Reviewer's report

Title: Data sharing in clinical research is a moral imperative

Version: 1 Date: 5 July 2011

Reviewer: David Ransohoff

Reviewer's report:

OVERALL
While there are good ideas in this paper, it also risks being diffuse, including substantial problems in important definitions (what exactly is data sharing, for example) and implications. It might be more useful to focus on one or two specific problems and next-step reachable solutions.

As currently written, it sounds like it is calling for everything to be publicly available, including all protocols and raw data of every study. If that's the case, the argument is not well-made. If that's not the case, then details need to be specified.

As an example of lack of focus, some "calls for data sharing" on page 17ff are not about sharing raw data but are about establishing clinical trials registries. The author's own interpretation of "calls for data sharing" risks looking skewed to bolster the author's argument which appears to be "always release everything."

MAJOR SPECIFIC POINTS

1. Definition of "data sharing"
The definition of "data sharing" is, perhaps surprisingly, not really clear. The problems identified in the first part of the manuscript are of several different sorts, including (this list is in no particular order and is not comprehensive; it is meant to illustrate that "data sharing" includes quite different things):

a. trial registries (to reduce problems of selective reporting) [This is the subject of many of the problems listed, and is the main topic in the next to last paragraph of the Abstract.]
b. academic authors should have access to all data/results (to avoid hiding of problems)
c. anyone should have access to all data/results

It would seem to me that serious efforts at 1 and 2 would be hard enough, for logistic and political reasons, but may be worthwhile enough to seriously entertain as "next steps."

As it stands, it sounds like you are recommending "release all data for every study." Couldn't that situation be chaotic, in that anyone with an agenda could
selectively utilize or interpret complicated data in a way that advances a cause (e.g. as done by plaintiffs' lawyers, say, for silicone breast implants)? Currently we have institutions (academic authors, regulatory agencies like FDA, NICE, etc) with responsibility for interpreting complexity and preventing distortion. They may not be doing their jobs in the best way all the time, but would total transparency all the time (Is that what you are proposing?) really improve the situation? In contrast are there incremental ways - yet substantive - to make institutions do a better job so that we can trust them?

2. Tone

The tone of the piece ("moral imperative" and "appropriate punitive measures" in Abstract) risks being self-righteous and not the right metaphor. If you are going to treat something as a crime (punitive measures), you need to be very specific about what the crime is. In the body of the piece you list very different problems or "crimes," so it is not exactly clear what you are calling for. Your argument about "moral imperative" is undermined when, in the "call for data sharing" you list so many different kinds of things without a realistic or clear enough argument (including about others' support) for each.

3. Title

"Data sharing" should be used cautiously in a title if it's not really clear what data sharing means.

Similarly, "moral imperative" should be used cautiously if it applies to multiple things.

The title is engaging, but it risks not being a fair description of the content.

4. Focus

There are a number of problems reflecting lack of focus:

a) On page 3, in the top section, you discuss how the drug industry can compare new drugs with placebo rather than best-existing drugs or instead of head-to-head comparison. This may be true, but is this problem solved by "data sharing"? Much of the last 2 paragraphs of the 4-paragraph section titled "The problem" seem to be about topics note related to data-sharing.

b) On page 3, "selective reporting", first paragraph, is clearly important. Isn't this addressed by trial registries? Is the "per protocol" problem solved by "data sharing"?

c) On page 6, you describe how papers are written not by academic authors but by companies. This may be a serious problem when it occurs, but is the only solution of the best solution "data sharing?"

5. Solutions

On page 12 you describe legislation, but it is not clear exactly what is being called for:
- What is "protocol-related information"? What is the initial experience with this legislation?
- Registration of clinical trials: This seems sensible; why hasn't this been adopted in more countries?

**Level of interest:** An article of importance in its field

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**

I declare that I have no competing interests' below.