Risk of bias in industry-funded oseltamivir trials
Comparison of core reports versus full clinical study reports
Jefferson, Tom; Jones, Mark A.; Doshi, Peter; Del Mar, Chris B.; Hama, Rokuro; Thompson, Matthew J.; Onakpoya, Igho; Heneghan, Carl J.
Published in:
BMJ Open

DOI:
10.1136/bmjopen-2014-005253

Published: 01/01/2014

Document Version:
Publisher's PDF, also known as Version of record

Licence:
CC BY-NC

Link to publication in Bond University research repository.

Recommended citation (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

For more information, or if you believe that this document breaches copyright, please contact the Bond University research repository coordinator.
Risk of bias in industry-funded oseltamivir trials: comparison of core reports versus full clinical study reports

Tom Jefferson, Mark A Jones, Peter Doshi, Chris B Del Mar, Rokuro Hama, Matthew J Thompson, Igho Onakpoya, Carl J Heneghan

INTRODUCTION

The risk of bias tool in Cochrane reviews of randomised trials is routinely used to assess essential items pertaining to validity of trial design such as random sequence generation, allocation concealment, attrition and performance biases. There are six standard bias elements, each rated at a 'high', 'low' or 'unclear' risk of bias.

As Cochrane reviews are typically based on synthesising studies based on reports published in the scientific literature, the risk of bias tool is traditionally applied to journal publications. To the best of our knowledge, the ways in which risk of bias judgements change when they are based on more detailed reports of trials, such as those contained in clinical study reports, have not been previously investigated.

Clinical study reports are considered the most exhaustive summaries of randomised controlled trials of pharmaceuticals. Clinical study reports are highly structured and detailed documents that follow an outline format agreed between regulators and manufacturers in 1995, described in the ICH E3
document. Recent transparency policies adopted by the European Medicines Agency (EMA), as well as announcements by some pharmaceutical companies to make clinical study reports more readily available, suggest that clinical study reports may increasingly be incorporated into systematic reviews and other forms of evidence synthesis.

Although there is some variation in the structure and content of clinical study reports, they are usually composed of a core report of the trial and appendices. A core report (sections 1–15 of the ICH E3 document) is structured in the Introduction, Methods Results and Discussion (IMRAD) style. The numerous appendices (section 16 of ICH E3) contain important online supplementary data needed to understand and interpret the trial, its context and history. These appendices include such documents as the trial protocol, protocol amendments, statistical analysis plan, blank case report forms, certificates of analysis, randomisation lists and consent forms. For the purposes of this paper, the core report plus all its appendices will be known as the full clinical study report (see online supplementary appendix 1 for the table of contents of a typical oseltamivir clinical study report and http://dx.doi.org/10.5061/dryad.77471 for a free download of all the clinical study reports used in our review and featured in this paper. The core report was known as Module 1 in oseltamivir clinical study reports, and appendices were found in Modules 2–5). Core reports and full clinical study reports theoretically can help reduce uncertainty in judging risk of bias.

In 2012, we published an update of our Cochrane review of neuraminidase inhibitors which included a total of 32 oseltamivir trials. Unlike most Cochrane reviews, this review was based only on core reports, and risk of bias assessments were therefore based on each core report. Subsequently, in 2013, we obtained full clinical study reports from Roche and, as part of a further systematic review update, carried out new risk of bias assessments of the same trials based on the full clinical study reports.

Our overall aim was to investigate whether the level of detail contained in reports of trials affects judgements about risk of bias. We planned to achieve this by comparing documents which contain increasingly detailed information on each trial included in our review, namely journal publications, core reports and full clinical study reports. As well as using the standard Cochrane risk of bias tool, we developed an additional list of study elements that we wanted to extract in order to allow improved assessments of each trial’s design and conduct and facilitate the organisation of large quantities of information now available to us.

In this report, we describe our use of these tools to address three specific questions:
1. Do core reports change the risk of bias evaluation compared to published papers?
2. Do full clinical study reports change the risk of bias evaluation compared to core reports?
3. Do full clinical study reports change the risk of bias evaluation compared to published papers?

