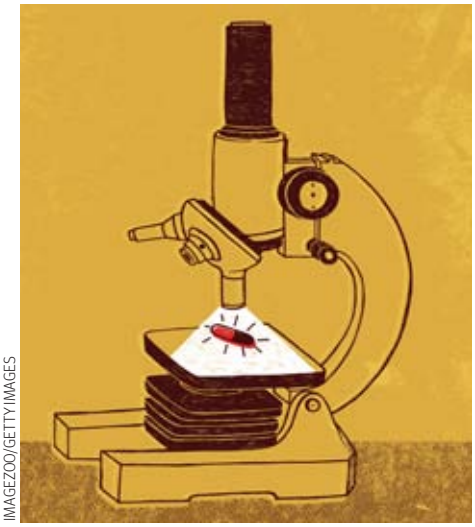
bmj.com archive Also by Tom Jefferson
- The Spanish influenza pandemic seen through the *BMJ*'s eyes: observations and unanswered questions (*BMJ* 2009;339:b5313)
- Neuraminidase inhibitors for preventing and treating influenza in healthy adults: systematic review and meta-analysis (*BMJ* 2009;339:b5106)
- Physical interventions to interrupt or reduce the spread of respiratory viruses: systematic review (*BMJ* 2009;339:b3675)



IMAGEZOO/GETTY IMAGES

Ensuring safe and effective drugs: who can do what it takes?

Drawing on their experience in producing a Cochrane review of neuraminidase inhibitors for influenza, **Tom Jefferson and colleagues** discuss how to improve the reliability of systematic reviews

In the midst of the H1N1 flu "pandemic," the Australian and UK governments commissioned an update of our longstanding Cochrane review on neuraminidase inhibitors for influenza in (otherwise) healthy adults. The review had first been published in 1999 with updates in 2006 and 2008. While preparing the 2009 update, we received a comment from a Japanese paediatrician. He questioned our conclusion that oseltamivir (Tamiflu) reduces the risk of complications (such as pneumonia) and pointed out that the evidence underlying this conclusion in our 2006 review was based on a single paper—a manufacturer funded meta-analysis[1] of 10 manufacturer trials, of which only two had been published in the peer reviewed literature.[2 3] To verify the quality and reliability of our previous conclusions, we wrote to the lead author of the meta-analysis, Laurent Kaiser, to obtain original data. We also wrote to Nicholson and Treanor, authors of the two published trials in the Kaiser meta-analysis. Months later, and despite additional correspondence with Roche, oseltamivir's manufacturer, we remained unable to access any data. In our updated 2009 review we did not include the unpublished trial data from the Kaiser paper, and concluded that the ability of oseltamivir to reduce complications was unknown.

Although our review was praised by some for highlighting important questions about the evidence base of a global public health drug,[4] we were left feeling that conclusions drawn from only a proportion of all existing trials (that is, just the published ones) are wholly inadequate. The extent to which unpublished data are included in evidence syntheses is low; a recent survey found that less than 10% of Cochrane reviews did.[5]

Since our review, and perhaps in response to the enormous publicity generated by the joint *BMJ-Channel 4 News* investigation of oseltamivir,[6] Roche publicly pledged to make its unpublished full clinical study reports available (box).[7] We expected that these reports would provide sufficient detail to verify the findings of the Kaiser meta-analysis. However, what Roche provided was not the full study reports of the 10 trials but module 1 of seven trial reports. The tables of contents showed that the full reports probably comprise four or five modules. Unfortunately, module 1 does not include the analysis plan, randomisation schedule details, the study protocol with a list of deviations, or detailed

