Reviewer's report

Title: Breast Cancer Awareness - The effect of the October campaign on Internet search activity

Version: 1 Date: 25 April 2011

Reviewer: Parisa Tehranifar

Reviewer's report:

This manuscript uses data on internet use to indirectly assess the impact of breast cancer awareness campaigns that have been taking place for nearly two decades. The results are significant in providing feedback on the effectiveness of this type of public health efforts in stimulating interest and hopefully actions. My suggestions to the authors are listed below and are intended to help clarify the methods and results and to improve on the discussion of the results.

1. Please provide a brief description of what is meant by “Google domains”.
2. Since many readers will not be familiar with the measure of internet activity, more detail on the data and measurement would be very helpful. For example is the scale of 0-100 on an integer scale? Are there multiple search terms with the same score? What is the range of number of search terms that may have scores 0-100? This type of information will also provide some context for the data presented in the results section.
3. Is the “average search activity” the arithmetic mean of the search scores – for example for each month?
4. Can the authors provide some meaningful data on the magnitude of differences observed – for example, for differences across the three cancer sites? Beyond statistical chance and p-values, the results are very limited in giving a good picture of what the differences really mean. This goes back to point #2 made above with regard to some context to the data presented.
5. The authors need to better orient the readers to figure 4. Perhaps they can indicate to which figure panel each of the 5 p-values presented in the last paragraph of the results sections refer.
6. Can the authors offer any speculations as to why a downward trend is observed for all three cancer sites across the years of the study? This seems to be quite drastic for lung cancer.
7. Paragraph 2 and 3 of the discussion: I don’t believe the explanation offered regarding differences in the prevalence of cancers is very convincing. It seems that this explanation could also be easily tested by considering a number of cancer sites with varying prevalence. In light of the analysis of the monthly data, a more likely explanation point to the breast cancer campaigns as a stimulating factor for increased internet activity for this cancer. Differences in the extent and intensity of campaigns across the three cancer sites may be a plausible explanation. There are also several important differences among the three
cancer sites that may potentially be significant in interpreting the results. For example, there are more knowledge and medical interventions available for breast cancer than for the other cancer sites, including genetic susceptibility testing, early detection methods, and primary preventive measures. Thus, some of the searches may reflect information seeking on these areas which may not have been captured in studies of internet use among patients with breast, prostate and lung cancer as the authors cite. Of course, these explanations remain speculative but I believe the authors can expand their discussion in this area.

8. Is it really accurate that lung cancer affects an older cohort of patients than prostate cancer?

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable

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