of CSRs. Through familiarity with CSRs for oseltamivir and zanamivir, which were included in one of our Cochrane reviews,[12] we developed and piloted a data extraction sheet designed to capture the salient characteristics of CSRs . We created a list of around 40 potential sections we expected to find, generated directly from elements specified in E3. For each element in the list, we checked whether the obtained CSR included that section (confirmed either by direct identification of the section or an indication the section existed based on the CSR's table of contents), whether we had access to it, and its page length. Because of previous difficulties we had accessing CSR appendices, we also recorded whether sections

For peer review only

were listed as appendices or not. Page length was calculated either by directly counting the pages or by estimating their size from the table of contents of each report, and was used as a crude proxy for the level of detail available. Page lengths were rounded up to the next integer, and were summarized by reporting medians and ranges. We also included questions relating to trial registration and authorship. Our (blank) data extraction sheet is in Appendix 2.

All variables from CSRs were first extracted in single. We subsequently audited each other's extractions, checking the accurateness of the information. We chose to present elements analogous with those that typically appear in trials reported in scientific journals including the study Synopsis (a brief summary of the study), the study Protocol (written prospectively, describing the study methods), Efficacy and Safety Evaluations (a narrative summary of the efficacy and safety results of the study, including tables and figures), as well as attached tables. We also included elements rarely found in journal publications: sample (blank) and completed case report forms (CRFs are paper or electronic forms designed to capture pre-specified efficacy and safety related information for each study participant), the statistical analysis plan (a prospectively written narrative and/or statistical code indicating how trial data will be analyzed), and individual participant efficacy and safety listings. The corresponding E3 section numbers are listed in Table 2. Disagreements were resolved by discussion.

Our uncorrected (original) and corrected extraction sheets as well as audit records are available upon request from the corresponding author.

We calculated a compression factor for published trials which we defined as the ratio of CSR page length compared to the page length of the same trial as published in scientific journals. The objective of this metric was to convey a rough sense of how much information present in CSRs may be being condensed ("compressed") in short journal publications, in consideration of CSRs' far greater length and level of detail. Size (page length) reflects the level of detail as well as the presence of many elements such as protocols and their amendments, randomization lists, statistical analysis plans, certificates of analysis and extra data on subpopulations. We have demonstrated[12] that these elements are essential for understanding and appraising a trial. The compression factor is a crude measure of how much is compressed or simply left out of each publication which will affect the reliability of the appraisal and interpretation of trials. Trial publications were searched for in multiple sources: clinical trial registers, published systematic reviews, and correspondence with sponsors. Because in most cases we could not access all parts of all CSRs (and therefore do not know their complete page length), we calculated "conservative" compression factors as well as "realistic" compression factors. "Conservative" compression factors were calculated on a trial by trial basis using the total number of pages in

For peer review only

CSRs available to us divided by the length of journal reports for that same trial, while "realistic" compression factors were based on the true total page length of the CSR.

Results

We identified 84 documents believed to be CSRs for 14 compounds. These covered therapeutic and biological interventions including antipsychotics, antidepressants, antivirals, natural antiarthritics, anti-inflammatory agents, pandemic influenza vaccines, statins, erythropoietins, and anti-platelet compounds. We included English-language summaries of two Japanese oseltamivir studies (JV15823, JV15824) as they had been presented to EMA in this form. We excluded documents which were sections of CSRs but nonetheless contained insufficient information to understand the overall content of the CSR (olanzapine F1D-LC-HGAV, F1D-MC-HGAJ and F1D-MC-HGAO) and 3 documents which we had originally classified as CSRs but were not (reboxetine 14, 22 and 37). This left 78 CSRs (144,610 pages) (Figure 2). The median pages obtained per CSR was 644 (range 9 to 15,440). Only 4 of 78 CSRs (reboxetine 8, 16, 17, and 91) were written prior to November 30 1995 when ICH E3 was approved. Table 1 summarizes the pharmaceutical, manufacturer, date and provenance of the CSRs in our review. EMA reported not holding studies for esomepazole magnesium (Nexium), Advair diskus, quetiapine fumarate (Seroquel), montelukast sodium (Singulair), epoetin alfa (Epogen), and simvastatin.

All of the 78 included CSRs contained a synopsis (median page length 5 pages). The efficacy evaluation was identifiable and directly accessible in 76 (97%, median length 13.5 pages) and safety in 77 (99%; median length 17 pages). Attached tables were likewise present in 63 (81%) CSRs, and were a median of 337 pages long (range: 1 to 3665). Seventy-three CSRs (94%) reported including the study protocol. In the 40 we could access, the median page length was 62. We found blank CRFs included in 68 (87%) CSRs. Of the 33 we could directly access, the median length was 133 pages (range 14 to 981). For completed CRFs, 16 (21%) reports made direct mention of a section on completed CRFs, but we had access to completed CRFs in only 1 case (Arthronat; length 765 pages).

Fifty-five (71%) of 78 included CSRs included a statistical analysis plan in some form. Of those for which we could directly access the content (n=37), the median page length was 15 (range 3 to 85). Individual efficacy and safety listings were included in 53 (69%) and 62 (81%) CSRs respectively. The median page length was 447 (range 15 to 21,698) for efficacy and 109.5 (range 2 to 10,954) for safety.

For peer review only

A summary is presented in Table 2.

All trial reports in our review were sponsored by industry.

Median conservative compression factors ranged between 1 and 1221. The realistic compression factors, calculated for the Arthronat, paroxetine, and clopidogrel CAPRIE trials, were 379, 1021, and 8805, respectively. (Table 3)

Discussion

We collected and described a sizeable number of CSRs written in the last two decades. All CSRs contained a table of contents (as specified in E3 section 3); this, together with optical character recognition (to enable searching the full text of the scanned documents) and the occasional need to combine multiple files to create a single document, substantially improved the ease of navigating CSRs.

Despite the size of our non-random sample, it is unclear whether our conclusions are generalisable to all other CSRs. This is because we have extremely limited knowledge about the total population of CSRs in regulators' and sponsors' possession. Nevertheless, within our sample spanning different manufacturers, therapeutic classes, and times, we found that the structure of CSRs was, within different house styles of presentation, strikingly similar, probably due to the guidance by ICH E3.[37] This suggests that the structure and content of other CSRs is likely to be similar.

The future basic currency of research synthesis?

The median length of 644 pages for reports in this study, as well as CSRs' routine inclusion of trials' protocol, statistical analysis plans, and blank case report forms, strongly suggests that CSRs are the most detailed and complete, integrated form of reporting of the design, conduct, and results of clinical trials. In a study that directly compared the adequacy of reporting between journal articles and CSRs, the authors found that complete information regarding greater than 40% of methods items were only available in CSRs.[22] The level of detail found in CSRs thus far surpass the level of detail available in journal publications, and as such they are prime candidates for the next basic currency of evidence synthesis and appraisal of a trial. Given the EMA's new policy making such documents publicly available, access to these documents is now relatively straightforward.[25] However including CSRs in systematic reviews is labor-intensive, given their size and complexity.[12]

For peer review only

Accessing complete CSRs

While CSRs may trump other forms of trial reporting in the public domain (such as conference abstracts or journal publications), serious limitations remain. Despite obtaining 144,610 pages for 78 CSRs, in almost all instances, we lacked full access to the CSRs' numerous appendices. Even for the sole complete CSR we obtained (Arthronat MA-CT-10-002), case report forms were provided for only 20% of participants. The Arthronat text does not provide a reason for this omission, but it reflects the vagueness of the relevant section of the E3 guidance (16.3.2) which does not define "Other CRFs submitted." Also, we could only access the original trial protocol in 40 (51%) of 78 CSRs obtained. This is important because trial protocols, written prior to patient enrollment in a trial, are an important way to guard against reporting biases.[26,27]

We could obtain individual patient listings in only a minority of cases despite confirming their inclusion in the majority of CSRs (Table 2). This may be a significant limitation, as the E3 specifies that "the report with its appendices should also provide enough individual patient data, including the demographic and baseline data, and details of analytical methods, to allow replication of the critical analyses…"[19] Unavailability was possibly due to the fact that EMA allows manufacturers to submit CSRs omitting a number of appendices including individual patient data and case report forms (which EMA states should be available within 48 hours if requested).[28] In the case of oseltamivir, the subject of a Cochrane review we conducted,[12] the manufacturer refused to share with us report appendices not submitted to EMA,[29] and EMA declined requesting them on our behalf.[8] Although FDA likely possesses more complete CSRs and patient level data, it historically has treated such data as trade secret and/or confidential.[30–32] EMA is therefore at present the only reliable source of obtaining CSRs. As such, despite European regulators' progressive stance—announcing that "clinical trial data should not be considered commercial confidential information"[33]—the completeness gap is unlikely to be filled any time soon.

Another significant limitation is that CSRs are only written for therapeutic, prophylactic, or diagnostic agents, and therefore inadequacies remain in evidence synthesis of other types of interventions such as surgical or behavioral interventions.

Individual participant listings

Individual participant listings—which identify participants by a unique ID—were accessible in 29 of the 78 CSRs we reviewed. But these data are difficult to analyze because they are presented as database printouts rather than in electronic form. This is understandable considering that CSRs are a written/archival format, but because EMA does not accept SAS datasets,[34,35] the industry standard, third-party access to databases of patient-level data remains elusive. We

For peer review only

see no compelling reason why all regulators should not request these from sponsors and make them publicly available. Whether availability of individual listings and CRFs, with its attendant laborious analysis, would increase our understanding of the trial and its results is unclear. But there is at least one case where the re-analysis of CRFs added invaluable knowledge to that already available in CSRs.[36]

The public-private debate

One manufacturer has claimed that the non-release of case report forms is motivated by concerns over protecting participant confidentiality.[38] Nothing we have seen so far corroborates this claim, however an ongoing EMA working group is specifically discussing issues related to protecting participant confidentiality. Based on current document releases and position statements, however, it appears that EMA has deemed case report forms and individual patient listings to be, in principle, releasable in their entirety (after a preliminary review).[39] Furthermore, individual patient listings are intended to duplicate information contained in filled case report forms. The release of case report forms would ensure the accuracy of individual patient listings with little additional risk to patient confidentiality. Moreover, extra checks such as registration of protocols by bona fide research groups could deter any inappropriate use. We also believe that the sheer bulk of the forms acts as a deterrent against malice.

Size matters

Our range of compression factors show the scale of selection and synthesis which must (consciously or unconsciously) occur in the process of transforming CSRs into journal-length articles. We found a strong resemblance in detail, page length, structure, and purpose between the short Synopsis section of CSRs and reports of trials as published in scientific journals. In some cases essential items of information such as the trial protocol and its subsequent amendments are simply not included in journal articles or are replaced by methods written *post facto*. In other cases of items essential for the interpretations of the trial results (such as the statistical analysis plan), tens of pages are reduced to a paragraph on sample size calculation in the journal report, underscoring the lack of detail (and its attendant problems) common to public forms of trial reporting. For example, the ratio of words in Protocol of the CSR for Aripiprazole CN138135 to the Methods section for published journal article of the same trial is 30.5 (53,713 words in the CSR Protocol versus 1,763 words in the journal article). For the oseltamivir WP16263 trial, the ratio was 22.7 (26,761 words in the CSR Protocol and amendments versus 1,177 words in the journal article).

This compression of information also occurs in databases not restricted by length, such as ClinicalTrials.gov.[40]

For peer review only

Our study raises the question of why the medical community has accepted the low (summary, aggregate) level of detail found in most peer-reviewed journal publications compared to the depth of detail available in CSRs. European regulators recently noted: "Documents that provide critical information on a study, such as the protocol (16.1.1), statistical methods (16.1.9), list of investigators and study sites and sample case report forms, would always be needed by reviewers assessing a study"[41] Why have those outside of the regulatory world tolerated journal publications lacking such details?

One possibility may be that while the clinical trial enterprise has changed dramatically in the last half century, the scientific journal publication model has not. Since the 1950s, there have been considerable transformations in the political economy of clinical trials driven by the increasingly commercialized and global nature of the pharmaceutical industry, the rise in academic-industry "partnerships" in medicine, and increased communication among regulators. It is now common to find trials with study centers scattered around the globe. This increasing complexity and the need to provide an audit record is reflected in the comprehensive tomes documenting the trials—CSRs—but trial reporting in scientific journals remains limited to summary and aggregate details. It should be noted, however, that many journals now have websites which enables them to make available extended content beyond what traditionally appears in the printed journal.

Authorship or Contributorship?

Examination of CSRs revealed scores of important technical contributions to the design, conduct, and reporting of each trial. These included contributions from database programmers, records officers, and CSR writers, often invisible in the published journal article. In some cases, we found no mention in CSRs of individuals who figured as authors of subsequent published trial reports while individuals named as CSR authors went unacknowledged in journal publications. Current ICMJE guidelines on authorship and contributorship are largely focused on ensuring those placed on by-lines deserve to be authors. But the guidelines also suggest that "all contributors who do not meet the criteria for authorship should be listed in an acknowledgments section."[42] Given the complexity of clinical trials, the ICMJE should call for itemized contributorship: the names of all contributors to be specified along with their role in the design, conduct, analysis, or reporting of the trial. If the contribution of most people goes unrecorded, so does their individual responsibility for what is produced. Itemized contributorship records, to all phases of a trial, could be piloted in trial registers.

