

----- Forwarded Message -----

From: "eloder@bmj.com" <eloder@bmj.com>
To: jon.jureidini@health.sa.gov.au
Cc: jo.lenoury@btopenworld.com; nardo.mickey@gmail.com;
david.healy54@googlemail.com; jon.jureidini@health.sa.gov.au;
melissa.raven@flinders.edu.au; catalin.tufanaru@adelaide.edu.au;
elia.abi.jaoude@utoronto.ca

Sent: Monday, 3 November 2014, 16:14

Subject: BMJ - Decision on Manuscript ID BMJ.2014.022376

03-Nov-2014

Dear Prof. Jureidini,

Re: Manuscript ID BMJ.2014.022376 entitled "A randomized, controlled trial of the efficacy and harms of paroxetine and imipramine in the treatment of adolescent major depression: Restoring Study 329"

Thank you very much for sending us this paper. We consider the RIAT initiative to be very important, and we recognise the controversy that has surrounded this study. We sent the paper for external peer review and have discussed it at a recent manuscript meeting with editors and a consulting statistician in attendance.

We recognise the value of this paper but we have not yet reached a final decision on it. We believe the paper needs extensive revision and clarifications in response to a number of matters identified by the peer reviewers and editors. We hope very much that you will be willing and able to revise your paper as explained below in the report from the manuscript committee meeting, so that we will be in a better position to make a decision about publishing it.

Many thanks again. We ordinarily ask to have revised articles back within a month but we are very willing to give you additional time if you need more than a month to work on the revision. This paper will set a precedent for other RIAT papers that follow, and we want to get things right.

**** THE REPORT FROM THE MANUSCRIPT COMMITTEE MEETING, REVIEWERS' REPORTS, AND THE BMJ'S GENERAL REQUIREMENTS FOR RESEARCH PAPERS ARE AVAILABLE AT THE END OF THIS LETTER.****

First, however, please read these four important points about sending your revised paper back to us:

1. **Deadline:** Your revised manuscript should be returned within one month -- but as I mentioned above, please do contact me if you require additional time.
2. **Online and print publication:** All original research in The BMJ is published with open access. The full text online version of your article, if accepted after revision, will be the

indexed citable version (full details are at <http://resources.bmj.com/bmj/about-bmj/the-bmjs-publishing-model>), while the print and iPad BMJ will carry an abridged version of your article, usually a few weeks afterwards. This abridged version of the article is essentially an evidence abstract called BMJ pico, which we would like you to write using a template and then email it to papersadmin@bmj.com (there are more details below on how to write this using a template). Publication of research on bmj.com is definitive and is not simply interim "epublication ahead of print", so if you do not wish to abridge your article using BMJ pico, you will be able to opt for online only publication. Please let us know if you would prefer this option.

If/when your article is accepted we may invite you to submit a video abstract, lasting no longer than 4 minutes, and based on the information in your paper's BMJ pico evidence abstract. The content and focus of the video must relate directly to the study that has been accepted for publication by The BMJ, and should not stray beyond the data.

3. Open access publication fee: The BMJ is committed to keeping research articles Open Access (with Creative Commons licences and deposit of the full text content in PubMedCentral as well as fully Open Access on bmj.com). To support this we are now asking all authors to pay an Open Access fee of £3000 on acceptance of their paper. If we accept your article we will ask you to pay the Open Access publication fee; we do have a waiver policy for authors who cannot pay. Consideration of your paper is not related to whether you can or cannot pay the fee (the editors will be unaware of this), and you need do nothing now.

4. How to submit your revised article: Log into <http://mc.manuscriptcentral.com/bmj> and enter your Author Center, where you will find your manuscript title listed under "Manuscripts with Decisions." Under "Actions," click on "Create a Revision." Your manuscript number has been appended to denote a revision.

You may also click the below link to start the revision process (or continue the process if you have already started your revision) for your manuscript. If you use the below link you will not be required to login to ScholarOne Manuscripts.

http://mc.manuscriptcentral.com/bmj?URL_MASK=f73d5ae9d29844dc976639d206d53db0

You will be unable to make your revisions on the originally submitted version of the manuscript. Instead, revise your manuscript using a word processing program and save it on your computer.

Once the revised manuscript is prepared, you can upload it and submit it through your Author Center. When submitting your revised manuscript, you will be able to respond to the comments made by the reviewer(s) and Committee in the space provided. You can use this space to document any changes you make to the original manuscript and to explain your responses. In order to expedite the processing of the revised manuscript, please be as specific as possible in your response to the reviewer(s).

Very truly yours,

Elizabeth Loder, MD, MPH
BMJ Editorial Team

As well as submitting your revised manuscript, we also require a copy of the manuscript with changes highlighted. Please upload this as a supplemental file with file designation 'Revised Manuscript Marked copy'.

IMPORTANT: Your original files are available to you when you upload your revised manuscript. Please delete any redundant files before completing the submission.

INFORMATION ON REVISING THE CONTENT AND FORMAT OF YOUR ARTICLE

****Report from The BMJ's manuscript committee meeting of 30 October 2014****

These comments are an attempt to summarise the discussions at the manuscript meeting. They are not an exact transcript.

Members of the committee were:

Elizabeth Loder (chair);
Angela Wade (statistician);
Jose Merino;
Wim Weber;
Tiago Villanueva;
Emma Parish

Decision: Put points. Revision to go back to the statistician and peer reviewers once returned.

* This is the first RIAT paper of a pharmaceutical trial and will set a precedent for similar papers that may follow. We want to be certain that we are completely satisfied with the presentation. We have several procedural questions about this reanalysis: 1) did you register the study in an approved trial registry? 2) how many versions of the protocol are there, and if there was more than one, how did you choose which one to follow?

* We agree with several of the reviewers that the problem of potential bias and conflict of interest needs more attention. We would like to hear your thoughts about these matters and we think some comment in the paper itself might be necessary.

* We are particularly troubled by the recoding of adverse events. We did not agree, for example, that it makes sense to move symptoms such as dizziness and headache out of the nervous system cluster. I am afraid this makes us worry about other decisions that were made in the process of recoding. We agree with reviewers that coding of adverse events needs to be redone by people who are independent of your group. We also agree with several of the reviewers that extrapolation of AEs from the non-random sample of CRFs is unwise. This analysis should be removed from the paper. (Table 6)

* Please present a true ITT analysis (in other words, analyze all subjects in the groups to which they were randomised, regardless of whether they received the study drug or not). Our statistician suggests that you consider having several columns in your results table. The first would present an ITT analysis using LOCF, the second using imputation and correcting for strata (12 centres). The third column could show the per protocol or complete case analysis using LOCF and the fourth the per protocol or complete case analysis using imputation. This would allow readers to judge for themselves the effects, if any, of using more modern methods of analysis, while still showing the originally intended efficacy analysis.

* We would also like to see the results of pairwise comparisons.

* Can you please also include a table that contrasts all of your findings with those of the original paper? You do this for AEs but not for the efficacy outcomes. Many editors commented that it was difficult to understand how and where the reanalyses differ from the original ones.