THE SCOPE OF CLINICAL TRIAL DATA

- *Raw data* might be considered the most broad, detailed, and comprehensive type of data, comprising any records created in preparing for and carrying out clinical trials—trial methods with protocol, investigator notes, individual patient data, ethics committee reports, clinical case notes, management committee's minutes, transcripts/videos of meetings, contracts, book keeping, and so on. Access to raw data "enables data checking, thorough exploration, and re-analysis of the data in a consistent way"[8]
- *Clinical study reports* are a distillation and summary of the raw data from a given individual trial, but, importantly, are unabridged reports that (depending on study size) can be thousands of pages in length. They should report a trial's background and rationale, methods, results, and discussion and also include important study documents such as the analysis plan, randomisation schedule, study protocol (with a list of deviations and amendment history), detailed case histories for patients who have adverse events, example case report forms, and list of ethics committees who approved the research
- *Published trials* are those trials for which a primary publication appears in the scientific press, typically written in the structure of Introduction, Methods, Results, and Discussion (IMRAD). Published trials usually fit on a few sheets of paper and as such represent a summary of the much larger clinical study report
- *Unpublished trials* are those trials for which no primary publication appears in the scientific press. Information and details about unpublished trials may however appear as secondary publications in the so called grey literature, scientific congresses, conference posters, or other abstracted form
- *Abstracts* are IMRAD structured extreme syntheses of a trial or a group of trials, usually not longer than 200 words
- *Regulatory data* contextualise other types of data about a trial. When the correspondence between a regulator and drug manufacturer is accessible to third parties, independent researchers gain access to an additional level of data about a trial (or trials) relevant to systematic reviews but not directly produced from a trial. This and other regulatory data such as regulators' medical and statistical reviews can be essential to understanding individual trials within the context of a trial programme

To make sense of the evidence we must look not at single trials individually but at the whole trial programme

Reporting bias testing framework for comparing evidence from multiple regulators, manufacturer clinical study reports, trial registries, and published trials

Null hypothesis	Definition	Potential effect	Framework to test hypothesis
There is no under-reporting (overview hypothesis)	Under-reporting is an overall term including all types of bias when there is an association between results and what is presented to the target audience	Tailoring methods and results to the target audience may be misleading. The direction of the effect could change, or the statistical significance of the effect could change, or the magnitude of the effect could change from clinically worthwhile to not clinically worthwhile and vice versa	1. Is there evidence of under-reporting? 2. What types of under-reporting are apparent (list and describe them)? 3. What is the overall effect of under-reporting on the results of a meta-analysis (compare estimates of effects using (under) reported data and all data)? 4. What is the effect of under-reporting on the conclusions of a meta-analysis—are conclusions changed when all data are reported?
There is no difference between analysis plan in the protocol and final report (or the differences are listed and annotated)	When protocol violations, especially if not reported and justified, are not associated with study results	Post hoc analyses and changes of plan lead to manipulation of reporting and choice of what is and is not reported	1. List any discrepancies between what is pre-specified in protocol and what was actually done. 2. Can these discrepancies be explained by documented changes or amendments to the protocol? 3. Were these changes made before observing the data? 4. What is the perceived effect of these changes on the results and conclusions?
There is no difference between published and unpublished conclusions of the same study	A specific bias relating to the selective reporting of data in association with target audience	Results have been tailored to the intended recipient audience	1. Compare reporting of important outcomes (harms, complications) between published reports and other reports such as those to regulatory bodies. 2. Document any differences in conclusions based on separate reports of the same studies
Presentation of same dataset is not associated with differences in spelling, incomplete, discrepant, contradictory, or duplicate entries	Different versions of the same dataset are associated with discrepancies	Raises question whether these discrepancies are mistakes or deliberate?	1. Document any differences or similarities in separate reports of important outcomes (harms, complications) based on the same studies. 2. Report any discrepancies to the manufacturer and ask it to clarify and correct any errors. 3. What is the effect on evidence base of including or excluding material with similar discrepancies?
There is no evidence of publication bias	Publication status is not associated with size and direction of results	Negative or positive publication bias can have major effect on the interpretation of the data at all levels	1. Are there studies that have not been published (yes/no)? 2. How many studies have not been published (number and proportion of trials not published and proportion of patients not published)? 3. Construct a list of all known studies indicating which are published and which are not. 4. What is the impact on the evidence base of including or excluding unpublished material?
There is no evidence of outcome emphasis bias	When overemphasis or underemphasis of outcomes is not associated with size or direction of results	Can lead to wrong conclusions because overemphasis on certain outcomes	1. Are all of the prespecified outcomes in the study protocol reported? 2. Are the outcomes reported in the same way as specified in the study protocol? 3. Are there any documented changes to outcome reporting listed in the study protocol? 4. What is the impact on the evidence base of including or excluding emphasised outcomes?
There is no evidence of relative *v* absolute measure bias	When choice of effect estimates is not associated with size or direction of results	Can lead to wrong conclusions because of apparent underestimation or overestimation of effects (eg, in the use of relative instead of absolute measures of risk)	1. Are both relative and absolute measures of effect size used to report the results? 2. Is the incidence of each event reported for each treatment group? 3. What is the effect on the evidence base of including estimates of effect expressed either in relative and absolute measures?
There is no evidence of follow-up bias	When there is no evidence that length of follow-up is related to size and direction of results	Can lead to wrong conclusions owing to overemphasis or underemphasis of results	1. Are reported results based on the complete follow-up of each patient? 2. Are important events (harms, complications) unreported because they occurred in the off-treatment period? 3. What is the effect on the evidence base of including or excluding material with complete follow-up?
There is no evidence of data source bias	There is no difference between the evidence base presented to regulators (for approval for an indication) and that produced by or in possession of the manufacturer	Can lead to approved indications inconsistent with full dataset	1. Have regulators been presented with all data from trials sponsored by the drug's manufacturer? 2. Have all national regulatory agencies been presented with the same trial data? 3. Can differences between national regulatory agencies be explained by access to different data?