For peer review only

E3 guidance

The E3 guideline set an excellent standard, but it needs formal updating and further development. For example, there should be a self-standing set of definitions for terms such as "case report forms" and "Other CRF's submitted," (section 16.3.2) and a description of how a particular trial fits within a sponsor's trial programme of pharmaceutical development. Apparently forgotten items such as certificates of analysis (describing the appearance and content of the interventions being tested) and post-1995 details such as trial registration numbers should be mentioned.

We hope our review has given CSRs what they have lacked so far: visibility. CSRs represent a largely untapped source of detailed data that we believe can serve as a means of addressing the ravages of reporting bias in all its forms, leading to a more accurate understanding of the effects of medicines.

Conflicts of interest statement

All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare that:

Both authors are co-recipients of a UK National Institute for Health Research grant to carry out a Cochrane review of neuraminidase inhibitors (http://www.hta.ac.uk/2352).

Tom Jefferson was an ad hoc consultant for F. Hoffman-La Roche Ltd in 1998-1999. He receives royalties from his books published by Blackwells and Il Pensiero Scientifico Editore, none of which are on clinical study reports. He is occasionally interviewed by market research companies for anonymous interviews about Phase 1 or 2 products unrelated to products in this review. In 2011-12 he has acted as an expert witness in a litigation case related to one of the compounds in the review (oseltamivir). He is on a legal retainer for expert advice on litigation for influenza vaccines in health care workers.

Peter Doshi received €1500 from the European Respiratory Society in support of his travel to the society's September 2012 annual congress where he gave an invited talk on oseltamivir.

Both authors' spouses and children have no financial relationships that may be relevant to the submitted work.

For peer review only

Data sharing statement

The original extraction forms and audit record are available on request from the corresponding author.

Acknowledgements

We thank Drs Vallance and Kraus of GlaxoSmithKline for making public selected report appendices from the 9 paroxetine trials. We also thank Daniel Coyne for sharing the CSR that FDA sent him in response to his Freedom of Information request, and Iain Chalmers for guidance.

Peter Doshi is funded by an institutional training grant from the Agency for Healthcare Research and Quality #T32HS019488. AHRQ had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

For peer review only

References

1. Rising K, Bacchetti P, Bero L. Reporting Bias in Drug Trials Submitted to the Food and Drug Administration: Review of Publication and Presentation. Ioannidis J, editor. Plos Med. 2008;5(11):e217.

2. Doshi P, Jones M, Jefferson T. Rethinking credible evidence synthesis. BMJ. 2012 Jan 17;344(jan17 2):d7898.

3. Chan A-W, Hróbjartsson A, Haahr MT, Gøtzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. JAMA. 2004 May 26;291(20):2457–65.

4. Gøtzsche PC. Why we need easy access to all data from all clinical trials and how to accomplish it. Trials. 2011;12(1):249.

5. Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R. Selective publication of antidepressant trials and its influence on apparent efficacy. N. Engl. J. Med. 2008 Jan 17;358(3):252–60.

6. Dickersin K, Min YI. Publication bias: the problem that won't go away. Ann. N. Y. Acad. Sci. 1993 Dec 31;703:135–146; discussion 146–148.

7. Hopewell S, Loudon K, Clarke MJ, Oxman AD, Dickersin K. Publication bias in clinical trials due to statistical significance or direction of trial results. Cochrane Database Syst Rev. 2009;(1):MR000006.

8. Melander H. Evidence b(i)ased medicine--selective reporting from studies sponsored by pharmaceutical industry: review of studies in new drug applications. BMJ. 2003 May 29;326:1171–3.

9. McGauran N, Wieseler B, Kreis J, Schuler Y-B, Kolsch H, Kaiser T. Reporting bias in medical research - a narrative review. Trials. 2010;11(1):37.

10. Sismondo S, Doucet M. Publication ethics and the ghost management of medical publication. Bioethics. 2009 Feb;24(6):273–83.

11. Vedula SS, Bero L, Scherer RW, Dickersin K. Outcome reporting in industry-sponsored trials of gabapentin for off-label use. N. Engl. J. Med. 2009 Nov 12;361(20):1963–71.

12. Jefferson T, Jones MA, Doshi P, Del Mar CB, Heneghan CJ, Hama R, et al. Neuraminidase inhibitors for preventing and treating influenza in healthy adults and children. Cochrane Database of Systematic Reviews. 2012;(1):CD008965.

13. European Medicines Agency. European Medicines Agency policy on access to documents (related to medicinal products for human and veterinary use) POLICY/0043 [Internet]. 2010 [cited 2012 May 14]. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Other/2010/11/WC500099473.pdf

For peer review only

14. European Medicines Agency. Output of the European Medicines Agency policy on access to documents related to medicinal products for human and veterinary use [Internet]. 2010 [cited 2011 Sep 25]. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2010/11/WC500099472.pdf

15. Gøtzsche PC, Jørgensen AW. Opening up data at the European Medicines Agency. BMJ. 2011;342:d2686.

16. Jefferson T, Doshi P, Thompson M, Heneghan C. Ensuring safe and effective drugs: who can do what it takes? BMJ. 2011;342:c7258.

17. Chan A-W. Out of sight but not out of mind: how to search for unpublished clinical trial evidence. BMJ. 2012 Jan 3;344(jan03 2):d8013–d8013.

18. Rodwin MA, Abramson JD. Clinical trial data as a public good. JAMA. 2012 Sep 5;308(9):871–2.

19. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. Structure and Content of Clinical Study Reports: E3 [Internet]. 1995 [cited 2012 Jul 8]. Available from: http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E3/E3_Guideline.pdf

20. Grens K. Data Diving. The Scientist [Internet]. 2012 Apr [cited 2012 May 3]; Available from: http://the-scientist.com/2012/05/01/data-diving/

21. Eyding D, Lelgemann M, Grouven U, Harter M, Kromp M, Kaiser T, et al. Reboxetine for acute treatment of major depression: systematic review and meta-analysis of published and unpublished placebo and selective serotonin reuptake inhibitor controlled trials. BMJ. 2010 Oct 12;341(oct12 1):c4737–c4737.

22. Wieseler B, Kerekes MF, Vervoelgyi V, McGauran N, Kaiser T. Impact of document type on reporting quality of clinical drug trials: a comparison of registry reports, clinical study reports, and journal publications. BMJ. 2012 Jan 3;344(jan03 1):d8141–d8141.

23. IMS Institute for Healthcare Informatics. The Use of Medicines in the United States: Review of 2010 [Internet]. 2011 [cited 2012 Jul 30]. Available from: http://www.imshealth.com/imshealth/Global/Content/IMS%20Institute/Documents/IHII_UseOfMed_report%20.pdf

24. Coyne DW. The health-related quality of life was not improved by targeting higher hemoglobin in the Normal Hematocrit Trial. Kidney International [Internet]. 2012 Mar 21 [cited 2012 Apr 13]; Available from: http://www.nature.com/ki/journal/vaop/ncurrent/full/ki201276a.html

25. Doshi P, Jefferson T. The first 2 years of the European Medicines Agency's policy on access to documents: secret no longer. Arch Intern Med. 2012 Dec 19;1–2.

26. Chan A-W. Bias, Spin, and Misreporting: Time for Full Access to Trial Protocols and Results. Plos Med. 2008;5(11):e230.

For peer review only

27. Miller JD. Registering clinical trial results: the next step. JAMA. 2010 Feb 24;303(8):773–4.

28. European Medicines Agency. Note for guidance on the inclusion of appendices to clinical study reports in marketing authorisation applications [Internet]. 2004 [cited 2012 Jul 30]. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003638.pdf

29. Doshi P, Jefferson T, Del Mar C. The Imperative to Share Clinical Study Reports: Recommendations from the Tamiflu Experience. PLoS Med. 2012 Apr 10;9(4):e1001201.

30. Unlu M. It is time: why the FDA should start disclosing drug trial data. Mich. Telecomm. Tech. L. Rev. 2010;16:511–45.

31. Kesselheim AS, Mello MM. Confidentiality laws and secrecy in medical research: improving public access to data on drug safety. Health Aff (Millwood). 2007 Apr;26(2):483–91.

32. Halperin RM. FDA Disclosure of Safety and Effectiveness Data: A Legal and Policy Analysis. Duke Law Journal. 1979 Feb 1;1979(1):286–326.

33. Eichler H-G, Abadie E, Breckenridge A, Leufkens H, Rasi G. Open Clinical Trial Data for All? A View from Regulators. PLoS Med. 2012 Apr 10;9(4):e1001202.

34. European Medicines Agency. EMEA implementation of electronic-only submission and eCTD submission: Practical guidelines relating to non-eCTD electronic submissions [Internet]. 2008 [cited 2012 Jul 31]. Available from: http://www.emea.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2009/10/WC500004100.pdf

35. European Medicines Agency. EMEA implementation of electronic-only submission and eCTD submission: Questions and answers relating to practical and technical aspects of the implementation [Internet]. 2008 [cited 2012 Jul 31]. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2009/10/WC500004095.pdf

36. Psaty BM, Prentice RL. Minimizing bias in randomized trials: the importance of blinding. JAMA. 2010 Aug 18;304(7):793–4.

37. Molzon JA, Giaquinto A, Lindstrom L, Tominaga T, Ward M, Doerr P, et al. The value and benefits of the International Conference on Harmonisation to drug regulatory authorities: advancing harmonization for better public health. Clin. Pharmacol. Ther. 2011 Apr;89(4):503–12.

38. GlaxoSmithKline. Paroxetine and pediatric and adolescent patients [Internet]. 2012 [cited 2012 Jul 30]. Available from: http://www.gsk.com/media/paroxetine.htm

39. HMA/EMA Working Group on Transparency. HMA/EMA guidance document on the identification of commercially confidential information and personal data within the structure of the marketing authorisation (MA) application - release of information after the granting of a marketing authorisation [Internet]. 2012 [cited 2012 Jul 30]. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Other/2012/03/WC500124536.pdf

For peer review only

40. Zarin DA, Tse T, Williams RJ, Califf RM, Ide NC. The ClinicalTrials.gov Results Database — Update and Key Issues. New England Journal of Medicine. 2011;364(9):852–60.

41. European Medicines Agency. ICH guideline E3 - questions and answers (R1) [Internet]. 2012 [cited 2012 Jul 30]. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2012/07/WC500130356.pdf

42. International Committee of Medical Journal Editors. Uniform Requirements for Manuscripts Submitted to Biomedical Journals: Ethical Considerations in the Conduct and Reporting of Research: Authorship and Contributorship [Internet]. 2009 [cited 2012 Jul 31]. Available from: http://www.icmje.org/ethical_1author.html

For peer review only

Table 1. Pharmaceutical, trials, producers, dates and sources of CSRs in the review.

Pharmaceutical and number (n) of assessed trial documents	Trial IDs	Manufacturer	Date of CSRs	Provenance in our study
Aripiprazole (Abilify) n=1	CN1368135	Bristol-Myers Squibb	2007	Freedom of Information request to EMA
Arthronat n=1	MA-CT-10-002	Rowtasha	2011	Manufacturer website http://arthronat.com/clinical-study.php
Atorvastatin (Lipitor) n=1	981-080	Pfizer	1999	Freedom of Information request to EMA
Clopidogrel (Plavix) n=5	CURE, CLARITY, COMMIT-CCS2, CAPRIE, PICOLO	Bristol-Myers Squibb	1997-2007	Freedom of Information request to EMA
Epoetin alfa (Epogen) n=1	930107	Amgen	1996	Freedom of Information request to FDA
H5N1 influenza vaccine n=1	H5N1-008, H5N1-011 EXT 008	GSK	2006	Freedom of Information request to EMA
H5N1 influenza vaccines n=2	V87P1, V87P6	Novartis	2008-2009	Freedom of Information request to EMA
Olanzapine (Zyprexa) n=3	F1D-LC-HGAV*, F1D-MC-HGAO*, F1D-MC-HGAJ*	Eli Lilly	1995†	Litigation http://zyprexalitigationdocuments.com/unsealed.php http://www.furiousseasons.com/zyprexadocs.html
Oseltamivir (Tamiflu) n=19	JV15823, JV15824, M76001, NP15757, NV16871, WP16263, WV15670, WV15671, WV15673 WV15697, WV15707, WV15708, WV15730, WV15758, WV15759 WV15871, WV15799, WV15812 WV15872, WV15819 WV15876 WV15978, WV15825, WV16193	Roche	1999-2004	Documents obtained as part of previous Cochrane review[12]
Paroxetine (Paxil, Aropax, Pexeva, Seroxat, Sereupin) n=9	329, 377, 453, 511, 676, 701, 704, 715, 716	GSK	1998-2002	Litigation (2004 legal settlement mandated release of clinical study reports on manufacturer's website of 9 studies on

For peer review only

				pediatric and adolescent patients) http://www.gsk.com/media/paroxetine.htm
Quetiapine (Seroquel) n=7	015, 041, 049, 125, 126, 127, 135	AstraZeneca	1996-2007	Litigaton http://psychrights.org/research/Digest/NLPs/Seroquel/UnsealedSeroquelStudies/
Reboxetine (Edronax, Norebox, Prolift, Solvex, Davedax, Vestra) n=24	8, 9, 13, 14*, 15, 16, 17, 22*, 32, 32a, 34, 35, 37*, 43, 45, 46, 47, 49, 50, 52, 71, 83, 91, 96	Pfizer	1991-2009	Health Technology Assessment website (The German IQWiG obtained CSRs as part of its health technology assessment work) https://www.iqwig.de/information-on-studies-of-reboxetine.980.en.html
Rofecoxib (Vioxx) n=1	78	Merck	2003	Litigation http://dida.library.ucsf.edu/
Zanamivir (Relenza) n=9	NAI30009, NAI300010, NAIA2005, NAIA3002, NAIA3005, NAIB2005, NAIB2007, NAIB3001, NAIB3002	GSK	1998-1999	Documents obtained as part of previous Cochrane review[12]