* Like the reviewers, we were disappointed that you did not examine the CRFs for all subjects. This seems a serious problem. It is, we understand, a major undertaking to review all of these documents, but seems necessary to set the record straight. After all, the trial itself was a major effort on the part of the original investigators.

* We believe the original investigators in the trial should be acknowledged in the paper. You mention that in some cases it was not clear what happened in the original study, for example why some secondary outcomes were changed. Did you make any attempt to ask the original investigators? If not, why not?

* You mention the original study was funded by the pharmaceutical company. Did you have any funding for this reanalysis?

* The abstract contains no numerical findings. Please present the figures for the principal study outcomes in the abstract.

* We thought that information about the alleged problems with the original study could be dealt with in a single paragraph in the introduction. Please be careful not to include ad hominem remarks. Has the previous paper been retracted? If not, how will readers of that paper know about this one?

* We thought you should comment on the matter of dropouts. These seemed higher in the placebo group. We also wondered whether 8 weeks is too soon to see any possible benefit of an antidepressant. Several editors who are practicing physicians and use these drugs thought that 8 weeks might be too soon to expect the drugs to diverge from placebo. Could you comment on this?

* The methods section should give more information about how subjects were recruited, number of centers involved in the study and how they were chosen. Who did the interviews? How were they trained? You say that children signed an informed consent form, but should this not be "assent?" Please explain how the decision was made to reduce the number of subjects from 300 to 275. In describing the intervention, please clarify the definition of "non responder." Although subjects could be titrated up to 60 mg paroxetine or 300 mg imipramine, how many actually did achieve these doses? Can you provide information about the mean final dose in each group and the range?

* How many subjects were screened for the study? Please show this in Figure 1. Figure 1 also needs to show the number analyzed for the complete case outcome at 8 weeks.

* In addition to responding to the comments above, please revise your paper to respond to all of the comments by the reviewers. Their reports are available at the end of this letter, below. We realise that there are a large number of comments from reviewers, but the reviews are of very high quality and raise many important matters.

IMPORTANT

When you revise and return your manuscript, please take note of all the following points about revising your article. Even if an item, such as a competing interests statement, was present and correct in the original draft of your paper, please check that it has not slipped out during revision.

- a. In your response to the reviewers and committee please provide, point by point, your replies to the comments made by the reviewers and the editors, and please explain how you have dealt with them in the paper. It may not be possible to respond in detail to all these points in the paper itself, so please do so in the box provided
- b. If your article is accepted it will then be edited, proofed, and - after your approval - published on bmj.com with open access. This open access Online First article will not be a pre-print. It will represent the full, citable, publication of that article. The citation will be year, volume, elocator (a unique identifier for that article): eg BMJ 2008;337:a145 — and this is what will appear immediately in Medline, PubMed, and other bibliographical indexes. We will give this citation in print and online, and you will need to use it when you cite your article.
- c. Please write an abridged version of the article for the print and iPad BMJ using the appropriate BMJ pico template for your study's design. Please be reassured that it doesn't take long to complete this. When your BMJ pico is ready please email it to papersadmin@bmjgroup.com. The templates for you to download are at <http://resources.bmj.com/bmj/authors/bmj-pico>
- d. Please include these items in the revised manuscript to comply with BMJ style:

Title: this should include the study design eg "systematic review and meta-analysis"

Abstract

structured abstract including key summary statistics, as explained below (also see <http://resources.bmj.com/bmj/authors/types-of-article/research>) for every clinical trial - and for any other registered study - the study registration number and name of register – in the last line of the structured abstract.

Introduction

this should cover no more than three paragraphs, focusing on the research question and your reasons for asking it now

Methods:

for an intervention study the manuscript should include enough information about the intervention(s) and comparator(s) (even if this was usual care) for reviewers and readers to understand fully what happened in the study. To enable readers to replicate your work or implement the interventions in their own practice please also provide (uploaded as one or more supplemental files, including video and audio files where appropriate) any relevant detailed descriptions and materials. Alternatively, please provide in the manuscript urls to openly accessible websites where these materials can be found

Results

please report statistical aspects of the study in line with the Statistical Analyses and Methods in the Published Literature (SAMPL) guidelines <http://www.equator-network.org/reporting-guidelines/sampl/>

summary statistics to clarify your message. Please include in the results section of your structured abstract (and, of course, in the article's results section) the following terms, as appropriate:

For a clinical trial:

- Absolute event rates among experimental and control groups
- RRR (relative risk reduction)
- NNT or NNH (number needed to treat or harm) and its 95% confidence interval (or, if the trial is of a public health intervention, number helped per 1000 or 100,000)

Discussion

please write the discussion section of your paper in a structured way, to minimise the risk of careful explanation giving way to polemic. Please follow this structure:
statement of principal findings of the study
strengths and weaknesses of the study
strengths and weaknesses in relation to other studies, discussing important differences in results and what your study adds. Whenever possible please discuss your study in the light of relevant systematic reviews and meta-analyses (eg Cochrane reviews)
meaning of the study: possible explanations and implications for clinicians and policymakers and other researchers; how your study could promote better decisions
unanswered questions and future research

Footnotes and statements

What this paper adds/what is already known box (as described at <http://resources.bmj.com/bmj/authors/types-of-article/research>)

ID of ethics committee approval and name of the ethics committee/IRB; or a statement that approval was not required (see <http://resources.bmj.com/bmj/authors/editorial-policies/guidelines>) and a statement that participants gave informed consent before taking part

a statement that any identifiable patients have provided their signed consent to publication. Please submit, as a supplemental file, the signed BMJ patient consent form giving consent to publication in The BMJ of any information about identifiable individual patients. Publication of any personal information about a patient in The BMJ, for example in a case report or clinical photograph, will normally require the signed consent of the patient.

competing interests statement (see <http://resources.bmj.com/bmj/authors/editorial-policies/competing-interests>)

contributorship statement+ guarantor (see <http://resources.bmj.com/bmj/authors/article-submission/authorship-contributorship>)

transparency statement: a statement that the lead author (the manuscript's guarantor) affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies are disclosed.

copyright statement/ licence for publication (see <http://www.bmj.com/about-bmj/resources-authors/forms-policies-and-checklists/copyright-open-access-and-permission-reuse>)

signed patient consent form(s), if the article gives enough personal information about any patient(s): this sometimes occurs even in research papers - for example in a table giving demographic and clinical information about a small subgroup in a trial or observational study,

or in quotes/tables in a qualitative study - (see http://resources.bmj.com/bmj/authors/editorial-policies/copy_of_patient-confidentiality)

a data sharing statement declaring what further information and data you are willing to make available, over and above the results reported in the paper. Suggested wording: "Data sharing: technical appendix, statistical code, and dataset [state whether any patient level data have been anonymised] are available at this repository or website OR from the corresponding author at ". If there are no such further data available, please use this wording: "Data sharing: no additional data available". For papers reporting the main results of trials of drugs or devices we require that the authors state, at a minimum, that the relevant anonymised patient level data are available on reasonable request from the authors.

