Reviewer’s report

Title: What Counts as Reliable Evidence for Public Health Policy: the Case of Circumcision for Preventing HIV Infection

Version: 1 Date: 1 November 2010

Reviewer: Susan L Norris

Reviewer’s report:

Comments for the authors
Major essential revisions
1. This is an interesting and well-written manuscript that discusses an important topic in comparative effectiveness: the relative merits of observational studies and RCTs. The authors focus on a great example and present their arguments well.

2. My main concern about this paper is a mixing of the construct of confounding with both bias and with external validity. Confounding refers to a factor that is associated with circumcision and also causes (is a risk factor for) HIV infection. Bias relates to the design and performance of a study (observational or RCT) and every effort needs to be made to minimize it and discuss its potential effect. Confounding can be adjusted for and addressed. Bias can’t be corrected for post hoc. The authors refer to “a lack of blinding” as a confounder. A lack of blinding is a potential source of bias, and is not a confounder. The same for differential reporting (page 9).

3. Behavioral change which results from circumcision is therefore not a confounder, but rather an intermediate step in the causal chain between circumcision and HIV. (Behavioral change may be a confounder in other situations).

4. In the discussion of external validity (page 11), the authors refer to issues of selected recruitment into the studies as a confounder. It is not: rather it is an issue of generalizability of study results.

5. The authors mention that all 3 RCTs were stopped early and they provide a nice discussion about the ethics (or lack thereof) of doing so. Stopping trials early for benefit is a major potential source of bias and should be presented as such, and not just mentioned seemingly as a secondary issue under an Ethics section. I suggest the authors cite some of the recent literature on this bias and ethical issues (G. Guyatt and/or Bassler in the 2009 and 2010).

Minor essential revisions
1. The terms efficacy is used in a several places where I think the authors mean effectiveness. For example, page 3, Background, 1st paragraph: “However there is an ongoing controversy over the... in assessing efficacy...”. Isn’t it effectiveness that is important to health policy?

3. Paged 3, last paragraph: “All trials are subject to confounding and bias…”. I think you mean “all studies”, not just trials.

4. I think the terms “hard” and “soft” endpoints should not be used as they are imprecise and difficult to define. You might consider distal health outcomes, or outcomes which are objectively measured, depending on what you mean.

5. Page 9: “It is therefore important to control for this risk factors when evaluating the CAUSAL” (currently reads “casual”).

Minor discretionary revisions: none

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