519 * Subsequently excluded because of insufficient documentation

520 † H1D-MC-HGAO clinical study report date unknown

521 EMA = European Medicines Agency

522 FDA = Food and Drug Administration

523

For peer review only

Table 2. Key characteristics of the CSRs in the review

Section of CSR (corresponding section of E3)	Presence	Length	
	CSRs including section, n	CSRs with section length available, n	Median length (range), pages
Synopsis (E3 section 2)	78 (100%)	78	5 (1 - 15)
Efficacy evaluation (E3 sec. 11)	76 (97%)	77	13.5 (2 - 132)
Safety evaluation (E3 sec. 12)	77 (99%)	58	17 (2 - 188)
Attached tables not in report text (E3 sec. 14)	63 (81%)	76	337 (1 - 3665)
Protocol (E3 sec 16.1.1)	73 (94%)	41	62 (21 - 139)
Blank Case Report Form (CRF) (E3 sec. 16.1.2)	68 (87%)	33	133 (14 - 981)
Statistical Analysis Plan (E3 sec. 16.1.9)	55 (71%)	37	15 (3 - 85)
Individual participant efficacy listings (E3 sec. 16.2.6)	53 (69%)	19	447 (15 - 21698)
Individual participant safety listings (E3 sec. 16.2.7)	62 (81%)	26	109.5 (2 - 10954)
Completed CRFs (E3 sec. 16.3.2)	16 (21%)	1	765

For peer review only

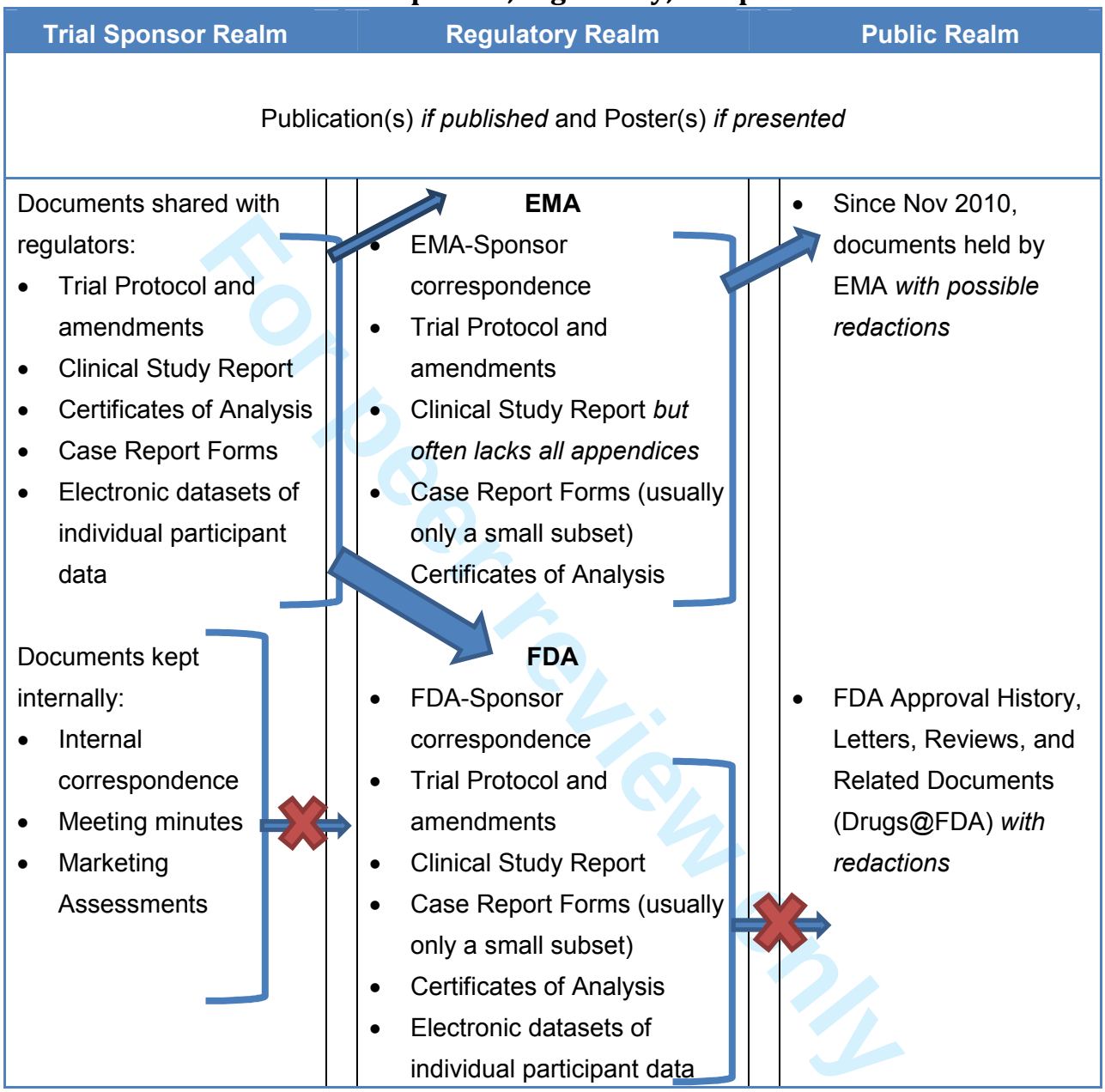
Table 3. Conservative and realistic compression factors. A ratio of CSR page length to corresponding journal publication page length.

Pharmaceutical	Studies published in journals, n	Mean compression factor (range)
Conservative compression factors		
Aripiprazole	1	672
Clopidogrel	5	11 (4 - 19)
Epoetin Alfa	1	41
Fluad	2	488 (367 - 609)
GSK H5N1 vaccine	1	19
Oseltamivir	12	195 (1 - 1221)
Quetiapine	2	578 (352 - 803)
Reboxetine	5	88 (9 - 245)
Zanamivir	8	54 (28 - 92)
Realistic compression factors		
Arthronat*	1	379
Clopidogrel	1	8805
Paroxetine	9	1021 (50 - 5473)

* The Arthronat trial has not yet been published. Compression factor calculation is based on the page length of a draft manuscript "to be published soon," according to Arthronat.com.

For peer review only

534 **Figure 1. Types of clinical trial data typically held within and transferred**
535 **between three realms: trial sponsor, regulatory, and public.**

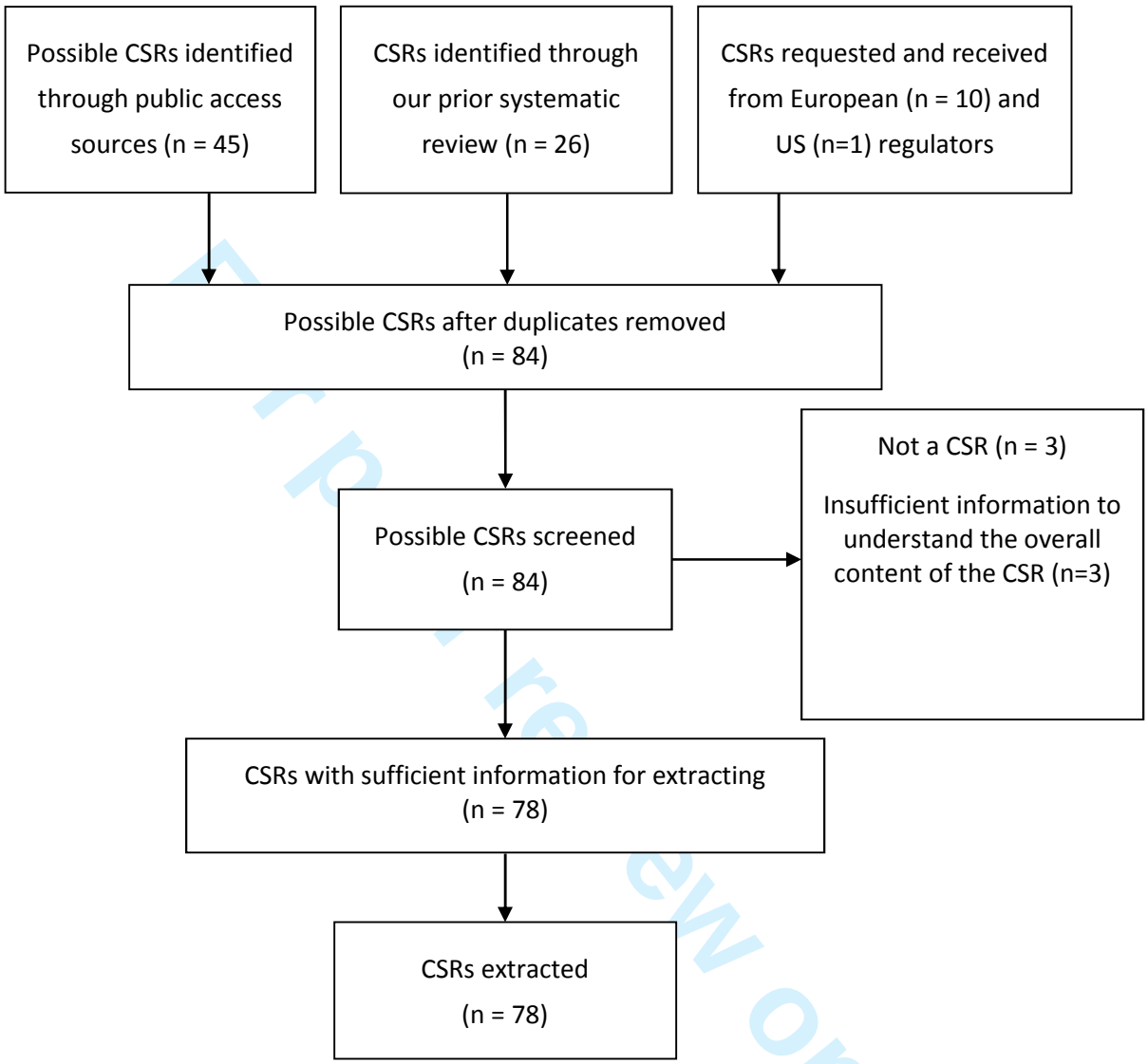


536

537

For peer review only

Figure 2. Study flow



For peer review only

Appendix 1. Elements specified ICH E3 "Structure and Content of Clinical Study Reports" (1995)[19]

1. TITLE PAGE
2. SYNOPSIS
3. TABLE OF CONTENTS FOR THE INDIVIDUAL CLINICAL STUDY REPORT
4. LIST OF ABBREVIATIONS AND DEFINITION OF TERMS
5. Ethics
    - 5.1. Independent Ethics Committee (IEC) or Institutional Review Board (IRB)
    - 5.2. Ethical conduct of the study
    - 5.3. Patient information and consent
6. INVESTIGATORS AND STUDY ADMINISTRATIVE STRUCTURE
7. INTRODUCTION
8. STUDY OBJECTIVES
9. INVESTIGATIONAL PLAN
    - 9.1. Overall study design and plan – description
    - 9.2. Discussion of study design, including the choice of control groups
    - 9.3. Selection of study population
        - 9.3.1. Inclusion criteria
        - 9.3.2. Exclusion criteria
        - 9.3.3. Removal of Patients from Therapy or Assessment
    - 9.4. Treatments
        - 9.4.1. Treatments Administered
        - 9.4.2. Identity of Investigational Product(s)
        - 9.4.3. Method of Assigning Patients to Treatment Groups
        - 9.4.4. Selection of Doses in the Study
        - 9.4.5. Blinding
        - 9.4.6. Prior and Concomitant Therapy
        - 9.4.7. Treatment Compliance
    - 9.5. Efficacy and safety variables
        - 9.5.1. Efficacy and Safety Measurements Assessed and Flow Chart
        - 9.5.2. Appropriateness of Measurements
        - 9.5.3. Primary Efficacy Variable(s)
        - 9.5.4. Drug Concentration Measurements
    - 9.6. Data quality assurance
    - 9.7. Statistical methods planned in the protocol and determination of sample size

For peer review only

9.7.1. Statistical and Analytical Plans
9.7.2. Determination of Sample Size
9.8. Changes in the conduct of the study or planned analyses
10. STUDY PATIENTS
10.1. Disposition of patients
10.2. Protocol deviations
11. EFFICACY EVALUATION
11.1. Data sets analyzed
11.2. Demographic and other baseline characteristics
11.3. Measurements of treatment compliance
11.4. Efficacy results and tabulations of individual patient data
11.4.1. Analysis of efficacy
11.4.2. Statistical/analytical issues
11.4.2.1. Adjustments for covariates
11.4.2.2. Handling of Dropouts or Missing Data
11.4.2.3. Interim Analyses and Data Monitoring
11.4.2.4. Multicentre Studies
11.4.2.5. Multiple Comparison/Multiplicity
11.4.2.6. Use of an "Efficacy Subset" of Patients
11.4.2.7. Active-Control Studies Intended to Show Equivalence
11.4.2.8. Examination of Subgroups
11.4.3. Tabulation of Individual Response Data
11.4.4. Drug Dose, Drug Concentration, and Relationships to Response
11.4.5. Drug-Drug and Drug-Disease Interactions
11.4.6. Drug Dose, Drug Concentration, and Relationships to Response
11.4.7. By-Patient Displays
12. SAFETY EVALUATION
12.1. Extent of exposure
12.2. Adverse events (AES)
12.2.1. Brief Summary of Adverse Events
12.2.2. Display of Adverse Events
12.2.3. Analysis of Adverse Events
12.2.4. Listing of Adverse Events by Patient
12.3. Deaths, other Serious Adverse Events and Other Significant Adverse Events