The BMJ has partnered with the Dryad Digital Repository datadryad.org to make open deposition easy and to allow direct linkage by doi from the dataset to The BMJ article and back - we encourage authors to use this option

funding statement (see

<http://resources.bmj.com/bmj/authors/article-submission/article-requirements>)

statement of the independence of researchers from funders (see

<http://resources.bmj.com/bmj/authors/article-submission/article-requirements>)

for studies funded or sponsored by industry (see

<http://resources.bmj.com/bmj/authors/article-submission/article-requirements>)

a statement describing the role of the study sponsor(s), if any, in study design; in the collection, analysis, and interpretation of data; in the writing of the report; and in the decision to submit the article for publication

assurance, in the cover letter, that a clinical trial funded by a pharmaceutical or other commercial company follows the guidelines on good publication practice (see

<http://resources.bmj.com/bmj/authors/article-submission/article-requirements>)

inclusion in the list of contributors the name(s) any professional medical writer(s), specifying in the formal funding statement for the article who paid the writer. Writers and authors must have access to relevant data while writing articles.

Patient centred research

for studies that are relevant to patients we expect authors to report in their articles the extent of their study's patient-centredness, as highlighted by these questions:

did you involve patients/service users/carers/lay people in the design of this study? Please state whether you did, and give details (Methods section)

was the development and/or selection of outcome measures informed by patients' priorities and experiences? Please give details (Methods section)

were patients/service users/carers/lay people involved in developing plans for participant recruitment and study conduct? If so, please specify how (Methods section)

have you planned to disseminate the results of the study to participants? If so how will this be done? (Describe in brief footnote)

are patients thanked in the contributorship statement or acknowledgements?

for articles reporting randomised controlled trials: did you assess the burden of the intervention on patients' quality of life and health? If so, what evaluation method did you use, and what did you find? (Methods and Results sections)

REFEREES COMMENTS

Reviewer: 1

Recommendation:

Comments:

This is an important paper and this study is emblematic of the RIAT initiative since it is a widely discussed and polemic study. It is in line with a previous report by and gives a perfect and detailed example of selective outcome reporting. It deserves thus to be published in the BMJ. The paper respects CONSORT statement as stated in the RIAT list given in the appendix.

I have nevertheless some comments for authors in order to improve their paper.

First, I fell not ease with the presence of comments in the method section and in the results section which are generally not the place to discuss choices and results. Please see for example:

- in the introduction : “Consequently, we have reanalysed Study 329 according to the RIAT statement.. To this end, we have used the Clinical Study Report (CSR; GSK's 'Final Clinical Report') available on the GSK website,[7] other publically available documents,[8] and the data access system SAS Solutions OnDemand,[9] on which GSK has posted some Study 329 documents (available only to users approved by GSK). Following negotiation,[10] GSK posted de-identified individual case report forms (CRFs) on that site. A table of sources of data consulted in preparing each part of this paper is available as Appendix 1.” This should appear in the method section;
- in the method section, authors state “These imipramine doses are high for adolescents. In the six comparator studies submitted by SKB as part of their 1991 Approval NDA for paroxetine in adults, the mean imipramine dose overall was 140mg, with a mean endpoint dose of 170mg”
- in the method section we can read “(we acknowledge differing opinions about this issue in the statistical literature).” This comment has no reference.
- in the result section “(with a difference of 4 points being pre-specified as clinically significant)” : it is in already in the method section and should not appear in the results which are descriptive ;
- in the result section ‘(Scores on the HAM-D can vary from zero to a maximum of 52)’ that should appear in the method section.
- in the result section “the protocol also listed the relapse rate in the continuation phase for responders as a secondary outcome variable. Our calculation differed from the CSR calculation because we included those whose HAM-D scores rose above the ‘response’ range and those who intentionally overdosed.”
- in the results section authors states that “alternative treatments of the data could give different results.” It must be in the discussion section and not in the results.
- I also think that, for clarity purpose, the information about changes in sample size can be presented after the sample size calculation for clarity purposes.

One of the crucial points is the question of efficacy.

If I understand, in the study, there were two pre-specified outcome variables, with three groups. Was there a correction for multiple comparisons mentioned in the protocol? If I understand, there was also a change of primary outcome criteria which was done a posteriori and after breaking the blind. These points must be detailed. Can authors give the date of:

- Breaking the blind ;
- Changes made in the outcomes criteria ;

It would be also helpful to list and compare all the outcomes reported in the published paper by Keller et al.

In the sentence: "Global impression scale?" please suppress the "?" and explain that it is the CGI (as reported in the table).

The primary efficacy variable reported in the statistical methods and in the primary outcome variables are not the same. Please explain or correct.

In Table 1: please legend (mean [SD]).

Figures are represented for OC analysis, please provide the data for W8 (endpoint) ITT analysis with LOCF which was defined as the principal population of analysis. Please also indicate the number of patient in each group under the figure for each time point.

The other crucial point is the question of adverse events:

I understand that it is time consuming and difficult, but I think that the analysis of CRF should be complete to avoid any misinterpretation. It is indeed important since this audit process gave rise to additional AEs. Indeed, since this analysis is not complete, and since it was not at random, it is a major limitations and one can be very critic on this point.

In tables where the CRF estimates are presented, I think that this estimates are highly speculative and that the data cannot be analysed in this way. I suggest to delete this column and to analyse all the CRF. As authors state (and here again, this must be in the discussion section and not in the results), "alternative treatments of the data could give different results."

Moreover, authors agree with this point of view when they state : "The post-audit estimated figures for rates of AEs in this table may be an overestimate, since the CRFs audited were those of participants who were withdrawn from the study or who were known to have become suicidal."

SAE have a specific definition in MEDRA. I'm not sure that it is strictly overlapping with the notion of severity. Thus the comparison with Keller's et al. paper is very difficult as stated by the authors. For MEDRA, a SAE is serious when it results in death, life-threatening, hospitalization (initial or prolonged), a disability or Permanent Damage, in a congenital Anomaly/Birth Defect, it required Intervention to Prevent Permanent Impairment, and for other Serious (Important Medical Events).

This last category is a crucial point and it is probably not stricly overlapping with the notion of severe AE (used by the authors) : it is when the event may jeopardize the patient and may require medical or surgical intervention (treatment) to prevent one of the other outcomes. Examples include allergic brochospasm (a serious problem with breathing) requiring treatment in an emergency room, serious blood dyscrasias (blood disorders) or seizures/convulsions that do not result in hospitalization. The development

of drug dependence or drug abuse would also be examples of important medical events.

When authors state that “The majority of patients stopped at this point were designated by GSK as lack of efficacy (see Table 11). Investigators in four centres reported lack of efficacy as a reason for stopping six placebo patients even though the HAM-D score was in the responder range and as low as 2 or 3 points in some instances.” I would like to see more details. Additionally, I think that the change of coding between Lack of Efficacy and Adverse Event is difficult and could be misleading. Many times, discontinuation occurs for both lack of efficacy and adverse events, since one can easily consider that adverse events like dry mouth can be more acceptable in the case of treatment efficacy. This point could be addressed in the discussion and I’m not sure that a a posteriori interpretation of the CRF can give a perfect information about the individual patient experience (even if it is very better than aggregated data of course...). Moreover, I also think that a lack of efficacy can be considered for patients even if they are responder upon the HDRS. Patients are not just a score on a scale. The authors’ a posteriori proposal for recoding this can be thus erroneous.