METHODS
Ten core reports (M76001; NV16871; WV15670; WV15671; WV15707; WV15720; WV15759/WV15871; WV15799; WV15812/WV15872; WV15819/WV15876/WV15978) were received in PDF files from Roche and EMA by 12 April 2011 (the date of time lock for our 2012 Cochrane review). The reporting of more than one trial in the same clinical study report was justified by Roche as a consequence of lower than expected participant recruitment due to low influenza circulation and consequently a need to pool studies.

The current Cochrane risk of bias tool consists of six domains; each may have more than one source of bias application, depending on the subject matter. Our applications were as follows: selection bias (random sequence generation and allocation concealment), performance bias (blinding of participants and personnel—all outcomes), detection bias (blinding of outcome assessment—all outcomes), attrition bias (influenza symptoms, complications and harms outcome data), reporting bias (selective reporting) and other bias. The identification of sources of other bias was left at the reviewers’ discretion.

Risk of bias assessments were performed following Cochrane methods and published in 2012.6 In that review, risk of bias was assessed by an external reviewer on the basis of data extracted from core reports. After 12 April 2011, we obtained the appendices of the clinical study reports included in our review. For most of the clinical study reports we requested, EMA had the protocol, protocol amendments, statistical analysis plan, blank case report forms and other appendices contained in what Roche terms the second ‘module’ of a full clinical study report (see online supplementary appendix 1). However, EMA did not possess—and therefore could not provide us with—full clinical study reports with the exception of trial WP16263.8 For approximately 3 years, Roche had repeatedly refused our requests for full clinical study reports.

In April 2013 in the course of carrying out these new extractions, Roche changed its policy on access to data and pledged to share with us 77 full clinical study reports (http://www.bmj.com/tamiflu/roche). Fifteen clinical study reports containing 20 trials were included in the analysis of our current review. As we were already in possession of core reports and appendices such as the protocol and statistical analysis plan for the 14 trials in this analysis, the additional data for other clinical study reports provided by Roche do not concern this paper. In the clinical study reports, Roche redacted information that they judged to be of ‘legitimate commercial interest’ or present a risk of trial participant reidentification. The redactions did not impede our analyses of risk of bias.

On the basis of our growing familiarity with clinical study reports, we designed and piloted a data extraction sheet to record how our understanding of the trials changed in the light of the availability of the additional
appendices. We realised that, in addition to the standard Cochrane risk of bias elements, we needed to organise the abundant material at our disposal and reconstruct a timeline of the trials. We used the Cochrane risk of bias tool to appraise clinical study reports and a data extraction sheet for recording information relevant to this appraisal. We added the following elements to our extraction sheets: date of participant enrolment, unblinding of the trial, protocol for which we had the full text, protocol amendments, statistical analysis plan for which we have the full text (and its amendments), patient consent form, randomisation list and certificate of analysis. Timeline reconstruction allowed us to conceptualise the design and conduct of the trials and appreciate their role in the trial programme with their strengths and limitations. In addition, following a timeline allows a judgement to be made on the integrity and temporal sequence of the documents. The finalised extraction sheet is in online supplementary appendix 2.

On the basis of access to the full clinical study reports, we carried out our final assessment of risk of bias. These were carried out by a single reviewer, checked by a second with final consensus reached through a face-to-face discussion among the entire group.

Since with full clinical study reports there should be no ambiguity, we only allowed ‘low’ or ‘high’ risk of bias judgements (ie, no ‘unclear’). We adopted the position that, unlike a publication which may have page limits, there was no reason why a full clinical study report should be missing details necessary for a third party to judge risk of bias. Therefore, when information that would have otherwise allowed us to judge a risk of bias as either ‘low’ or ‘high’ was missing, this would automatically be categorised as ‘high’ risk of bias. This decision to eliminate the ‘unclear’ option when assessing full clinical study reports was made following an initial assessment of the trials, which included ‘unclear’ judgements. On the basis of an earlier peer review of this paper, which suggested we analyse the data had we kept the ‘unclear’ category, we also carried out this post hoc analysis.