For peer review only

12.3.1. Listing of Deaths, other Serious Adverse Events and Other Significant Adverse Events
12.3.1.1. Deaths
12.3.1.2. Other Serious Adverse Events
12.3.1.3. Other Significant Adverse Events
12.3.2. Narratives of Deaths, Other Serious Adverse Events and Certain Other Significant Adverse Events
12.3.3. Analysis and Discussion of Deaths, Other Serious Adverse Events and Other Significant Adverse Events
12.4. Clinical laboratory evaluation
12.4.1. Listing of Individual Laboratory Measurements by Patient (16.2.8) and Each Abnormal Laboratory Value (14.3.4)
12.4.2. Evaluation of Each Laboratory Parameter
12.4.2.1. Laboratory Values Over Time
12.4.2.2. Individual Patient Changes
12.4.2.3. Individual Clinically Significant Abnormalities
12.5. Vital signs, physical findings and other observations related to safety
12.6. Safety conclusions
13. DISCUSSION AND OVERALL CONCLUSIONS
14. TABLES, FIGURES, AND GRAPHS REFERRED TO BUT NOT INCLUDED IN THE TEXT
14.1. Demographic data
14.2. Efficacy data
14.3. Safety data
14.3.1. Displays of Adverse Events
14.3.2. Listings of Deaths, Other Serious and Significant Adverse Events
14.3.3. Narratives of Deaths, Other Serious and Certain Other Significant Adverse Events
14.3.4. Abnormal Laboratory Value Listing (Each Patient)
15. REFERENCE LIST
16. APPENDICES
16.1. Study Information
16.1.1. Protocol and protocol amendments
16.1.2. Sample case report form (unique pages only)

For peer review only

16.1.3. List of IECs or IRBs (plus the name of the committee Chair if required by the regulatory authority) - Representative written information for patient and sample consent forms

16.1.4. List and description of investigators and other important participants in the study, including brief (1 page) CVs or equivalent summaries of training and experience relevant to the performance of the clinical study

16.1.5. Signatures of principal or coordinating investigator(s) or sponsor's responsible medical officer, depending on the regulatory authority's requirement

16.1.6. Listing of patients receiving test drug(s)/investigational product(s) from specific batches, where more than one batch was used

16.1.7. Randomisation scheme and codes (patient identification and treatment assigned)

16.1.8. Audit certificates (if available)

16.1.9. Documentation of statistical methods

16.1.10. Documentation of inter-laboratory standardisation methods and quality assurance procedures if used

16.1.11. Publications based on the study

16.1.12. Important publications referenced in the report

16.2. Patient Data Listings

16.2.1. Discontinued patients

16.2.2. Protocol deviations

16.2.3. Patients excluded from the efficacy analysis

16.2.4. Demographic data

16.2.5. Compliance and/or drug concentration data (if available)

16.2.6. Individual efficacy response data

16.2.7. Adverse event listings (each patient)

16.2.8. Listing of individual laboratory measurements by patient, when required by regulatory authorities

16.3. Case Report Forms

16.3.1. CRFs for deaths, other serious adverse events and withdrawals for AE

16.3.2. Other CRFs submitted

16.4. Individual Patient Data Listings (US Archival Listings)

For Peer review only

Clinical Study Reports of randomized controlled trials – an exploratory review of previously confidential industry reports

Authors:

Peter Doshi PhD
Johns Hopkins University School of Medicine, Baltimore, Maryland, USA

Tom Jefferson MD
The Cochrane Collaboration, Roma, Italy

Corresponding author: Peter Doshi <pnd@jhu.edu>

Word count: ~~2998~~3464

Tables: 3
Figures: 2
Appendices: 2
References: ~~41~~42

For peer review only

Patient consent statement No consent was necessary as no patients were involved

Ethics approval statement No ethical approval was necessary as no patients were involved and all data were aggregate or anonymized and publicly available.

Role of the sponsor statement As the review had no extramural funding, there was no sponsor.

Author Contributions: Doshi had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. *Study concept and design*: Doshi and Jefferson. *Acquisition of data*: Doshi and Jefferson. *Analysis and interpretation of data*: Doshi and Jefferson. *Critical revision of the manuscript for important intellectual content*: Doshi and Jefferson. *Statistical analysis*: Doshi.

For peer review only

Research questions or hypotheses addressed

What are Clinical Study Reports (CSRs)? What do they contain and how long are they?

Might CSRs help address reporting biases associated with the published literature, and improve the quality of evidence synthesis?

Key Messages (up to 3)

CSRs represent a hitherto hidden and untapped source of detailed RCT data (mean page length: 1,854 pages), increasingly becoming publicly available, and should form the basic unit for evidence synthesis to minimize the problem of reporting bias.

CSRs show that numerous individuals make important technical contributions to the design, conduct, and reporting of each trial, but journal publications often fail to record these details, resulting in a loss in individual responsibility for what is reported.

The E3 guideline to which most CSRs conform was published in 1995, and needs updating.

Strengths and Limitations

We cannot say whether our sample is representative and whether our conclusions are generalizable to an undefined and undefineable population of CSRs.

For peer review only

Abstract

Objective: To explore the structure and content of a non-random sample of clinical study reports (CSRs) to guide clinicians and systematic reviewers.

Search strategy: We searched public sources and lodged Freedom of Information requests for previously confidential CSRs primarily written by industry for regulators.

Selection criteria: CSR reporting sufficient information for extraction ("adequate")

Primary outcome measures: Presence and length of essential elements of trial design and reporting and compression factor (ratio of page length for CSR compared to its published counterpart in a scientific journal).

Data extraction: data were extracted on standard forms and cross-checked for accuracy

Results: We assembled a population of 7884 CSRs (covering 90 RCTs; 144,610 pages total) dated 1991-2011 of 14 pharmaceuticals. 78 were adequate. Report synopses had a median length of 5 pages, efficacy evaluation 13.5 pages, safety evaluation 17 pages, attached tables 337 pages, trial protocol 62 pages, statistical analysis plan 15 pages, and individual efficacy and safety listings had a median length of 447 and 109.5 pages, respectively. While 16 (21%) of CSRs contained completed case report forms, these were accessible to us in only one case (765 pages representing 16 individuals). Compression factors ranged between 1 and 8805.

Conclusions: Clinical study reports represent a hitherto mostly hidden and untapped source of detailed and exhaustive data on each trial. They should be consulted by independent parties interested in a detailed record of a clinical trial, and should form the basic unit for evidence synthesis as their use is likely to minimize the problem of reporting bias. We cannot say whether our sample is representative and whether our conclusions are generalizable to an undefined and undefineable population of CSRs.

Word count: 2725

Primary Funding Source: The review had no extramural funding.

For peer review only

Introduction

Systematic reviews are thought to provide one of the most robust ways to evaluate the effects of healthcare interventions. But the robustness of findings clearly rests upon reviewers' access to clinical trial information sufficient to critically evaluate and reproduce the original research. Research on reporting bias over the last decades has shown that trusting the published literature at face value, even peer-reviewed publications, can be fraught with difficulty—a problem that spans drug classes.[1–12]

Following the decision by the European regulator, European Medicines Agency (EMA) on 30 Nov 2010, to make available a broad spectrum of documents related to medicinal products for human and veterinary use,[13,14] attention is focusing on one particular type of regulatory document: clinical study reports (CSRs).[15–18] CSRs are usually written for regulators following guidelines developed by the industry-regulatory collaborative effort "International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use" (ICH). The ICH guidelines "Structure and Content of Clinical Study Reports"[19] (See Appendix 1) are known by the document code "E3". They were formalized in 1995 "to assist sponsors in the development of a report that is complete, free from ambiguity, well organised and easy [for regulators] to review."[19]. E3 has not been edited or changed since 1995.

CSRs are but one category of information that is transmitted from study sponsors to regulators (Figure 1), but are important as they contain substantially more information and detail on the intervention being tested than published versions of the same trial. The wealth of information may be sought with increasing frequency by researchers appraising single trials, entire trial programmes, or by those synthesizing evidence.[17,20] We are aware of two recent examples of systematic reviews of the effects of pharmaceuticals carried out using CSRs and other regulatory material.[12,21] One group also concluded that journal publications insufficiently report clinical trials.[22]

Despite CSRs' potential importance very little is known about their structure and content outside of those individuals with direct involvement in regulatory processes. This knowledge gap may hinder development of methods for fair and reliable appraisal of CSRs and their use in evidence synthesis. We are not aware of any instruments specifically designed for appraising CSRs. Lack of visibility may also hinder understanding of ~~conceal~~ the complexity of the organization and reporting of clinical trials.

We carried out an exploratory review to describe the structure and content of a non-random sample of clinical study reports. By describing the contents of CSRs, this research seeks to

For peer review only

transform CSRs from an obscure document only known to regulators and industry into a more widely known and accessible document. Our long-term intention is to improve the credibility of research synthesis by facilitating a move from the level of detail found in journal articles to the level of detail found in regulatory documents, thus guiding clinicians and other decision makers at all levels.

Methods

We obtained CSRs from public sources, as follows:

1. Requesting from EMA, under its freedom of information (FOI) policy, CSRs for manufacturer sponsored trials of the 10 best-selling prescription-bound products in the United States in 2010.[23]
2. Reusing CSRs from our own previous research (oseltamivir, zanamivir) [12]
3. Downloading CSRs openly available on the Internet. Search terms were not predefined, but sites searched included Google (http://www.google.com), the Drug Industry Document Archive (http://dida.library.ucsf.edu/), and IQWIG's library of reboxetine studies (https://www.iqwig.de/information-on-studies-of-reboxetine.980.en.html)
4. Corresponding with ~~other~~ one researcher~~s~~ who ~~have~~ obtained CSRs through a FOI request~~s~~ to FDA (epoetin alfa)
5. Requesting manufacturers fill any gaps in the completeness of reports that we believe are legally required to be publicly available (paroxetine).

To create as broad a database as possible, we did not apply restrictions in drug type or family or sponsor. We did not submit requests under the Freedom of Information Act to the Food and Drug Administration because such requests can take years to be fulfilled and—~~once~~ if fulfilled—may be heavily redacted.[24]

We did not draw a random sample of CSRs as there is no known sampling frame. No one knows how many reports have been written by intervention category as there is no central register of CSRs. Through familiarity with CSRs for oseltamivir and zanamivir, which were included in one of our Cochrane reviews,[12] we developed and piloted a data extraction sheet designed to capture the salient characteristics of CSRs . We created a list of around 40 potential sections we expected to find, generated directly from elements specified in E3. For each element in the list, we checked whether the obtained CSR included that section (confirmed either by direct identification of the section or an indication the section existed based on the CSR's table of contents), whether we had access to it, and its page length. Because of previous difficulties we had accessing CSR appendices, we also recorded whether sections

For peer review only

were listed as appendices or not. Page length was calculated either by directly counting the pages or by estimating their size from the table of contents of each report, and was used as a crude proxy for the level of detail available. Page lengths were rounded up to the next integer, and were summarized by reporting medians and ranges. We also included questions relating to trial registration and authorship. Our (blank) data extraction sheet is in Appendix 2.

All variables from CSRs were first extracted in single. We subsequently audited each other's extractions, checking the accurateness of the information. We chose to present elements analogous with those that typically appear in trials reported in scientific journals including the study Synopsis (a brief summary of the study), the study Protocol (written prospectively, describing the study methods), Efficacy and Safety Evaluations (a narrative summary of the efficacy and safety results of the study, including tables and figures), as well as attached tables. We also included elements rarely found in journal publications: sample (blank) and completed case report forms (CRFs are paper or electronic forms designed to capture pre-specified efficacy and safety related information for each study participant), the statistical analysis plan (a prospectively written narrative and/or statistical code indicating how trial data will be analyzed), and individual participant efficacy and safety listings. The corresponding E3 section numbers are listed in Table 2. Disagreements were resolved by discussion.

Our uncorrected (original) and corrected extraction sheets as well as audit records are available upon request from the corresponding author.

We calculated a compression factor for published trials~~.~~ which we defined as the ratio of CSR page length compared to the page length of the same trial as published in scientific journals. The objective of this metric was to convey a rough sense of how much information present in CSRs may be being condensed ("compressed") in short journal publications, in consideration of CSRs' far greater length and level of detail. Size (page length) reflects the level of detail as well as the presence of many elements such as protocols and their amendments, randomization lists, statistical analysis plans, certificates of analysis and extra data on subpopulations. We have demonstrated[12] that these elements are essential for understanding and appraising a trial. The compression factor is a crude measure of how much is compressed or simply left out of each publication which will affect the reliability of the appraisal and interpretation of trials. Trial publications were searched for in multiple sources: clinical trial registers, published systematic reviews, and correspondence with sponsors. Because in most cases we could not access all parts of all CSRs (and therefore do not know their complete page length), we calculated ~~both~~ "conservative" compression factors as well as ~~and~~ "realistic" compression factors. "Conservative" compression factors were calculated on a trial by trial basis using the total

For peer review only

number of pages ~~of~~ in CSRs available to us divided by the length of journal reports for that same trial, while "realistic" compression factors were based on the true total page length of the CSR~~, when known, even if inaccessible~~.