case histories for patients experiencing adverse events.

This additional material, although incomplete, raises an additional reason to doubt the integrity of the published evidence. For example, the first of the two published studies in the Kaiser meta-analysis does not mention serious adverse events,[9] and the second states that "there were no drug-related serious adverse events."[10] However, the partial study reports that Roche made available to us list 10 serious adverse events (in nine subjects) in the two trials, three of which were classified as possibly related to oseltamivir.

If published studies are incomplete and do not report important outcomes, the current process for conducting systematic reviews is not sufficiently rigorous, and in some cases it risks turning into unsolicited authoritative advertising for the drug industry. In other words, we need access to all unpublished data, even of trials published in the peer reviewed literature.

Change of methods

To make sense of the evidence we must look not at single trials individually but at the whole trial programme, a point recently made forcefully by Ioannidis and Karassa.[11] It is naive to think that single trials and narrow questions relevant to, say, 10 trials out of a total of 28 in the dataset would provide a manageably sized and reliable systematic review. A broad study question such as ours ("What are the effects of neuraminidase inhibitors?") requires evidence from the whole trial programme to answer it, and we have published a formal protocol detailing our new methodology.[12]

The first steps are fairly clear, if somewhat laborious: compiling a full list of trials carried out by the drug industry and non-industry

bmj.com/podcasts
Tom Jefferson discusses his call for access to all clinical data in a *BMJ* podcast at bmj.com/podcasts *BMJ's* eyes: observations and unanswered questions (*BMJ* 2009;339:b5313)

funders. Independent investigators must do this themselves before contacting drug companies to verify the completeness and accuracy of lists. For example, we identified a large oseltamivir trial by Roche Shanghai[13] (ML16369) that Roche Basel did not appear to know about—it was omitted from Roche's list of 101 sponsored and supported trials—despite the existence of an English language study report dated July 2001 (Mengzhao Wang, personal correspondence, 2 Oct 2009).

Next we must request full clinical study reports for each trial. However, even if the manufacturers agree, there is no guarantee that the reports are reliable. And what should we do when incomplete reports are provided, as we found with Roche? Regulatory material can help us answer the questions. The US Food and Drug Administration and Japan's regulatory body make available medical officers' reports written as part of drug evaluation. Although regulatory documents do not provide primary trial data, they can offer invaluable insights into trials. Regulatory materials, including correspondence and scientific reports, can potentially answer why, for example, US regulators today still require that oseltamivir's label clearly states that it has not been shown to reduce complications whereas the opposite is stated on European and Australian labels. If, as it appears, the trials did not use standardised definitions of secondary complications (such as pneumonia), this may preclude meta-analysis.[1]

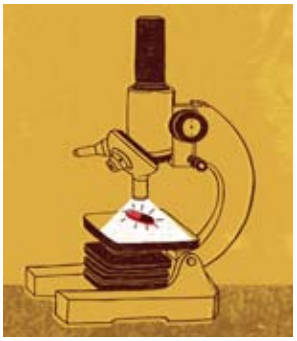
We then believe reviewers should construct a table of all the evidence to organise efforts and clarify what kind and amount of information exists for each of the trials included in a review. By doing this we found that the largest phase III treatment trial of oseltamivir (M76001) was not only never published but is little mentioned in regulatory documents. Why so little discussion would have occurred about such an important trial remains unclear.

