Reviewer's report

**Title:** Determinants of sexual activity and its relation to cervical cancer risk among South African women

**Version:** 1  **Date:** 7 July 2007

**Reviewer:** Paul Hewson

**Reviewer's report:**

**General**

This is an interesting paper which examines a number of risk factors for cervical cancer.

-------------------------------------------------------------------------------------------------------------------------------

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

In a number of places, the paper is rather over-assertive. For example, it states that "The South African age-standardized incidence is 30/100,000 per year", whereas this is clearly an **estimate** of the incidence. Another example is that you find no evidence of a trend in time, I can't see anything in your results about that.

My main concern was in terms of the selection of cases and controls - specifically a wider range of hospitals were used than the two treating the cancer. Readers unfamiliar with the arrangements for health care in South Africa need to know a little more about the reasons for doing this, and particularly need to be reassured for example that the cases weren't taking a narrower set of patients than may have been the case with the control hospitals. It would also be useful to know a little more about how they were judged to be "independend of cervical cancer risk, contraceptive use and sexual practice" (some background statistics may help here), and some more detail on the matching is needed given the potential for over-matching here.

There's a fascinating result, in that you state age of first intercourse is associated with low education, and then go on to highlight infection risks being due to "higher education level and lower age of first intercourse". I can guess as to what's going on here, but it needs to be clearly explored in the context of the modelling results.

-------------------------------------------------------------------------------------------------------------------------------

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

In places you have used "data were", but in others you have used "data was". I prefer the former.
I'm always curious about forcing binary splits within data when you have a continuum (such as examining first intercourse below 16) - if we look at this continuously perhaps it may be the case that another age is more important here.

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

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