Results

We identified 84 documents believed to be CSRs for 14 compounds. These covered therapeutic and biological interventions including antipsychotics, antidepressants, antivirals, natural antiarthritics, anti-inflammatory agents, pandemic influenza vaccines, statins, erythropoietins, and anti-platelet compounds. We included English-language summaries of two Japanese oseltamivir studies (JV15823, JV15824) as they had been presented to EMA in this form. We excluded ~~CSRs~~ documents which were sections of CSRs but nonetheless contained insufficient information to understand the overall content of the CSR ~~were too fragmentary to evaluate~~ (olanzapine F1D-LC-HGAV, F1D-MC-HGAJ and F1D-MC-HGAO) and 3 documents which we had originally classified as CSRs ~~re not~~but were not ~~in fact CSRs~~ (reboxetine 14, 22 and 37). This left 78 CSRs (144,610 pages) (Figure 2). The median pages obtained per CSR was 644 (range 9 to 15,440). Only 4 of 78 CSRs (reboxetine 8, 16, 17, and 91) were written prior to November 30 1995 when ICH E3 was approved. Table 1 summarizes the pharmaceutical, manufacturer, date and provenance of the CSRs in our review. EMA reported not holding studies for esomepazole magnesium (Nexium), Advair diskus, quetiapine fumarate (Seroquel), montelukast sodium (Singulair), epoetin alfa (Epogen), and simvastatin.

All of the 78 included CSRs ~~comprised of~~contained a synopsis (median page length 5 pages). The efficacy evaluation was identifiable and directly accessible in 76 (97%, median length 13.5 pages) and safety in 77 (99%; median length 17 pages). Attached tables were likewise present in 63 (81%) ~~of~~ CSRs, and were a median of 337 pages long (range: 1 to 3665). Seventy-three CSRs (94%) reported including the study protocol. In the 40 we could access, the median page length was 62. We found blank CRFs included in 68 (87%) ~~of~~ CSRs. Of the 33 we could directly access, the median length was 133 pages (range 14 to 981). For completed CRFs, 16 (21%) reports made direct mention of a section on completed CRFs, but we had access to completed CRFs in only 1 case (Arthronat; length 765 pages).

Fifty-five (71%) of 78 included CSRs included a statistical analysis plan in some form. Of those for which we could directly access the content (n=37), the median page length was 15 (range 3 to 85). Individual efficacy and safety listings were included in 53 (69%) and 62 (81%) CSRs respectively. The median page length was 447 (range 15 to 21,698) for efficacy and 109.5 (range 2 to 10,954) for safety.

For peer review only

A summary is presented in Table 2.

All trial reports in our review were sponsored by industry.

Median conservative compression factors ranged between 1 and 1221. The realistic compression factors, calculated for the Arthronat, paroxetine, and clopidogrel CAPRIE trials, were 379, 1021, and 8805, respectively. (Table 3)

Discussion

We collected and described a sizeable number of CSRs written in the last two decades. All CSRs contained a table of contents (as specified in E3 section 3); this, together with optical character recognition (to enable searching the full text of the scanned documents) and the occasional need to combine multiple files to create a single document, substantially improved the ease of navigating CSRs.

Despite the ~~apparent~~ size of our non-random sample, ~~we are not sure~~it is unclear whether our conclusions are ~~broadly~~ generalisable to all other CSRs. This is because we have extremely limited knowledge about the total population of CSRs in regulators' and sponsors' possession. Nevertheless, within our sample spanning different manufacturers, therapeutic classes, and times, we found that the structure of CSRs was, within different house styles of presentation, strikingly similar ~~across medical products and sponsors~~, probably ~~thanks~~ due to the guidance by ICH~~'s~~ E3.[37] This suggests that the structure and content of other CSRs is likely to be similar.

The future basic currency of research synthesis?

The median length of 644 pages for reports in this study, as well as CSRs' routine inclusion of trials' protocol, statistical analysis plans, and blank case report forms, ~~confirms~~ strongly suggests that CSRs are the most detailed and complete, integrated form of reporting of the design, conduct, and results of clinical trials. In a study that directly compared the adequacy of reporting between journal articles and CSRs, the authors found that complete information regarding greater than 40% of methods items were only available in CSRs.[22] ~~They~~ The level of detail found in CSRs thus far surpass the level of detail available in journal publications, and as such they are prime candidates for the next basic currency of evidence synthesis and appraisal of a trial. Given the EMA's new policy making such documents publicly available, access to these documents is now relatively straightforward.[25] However including CSRs in systematic reviews is labor-intensive, given their size and complexity.[12]

For peer review only

Accessing complete CSRs

While CSRs may trump other forms of trial reporting in the public domain (such as conference abstracts or journal publications), serious limitations remain. Despite obtaining 144,610 pages for 78 CSRs, in almost all instances, we lacked full access to the CSRs' numerous appendices. Even for the sole complete CSR we obtained (Arthronat MA-CT-10-002), case report forms were provided for only 20% of participants. The Arthronat ttext does not provide a reason for this omission, but it reflects the vagueness of the relevant section of the E3 guidance (16.3.2) which does not define "Other CRFs submitted." Also, we could only access the original trial protocol in 40 (51%) of 78 CSRs obtained. This is important because trial protocols, written prior to patient enrollment in a trial, are an important way to guard against reporting biases.[26,27]

We could obtain individual patient listings in only a minority of cases despite confirming their inclusion in the majority of CSRs (Table 2). This may be a significant limitation, as the E3 specifies that "the report with its appendices should also provide enough individual patient data, including the demographic and baseline data, and details of analytical methods, to allow replication of the critical analyses…"[19] Unavailability was possibly due to the fact that EMA allows manufacturers to submit CSRs omitting a number of appendices including individual patient data and case report forms (which EMA states should be available within 48 hours if requested).[28] In the case of oseltamivir, the subject of a Cochrane review we conducted,[12] the manufacturer refused to share with us report appendices not submitted to EMA,[29] and EMA declined requesting them on our behalf.[8] Although FDA likely possesses more complete CSRs and patient level data, it historically has treated such data as trade secret and/or confidential.[30–32] EMA is therefore at present the only reliable source of obtaining CSRs. As such, despite European regulators' progressive stance—announcing that "clinical trial data should not be considered commercial confidential information"[33]—the completeness gap is unlikely to be filled any time soon.

Another significant limitation is that CSRs are only written for therapeutic, prophylactic, or diagnostic agents, and therefore inadequacies remain in evidence synthesis of other types of interventions such as surgical or behavioral interventions.

Individual participant listings

Individual participant listings—which identify participants by a unique ID—were accessible in 29 of the 78 CSRs we reviewed. But these data are difficult to analyze because they are presented as database printouts rather than in electronic form. This is understandable considering that CSRs are a written/archival format, but because EMA does not accept SAS datasets,[34,35] the industry standard, third-party access to databases of patient-level data remains elusive. We

For peer review only

see no compelling reason why all regulators should not request these from sponsors and make them publicly available. Whether availability of individual listings and CRFs, with its attendant laborious analysis, would increase our understanding of the trial and its results is unclear. But there is at least one case where the re-analysis of CRFs added invaluable knowledge to that already available in CSRs.[36]

The public-private debate

One manufacturer has claimed that the non-release of case report forms is motivated by concerns over protecting participant confidentiality.[38] Nothing we have seen so far corroborates this claim, however an ongoing EMA working group is specifically discussing issues related to protecting participant confidentiality. ~~The~~ Based on current document releases and position statements, however, it appears that EMA has deemed case report forms and individual patient listings to be, in principle, releasable in their entirety (after a preliminary review).[39] Furthermore, individual patient listings are intended to duplicate information contained in filled case report forms. The release of case report forms would ensure the accuracy of individual patient listings with little additional risk to patient confidentiality. Moreover, extra checks such as registration of protocols by bona fide research groups could deter any inappropriate use. We also believe that the sheer bulk of the forms acts as a deterrent against malice.

Size matters

Our range of compression factors show the scale of selection and synthesis which must (consciously or unconsciously) occur in the process of transforming CSRs into journal-length articles. We found a strong resemblance in detail, page length, structure, and purpose between the short Synopsis section of CSRs and reports of trials as published in scientific journals. ~~.~~ In some cases essential items of information such as the trial protocol and its subsequent amendments are simply not included in journal articles or are replaced by methods written *post facto*. In other cases of items essential for the interpretations of the trial results (such as the statistical analysis plan), tens of pages are reduced to a paragraph on sample size calculation in the journal report, underscoring the lack of detail (and its attendant problems) common to public forms of trial reporting. For example, the ratio of words in Protocol of the CSR for Aripiprazole CN138135 to the Methods section for published journal article of the same trial is 30.5 (53,713 words in the CSR Protocol versus 1,763 words in the journal article). For the oseltamivir WP16263 trial, the ratio was 22.7 (26,761 words in the CSR Protocol and amendments versus 1,177 words in the journal article).

This compression of information ~~is true even~~also occurs in databases not restricted by length, such as ClinicalTrials.gov.[40]

For peer review only

Our study raises the question of why the medical community has accepted the low (summary, aggregate) level of detail found in most peer-reviewed journal publications compared to the depth of detail available in CSRs. European regulators recently noted: "Documents that provide critical information on a study, such as the protocol (16.1.1), statistical methods (16.1.9), list of investigators and study sites and sample case report forms, would always be needed by reviewers assessing a study"[41] Why have those outside of the regulatory world tolerated journal publications lacking such details?

One possibility may be that while the clinical trial enterprise has changed dramatically in the last half century, the scientific journal publication model has not. Since the 1950s, there have been considerable transformations in the political economy of clinical trials driven by the increasingly commercialized and global nature of the pharmaceutical industry, the rise in academic-industry "partnerships" in medicine, and increased communication among regulators. It is now common to find trials with study centers scattered around the globe. This increasing complexity and the need to provide an audit record is reflected in the comprehensive tomes documenting the trials—CSRs—but trial reporting in scientific journals remains limited to summary and aggregate details. It should be noted, however, that many journals now have websites which enables them to make available extended content beyond what traditionally appears in the printed journal.

Authorship or Contributorship?

Examination of CSRs revealed scores of important technical contributions to the design, conduct, and reporting of each trial. These included contributions from database programmers, records officers, and CSR writers, often invisible in the published journal article. In some cases, we found no mention in CSRs of individuals who figured as authors of subsequent published trial reports while individuals named as CSR authors went unacknowledged in journal publications. Current ICMJE guidelines on authorship and contributorship are largely focused on ensuring those placed on by-lines deserve to be authors. But the guidelines also suggest that "all contributors who do not meet the criteria for authorship should be listed in an acknowledgments section."[42] Given the complexity of clinical trials, the ICMJE should call for itemized contributorship: the names of all contributors to be specified along with their role in the design, conduct, analysis, or reporting of the trial. If the contribution of most people goes unrecorded, so does their individual responsibility for what is produced. Itemized contributorship records, to all phases of a trial, could be piloted in trial registers.

For peer review only

E3 guidance

The E3 guideline set an excellent standard, but it needs formal updating and further development. For example, there should be a self-standing set of definitions for terms such as "case report forms" and "Other CRF's submitted," (section 16.3.2) and a description of how a particular trial fits within a sponsor's trial programme of pharmaceutical development. Apparently forgotten items such as certificates of analysis (describing the appearance and content of the interventions being tested) and post-1995 details such as trial registration numbers should be mentioned.

We hope our review has given CSRs what they have lacked so far: visibility. CSRs represent a largely untapped source of detailed data that we believe can serve as a means of addressing the ravages of reporting bias in all its forms, leading to a more accurate understanding of the effects of medicines.

Conflicts of interest statement

All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare that:

Both authors are co-recipients of a UK National Institute for Health Research grant to carry out a Cochrane review of neuraminidase inhibitors (http://www.hta.ac.uk/2352).

Tom Jefferson was an ad hoc consultant for F. Hoffman-La Roche Ltd in 1998-1999. He receives royalties from his books published by Blackwells and Il Pensiero Scientifico Editore, none of which are on clinical study reports. He is occasionally interviewed by market research companies for anonymous interviews about Phase 1 or 2 products unrelated to products in this review. In 2011-12 he has acted as an expert witness in a litigation case related to one of the compounds in the review (oseltamivir). He is on a legal retainer for expert advice on litigation for influenza vaccines in health care workers.

Peter Doshi received €1500 from the European Respiratory Society in support of his travel to the society's September 2012 annual congress where he gave an invited talk on oseltamivir.

Both authors' spouses and children have no financial relationships that may be relevant to the submitted work.

For peer review only

Data sharing statement

The original extraction forms and audit record are available on request from the corresponding author.

Acknowledgements

We thank Drs Vallance and Kraus of GlaxoSmithKline for making public selected report appendices from the 9 paroxetine trials. We also thank Daniel Coyne for sharing the CSR that FDA sent him in response to his Freedom of Information request, and Iain Chalmers for guidance.

Peter Doshi is funded by an institutional training grant from the Agency for Healthcare Research and Quality #T32HS019488. AHRQ had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

For peer review only

References

1. Rising K, Bacchetti P, Bero L. Reporting Bias in Drug Trials Submitted to the Food and Drug Administration: Review of Publication and Presentation. Ioannidis J, editor. Plos Med. 2008;5(11):e217.