To assess the integrity of a full trial programme (as opposed to that of individual trials, for which Cochrane methods probably suffice) we need new tools. Using insights from the growing literature on reporting bias,[14] and guided by our experience with neuraminidase inhibitors, we constructed a table of null hypotheses to test for the presence of biases potentially affecting trial programmes and their conclusions (table). The series of questions based on documented cases allow a critical overview of the trial programme.

Finally, a decision must be reached about whether a traditional review (including quantitative meta-analysis) can be done or whether the full trial programme and its dataset is of insufficient integrity to allow quantitative synthesis. We believe that you can pool data only if the evidence base is reasonably complete and sound. If a quantitative analysis is not possible, independent investigators must report the reasons why they are unable to satisfactorily assess the effects of a drug. Saying why you cannot do what you set out to do may prove as valuable as providing numerical results.

Publication and role of medical journals

By looking at some of the earlier versions of the industry sponsored reviews of neuraminidase inhibitors we discovered that the report by Kaiser et al was not the first meta-analysis mixing published and unpublished material; GlaxoSmithKline had done the same for zanamivir in a 1999 review in which only four of the seven trials included were published.[15] Clearly the inclusion criteria for our 2006 Cochrane review[16] were inconsistent as we included the Kaiser[1] meta-analysis but not the zanamivir meta-analysis. In retrospect, Kaiser's paper should never have been included.



Both our review and the Kaiser meta-analysis were published in two of the world's most prestigious peer reviewed medical journals.[1 17] The fact that they included data from unpublished randomised trials shows the extent to which trust underlies current practice, but is this any longer acceptable? We cannot expect busy doctors to be aware that trials and meta-analyses of drugs in respected publications are heavily influenced by drug companies' marketing decisions on what is and isn't published. We believe we need to change the way information is identified, appraised, and synthesised and regard any industry sponsored trial published in journals as marketing, unless proved otherwise.

Who can ensure safe and effective drugs?

It is important to not lose sight of the goal: providing doctors, patients, and policy makers with unbiased systematic assessments of drugs. Governments currently have this responsibility, but regulators are under-resourced[18] and powerful disincentives for rigorous review exist because candid analyses may undermine current policy. In the case of oseltamivir, regulators also disagree about whether it reduces complications, and agencies or departments within a single country may make inconsistent claims.[3]

The case of the diabetes drug rosiglitazone also shows that regulators are fallible. It took a legal settlement in 2004 requiring GlaxoSmithKline to post results of all of its clinical trials[18] before Nissen et al could do a meta-analysis of all 42 clinical trials of rosiglitazone, 35 of which were unpublished.[19] Their analyses showed a substantial increase in risk of myocardial infarction and cardiovascular death, results that were later replicated by the FDA using patient level data.[19 20] Thus independent systematic analyses remain vital.

How can robust independent assessment of drugs be carried out in a world where data are privately owned? The answer is to make the data freely available: we should accept nothing less than a full dataset. Before licensing a drug—and certainly before large purchase decisions are made—our governments and policy makers should ensure that all researchers can access data in sufficient detail to allow for the independent exploration and re-analysis of trials.

Researchers, the public, and lay and scientific media will need to work together to put pressure on industry to embrace the ethical responsibility to release data in the public interest. Currently legal action is often required to achieve this (as happened with rosiglitazone). Ideally, regulators would make full clinical study reports publicly available once a regulatory decision is reached. This is a daunting task considering that the FDA submission for oseltamivir was at least 363 volumes (although we do not know the number of pages).

