Reviewer's report

Title: Celecoxib decreases prostaglandin E2 concentrations in nipple aspirate fluid from high risk women and women with breast cancer

Version: 1 Date: 6 July 2006

Reviewer: Margaret Wrensch

Reviewer's report:

1. The authors present relevance and importance of the research question. The authors evaluate the influence of taking 400 mg of celecoxib twice a day for two weeks on prostaglandin levels in nipple aspirates of breast fluids and plasma in women at high risk for breast cancer and in women with breast cancer. The research question is relevant because there are few options available for breast cancer prevention among high-risk women and the available options themselves have potentially detrimental side effects. Furthermore, high prostaglandin levels have been associated with poorer survival from breast cancer. Celecoxib is of interest as a potential breast cancer preventive agent because it (1) has been shown to reduce numbers of polyps in people at high risk of colonic polyps, (2) was less likely to have been used by controls than breast cancer patients in one hospital based study, and (3) was shown to reduce numbers of chemically induced breast tumors in rats. Decreased prostaglandin levels are thought to be a causal mechanism or early biomarker of these beneficial effects of celecoxib on tumor or polyp reduction. In a previous study, the authors had not found an effect of 200mg of celecoxib on NAF or serum prostaglandin levels, but the colon polyp study had found an effect for people taking 400mg but not 100mg.

2. Methods:
The general design is a two or three stage study in which prostaglandin levels are evaluated pre-treatment with celecoxib, two-weeks after taking celecoxib, and in the high-risk women, but not the women with breast cancer, after another two-week wash out period. Serum was collected from each woman at the two or three time points. High risk woman contributed NAF from one breast for three time points, while women with breast cancer contributed NAF from both the affected and unaffected breast (prior to surgery) at two time points.

a. A study strength is that clear criteria are provided for definition of high-risk and inclusion/exclusion criteria are clearly stated. However, a confusing part of the methodology is including results from the unaffected breast from women with breast cancer in with the results from the high-risk women. This leads to analytic difficulties in comparing the prostaglandin levels in NAF from high-risk women versus in those with breast cancer because the measures from unaffected and affected breasts in the same woman are not independent (this lack of independence applies to 25% (9/35) of the “high risk” subjects. The analytic methods used do not account for this lack of independence. The authors might consider creating a separate category for these observations.

b. Also, while the Wilcoxon rank sum test is appropriate for comparisons of two independent groups (e.g. premenopausal vs postmenopausal), the Wilcoxon signed rank test is appropriate for comparisons of paired samples (e.g. pretreatment vs posttreatment). The authors state that they used the Wilcoxon rank sum test for both situations. The authors also need to provide references for the statistical tests and data analysis software used. Did the authors try data transformation that may have facilitated use of parametric methods?

c. The lack of a placebo group also limits study findings, which the authors need to acknowledge. However, the authors do present the study as one that will be used to motivate further studies and not to recommend celecoxib as a preventive therapy based on these study results.

d. This reviewer is not sufficiently familiar with the prostaglandin assay to evaluate its suitability.

e. How was subject compliance assessed or monitored?

3. Results:

a. The methods state that 27 women were eligible and enrolled out of 158 screened women, this should be stated in the results rather than the methods and reasons for ineligibility summarized. Why are 26+13=39
women given in the description of results on page 8 and in Table 1? The methods and results need to be clarified to explain or rectify this discrepancy. Some additional results also are inappropriately given in the methods, such as the number of women from whom NAF was successfully collected from the same breast; this information could be provided in Table 1.

b. As mentioned above, the p-values given in table 2 probably are not correct because they do not account for the paired nature of the samples before and after treatment or for the overlap of subjects in the high risk and cancer groups. The same criticism applies to the p-values for paired comparisons in tables 3-5.

c. Although the authors consider the effects of menopause on the PGE levels, they don’t consider the effects of age. Was this examined?

d. While it is clear that the NAF levels are higher than plasma levels, the authors do not indicate whether they are correlated; a scatter plot of NAF vs plasma levels might be helpful.

e. The two figures are not cited in the text. Also, the figures could be more illuminating; each woman’s values could be plotted for the two or three periods and connected by a line; thus the graph would be a collection of lines; there maybe some women who show an effect of treatment and others who do not. As it is, the figures provide only a little information above that provided in the tables. And as noted above the authors might consider transformation to normalize the data. Figure 1 panel B does not support the idea of a significant decrease in NAF or serum PGE levels after treatment. Figure 2 panels A and B show a tendency toward less variability in values in time period 3 for the high risk group and in time period 2 for the cancer group; but this could be due to either loss of women with the more extreme values, the effect of the drug, or some other reason. Another complexity is that while the timing of specimen collections from high risk women was the same (14 days after treatment and after washout), the timing of collections the women with cancer was quite variable 10-24 days. Perhaps a plot of PGE levels by time since cessation of the drug might be help to clarify this problem.

---

**Major Compulsory Revisions** (that the author must respond to before a decision on publication can be reached)

**Summary of compulsory modifications**
1. Use analytic methods that account for the lack of independence of observations in paired comparisons (pre vs post treatment) or in comparisons of affected and unaffected breasts in the same women.
2. Use graphical methods appropriate to the study design and cite figures in results section.
3. Consider data transformation that could allow use of parametric methods, which would then allow multivariate modeling of PGE levels with age, menopausal status, study time period, etc.
4. Explain discrepancies in numbers of women enrolled as stated in the methods and in the results sections.
5. Examine how subject compliance assessed or monitored.
6. Examine and present correlation, if any, of NAF and plasma PGE levels.

---

**Minor Essential Revisions** (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

---

**Discretionary Revisions** (which the author can choose to ignore)

---

**What next?:** Unable to decide on acceptance or rejection until the authors have responded to the major compulsory revisions

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

**Quality of written English:** Acceptable

**Statistical review:** Yes

**Declaration of competing interests:**
I declare that I have no competing interests.