Please explain, in the discussion, for readers that the interpretation of qualitative information in CRF is very subjective and prone to an interpretation bias (including for the first manuscript and for this one). Please explain why it is not possible to collect AE in an otherway (or explain how they should be collected) and the interest of MEDRA.

Table 5 can be deleted since it presents results that are also presented in table 6.

Legend of table 6 is missing (SOC*).

In table 11, please legend what is “RIAT proposed” ?

It is stated that “Roughly 1000 pages were missing from the CRFs audited”. Can authors precise why?

In the box Patient 00039, please detail wether it was AE or SAE.

In the discussion section, when authors state that “The RIAT approach [...] outcome variables.” It must be recalled that the message is very different since the Keller’s report state in the abstract that “Paroxetine demonstrated significantly greater improvement compared with placebo in HAM-D total score $< \text{ or } = 8$, HAM-D depressed mood item”.

When they state “In our opinion, statistically significant or not, all relevant primary and secondary outcomes, and harms outcomes, should be explicitly reported”. I’m not sure that it is only the opinion of this paper’authors. RCTs are often underpowered for detecting these changes.

The URL www.xxx is not exactly the good URL... Please do not test... and correct...

Where they state : “They reveal evidence consistent with dependence on and withdrawal from paroxetine.” I would nuance, “with possible dependence”.

Please xcuse my englilh but I tried to be very rapid.

Additional Questions:

Please enter your name: Florian Naudet

Job Title: MD, PhD

Institution: Rennes 1 university

Reimbursement for attending a symposium?: Yes

A fee for speaking?:

A fee for organising education?: No

Funds for research?:

Funds for a member of staff?: No

Fees for consulting?: No

Have you in the past five years been employed by an organisation that may in any way gain or lose financially from the publication of this paper?: Yes

Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this paper?: No

If you have any competing interests <A

HREF=<http://www.bmj.com/sites/default/files/attachments/resources/2011/07/bmipolicyondeclarationofinterestsmarch2014.pdf> target='_new'>(please see BMJ policy) please declare them here:

I have sat on a scientific board for Bristol-Myers Squibb (with a maximum of 10 sessions) and has received expenses for travel or accommodation from Servier, Lundbeck, and Janssen.

I have met David Healy in a symposium in France in 2013.

Reviewer: 2

Recommendation:

Comments:

Thank you for inviting my views on this important manuscript. I have been speaking with the RIAT authors for the past year or so, offering advice from time to time as to how to apply the RIAT concept to paroxetine study 329. As the first author of the RIAT declaration and as a BMJ editor that has been watching and chronicling the data transparency movement, their work has been of particular interest to me, and I am very happy to read their submission.

As RIAT declaration author, my aim is to have them produce the most robust and solid analysis possible in terms of the form of the RIAT reanalysis. Given the authors access to completed case report forms, their analysis ended up being far more in-depth than I think any of the RIAT declaration authors had ever expected RIAT papers would be. I believe this unprecedented level of access was due, at least in part, to The BMJ's coverage of their struggle for access to data e.g. see <http://www.bmj.com/content/347/bmj.f6754> In doing so, they have broken new ground, methodologically, particularly in terms of their analysis of harms data, an area that generally speaking lacks robust methods, and this, I think, is one of the most exciting parts of this paper.

Below, I provide some general thoughts on the content of the paper, places that are unclear, suggested edits, thoughts on where to host the data (a key RIAT issue), and my own competing interests relevant to this paper.

Some general issues

- Organizational issue. I think that in general the authors do not need to mention the Keller et al. publication in the Methods or Results sections of this RIAT manuscript. The misreporting of study 329 in the Keller manuscript has been well documented by the authors elsewhere. The primary purpose of this manuscript, as I see it, is on presenting an honest and accurate report of the study 329 results than it is to further document misreporting of Keller et al. If additional aspects of misreporting in Keller et al. were discovered in the process of RIATing study 329, this is important and I think the authors can include this information, but I think it would be best to keep this to the Introduction and Discussion sections. If that does not do sufficient justice, it might also form an independent paper itself.
- Audit of non-random sample of AEs. The RIAT authors carried out an audit of the adverse event section of case report forms (CRFs) for a non-random sample of 93 of the total 275 trial participants. The authors are very clear throughout the manuscript to indicate that this was a non-random sample. It would have been better of course if 100% of CRFs were audited, but given the number of hours it took to audit 93 (approx. 1000 hours they say in the text), a full audit likely only will happen if another group picks up the baton. I think the authors are correct to include analyses and tables that show the pre-audit and post-audit tallies of AEs. However I do not think it wise for the authors to extrapolate and present estimates, based on findings from their non-random sample, of the number of additional AEs they would have discovered had they been able to audit all

275 CRFs. (This might be OK if it were a random sample but it is not.) But here in particular, I do not think it wise because my impression of the non-random sample – of all participants that withdrew from the study (85) plus 8 children known to have become suicidal – is that it is a sample more likely to have problems in the transfer of information from CRF to CSR.

- I didn't see a COI statement for the authors in any of the manuscript and appendix files?

Places that need further clarification

1. Methods. Can the authors explain why they chose to follow the 1994 protocol instead of the 1993 or 1996 versions? Which version of the protocol was the last version before participant recruitment began in April 1994? Which versions do the authors have the full text for?
2. Methods. "Where relevant, we have referred to these variations." What does this mean?
3. Methods/Participants. "The protocol called for 300 subjects, but this was reduced to 275." Can this be clarified? So the 1993 protocol called for 300 subjects but this was revised to 275 in the 1994 protocol?
4. Methods/sources of harms data. "Roughly 1000 pages were missing from the CRFs audited." Can the authors explain how they knew pages were missing and can conclude this? (e.g. numbered pages indicating missing pages etc.) Were all 93 participants whose CRFs were audited missing the same pages/sections? Also, did they alert GSK to this and if so what was GSK's response?
5. Methods/coding of AEs. In the paragraph beginning, "Classifying a problem..." can the authors clarify if MedDRA puts 'sore through' in the central nervous system bucket?
6. Box 1. "Most recoding issues were clear-cut." What is meant by 'clear-cut'?
7. Competing interests statement appears missing. The authors say "as attached" but I could not find the attachment.
8. Methods/analysis of harms data. The authors chose to analyze MedDRA SOC classes psychiatric, cardiovascular, gastrointestinal, respiratory and place all other AEs in "other". After looking at the results tables, these look like reasonable choices to me, but can the authors include a sentence that explains how they made this choice?
9. Methods/patient withdrawal. In the paragraph beginning "The CSR states that the primary reason..." it mentions "CSR Appendix G". Can the authors say here briefly what Appendix G contains?
10. Methods/blinding. Could the authors also mention whether they reviewed the Certificates of Analysis for the study medications to double-check whether they appeared to have been correctly formulated to ensure blinding?
11. Methods/statistical methods. The authors write, "We followed the methodology of the a priori 1994 study protocol." Why is the 1994 protocol labeled "a priori"? Was it the last version prior to participant enrollment?
12. Methods/statistical methods. In the paragraph beginning "The primary efficacy variable", there are two sentences with the phrase "primary efficacy variable". I suppose this is a reflection of the trial having two outcomes prespecified as "primary"?
13. Discussion. Does the following text refer to Keller et al. or the CSR: "The authors/sponsors departed from protocol by performing pairwise comparisons of two of the three groups when the omnibus ANOVA showed no significance in either the continuous or dichotomous variables." This should be clarified. If this refers to the CSR, then to some extent there is a discovery among the RIAT authors that they have found reporting bias within the CSR itself, and I think this is an important finding which they should highlight as such.

