Reviewer's report

Title: Estimating the Cumulative Risk of False Positive Cancer Screenings

Authors:
Dr Stuart G Baker (sb16i@nih.gov)
Diane Erwin (d.erwin@btinternet.com)
Barnett S Kramer (KramerB@OD.NIH.GOV)

Version: 1 Date: 2 May 2003

Reviewer: Hans Reitsma

Level of interest: A paper whose findings are important to those with closely related research interests

Advice on publication: Unable to decide on acceptance or rejection until the authors have responded to the compulsory revisions

General opinion

The authors propose and work out a generalisation of the approach by Gelfand and Wang (Stat Med 200;19:1865-79) to estimate the cumulative risk of one or more false positive test results during a screening programme on a periodic basis.

Their approach and equations are valid and correct as far as I am able to review this rather technical paper (no formal statistical background). This is directly my main concern with this paper: what should be the targeted audience? This paper is now limited to persons with a formal statistical background. This is clearly a disadvantage because the topic is relevant to larger audience.

Compulsory revisions
1. The authors should make a better effort to explain and illustrate the difference between their approach and that of Gelfand, including:
   * why, how, and in which situations will their approach differ.
   * an empirical comparison showing the numerical differences between their approach and that of Gelfand by re-analysing their data by the method of Gelfand (or drop the assumption that prior that dropout depends on previous false positive(s)).

2. In my view, the cumulative risk for a first false positive is the most important outcome to model, as the first false positive will generate the most anxiety. Women with more 'difficult breasts' will be informed about the higher likelihood of false positives in the future. This is clearly a different group. To my understanding, the current approach is not fundamentally different from that of Gelfand with respect to the risk of at least one false positive. Is this is correct, and if so, this should be stated more clearly.

3. Provide more information about the dataset that has bee used as an example, including: the general outline of the screening study, how many women were included, the mean age of these women, the dropout rate at each visit, the number of women having had all 4 screening, three, two, one screening, the number of women with one, two, three or more false positives (power of the study), etc.
4. A more thorough discussion about the assumption that dropout does not depend on future false positives. I still believe that this is a strong assumption. For instance, women detecting a lump by self examination will visit their regular physician, thereby receiving imaging tests outside the screening programme. These women are more likely to dropout of the screening programme and are more likely to have a ‘false positive’ test. Is there any information on the rate and outcomes of imaging tests outside the screening programme?

**Competing interests:**

None declared.