Reviewer's report

Title: Are hospital discharge administrative data an appropriate source of information to identify incident cases of cancer? Some insights from prostate cancer

Version: 2 Date: 6 March 2009

Reviewer: Heather Gold

Reviewer's report:

This is an interesting manuscript that finds what we'd expect, that prostate cancer can't be found just by looking at diagnoses codes and surgical procedures. The authors may want to suggest other ways to identify prevalent cases other than V codes, for example, by looking for previous prostate cancer surgery or prostate procedures found in discharges prior to the year of cancer diagnosis.

Major compulsory revisions:
1. Needs major editing for language. Certain word usage makes this manuscript hard to understand. For example, in the Results section, the use of the word "Figures" is hard to follow. Do you mean specificity in some places, percentage in others? Very unclear.

2. Should state up front that only looking at surgical procedures might hinder the ability to identify prostate cancer cases. That was my first expectation when I read this paper.

3. Discuss the limitation or implication of not identifying any true negatives in your sample, at least according to the included figure.

Minor Essential Revisions:
1. Word choice/acronym issues: Use the term "registry," instead of PCR. Use "malignant," not malign. What are "record-bases," found on p.5? There are many uses of the word "form" instead of "from." Don't use acronyms, eg, "ZCR" without introducing them first. Avoid acronyms for sensitivity or specificity - awkward.

2. On p. 7, Instead of "All this process was programmed in Stata9 language." write "All analyses were conducted in Stata v.9."

3. Table 1, note at bottom is unclear in definition of confidence interval.

4. Box 1, instead of CIE, use ICD. What are "MC codes"?

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

Quality of written English: Not suitable for publication unless extensively edited.
**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.