14. Box 3. "The inability to access all CRFs may have introduced some error." Not sure what is meant by this. Are the authors talking about their inability due to time/resources to audit everything? Is this a reference to the difficulty to use portal for accessing the study data? Or is this a reference to the approximately 1000 pages that were missing from the CRFs that GSK made available through their portal?
15. RIATAR. Why are some items so long? For example, so many sources are given for Funding (#25).

Suggested changes in wording:

1. Abstract/Results. Suggest changing, if appropriate, "for any measure" to "for any primary or secondary [efficacy] outcome."
2. Background. "RIAT publication of Study 329 which was funded by..." Change to "RIAT publication of Study 329. The original study was funded by..."
3. Background. "On 14 June 2013, the RIAT researchers notified GSK that Keller et al. appeared ... Study 329." This refers to a letter I sent GSK. We did not specifically mention study 329 in this email. In order to make the sentence accurate, I suggest rewording: "On 14 June 2013, the RIAT researchers asked GSK whether it had any intention to restore any of the trials it sponsored."
4. Similarly, change "GSK did not signal any intent to publish a corrected version of the article." to "GSK did not signal any intent to publish a corrected version of any of its trials."
5. Methods/Secondary Efficacy Variables. "We could not find any document that provided any scientific rationale for these post-hoc changes..." Did you find any "non-scientific" rationale? If not, perhaps delete "scientific".
6. Methods/Outcomes. The headings 1. Principal Endpoints for Efficacy and 2. Principal Endpoints for Harms. I think this is slightly confusing with the language of "primary" and "secondary" efficacy variables. How about just labeling the sections "Efficacy Endpoints" and "Harms Endpoints"?
7. Methods/Harms. I think the "(p. 18)" at the end of the quoted paragraph is a typo as it is also stated above.
8. Box 1. "At the week 6 visit ... GSK..." Do the authors mean SKB?
9. A variety of terms are used to represent the provenance of AE data e.g. "CSR recoded" and "CRF audit" from table 7, "AEs in Appendix D" from table 9, and "AEs reported (CSR check)" in table 12. I wonder if better terms can be used to make the meaning more transparent. Perhaps some variant of "SKB/GSK coded", "RIAT recoded", and "RIAT recoded plus CRF audit"? Another thought is to use terms like ADECS and MedDRA e.g. "SKB/GSK ADECS coded", "RIAT MedDRA recoded", and "RIAT MedDRA recoded plus CRF audit discovered additional AEs". I realize that some of my proposed titles are long and won't fit in the space of a tight table, but my suggestion to remove the Keller columns as well as the "CRF estimated" i.e. extrapolated AE count columns from the Results section will hopefully free up some space.
10. Results/Discontinuations. "Consort" should be "CONSORT".
11. Results/Discontinuations. "GSK regarded these patients as participants in the continuation phase..." Should this be SKB?
12. Box 2/section 8. "... because it became clear that the blind had been broken..." Can you just be clear whose blind you are talking about? I.e. I think this is SKB's blind, but I'm not 100% sure as part of the RIATers recoding happened blind while other parts did not.
13. Discussion section/two paragraphs before Conclusion. "... analysis of adverse events requires access to individual patient level data (CRFs)." I would reword the ending to "...requires access to individual patient level data in the form of CRFs."
14. Conclusion. "Study 329 showed no advantage ... on any of the specified

parameters.” Would using the word “pre-specified” be better than “specified”?

Moving text around:

1. Methods/Interventions. “These imipramine doses are high for adolescents. In the six comparator studies submitted by SKB as part of their 1991 Approval NDA for paroxetine in adults, the mean imipramine dose overall was 140mg, with a mean endpoint dose of 170mg.[14]” I think this should go to the Discussion section unless it was part of the original methods.
2. Methods/Source of harms data. Suggest moving the following to Results: “Of the eleven paroxetine patients with AEs designated as serious, nine discontinued because of AEs. A large number of other patients discontinued because of AEs that were not regarded as serious, or for lack of efficacy or protocol violations (see Figure 1). None of these latter discontinuations led to patient narratives.”
3. Box 1 looks like it belongs in Results, not Methods.
4. Table 8 is great, but perhaps should go in the Discussion?

Hosting of data

1. I would favor the authors deposit a copy of all public data (i.e. all the data that they legally can make public) with The BMJ for us to post online as appendices. This will represent a contemporaneous copy of what material they had to work with (that was public). In addition to this, I think it's fine to also make the data available in other places, such as a dedicated website for this study, in ways that might make it easier for others to work with.
2. The rationale for my suggestion is that part of the RIAT vision was to publish articles WITH data i.e. having the data available alongside the article, for all readers. At present, the authors cannot make all the data they accessed available because of GSK's terms of use for access to electronic individual participant data and completed case report forms which the authors accessed through GSK's secure web-based system. This system is specifically designed to enable access to but prevent downloading of the data. The study 329 data in the authors' hands that can be made available includes, as I understand it, the CSR as well as earlier (and later?) versions of the trial protocol plus other memos and other documents that were made public through the discovery process of past lawsuits.
3. Another option instead of The BMJ website would be to place the documents on Dryad. This is fine and what we did with the neuraminidase inhibitors CSRs which were about 1.3GB. With study 329, however, I think the file sizes are a lot smaller since it is just 1 trial.
4. Although it is an option, I do not think pointing readers to a link to the CSR that is available on GSK's website is a good idea. The 2004 settlement between GSK and the New York State Attorney General's office only stipulates that GSK make these data available for ten years after Feb 1, 2005. Therefore it is possible GSK may take the CSRs down or move them elsewhere.

Competing interests

My own competing interests in this case are complex but should be spelled out for clarity.

- I am the first author of the RIAT declaration.
<http://www.bmj.com/content/346/bmj.f2865>
- One of the RIAT declaration co-authors is also a co-author of the paper under review

(David Healy)

- The RIAT declaration was published in The BMJ before I came on staff at The BMJ.
- I have written and spoken about the misreporting of paroxetine study 329.

Thanks again for inviting my comments.

