Reviewer's report

Title: Estimation of Hospital Emergency Room Data Using OTC Pharmaceutical Sales and Least Mean Square Filters

Version: 1 Date: 5 December 2003

Reviewer: Fernando Pineda

Reviewer's report:

General

The manuscript lends further support to the observation that OTC data is correlated with ER clinical data. The mathematical analysis appears correct. This is an interesting, but incomplete manuscript that appears to have been rushed to submission. The mathematical analysis is sound, but it does not support the conclusion that this might be useful as an EARLY indicator of human disease, thus the conclusion is unsound. The research and the manuscript needs more work before it can be recommended for publication.

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

(1) The manuscript would benefit from a more careful and thoughtful presentation of the previous work and the context. The description of previous work is inadequate and poorly referenced. There is nothing indicating how this work extends previous published and unpublished work (e.g. some previous work supporting a relationship between OTC data and ER visits is provided in "A systems overview of the Electronic Surveillance System...", which has S. F. Magruder as a co-author, but this is not mentioned).

The references themselves are incomplete. Indeed, only one of the references has an author list -- and that one is incomplete! Key statements are unsupported by references, e.g. 5th paragraph on p. 2 states: "Product sales from some of these product groups are known to be good indicators of the corresponding clinical data...". But these statements unsupported by data or a reference.

(2) Nowhere is there a description or overview of the primary data. This is a critical flaw in the manuscript. The only mention is in the first paragraph of the "Results and Discussion section", which states: "We have used emergency room visit data..."

(3) Is the FIR filter any better than a naive predictor?, e.g. how much of the variance would the following naive predictor account for: Tomorrow's ER data will be the same as today's ER data. If the FIR filter does not do significantly better than this, then it is highly questionable whether the OTC data will be useful as an EARLY indicator.

(4) I believe the authors have shown that the OTC data correlates with the ER data, but in the absence of comparisons to a naive predictor or some kind of robustness analysis, there is no evidence that there is any temporal predictive ability in the OTC data. (It's even hard to believe just one day ahead, since the OTC data from the same day is used to predict the ER data).

Granted, the authors don't claim any predictive power, on the other hand, they conclude that: "These results tend to strengthen the hypothesis that OTC product sales might be used as an early indicator...if we are successful in extending the estimation algorithms...arbitrary time lags." I don't
understand how these results strengthen the hypothesis.

**What next?:** Reject because scientifically unsound

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

**Quality of written English:** Acceptable

**Statistical review:** No

**Declaration of competing interests:**

I worked on this project about 5 years ago while I was employed at the Applied Physics Laboratory (the same organization as the authors) and I am familiar with unpublished results by (myself and others) that is rather similar to those presented by these authors. I know the first author personally, but I do not believe I am biased either for or against these authors.