For the present, we call on journals to require, as a condition for consideration of publication of a randomised trial, submission of the most detailed report available (anonymised to protect patient privacy), in addition to the summary manuscript for publication. This concept is not dissimilar to the format of a short print version with full text online that many journals already use. Access to both parts would allow peer reviewers to carry out random checks and to compare detailed reports with submitted manuscripts. If patient privacy can be assured, posting the detailed report as an online supplement will also improve post-publication review. Had this happened 10 years ago, the omission of

How can robust independent assessment of drugs be carried out in a world where data are privately owned?

serious adverse events possibly related to oseltamivir may have been noticed and corrected before publication. An overview of the trial programme should also be submitted explaining the rationale and findings of each trial. We need to understand how and why a particular trial was designed and conducted—that is, how every trial can potentially advance our knowledge.

Anyone conducting experiments on humans has obligations transcending patents and commercial confidentiality. Physician involvement in any type of reporting bias may be unethical, which carries implications for professional misconduct. We must remember that trial participants are performing a service to humanity, entering a potentially risky situation for the sake of determining the toxicity and effectiveness of a new drug. Withholding the results of such trials, and in some cases archiving them in such a way that they come to light only after prolonged and detailed investigation, seems ethically dubious. Ethical committees should be among the most vocal calling for the freeing of data.

We do not yet know whether sound scrutiny is feasible with a journal's resources, but we intend to use the proposed review of neuraminidase inhibitors to attempt to quantify these tasks. We will be keeping a journal of our review with a resource tally, and proposing methods for in-depth scrutiny of a whole trial programme of a new drug. It is time the media, the Cochrane Collaboration, and any reader interested in knowing what they are prescribing or are being prescribed increase the pressure on policy makers. If you swallow a medication, you need to know how it works—for real.

Tom Jefferson researcher, Acute Respiratory Infections Group, Cochrane Collaboration, Rome, Italy
Peter Doshi doctoral student, Program in History, Anthropology, Science, Technology and Society, Massachusetts Institute of Technology, Cambridge, MA, USA
Matthew Thompson senior clinical scientist, Department of Primary Health Care, University of Oxford, Oxford, UK and Department of Family Medicine, Oregon Health and Sciences University, Portland, Oregon, USA
Carl Heneghan clinical reader and director CEBM, Department of Primary Health Care, University of Oxford, Oxford, UK
Cochrane Acute Respiratory Infections Group
Correspondence to: P Doshi **pnd@mit.edu**

Accepted: 7 November 2010
We thank Chris Del Mar, Mark Jones, and Rokuro Hama. Thanks also to the Australian NHMRC and the UK NHS Research and Development fund for grants to enable the 2009 healthy adults review update, and to the UK NIHR for granting funding for our current review update. Also thanks to the Hugh Hampton Young fellowship committee at MIT for its generosity. The NIHR school of Primary Care Research provides financial support for Dr Carl Heneghan and funding for an investigators' meeting in Oxford (UK).
Contributors and sources: All authors were authors of the separate relevant Cochrane reviews. All authors contributed to the writing of this essay.
Competing interests: 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 no support from any organisation for the submitted work. In 1998-9, TJ was an ad hoc consultant for Hoffman-La Roche; all authors are co-applicants for a NIHR grant to carry out the Cochrane review.