Peter Doshi

Additional Questions:

Please enter your name: Peter Doshi

Job Title: associate editor

Institution: The BMJ

Reimbursement for attending a symposium?: Yes

A fee for speaking?: No

A fee for organising education?: No

Funds for research?: No

Funds for a member of staff?: No

Fees for consulting?: No

Have you in the past five years been employed by an organisation that may in any way gain or lose financially from the publication of this paper?: No

Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this paper?: No

If you have any competing interests (please see BMJ policy) please declare them here:

I spoke in Naples and Rome last year on the topic of RIAT and used paroxetine study 329 as a prominent example of misreporting, highlighting the work that Jureidini and colleagues were doing to correct the record through RIAT. I was reimbursed for my travel to Italy.

Reviewer: 3

Recommendation:

Comments:

Review of the Manuscript entitled: 'A randomized, controlled trial of the efficacy and harms of paroxetine and imipramine in the treatment of adolescent major depression: Restoring study 329.' BMJ.2014.022376.

Journal: The BMJ

Reviewer: Hilde PA van der Aa, VU University Medical Centre, Amsterdam The Netherlands

This study describes the reanalyses of a double-blind RCT conducted from 1994 to 1998 in 275 adolescents with major depression on the effectiveness of paroxetine treatment and imipramine treatment versus placebo. The primary outcome was the difference in depression scores at baseline and at the end of the acute treatment phase. In the new analyses both treatment types were not clinically or statistically different from placebo and were both poorly tolerated based on serious adverse events. Though the RIAT initiative proved to be a lot of work, it is very important that it was performed as this study produced relevant outcomes that until now have been incorrectly analysed and described and used for clinical practice. Moreover, it draws attention to the impact of publication bias which is an important problem in current scientific research. Still I believe the results of this study should be interpreted with caution. Therefore, some remarks can be made.

Major issues:

1) The authors followed the methodology as stated in the pre-specified protocol of 1994, in which proposed statistical approaches or statistical assumptions were not justified. Outdated techniques were used to analyse the data, leading to more uncertainty. I would recommend authors to (also) include modern techniques of data-analysis or at least mention this 'limitation' in the discussion part of the paper:

- One of the limitations of this trial is the large number of dropouts. Therefore, a linear mixed models approach to analyse the data with a maximum likelihood assumption is better suited to estimate effects than the chosen ANOVA and GLM.
- If authors, however, do decide to use ANOVA and GLM multiple imputation would be a better way to handle missing data than the currently used LOCF, see for example the paper by Beunckens et al. 2005 [1].
- Authors described that they did not correct for attrition and non-compliance in the sample size calculation. In addition, they also did not correct for the different strata in their sample (12 centres included). This should also be reported.

2) Limitations of the current study should be described in more detail. The limitations of the statistical analysis (as mentioned above) should be mentioned. In addition, the authors state that 'The inability to access all CRFs may have introduced some error.' (page 27, line 25). This should be explained in more detail.

3) At the beginning of the discussion authors state to draw minimal conclusions regarding efficacy and harms, inviting others to offer their own analysis. I think this is a just conclusion based on previously mentioned limitations. However, this cautious approach

of interpreting the results of the RIAT study should also be reflected in the conclusion part of the abstract and the discussion.

Minor issues:

4) Throughout the whole paper authors describe the 'new study' compared to the 'old study' of Keller et al. This makes it difficult to read and to distinguish the methods used in the RIAT trial. Though it is important to report these differences, they might for instance be collected in boxes or reported in italic or combined in the methods section of the paper.

5) The abstract does not follow the standard style of 'The BMJ' for research articles: objectives, design, setting, participants, intervention, main outcomes, results and conclusion.

References

1. Beunckens C, Molenberghs G, Kenward MG. Direct likelihood analysis versus simple forms of imputation for missing data in randomized clinical trials. *Clinical Trial*, 2005; 2: 379-86.

Additional Questions:

Please enter your name: Hilde PA van der Aa

Job Title: Clinical researcher

Institution: VU University Medical Centre

Reimbursement for attending a symposium?: No

A fee for speaking?: No

A fee for organising education?: No

Funds for research?: No

Funds for a member of staff?: No

Fees for consulting?: No

Have you in the past five years been employed by an organisation that may in any way gain or lose financially from the publication of this paper?: No

Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this paper?: No

If you have any competing interests <A

href=<http://www.bmj.com/sites/default/files/attachments/resources/2011/07/bmjpolicyondeclarationofinterestsmarch2014.pdf> target=' new'>(please see BMJ policy) please declare them here:

Reviewer: 4

Recommendation:

Comments:

There did not appear to be any score sheet available. My comments are below.

Originality

It is difficult to know on what to base the judgment regarding originality. On the one hand the RIAT initiative is innovative and critically important initiative in terms of ensuring greater transparency of reporting of trials. And as authors point out, it is the first time a trial has been reported for a second time in the peer reviewed literature. On the other hand that paroxetine shows no statistically significant difference from placebo is not a new finding. For example, the Cochrane systematic review of antidepressants for children and adolescents (original publication and 2012 update) includes data from Keller, and indeed from the SmithKleine website with pooled analysis highlighting that lack of statistically significant difference between placebo and paroxetine. The correspondence between Jon Jureidini and Martin Keller in the Journal of the American Academy of Child and Adolescent Psychiatry (letter section) also ensured that findings according to the a priori outcomes were reported in 2003. These letters, and the issue of reporting bias are outlined in the Cochrane review.

Importance

I think that it is extremely important to ensure the issue of reporting bias is highlighted and measures are taken to prevent its occurrence. In this regard the RIAT initiative appears to be an important one and having publications from this group supported by a journal like BMJ might be very important to seeing more honest reporting from trialists.

It would certainly be beneficial to have journals who are able and willing to publish all of the associated data with a trial – making it easier to obtain it (for those who have access to the journal) and analyse data in alternative ways or use it in systematic reviews and meta-analyses.

Scientific Readability

This paper was well written and presented data in a clear, thorough and meaningful way. I particularly appreciated the tables of outcome data that clearly labeled the OC and LOCF data and what the data represented i.e. output from modeling (LSMeans and SEM).

It's hard to know exactly what should be in the background, or indeed what the objectives are or how a paper like this should be written up. On one hand it is simply the description of a trial, but on the other hand it has several important other objectives I think: first, to correct errors of the previous write-up; second, to highlight the issue of reporting bias. I am not 100% sure that the second objective was clearly articulated or achieved, and perhaps this is the objective of RIAT but not necessarily of this paper as such. My personal opinion is that more could be made of it in this paper (and perhaps this would address my concerns about originality made above) and that the background appears to indicate that that correcting errors and highlighting the issue of reporting bias is what the paper is about.