To allow for a comparison of risk of bias judgements based on published reports of trials and risk of bias judgements based on clinical study reports (either core reports alone or full clinical study reports), we used our previous risk of bias judgements for the same trials in the relevant Cochrane reviews that had been based on publications. Since with full clinical study reports there should be no ambiguity, we only allowed ‘low’ or ‘high’ risk of bias judgements (ie, no ‘unclear’). We adopted the position that, unlike a publication which may have page limits, there was no reason why a full clinical study report should be missing details necessary for a third party to judge risk of bias. Therefore, when information that would have otherwise allowed us to judge a risk of bias as either ‘low’ or ‘high’ was missing, this would automatically be categorised as ‘high’ risk of bias. This decision to eliminate the ‘unclear’ option when assessing full clinical study reports was made following an initial assessment of the trials, which included ‘unclear’ judgements. On the basis of an earlier peer review of this paper, which suggested we analyse the data had we kept the ‘unclear’ category, we also carried out this post hoc analysis.

To allow for a comparison of risk of bias judgements based on published reports of trials and risk of bias judgements based on clinical study reports (either core reports alone or full clinical study reports), we used our previous risk of bias judgements for the same trials in the relevant Cochrane reviews that had been based on publications. Since with full clinical study reports there should be no ambiguity, we only allowed ‘low’ or ‘high’ risk of bias judgements (ie, no ‘unclear’). We adopted the position that, unlike a publication which may have page limits, there was no reason why a full clinical study report should be missing details necessary for a third party to judge risk of bias. Therefore, when information that would have otherwise allowed us to judge a risk of bias as either ‘low’ or ‘high’ was missing, this would automatically be categorised as ‘high’ risk of bias. This decision to eliminate the ‘unclear’ option when assessing full clinical study reports was made following an initial assessment of the trials, which included ‘unclear’ judgements. On the basis of an earlier peer review of this paper, which suggested we analyse the data had we kept the ‘unclear’ category, we also carried out this post hoc analysis.

To allow for a comparison of risk of bias judgements based on published reports of trials and risk of bias judgements based on clinical study reports (either core reports alone or full clinical study reports), we used our previous risk of bias judgements for the same trials in the relevant Cochrane reviews that had been based on publications.
See tables 3 and 4. Twenty-nine per cent of previously certain judgements (ie, ‘high’ or ‘low’ risk of bias) based on core reports became ‘unclear’ with full clinical study reports.

An example of the kind of detail available in full clinical study reports, and the importance of the trial timeline in assessing the presence of bias, is the observation that of the clinical study reports for the 14 trials, only 1 contained a protocol which predated the beginning of participant enrolment, only 2 had statistical analysis plans which clearly predated participants enrolment and 3 had clearly dated protocol amendments. No clinical study report reported a clear date of unblinding. Completed extraction sheets with risk of bias comparisons and rationales are available on request from the corresponding author.

DISCUSSION

We used the Cochrane six-item risk of bias instrument to assess bias from two different levels of detail of trial reports. Owing to the unrestricted access to full clinical study reports, we took the view that all information needed to judge risk of bias for each of the six domains of the Cochrane risk of bias should be present. When the information was not available, we judged the corresponding risk of bias element as being ‘high’. Therefore, the availability of full clinical study reports decreased the uncertainty and allowed clearer judgements to be made. Risk of bias previously assessed as ‘unclear’ based on core reports became a more certain ‘low’ or ‘high’ risk of bias. When the information was not available, our judgements changed because we found gaps in the availability of information and inconsistent information. Whether the full study reports represent an exhaustive and coherent source of trial narrative and data remain unclear.

Throughout our study, we were assessing two different types of material within the clinical study reports: those that were created or written prior to patient enrolment (eg, trial protocols), and those written after (eg, core reports).