1 Kaiser L, Wat C, Mills T, Mahoney P, Ward P, Hayden F. Impact of oseltamivir treatment on influenza-related lower respiratory tract complications and hospitalizations. *Arch Intern Med* 2003;163:1667-72.
2 Jefferson T, Jones M, Doshi P, Del Mar C. Neuraminidase inhibitors for preventing and treating influenza in healthy adults: systematic review and meta-analysis. *BMJ* 2009;339:b5106.
3 Doshi P. Neuraminidase inhibitors—the story behind the Cochrane review. *BMJ* 2009;339:b5164.
4 Godlee F. We want raw data, now. *BMJ* 2009;339:b5405.
5 Van Driel ML, De Sutter A, De Maeseneer J, Christiaens T. Searching for unpublished trials in Cochrane reviews may not be worth the effort. *J Clin Epidemiol* 2009;62:838-44.e3.
6 Cohen D. Complications: tracking down the data on oseltamivir. *BMJ* 2009;339:b5387.
7 Smith J, on behalf of Roche. Point-by-point response from Roche to BMJ questions. *BMJ* 2009;339:b5374.
8 Higgins JPT, Green S. Cochrane handbook for systematic reviews of interventions. Wiley, 2008. www.cochrane-handbook.org.
9 Nicholson KG, Aoki FY, Osterhaus AD, Trottier S, Carewicz O, Mercier CH, et al. Efficacy and safety of oseltamivir in treatment of acute influenza: a randomised controlled trial. Neuraminidase Inhibitor Flu Treatment Investigator Group. *Lancet* 2000;355:1845-50.
10 Treanor JJ, Hayden FG, Vrooman PS, Barbarash R, Bettis R, Riff D, et al. Efficacy and safety of the oral neuraminidase inhibitor oseltamivir in treating acute influenza: a randomized controlled trial. US Oral Neuraminidase Study Group. *JAMA* 2000;283:1016-24.
11 Ioannidis JPA, Karassa FB. The need to consider the wider agenda in systematic reviews and meta-analyses: breadth, timing, and depth of the evidence. *BMJ* 2010;341:c4875.
12 Cochrane Editorial Unit. Neuraminidase inhibitors for influenza—HTA. www.editorial-unit.cochrane.org/neuraminidase-inhibitors-influenza-hta-project.
13 Li L, Cai B, Wang M, Zhu Y. A double-blind, randomized, placebo-controlled multicenter study of oseltamivir phosphate for treatment of influenza infection in China. *Chin Med J* 2003;116:44-8.
14 McGauran N, Wieseler B, Kreis J, Schuler Y, Kolsch H, Kaiser T. Reporting bias in medical research—a narrative review. *Trials* 2010;11:37.
15 Monto AS, Webster A, Keene O. Randomized, placebo-controlled studies of inhaled zanamivir in the treatment of influenza A and B: pooled efficacy analysis. *J Antimicrob Chemother* 1999;44(suppl 2):23-9.
16 Jefferson TO, Demicheli V, Di Pietrantonj C, Jones M, Rivetti D. Neuraminidase inhibitors for preventing and treating influenza in healthy adults. *Cochrane Database Syst Rev* 2006;3:CD001265.
17 Jefferson T, Demicheli V, Rivetti D, Jones M, Di Pietrantonj C, Rivetti A. Antivirals for influenza in healthy adults: systematic review. *Lancet* 2006;367:303-13.
18 Cohen D. Rosiglitazone: what went wrong? *BMJ* 2010;341:c4848.
19 Nissen S. Rosiglitazone: a critical overview. 2010 Jul 9 www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/EndocrinologicandMetabolicDrugsAdvisoryCommittee/UCM218483.pdf.
20 Nissen SE, Wolski K. Effect of rosiglitazone on the risk of myocardial infarction and death from cardiovascular causes. *N Engl J Med* 2007;356:2457-71.

Cite this as: *BMJ* 2011;342:c7258

See, EDITORIAL, p 117, BLOG, p 135 and RESEARCH, p 155

ANSWERS TO ENDGAMES, p 177. For long answers go to the Education channel on bmj.com

STATISTICAL QUESTION

Meta-analyses I

Answers *a*, *b*, *c*, and *d* are all true.

ON EXAMINATION QUIZ

Acute pancreatitis

Answer C is correct.

PICTURE QUIZ

Nail changes in a patient with leukaemia

1 Greyish brown pigmentation can be seen on the proximal nail plate of all fingers. These nail changes are often described as melanonychia (nail hyperpigmentation as a result of melanin accumulation in the epithelium). Onycholysis of the left little finger nail is also seen.
2 Melanonychia is most often caused by drugs (especially chemotherapeutic agents), nail apparatus melanoma, naevus, ethnic nail hyperpigmentation, HIV infection, and subungual haemorrhage. The most likely cause of melanonychia in this case is the drug hydroxyurea.
3 In the absence of a specific drug or disease history, melanonychia should be evaluated with dermoscopic examination and nail biopsy, especially when lesions are solitary, pigmentation extends to the eponychium, and lesions enlarge rapidly.
4 Because the features are not suspicious of melanoma, a nail biopsy is not needed. The patient can continue to take hydroxyurea because the pigmentation caused by this drug is benign.