Some specific queries I have (not necessarily that need to be addressed in the paper, but were questions that came to mind for me) and comments follow:

1. Should the background include something about letter by Jon Jureidini and Martin Kellers response in 2003? This saw the correction of findings to a certain extent.
2. I was interested to know whether the reader should just believe that the Keller 2001 paper was ghost written or whether there is some kind of proof of this? How did the authors find this out/know?
3. In the fourth paragraph the authors refer to the RIAT statement, but I wasn't clear what this was?
4. In the fifth paragraph the authors outline the objectives of the original study but don't state where these objectives were derived from? The Keller paper, or from the SKB reports?
5. It wasn't clear to me how patients were identified: obviously authors have stated that telephone screening was undertaken, but was this of a particular population? It also wasn't clear what happened during the screening phase that enabled investigators to know that symptoms were stable i.e. was the K-SADS or HAM-D administered twice over and at what time points. Was there a placebo lead-in phase? I think this information should be included.
6. Again, because the objectives were slightly unclear (or mixed?) I think the write-up is missing some detail about the methods (if one of the objectives is to publish a sound write-up of this trial). This includes details about how allocation was concealed (i.e. states that patients were assigned treatment numbers in strict sequential order, but where the treatment numbers in sealed opaque envelopes?), who was blinded and how i.e. from the write up we can assume that the patient and the person providing the pills to the patient were blinded, but were all the investigators, were the people giving the supportive counseling (who were these), was the statistician doing the analysis?
7. ITT analysis includes all those randomized not all those who receive at least one dose of medication and have at least one post-baseline efficacy assessment.
8. I wonder if the authors have thought about undertaking the analysis using more modern and robust methods of imputing the missing data e.g. multiple imputation? I know the authors have indicated that they have provided the data and that therefore others can undertake the analysis as they wish; and that their intentions were to analyse as per the original protocol. But it would be interesting to know what difference a more robust method of imputation makes to the outcomes. In the Cochrane systematic review, undertaking the analysis using LOCF vs OC data made little important difference to the outcomes.
9. I do wonder if the authors should highlight the possible overestimation of the AE figures as a limitation in the discussion and highlight this in the abstract; or I wondered if indeed, given the way in which AEs have been derived and that there is no analysis (and certainly no a priori planned analysis), that this finding should not be stated in the abstract at all. The abstract should perhaps be a clean reporting exactly according to the objectives and pre-planned analysis.
10. Having said that (9), the results with regard to drop outs and AEs is long and complicated and includes long tables with a lot of information; it is hard to know whether readers will take much note or be able to follow it. I think following through each step is important i.e. the author should try to do some synthesis by pulling the AE's into groups). Whether further analysis or synthesis could be helpful is unclear; perhaps it is more useful for those undertaking systematic reviews and meta-analyses to think about what to do with this data.

11. Some of the paper appears not to be finished e.g. there is a question mark after the dot point "Global Impression Scale" (pg 5) and xxx used to indicate some websites (pg 23-25).

12. In Box 3, authors state that trial participants had relatively chronic depression; which may be true but isn't clearly reported in the results. I'm not entirely convinced that many adolescents have shorter durations of depression. Previous studies suggest that the duration depression might range from 6 to 9 months; but that up to 50% of children and adolescents can still be at 12 months, and 20 to 40% at 24 months (Kovacs, Feinberg et al. 1984; Birmaher, Ryan et al. 1996; Harrington 2001). The trials included in the Cochrane review demonstrated this with a large range of duration of current episode from 10 or 15 weeks to 100 or 108 weeks.

Additional Questions:

Please enter your name: Sarah Hetrick

Job Title: Senior Research Fellow

Institution: Orygen Youth Health Research Centre, Centre for Youth Mental Health, University of Melbourne

Reimbursement for attending a symposium?: No

A fee for speaking?: No

A fee for organising education?: No

Funds for research?: No

Funds for a member of staff?: No

Fees for consulting?: No

Have you in the past five years been employed by an organisation that may in any way gain or lose financially from the publication of this paper?: No

Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this paper?: No

If you have any competing interests <A HREF="<http://www.bmj.com/sites/default/files/attachments/resources/2011/07/bmjpolicyondeclarationofinterestsmarch2014.pdf>" target="_new">(please see BMJ policy) please declare them here:

I am an author on the Cochrane review of Antidepressants for Children and Adolescents.

I am a trial investigator on a trial funded by the National Health and Medical Research Council to investigate the effectiveness of fluoxetine plus CBT vs placebo plus CBT.

Reviewer: 5

Recommendation:

Comments:

This paper is an important milestone in the history of medical research.

Study 329 has been the subject of considerable controversy and speculation as to what the 'missing data' may reveal if subjected to independent scrutiny. This has now been addressed by Jureidini et al, and answers a number of questions of primary importance to parents, patients and carers.

The relatively 'benign' reputation of SSRI's and Paroxetine in particular, in the treatment of adolescent major depression is now shown to be highly questionable. Despite the extremely difficult methods of access imposed upon the team and the tedium endured in producing this paper, they have come to conclusions that have been expressed as opinion in the past, by a number of critics: that this type of medication has considerable potential for harm when administered to our children.

It also seems somewhat controversial that the tricyclic medication chosen to be a control was given in doses that were highly likely to be poorly tolerated, being two to three times normal dose; leading one to the conclusion that the original trial design was biased in favour of Paroxetine.

Due to the huge amount of data this team had to evaluate they were only able to fully evaluate one third of the individual patient information but even this allowed a conclusion to be drawn that Paroxetine was of no significant worth for the treatment of adolescent major depression and that it had the capability of considerable harm.

This paper is important to any parent or patient cohort considering the use of anti-depressants for treatment and provides invaluable information for the exercising of any informed consent.

I would, as someone used to the deciphering the world of acronyms in my own sphere, make a heartfelt plea to reduce their scope and volume: they are intimidating to the layperson attempting to understand medical information and trials and often baffling to patients.

Additional Questions:

Please enter your name: Ernest Berry

Job Title: Safety & Security Consultant

Institution: Self Employed

Reimbursement for attending a symposium?: No

A fee for speaking?: No

A fee for organising education?: No

Funds for research?: No

Funds for a member of staff?: No

Fees for consulting?: Yes

Have you in the past five years been employed by an organisation that may in any way gain or lose financially from the publication of this paper?: No

Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this paper?: No

If you have any competing interests (please see BMJ policy) please declare them here:

I am a member of a pressure group seeking the establishment of a legal duty of candour for all health care professionals.

Reviewer: 6

Recommendation:

Comments:

I think this is a potentially important report. However I believe that it needs extensive revision. In particular, I don't think the authors have taken adequate steps to manage their potential professional conflicts of interest. I believe that if this goes forward the revision should include a retrieval of all of the clinical report forms, masking of the CRFs to remove any clues as to the drug being taken and independent re-coding of the adverse event reports by individuals not previously involved in criticism and re-analysis of this trial.

The authors report a re-analysis of a randomised trial comparing paroxetine and imipramine with placebo that was led by Dr Martin Keller and took place in the 1990s. The importance of this trial is that it shaped early attitudes towards the use of SSRI in childhood and adolescent depression and early enthusiasm for this class of drugs in younger patients. The authors' re-analysis and restoration of Trial 329 could be considered a landmark effort. It represents a massive amount of work and the full report and additional documentation (if accurate) provide a repository and also a reality check for groups planning an undertaking of this magnitude.