2. Doshi P, Jones M, Jefferson T. Rethinking credible evidence synthesis. BMJ. 2012 Jan 17;344(jan17 2):d7898.

3. Chan A-W, Hróbjartsson A, Haahr MT, Gøtzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. JAMA. 2004 May 26;291(20):2457–65.

4. Gøtzsche PC. Why we need easy access to all data from all clinical trials and how to accomplish it. Trials. 2011;12(1):249.

5. Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R. Selective publication of antidepressant trials and its influence on apparent efficacy. N. Engl. J. Med. 2008 Jan 17;358(3):252–60.

6. Dickersin K, Min YI. Publication bias: the problem that won't go away. Ann. N. Y. Acad. Sci. 1993 Dec 31;703:135–146; discussion 146–148.

7. Hopewell S, Loudon K, Clarke MJ, Oxman AD, Dickersin K. Publication bias in clinical trials due to statistical significance or direction of trial results. Cochrane Database Syst Rev. 2009;(1):MR000006.

8. Melander H. Evidence b(i)ased medicine--selective reporting from studies sponsored by pharmaceutical industry: review of studies in new drug applications. BMJ. 2003 May 29;326:1171–3.

9. McGauran N, Wieseler B, Kreis J, Schuler Y-B, Kolsch H, Kaiser T. Reporting bias in medical research - a narrative review. Trials. 2010;11(1):37.

10. Sismondo S, Doucet M. Publication ethics and the ghost management of medical publication. Bioethics. 2009 Feb;24(6):273–83.

11. Vedula SS, Bero L, Scherer RW, Dickersin K. Outcome reporting in industry-sponsored trials of gabapentin for off-label use. N. Engl. J. Med. 2009 Nov 12;361(20):1963–71.

12. Jefferson T, Jones MA, Doshi P, Del Mar CB, Heneghan CJ, Hama R, et al. Neuraminidase inhibitors for preventing and treating influenza in healthy adults and children. Cochrane Database of Systematic Reviews. 2012;(1):CD008965.

13. European Medicines Agency. European Medicines Agency policy on access to documents (related to medicinal products for human and veterinary use) POLICY/0043 [Internet]. 2010 [cited 2012 May 14]. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Other/2010/11/WC500099473.pdf

For peer review only

14. European Medicines Agency. Output of the European Medicines Agency policy on access to documents related to medicinal products for human and veterinary use [Internet]. 2010 [cited 2011 Sep 25]. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2010/11/WC500099472.pdf

15. Gøtzsche PC, Jørgensen AW. Opening up data at the European Medicines Agency. BMJ. 2011;342:d2686.

16. Jefferson T, Doshi P, Thompson M, Heneghan C. Ensuring safe and effective drugs: who can do what it takes? BMJ. 2011;342:c7258.

17. Chan A-W. Out of sight but not out of mind: how to search for unpublished clinical trial evidence. BMJ. 2012 Jan 3;344(jan03 2):d8013–d8013.

18. Rodwin MA, Abramson JD. Clinical trial data as a public good. JAMA. 2012 Sep 5;308(9):871–2.

19. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. Structure and Content of Clinical Study Reports: E3 [Internet]. 1995 [cited 2012 Jul 8]. Available from: http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E3/E3_Guideline.pdf

20. Grens K. Data Diving. The Scientist [Internet]. 2012 Apr [cited 2012 May 3]; Available from: http://the-scientist.com/2012/05/01/data-diving/

21. Eyding D, Lelgemann M, Grouven U, Harter M, Kromp M, Kaiser T, et al. Reboxetine for acute treatment of major depression: systematic review and meta-analysis of published and unpublished placebo and selective serotonin reuptake inhibitor controlled trials. BMJ. 2010 Oct 12;341(oct12 1):c4737–c4737.

22. Wieseler B, Kerekes MF, Vervoelgyi V, McGauran N, Kaiser T. Impact of document type on reporting quality of clinical drug trials: a comparison of registry reports, clinical study reports, and journal publications. BMJ. 2012 Jan 3;344(jan03 1):d8141–d8141.

23. IMS Institute for Healthcare Informatics. The Use of Medicines in the United States: Review of 2010 [Internet]. 2011 [cited 2012 Jul 30]. Available from: http://www.imshealth.com/imshealth/Global/Content/IMS%20Institute/Documents/IHII_UseOfMed_report%20.pdf

24. Coyne DW. The health-related quality of life was not improved by targeting higher hemoglobin in the Normal Hematocrit Trial. Kidney International [Internet]. 2012 Mar 21 [cited 2012 Apr 13]; Available from: http://www.nature.com/ki/journal/vaop/ncurrent/full/ki201276a.html

25. Doshi P, Jefferson T. The first 2 years of the European Medicines Agency's policy on access to documents: secret no longer. Arch Intern Med. 2012 Dec 19;1–2.

26. Chan A-W. Bias, Spin, and Misreporting: Time for Full Access to Trial Protocols and Results. Plos Med. 2008;5(11):e230.

For peer review only

27. Miller JD. Registering clinical trial results: the next step. JAMA. 2010 Feb 24;303(8):773–4.

28. European Medicines Agency. Note for guidance on the inclusion of appendices to clinical study reports in marketing authorisation applications [Internet]. 2004 [cited 2012 Jul 30]. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003638.pdf

29. Doshi P, Jefferson T, Del Mar C. The Imperative to Share Clinical Study Reports: Recommendations from the Tamiflu Experience. PLoS Med. 2012 Apr 10;9(4):e1001201.

30. Unlu M. It is time: why the FDA should start disclosing drug trial data. Mich. Telecomm. Tech. L. Rev. 2010;16:511–45.

31. Kesselheim AS, Mello MM. Confidentiality laws and secrecy in medical research: improving public access to data on drug safety. Health Aff (Millwood). 2007 Apr;26(2):483–91.

32. Halperin RM. FDA Disclosure of Safety and Effectiveness Data: A Legal and Policy Analysis. Duke Law Journal. 1979 Feb 1;1979(1):286–326.

33. Eichler H-G, Abadie E, Breckenridge A, Leufkens H, Rasi G. Open Clinical Trial Data for All? A View from Regulators. PLoS Med. 2012 Apr 10;9(4):e1001202.

34. European Medicines Agency. EMEA implementation of electronic-only submission and eCTD submission: Practical guidelines relating to non-eCTD electronic submissions [Internet]. 2008 [cited 2012 Jul 31]. Available from: http://www.emea.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2009/10/WC500004100.pdf

35. European Medicines Agency. EMEA implementation of electronic-only submission and eCTD submission: Questions and answers relating to practical and technical aspects of the implementation [Internet]. 2008 [cited 2012 Jul 31]. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Regulatory_and_procedural_guideline/2009/10/WC500004095.pdf

36. Psaty BM, Prentice RL. Minimizing bias in randomized trials: the importance of blinding. JAMA. 2010 Aug 18;304(7):793–4.

37. Molzon JA, Giaquinto A, Lindstrom L, Tominaga T, Ward M, Doerr P, et al. The value and benefits of the International Conference on Harmonisation to drug regulatory authorities: advancing harmonization for better public health. Clin. Pharmacol. Ther. 2011 Apr;89(4):503–12.

38. GlaxoSmithKline. Paroxetine and pediatric and adolescent patients [Internet]. 2012 [cited 2012 Jul 30]. Available from: http://www.gsk.com/media/paroxetine.htm

39. HMA/EMA Working Group on Transparency. HMA/EMA guidance document on the identification of commercially confidential information and personal data within the structure of the marketing authorisation (MA) application - release of information after the granting of a marketing authorisation [Internet]. 2012 [cited 2012 Jul 30]. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Other/2012/03/WC500124536.pdf

For peer review only

40. Zarin DA, Tse T, Williams RJ, Califf RM, Ide NC. The ClinicalTrials.gov Results Database — Update and Key Issues. New England Journal of Medicine. 2011;364(9):852–60.

41. European Medicines Agency. ICH guideline E3 - questions and answers (R1) [Internet]. 2012 [cited 2012 Jul 30]. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2012/07/WC500130356.pdf

42. International Committee of Medical Journal Editors. Uniform Requirements for Manuscripts Submitted to Biomedical Journals: Ethical Considerations in the Conduct and Reporting of Research: Authorship and Contributorship [Internet]. 2009 [cited 2012 Jul 31]. Available from: http://www.icmje.org/ethical_1author.html

For peer review only

Table 1. Pharmaceutical, trials, producers, dates and sources of CSRs in the review.

Pharmaceutical and number (n) of assessed trial documents	Trial IDs	Manufacturer	Date of CSRs	Provenance in our study
Aripiprazole (Abilify) n=1	CN1368135	Bristol-Myers Squibb	2007	Freedom of Information request to EMA
Arthronat n=1	MA-CT-10-002	Rowtasha	2011	Manufacturer website http://arthronat.com/clinical-study.php
Atorvastatin (Lipitor) n=1	981-080	Pfizer	1999	Freedom of Information request to EMA
Clopidogrel (Plavix) n=5	CURE, CLARITY, COMMIT-CCS2, CAPRIE, PICOLO	Bristol-Myers Squibb	1997-2007	Freedom of Information request to EMA
Epoetin alfa (Epogen) n=1	930107	Amgen	1996	Freedom of Information request to FDA
H5N1 influenza vaccine n=1	H5N1-008, H5N1-011 EXT 008	GSK	2006	Freedom of Information request to EMA
H5N1 influenza vaccines n=2	V87P1, V87P6	Novartis	2008-2009	Freedom of Information request to EMA
Olanzapine (Zyprexa) n=3	F1D-LC-HGAV*, F1D-MC-HGAO*, F1D-MC-HGAJ*	Eli Lilly	1995†	Litigation http://zyprexalitigationdocuments.com/unsealed.php http://www.furiousseasons.com/zyprexadocs.html
Oseltamivir (Tamiflu) n=19	JV15823, JV15824, M76001, NP15757, NV16871, WP16263, WV15670, WV15671, WV15673 WV15697, WV15707, WV15708, WV15730, WV15758, WV15759 WV15871, WV15799, WV15812 WV15872, WV15819 WV15876 WV15978, WV15825, WV16193	Roche	1999-2004	Documents obtained as part of previous Cochrane review[12]
Paroxetine (Paxil, Aropax, Pexeva, Seroxat, Sereupin) n=9	329, 377, 453, 511, 676, 701, 704, 715, 716	GSK	1998-2002	Litigation (2004 legal settlement mandated release of clinical study reports on manufacturer's website of 9 studies on

For peer review only

				pediatric and adolescent patients) http://www.gsk.com/media/paroxetine.htm
Quetiapine (Seroquel) n=7	015, 041, 049, 125, 126, 127, 135	AstraZeneca	1996-2007	Litigaton http://psychrights.org/research/Digest/NLPs/Seroquel/UnsealedSeroquelStudies/
Reboxetine (Edronax, Norebox, Prolift, Solvex, Davedax, Vestra) n=24	8, 9, 13, 14*, 15, 16, 17, 22*, 32, 32a, 34, 35, 37*, 43, 45, 46, 47, 49, 50, 52, 71, 83, 91, 96	Pfizer	1991-2009	Health Technology ~~Assessment~~Assessment website (The German IQWiG obtained CSRs as part of its health technology assessment work) https://www.iqwig.de/information-on-studies-of-reboxetine.980.en.html
Rofecoxib (Vioxx) n=1	78	Merck	2003	Litigation http://dida.library.ucsf.edu/
Zanamivir (Relenza) n=9	NAI30009, NAI300010, NAIA2005, NAIA3002, NAIA3005, NAIB2005, NAIB2007, NAIB3001, NAIB3002	GSK	1998-1999	Documents obtained as part of previous Cochrane review[12]

520 * Subsequently excluded because of insufficient documentation

521 † H1D-MC-HGAO clinical study report date unknown

522 EMA = European Medicines Agency

523 FDA = Food and Drug Administration

524

Field Code Changed

For peer review only

525 **Table 2. Key characteristics of the CSRs in the review**

526

Section of CSR (corresponding section of E3)	Presence	Length	
	CSRs including section, n	CSRs with section length available, n	Median length (range), pages
Synopsis (E3 section 2)	78 (100%)	78	5 (1 - 15)
Efficacy evaluation (E3 sec. 11)	76 (97%)	77	13.5 (2 - 132)
Safety evaluation (E3 sec. 12)	77 (99%)	58	17 (2 - 188)
Attached tables not in report text (E3 sec. 14)	63 (81%)	76	337 (1 - 3665)
Protocol (E3 sec 16.1.1)	73 (94%)	41	62 (21 - 139)
Blank Case Report Form (CRF) (E3 sec. 16.1.2)	68 (87%)	33	133 (14 - 981)
Statistical Analysis Plan (E3 sec. 16.1.9)	55 (71%)	37	15 (3 - 85)
Individual participant efficacy listings (E3 sec. 16.2.6)	53 (69%)	19	447 (15 - 21698)
Individual participant safety listings (E3 sec. 16.2.7)	62 (81%)	26	109.5 (2 - 10954)
Completed CRFs (E3 sec. 16.3.2)	16 (21%)	1	765