This approach is not possible when assessing trials reported in journal publications, in which articles necessarily reflect post hoc reporting with a far more sparse level of detail. We suggest that when bias is so limiting as
to make meta-analysis results unreliable, either it should not be carried out or a prominent explanation of its clear limitations should be included alongside the meta-analysis. We found the Cochrane risk of bias tool to be difficult to apply to clinical study reports. We think this is not because the tool was constructed to assess journal publications but, as with all list-like instruments, its use lends itself to a checklist approach (in which each design item is sought and, if found, eliminated from the bias equation rather than with thought and consideration). Similarly, the extraction sheet we assembled needs to be applied with thought and consideration—an approach that does not lend itself to reviewing under time pressure. However, more focus should be devoted to bias itself and its effects rather than theoretical risk of bias. Many of the variables we found to be important when assessing the trial (eg, date of trial protocol, date of unblinding, date of participant enrolment) are simply not captured in the risk of bias tool when used in a routine way or to review publications. We were also often unsure how to judge the risk of bias when bias itself can actually or potentially be measured with reviewers’ access to full clinical study reports and individual participant data. If, for example, the original trial protocol is available, one can judge whether reporting bias occurred. Reviewers need not guess at bias (ie, make a judgement of ‘risk’) but can judge bias directly. However, even with individual participant data, some forms of bias, such as attrition bias, may still be difficult to quantify, and one can only judge the risk (ie, potential) of bias. Therefore, access to detailed information and participant level data sometimes found in full clinical study reports provides an opportunity to consider both actual as well as risk of biases.

**Box 1** shows examples of the types of information found in clinical study reports that led to risk of bias assessment changes. While the judgements of ‘low’ or ‘high’ risk of bias may imply certainty, particularly when based on the reading of a full clinical study report, we found ourselves often in lengthy debate and discussion over the proper level of risk of bias before arriving at a consensus. We found the risk of bias judgements themselves to carry a high level of subjectivity, in which different judgements can be justified in different ways. The real strength of the risk of bias tool appears not to be in the final judgements it enables, but rather in the process it helps facilitate: critical assessment of a clinical trial.

<table>
<thead>
<tr>
<th>Trial (s)</th>
<th>Risk of bias, full clinical study reports</th>
<th>Risk of bias assessment performed based on</th>
</tr>
</thead>
<tbody>
<tr>
<td>M76001</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>NV16871</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>WV15670</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>WV15671</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>WV15707</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>WV15730</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>WV15759</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>WV15799</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>WV15812</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>WV15872</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>WV15819</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>WV15978</td>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>

**Table 1** Risk of bias assessments performed by trial, 2009–2014

<table>
<thead>
<tr>
<th>Trial (s)</th>
<th>Risk of bias, full clinical study reports</th>
<th>Risk of bias assessment performed based on</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>26 (20)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Unclear</td>
<td>28 (22)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Low</td>
<td>34 (26)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Total</td>
<td><strong>88 (68)</strong></td>
<td><strong>0 (0)</strong></td>
</tr>
</tbody>
</table>

**Table 2** Change in overall (all elements) risk of bias judgments for 15 core reports of oseltamivir trials compared with full clinical study reports

<table>
<thead>
<tr>
<th>Risk of bias, core reports</th>
<th>Risk of bias, full clinical study reports</th>
<th>Risk of bias assessment performed based on</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>11 (8)</td>
<td>15 (12)</td>
</tr>
<tr>
<td>Unclear</td>
<td>1 (1)</td>
<td>27 (21)</td>
</tr>
<tr>
<td>Low</td>
<td>12 (9)</td>
<td>22 (17)</td>
</tr>
<tr>
<td>Total</td>
<td><strong>24 (18)</strong></td>
<td><strong>64 (49)</strong></td>
</tr>
</tbody>
</table>

**Table 3** Change in overall (all elements) risk of bias judgments for 15 core reports of oseltamivir trials compared with full clinical study reports including unclear assessments
Another aspect to emerge is that tools based on publications are designed to detect the presence, absence or uncertainty regarding elements in a very restricted number of places in the text. The availability of full clinical study reports allows reviewers to follow consistency across chapters and appendices, creating a need for far more interaction with the text. An example of this active engagement is the cross-checking of active principle and placebo batches used across trials and their connection with a visual description of their properties such as colour in a certificate of analysis. For example, once the presence of a differently coloured placebo capsule cap in trial WP16263 was identified through the clinical study report’s certificate of analysis, its potential impact on blinding was captured in the Cochrane instrument. The interpretation of such a finding is difficult, as the colours of the active principle and placebo capsule caps are close (ivory and light yellow). However, publication-based or core report only based assessments would not have identified the potential differences in colour as the descriptions are simply given as ‘placebo’ and ‘matching placebo’, respectively. Reviewing the complete clinical study reports and our assessment of bias was very time consuming, necessitating prolonged exchanges including a face-to face meeting given the novelty of what we were doing. However, this activity was not as difficult or as time consuming as the reconstruction of trial evidence programmes for oseltamivir, an activity which necessitated a whole time equivalent researcher for 6 months. However, owing to the threat of reporting bias, we can think of no alternative to the use of full clinical study reports.