Since publication Trial 329 has been discredited on a number of levels, including selective outcome reporting, manipulation and misclassification of the adverse event data, and ghost writing of the original manuscript draft. The authors have had an important role in revealing these problems in previous publications and in lobbying efforts directed at the manufacturer and the journal. I believe that at least one author has appeared on behalf plaintiffs taking legal action against the manufacturer. So the authors have a long and legitimate history of involvement in highlighting problems with the trial. In the light of their work, and information from other sources, the FDA and other major drug regulatory agencies recommended against the use of SSRIs in child and adolescent depression from around 2002.

The work performed by the authors required great commitment and they should be congratulated and thanked for their efforts. This same commitment led them to lobby against the trial investigators and the manufacturer for many years. There is a price to pay for such advocacy. It might be awkward for the authors (or indeed anyone in this situation) to perform a re-analysis whose results undermined a position that they had held for years and has helped them build their professional reputations.

I am not for a minute suggesting that their work is tainted. But recognizing the real and perceived professional conflicts of interest that their previous work has created, I think it would have been wise – indeed necessary - to take steps to minimise the impact of this on their work. I don't believe that they have taken adequate steps to do this.

In my view, the questions raised by this re-analysis of Trial 329 are:

- 1) Beyond 'setting the record straight', which may be important in its own right, does

the re-analysis of the trial contribute to our understanding of the efficacy and safety of these drugs in young people?

2) Were the authors the best people to conduct the re-analysis of Trial 329? While overseeing the work, should they have commissioned another group to carry out the more sensitive re-coding of outcomes? Did the techniques used by the authors guard adequately against bias that might be introduced by their expectations, shaped by their previous experience of this study and related advocacy efforts?

3) Do the data and analytical methods support the conclusions of the authors?

4) How should the methods and results of the two analyses of trial 329 be presented?

1) The authors wished to create a publicly accessible corrected and comprehensive set of summary documents relating to trial 329. This is a laudable aim and, as noted, has involved a great deal of work. Their results contradict the manufacturer's original claims of substantial efficacy and they point to some serious adverse events that may apply more widely than the study population in the trial. On the other hand the major regulatory agencies have advised against the use of SSRIs in children and youth and the FDA has advised against its use in younger people. It is currently contra-indicated for those under 18 years.

a. In the light of the current situation will the authors provide a clear indication of how they believe their re-analysis of trial 329 will further inform regulatory decisions and clinical practice?

b. What additional value will this exercise provide – for instance for others performing restoration of other important trials?

2) If we accept a need for the re-analysis, have the authors taken adequate steps to manage their professional conflicts of interest? The guarantor of the study has been active over a number of years, has published at least one critique of this study, has corresponded quite vigorously with the authors of the original report, and the journal editor, and has acted on behalf of plaintiffs taking legal action against the manufacturer. There is nothing wrong with any of these activities. The concern here is that the authors have adopted such a strong negative position on the drug, and this trial, that they could suffer loss of face if the results of the re-analysis went against their original strongly held position. A number of the decisions that they made required judgements and I am not satisfied that they have taken adequate measures to avoid bias in making these.

a. Will the authors provide more detail on the methods of blinding assessors of the written material that required subjective judgments?

b. How successful was blinding and did they consider asking a group independent of the study team to carry out this work with copies of reports from which key information (such as drug name) had been masked?

c. With more resources and time could they have retrieved all the clinical report forms in order to reclassify the adverse outcomes?

3) There are two main conclusions that define the benefit to harm relationship of the drug used for this indication: a) the authors conclude that there are no important difference in the efficacy data between their analysis and the original report, but their interpretation of the results is different; and b) Their re-analysis reveals a disturbing excess of suicidal ideation and episodes of self-harm.

a. Regarding the analytical approach to the efficacy data, I understand the logic of what they propose – which is to carry out ANOVA and only do pairwise analyses if the overall analysis reaches a statistical threshold. I am not a methodologist, but I feel this is excessively conservative in this case with two active treatments being compared with placebo. Based on prior evidence it was possible to construct separate hypotheses for each drug that would justify pairwise comparisons. As the authors say this is controversial, but their stance could add to the impression that they did not start this re-analysis from a position of equipoise. I believe that for the record they should present and interpret these analyses – as readers will try to do them anyway.

b. The main differences revealed by the re-coding and re-analysis of Trial 329 are in the adverse event data. A crucial part of this is the audit of 93 cases with adverse outcomes. To quote the authors “This audit comprised all 85 participants identified in CSR Appendix H who were withdrawn from the study, along with 8 further participants who were known from prior inspection of the CSRs to have become suicidal. “ As noted elsewhere this recoding was largely carried out un-blinded and the authors use crude multipliers to estimate possible numbers for the trial populations. As they say this scaling up from a non-representative sample may have over-estimated the numbers for key adverse outcomes. The exercise led to a relatively large increase in CNS adverse events with paroxetine - in particular suicidal ideation and suicide attempt – plus depression worsening and aggression. As this is arguably the key finding of the re-analysis I think greater efforts should have been made to obtain all the CRFs and to mask them by manually screening out drug names or other clues in the text and to have the recoding carried out by people who were not involved directly in the study and had not been involved in previous efforts by the authors to discredit the trial.

4) Presenting the results:

a. Considering efficacy, I think the authors should present their re-analyses alongside the originals. For the primary outcomes, as far as I can see for the HAMD >50% drop or <8 the authors results for imipramine and placebo are identical to those in the original report by Keller et al. However, the proportional response with paroxetine in the re-analysis (65.6%) in the LOCF analysis is slightly lower than the original figure (66.7%). They may have reclassified a responder – can that be clarified? For the drop in HAMD scores the authors have presented the LSMeans from their modeling, whereas I think the original paper presents the differences in arithmetic means. The differences are small and probably don't affect the overall conclusion, but this point should be clarified.

b. The authors have presented the adverse event data in a series of tables. These are quite clear (except for doubts about the total estimated numbers from the incomplete audit). However it feels like they are scattered across several tables. I think the authors should try to produce a summary table where the major outcomes – efficacy and adverse events are summarized from the original trial report and from their re-analyses are presented.

c. The overall report is long – 127 pages, of which 32 constitute the main trial report. The remainder can be handled as supplementary material and as the authors state may be valuable to others. But the main part of the report is quite long and tends to editorialise in almost every section. I think a more tightly written report that sticks to a description of what was done, what was found and how the findings differ from the original would be more readable

d. I would like to highlight text box 2 in the Discussion entitled “Potential confounders of accurate reporting of harms”. This is particularly valuable and is a very useful guide to reporting adverse events in clinical trial reports.

Additional Questions:

Please enter your name: David Henry

Job Title: Professor

Institution: University of Toronto

Reimbursement for attending a symposium?: No

A fee for speaking?: No

A fee for organising education?: No

Funds for research?: No

Funds for a member of staff?: No

Fees for consulting?: No

Have you in the past five years been employed by an organisation that may in any way gain or lose financially from the publication of this paper?: No

Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this paper?: No

If you have any competing interests <A

href='http://www.bmj.com/sites/default/files/attachments/resources/2011/07/bmjpolicyo
ndeclarationofinterestsmarch2014.pdf'target=' new'>(please see BMJ policy) please

declare them here:

END