527

528

For peer review only

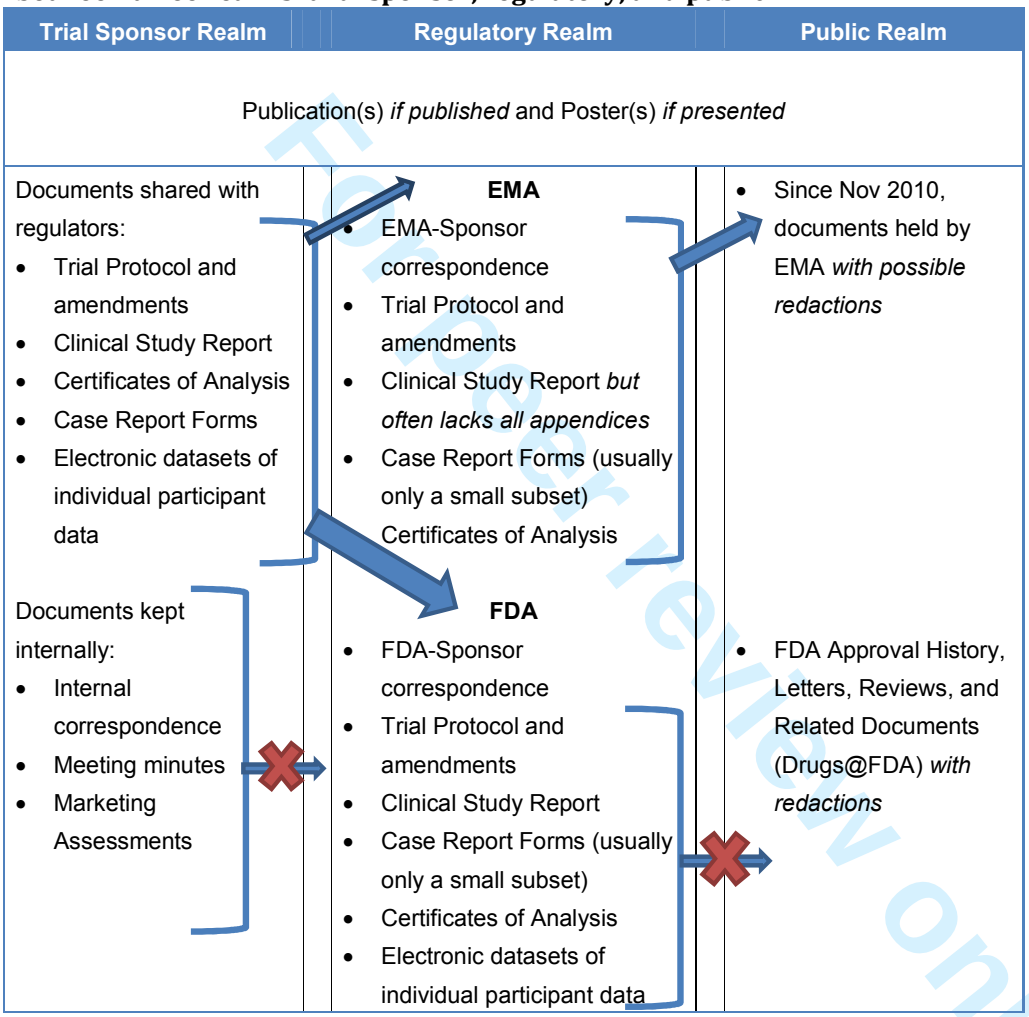
Table 3. Conservative and realistic compression factors. A ratio of CSR page length to corresponding journal publication page length.

Pharmaceutical	Studies published in journals, n	Mean compression factor (range)
Conservative compression factors		
Aripiprazole	1	672
Clopidogrel	5	11 (4 - 19)
Epoetin Alfa	1	41
Fluad	2	488 (367 - 609)
GSK H5N1 vaccine	1	19
Oseltamivir	12	195 (1 - 1221)
Quetiapine	2	578 (352 - 803)
Reboxetine	5	88 (9 - 245)
Zanamivir	8	54 (28 - 92)
Realistic compression factors		
Arthronat*	1	379
Clopidogrel	1	8805
Paroxetine	9	1021 (50 - 5473)

* The Arthronat trial has not yet been published. Compression factor calculation is based on the page length of a draft manuscript "to be published soon," according to Arthronat.com.

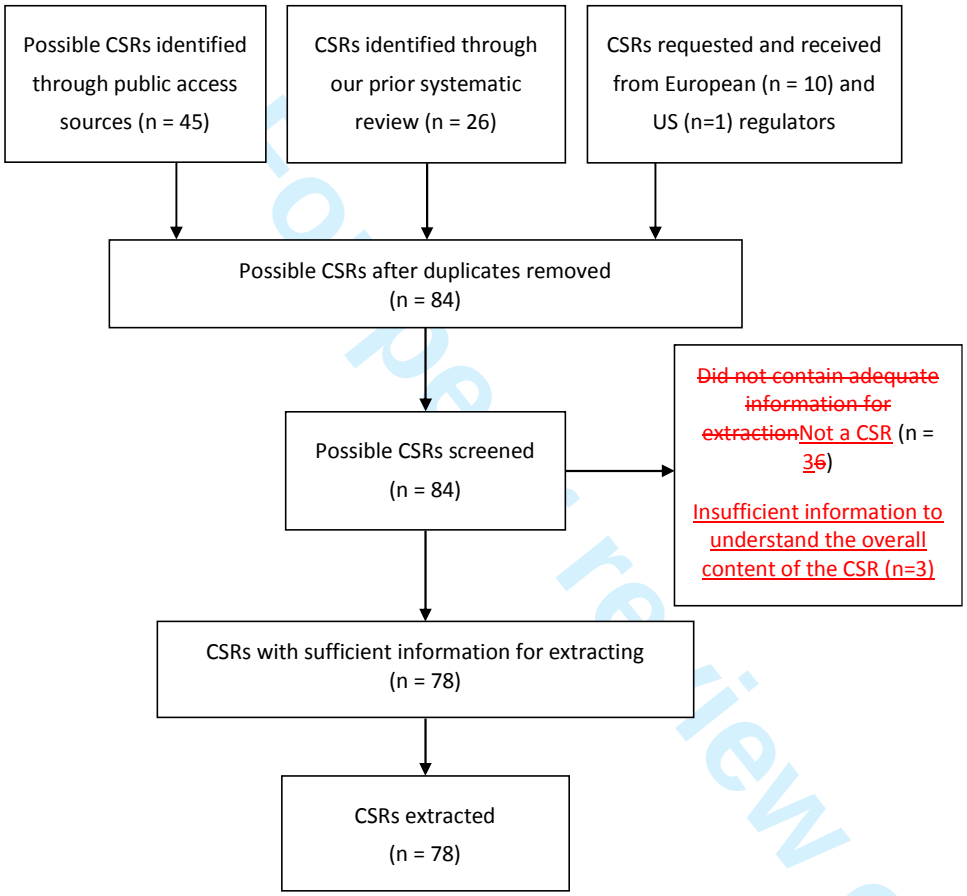
For peer review only

Figure 1. Types of clinical trial data typically held within and transferred between three realms: trial sponsor, regulatory, and public.



For peer review only

Figure 2. Study flow



For peer review only

Appendix 1. Elements specified ICH E3 "Structure and Content of Clinical Study Reports" (1995)[19]

1. TITLE PAGE
2. SYNOPSIS
3. TABLE OF CONTENTS FOR THE INDIVIDUAL CLINICAL STUDY REPORT
4. LIST OF ABBREVIATIONS AND DEFINITION OF TERMS
5. Ethics
    - 5.1. Independent Ethics Committee (IEC) or Institutional Review Board (IRB)
    - 5.2. Ethical conduct of the study
    - 5.3. Patient information and consent
6. INVESTIGATORS AND STUDY ADMINISTRATIVE STRUCTURE
7. INTRODUCTION
8. STUDY OBJECTIVES
9. INVESTIGATIONAL PLAN
    - 9.1. Overall study design and plan – description
    - 9.2. Discussion of study design, including the choice of control groups
    - 9.3. Selection of study population
        - 9.3.1. Inclusion criteria
        - 9.3.2. Exclusion criteria
        - 9.3.3. Removal of Patients from Therapy or Assessment
    - 9.4. Treatments
        - 9.4.1. Treatments Administered
        - 9.4.2. Identity of Investigational Product(s)
        - 9.4.3. Method of Assigning Patients to Treatment Groups
        - 9.4.4. Selection of Doses in the Study
        - 9.4.5. Blinding
        - 9.4.6. Prior and Concomitant Therapy
        - 9.4.7. Treatment Compliance
    - 9.5. Efficacy and safety variables
        - 9.5.1. Efficacy and Safety Measurements Assessed and Flow Chart
        - 9.5.2. Appropriateness of Measurements
        - 9.5.3. Primary Efficacy Variable(s)
        - 9.5.4. Drug Concentration Measurements
    - 9.6. Data quality assurance
    - 9.7. Statistical methods planned in the protocol and determination of sample size

For peer review only

9.7.1. Statistical and Analytical Plans
9.7.2. Determination of Sample Size
9.8. Changes in the conduct of the study or planned analyses
10. STUDY PATIENTS
10.1. Disposition of patients
10.2. Protocol deviations
11. EFFICACY EVALUATION
11.1. Data sets analyzed
11.2. Demographic and other baseline characteristics
11.3. Measurements of treatment compliance
11.4. Efficacy results and tabulations of individual patient data
11.4.1. Analysis of efficacy
11.4.2. Statistical/analytical issues
11.4.2.1. Adjustments for covariates
11.4.2.2. Handling of Dropouts or Missing Data
11.4.2.3. Interim Analyses and Data Monitoring
11.4.2.4. Multicentre Studies
11.4.2.5. Multiple Comparison/Multiplicity
11.4.2.6. Use of an "Efficacy Subset" of Patients
11.4.2.7. Active-Control Studies Intended to Show Equivalence
11.4.2.8. Examination of Subgroups
11.4.3. Tabulation of Individual Response Data
11.4.4. Drug Dose, Drug Concentration, and Relationships to Response
11.4.5. Drug-Drug and Drug-Disease Interactions
11.4.6. Drug Dose, Drug Concentration, and Relationships to Response
11.4.7. By-Patient Displays
12. SAFETY EVALUATION
12.1. Extent of exposure
12.2. Adverse events (AES)
12.2.1. Brief Summary of Adverse Events
12.2.2. Display of Adverse Events
12.2.3. Analysis of Adverse Events
12.2.4. Listing of Adverse Events by Patient
12.3. Deaths, other Serious Adverse Events and Other Significant Adverse Events

For peer review only

12.3.1. Listing of Deaths, other Serious Adverse Events and Other Significant Adverse Events
12.3.1.1. Deaths
12.3.1.2. Other Serious Adverse Events
12.3.1.3. Other Significant Adverse Events
12.3.2. Narratives of Deaths, Other Serious Adverse Events and Certain Other Significant Adverse Events
12.3.3. Analysis and Discussion of Deaths, Other Serious Adverse Events and Other Significant Adverse Events
12.4. Clinical laboratory evaluation
12.4.1. Listing of Individual Laboratory Measurements by Patient (16.2.8) and Each Abnormal Laboratory Value (14.3.4)
12.4.2. Evaluation of Each Laboratory Parameter
12.4.2.1. Laboratory Values Over Time
12.4.2.2. Individual Patient Changes
12.4.2.3. Individual Clinically Significant Abnormalities
12.5. Vital signs, physical findings and other observations related to safety
12.6. Safety conclusions
13. DISCUSSION AND OVERALL CONCLUSIONS
14. TABLES, FIGURES, AND GRAPHS REFERRED TO BUT NOT INCLUDED IN THE TEXT
14.1. Demographic data
14.2. Efficacy data
14.3. Safety data
14.3.1. Displays of Adverse Events
14.3.2. Listings of Deaths, Other Serious and Significant Adverse Events
14.3.3. Narratives of Deaths, Other Serious and Certain Other Significant Adverse Events
14.3.4. Abnormal Laboratory Value Listing (Each Patient)
15. REFERENCE LIST
16. APPENDICES
16.1. Study Information
16.1.1. Protocol and protocol amendments
16.1.2. Sample case report form (unique pages only)

For peer review only

16.1.3. List of IECs or IRBs (plus the name of the committee Chair if required by the regulatory authority) - Representative written information for patient and sample consent forms

16.1.4. List and description of investigators and other important participants in the study, including brief (1 page) CVs or equivalent summaries of training and experience relevant to the performance of the clinical study

16.1.5. Signatures of principal or coordinating investigator(s) or sponsor's responsible medical officer, depending on the regulatory authority's requirement

16.1.6. Listing of patients receiving test drug(s)/investigational product(s) from specific batches, where more than one batch was used

16.1.7. Randomisation scheme and codes (patient identification and treatment assigned)

16.1.8. Audit certificates (if available)

16.1.9. Documentation of statistical methods

16.1.10. Documentation of inter-laboratory standardisation methods and quality assurance procedures if used

16.1.11. Publications based on the study

16.1.12. Important publications referenced in the report

16.2. Patient Data Listings

16.2.1. Discontinued patients

16.2.2. Protocol deviations

16.2.3. Patients excluded from the efficacy analysis

16.2.4. Demographic data

16.2.5. Compliance and/or drug concentration data (if available)

16.2.6. Individual efficacy response data

16.2.7. Adverse event listings (each patient)

16.2.8. Listing of individual laboratory measurements by patient, when required by regulatory authorities

16.3. Case Report Forms

16.3.1. CRFs for deaths, other serious adverse events and withdrawals for AE

16.3.2. Other CRFs submitted

16.4. Individual Patient Data Listings (US Archival Listings)

For peer review only

Basic Extraction Information

Questions	Answer	Notes
1. Drug common name:		
2. Trial ID:		
➔ Now, fill in the drug and trial ID in the bottom-right corner the page.	E.g. "Tamiflu, WV15670"	
➔ Now, save this file under a new filename	Use the naming convention "*Drugname Trial ID - Extractor's initials - YYYYMMDD.docx*", e.g. "Seroquel 015 - TJ - 20120311.docx"	
3. Report/CSR ID (if different from Trial ID):		
4. Extractor's name (Initials)		
5. Date of extraction		

Notes to extractor:

- Page numbers should be referred to by the format p.(page # as printed)/PDFp.(PDF page number, possibly indicating volume), e.g.
  - p.V-235/PDFp.945 = page "V-235", on PDF page 945
  - p.234/PDF(3)p.18 = page "234", on the 3rd PDF for this CSR, PDF page 18
- Most questions can be answered with a Y or N (indicating Yes or No) or a number (e.g. the number of PDF pages.
- Where specified as "Free form answer", the extractor may answer in his/her own words based on the extractor's reading of the CSR.