The main limitation of our study is our relative inexperience in dealing with large quantities of information and our lack of familiarity with certain trial documents such as randomisation lists. Randomisation lists appeared to be of two types. The first was a prerandomisation list of random codes with which participants’ IDs cannot be matched with the participant IDs used within other sections of the clinical study report. The second was a post hoc randomisation list to which individual participants can be matched, but the original generated codes are not shown. In both cases, the truly randomisation list of random codes with which participants’ IDs cannot be matched with the participant IDs used within other sections of the clinical study report. The second was a post hoc randomisation list to which individual participants can be matched, but the original generated codes are not shown. In both cases, the truly random generation of the sequence could not be properly assessed because either the original codes are not provided or they cannot be matched to patients. Another limitation of our study is that the instrument we have developed is for using with clinical study reports, and may not apply to non-industry trials (which may not have a clinical study report).

The background to our use of clinical study reports was our mistrust of journal publications of oseltamivir trials. Many trials were unpublished, and of those published, we found and documented examples of

| Table 4 | Change in overall (all elements) risk of bias judgments for 15 full clinical study reports reports of oseltamivir trials with and without allowing unclear assessments |
|---|---|---|---|---|
| **Risk of bias, full clinical study reports** | **Risk of bias, full clinical study reports allowing unclear assessments** | | | |
| | **High, n (%)** | **Unclear, n (%)** | **Low, n (%)** | **Total, n (%)** |
| High | 24 (18) | 64 (49) | 0 (0) | 88 (68) |
| Unclear | 0 (0) | 0 (0) | 0 (0) | 0 (0) |
| Low | 0 (0) | 0 (0) | 42 (32) | 42 (32) |
| Total | 24 (18) | 64 (49) | 42 (32) | 130 (100) |

Box 1 Examples of risk of bias assessment changes and other concerns

- In trial WV15708, the risk of bias related to allocation concealment went from ‘Unclear’ based on core reports to ‘High’ risk of bias based on full clinical study reports because the full clinical study report did not report sufficient details about the method of allocation concealment.
- In trial WV15707, the risk of bias related to random sequence generation went from ‘Unclear’ based on core reports to ‘High’ risk of bias based on full clinical study reports because a full description of the randomisation procedure was not provided.
- Prophylaxis trials WV15673 and WV15697 are described as ‘identical’, but this could not be verified as we had only one protocol (and the protocol we did have was dated after the study’s completion). In addition, the placebo event rates for influenza infection were very different between the two trials and their pooling, combined with the redaction of centre numbers, preventing from them being individually added to a meta-analysis. Therefore, our assessment of the ‘Other’ risk of bias item changed from ‘unclear’ based on core reports to ‘high’ based on full clinical study reports.
- In the treatment trials WV15819, WV15876 and WV15978, it was difficult to reconcile the total number of hospitalisations despite access to the full clinical study reports. One patient in the placebo arm who was hospitalised according to serious adverse event narratives does not appear in the hospitalisations table, and for a separate placebo patient who is listed in the serious adverse event narratives, no hospitalisation is described in this narrative, but the same patient was hospitalised according to the hospitalisations table. It was therefore unclear how many hospitalisations occurred in the trial, to whom and why.
- In prophylaxis trials WV15673 and WV15697, bias was assessed as low for selective reporting because the intention-to-treat population was described and reported in a table. However, when the full clinical study report became available, we realised that the original protocol was missing.

reporting bias. At least one trial publication was drafted by an unnamed medical writer. As evidence of reporting bias in industry trial publication mounts, 8 16–21 we believe Cochrane reviews should increasingly rely on clinical study reports as the basic unit of analysis. Sponsors and researchers both have a responsibility to

clinical study reports as the basic unit of analysis. Our experience suggests that more detailed extraction sheets that prompt reviewers to consider additional aspects of study may be needed. Until a more appropriate guide is developed, we offer our custom extraction sheets to Cochrane reviewers and others interested in assessing risk of bias using clinical study reports and encourage further development.

REFERENCES