Item	Content	Notes
Overview questions 6. Does the CSR list a ISRCTN/NCT or equivalent registration number for this trial?		
7. List CSR number of authors		
8. List CSR authors & trialists (Copy names if available; "redacted" if redacted; "not listed" if not listed)		
9. Total length of CSR obtained, in PDF pages		
10. List CSR completion date		
11. Is the trial published?		
12. If Y give publication citation		
13. If Y give publication size (in pages)		
14. Who appears to be responsible for CSR? (Free form answer)		
Trial programme questions 15. How many trials appear to be in the trial programme?		
16. Does CSR indicate where this trial fits in the trial programme? (Free form answer)		
17. Does CSR say how much of the trial programme is published?		
18. How many trials are in possession of a ISRCTN/NCT or equivalent registration number?		
Basic elements of the Clinical Study Report		

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

19.	Does the CSR contain a **table of contents**?		
20.	If Y, is the **table of contents** listed as an Appendix?		
21.	If Y, is the **table of contents** accessible to us?		
22.	If Y, how long is the **table of contents** (in pages)?		
23.	Does the table of contents list a **title page**?		
24.	If Y, is the **title page** listed as an Appendix?		
25.	If Y, is the **title page** accessible to us?		
26.	If Y, how long is the **title page** (in pages)?		
27.	Does the table of contents list a **synopsis**?		
28.	If Y, is the **synopsis** listed as an Appendix?		
29.	If Y, is the **synopsis** accessible to us?		
30.	If Y, how long is the **synopsis** (in pages)?		
31.	Does the CSR contain a **list of abbreviations and definitions**?		
32.	If Y, is the **list of abbreviations and definitions** listed as an Appendix?		
33.	If Y, is the **list of abbreviations and definitions** accessible to us?		
34.	If Y, how long is the **list of abbreviations and definitions** (in pages)?		
35.	Does the CSR contain an **ethics section**?		
36.	If Y, is the **ethics section** listed as an Appendix?		
37.	If Y, is the **ethics section** accessible to us?		
38.	If Y, how long is the **ethics section** (in pages)?		
39.	Does the CSR contain a **investigators and study administrative structure**?		
40.	If Y, is the **investigators and study administrative structure** listed as an Appendix?		
41.	If Y, is the **investigators and study administrative structure** accessible to us?		
42.	If Y, how long is the **investigators and study administrative structure** (in pages)?		
43.	Does the CSR contain an **introduction**?		
44.	If Y, is the **introduction** listed as an Appendix?		
45.	If Y, is the **introduction** accessible to us?		
46.	If Y, how long is the **introduction** (in pages)?		
47.	Does the CSR contain a section on study objectives?		
48.	If Y, is the **section on study objectives** listed as an Appendix?		
49.	If Y, is the **section on study objectives** accessible to us?		
50.	If Y, how long is the **section on study objectives** (in pages)?		
51.	Does the CSR contain an **investigational plan** (from IHR 1995 E3, PDF p.13)?		
52.	If Y, is the **investigational plan** listed as an Appendix?		
53.	If Y, is the **investigational plan** accessible to us?		
54.	If Y, how long is the **investigational plan** (in pages)?		
55.	Does the CSR contain a section on **study patients**?		
56.	If Y, is the **study patients** listed as an Appendix?		
57.	If Y, is the **study patients** accessible to us?		
58.	If Y, how long is the **study patients** (in pages)?		

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

59.	If Y, does it include a list of protocol deviations?		
60.	Does the CSR contain a section on **efficacy evaluation**?		
61.	If Y, is the **efficacy evaluation** listed as an Appendix?		
62.	If Y, is the **efficacy evaluation** accessible to us?		
63.	If Y, how long is the **efficacy evaluation** (in pages)?		
64.	Does the CSR contain a section on **safety evaluation**?		
65.	If Y, is the **safety evaluation** listed as an Appendix?		
66.	If Y, is the **safety evaluation** accessible to us?		
67.	If Y, how long is the **safety evaluation** (in pages)?		
68.	Does the CSR contain a **discussion and overall conclusions** section?		
69.	If Y, is the **discussion and overall conclusions** listed as an Appendix?		
70.	If Y, is the discussion and overall conclusions accessible to us?		
71.	If Y, how long is the **discussion and overall conclusions** (in pages)?		
72.	Does the CSR contain a section on **tables, figures and graphs referred to but not included in the text**?		
73.	If Y, is the **tables, figures and graphs referred to but not included in the text** listed as an Appendix?		
74.	If Y, is the **tables, figures and graphs referred to but not included in the text** accessible to us?		
75.	If Y, how long is the **tables, figures and graphs referred to but not included in the text** (in pages)?		
76.	Does the CSR contain a **references** section?		
77.	If Y, is the **references** listed as an Appendix?		
78.	If Y, is the **references** accessible to us?		
79.	If Y, how long is the **references** (in pages)?		
Appendices related questions			
80.	Does the **table of contents** indicate that the CSR contains **appendices**?		
81.	If Y, does the **table of contents** list the **titles of the appendices**?		
82.	Does the CSR include the **study Protocol**?		
83.	If Y, is the **study Protocol** accessible to us?		
84.	If Y, how long is the **study Protocol** (in pages)?		
85.	Does the CSR contain a section on **Protocol amendments**?		
86.	If Y, is the section on **Protocol amendments** accessible to us?		
87.	If Y, how long is the section on **Protocol amendments** (in pages)?		
88.	Does the CSR contain a section on **Sample case report form (unique pages only)**?		
89.	If Y, is the section on **Sample case report form (unique pages only)** accessible to us?		
90.	If Y, how long is the section on **Sample case report form (unique pages only)** (in pages)?		
91.	Does the CSR contain a section on **List of IECs or IRBs (plus the name of the committee Chair if required by the**		

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

	regulatory authority) - Representative written information for patient and sample consent forms?		
92.	If Y, is the section on **List of IECs or IRBs (plus the name of the committee Chair if required by the regulatory authority) - Representative written information for patient and sample consent forms** accessible to us?		
93.	If Y, how long is the section on **List of IECs or IRBs (plus the name of the committee Chair if required by the regulatory authority) - Representative written information for patient and sample consent forms** (in pages)?		
94.	Does the CSR contain a section on **List and description of investigators and other important participants in the study, including brief (1 page) CVs or equivalent summaries of training and experience relevant to the performance of the clinical study**?		
95.	If Y, is the section on **List and description of investigators and other important participants in the study, including brief (1 page) CVs or equivalent summaries of training and experience relevant to the performance of the clinical study** accessible to us?		
96.	If Y, how long is the section on **List and description of investigators and other important participants in the study, including brief (1 page) CVs or equivalent summaries of training and experience relevant to the performance of the clinical study** (in pages)?		
97.	Does the CSR contain a section on **Signatures of principal or coordinating investigator(s) or sponsor's responsible medical officer, depending on the regulatory authority's requirement**?		
98.	If Y, is the section on **Signatures of principal or coordinating investigator(s) or sponsor's responsible medical officer, depending on the regulatory authority's requirement** accessible to us?		
99.	If Y, how long is the section on **Signatures of principal or coordinating investigator(s) or sponsor's responsible medical officer, depending on the regulatory authority's requirement** (in pages)?		
100.	Does the CSR contain a section on **Listing of patients receiving test drug(s)/investigational product(s) from specific batches, where more than one batch was used**?		
101.	If Y, is the section on **Listing of patients receiving test drug(s)/investigational product(s) from specific batches, where more than one batch was used** accessible to us?		
102.	If Y, how long is the section on **Listing of patients receiving test drug(s)/investigational product(s) from specific batches, where more than one batch was used** (in pages)?		
103.	Does the CSR contain a section on **Randomisation scheme and codes (patient identification and treatment assigned)**?		
104.	If Y, is the section on **Randomisation scheme and codes**		

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

	(patient identification and treatment assigned) accessible to us?		
105.	If Y, how long is the section on **Randomisation scheme and codes (patient identification and treatment assigned)** (in pages)?		
106.	Does the CSR contain a section on **Audit certificates (if available) (see Annex IVa and IVb of the guideline)**?		
107.	If Y, is the section on **Audit certificates (if available) (see Annex IVa and IVb of the guideline)** accessible to us?		
108.	If Y, how long is the section on **Audit certificates (if available) (see Annex IVa and IVb of the guideline)** (in pages)?		
109.	Does the CSR contain a section on **Documentation of statistical methods**?		
110.	If Y, is the section on **Documentation of statistical methods** accessible to us?		
111.	If Y, how long is the section on **Documentation of statistical methods** (in pages)?		
112.	If Y, is the **Documentation of statistical methods** dated?		
113.	If Y, what is the date of the **Documentation of statistical methods**?		
114.	Does the CSR contain a section on **Documentation of inter-laboratory standardisation methods and quality assurance procedures if used**?		
115.	If Y, is the section on **Documentation of inter-laboratory standardisation methods and quality assurance procedures if used** accessible to us?		
116.	If Y, how long is the section on **Documentation of inter-laboratory standardisation methods and quality assurance procedures if used** (in pages)?		
117.	Does the CSR contain a section on **Publications based on the study**?		
118.	If Y, is the section on **Publications based on the study** accessible to us?		
119.	If Y, how long is the section on **Publications based on the study** (in pages)?		
120.	Does the CSR contain a section on **Important publications referenced in the report**?		
121.	If Y, is the section on **Important publications referenced in the report** accessible to us?		
122.	If Y, how long is the section on **Important publications referenced in the report** (in pages)? Edfgyh+		
123.	Does the CSR contain a section on **Discontinued patients**?		
124.	If Y, is the section on **Discontinued patients** accessible to us?		
125.	If Y, how long is the section on **Discontinued patients** (in pages)?		
126.	Does the CSR contain a section on **Protocol deviations**?		
127.	If Y, is the section on **Protocol deviations** accessible to us?		
128.	If Y, how long is the section on **Protocol deviations** (in		

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

	pages)?		
129.	Does the CSR contain a section on **Patients excluded from the efficacy analysis**?		
130.	If Y, is the section on **Patients excluded from the efficacy analysis** accessible to us?		
131.	If Y, how long is the section on **Patients excluded from the efficacy analysis** (in pages)?		
132.	Does the CSR contain a section on **Demographic data**?		
133.	If Y, is the section on **Demographic data** accessible to us?		
134.	If Y, how long is the section on **Demographic data** (in pages)?		
135.	Does the CSR contain a section on **Compliance and/or drug concentration data (if available)**?		
136.	If Y, is the section on **Compliance and/or drug concentration data (if available)** accessible to us?		
137.	If Y, how long is the section on **Compliance and/or drug concentration data (if available)** (in pages)?		
138.	Does the CSR contain a section on **Individual efficacy response data**?		
139.	If Y, is the section on **Individual efficacy response data** accessible to us?		
140.	If Y, how long is the section on **Individual efficacy response data** (in pages)?		
141.	Does the CSR contain a section on Adverse event listings (each patient)?		
142.	If Y, is the section on **Adverse event listings (each patient)** accessible to us?		
143.	If Y, how long is the section on **Adverse event listings (each patient)** (in pages)?		
144.	Does the CSR contain a section on **Listing of individual laboratory measurements by patient, when required by regulatory authorities**?		
145.	If Y, is the section on **Listing of individual laboratory measurements by patient, when required by regulatory authorities** accessible to us?		
146.	If Y, how long is the section on **Listing of individual laboratory measurements by patient, when required by regulatory authorities** (in pages)?		
147.	Does the CSR contain a section on **Case Report Forms for deaths, other serious adverse events and withdrawals for AE**?		
148.	If Y, is the section on **Case Report Forms for deaths, other serious adverse events and withdrawals for AE** accessible to us?		
149.	If Y, how long is the section on **Case Report Forms for deaths, other serious adverse events and withdrawals for AE** (in pages)?		
150.	Does the CSR contain a section on **Other Case Report Forms** submitted?		
151.	If Y, is the section on **Other Case Report Forms submitted** accessible to us?		
152.	If Y, how long is the section on **Other Case Report Forms**		

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

	submitted (in pages)?		
153.	Does the CSR contain a section on **Individual patient data listings**?		
154.	If Y, is the section on **Individual patient data listings** accessible to us?		
155.	If Y, how long is the section on **Individual patient data listings** (in pages)?		

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml