Reviewer's report

**Title:** Cancer diagnosed by emergency admission in England: an observational study using the General Practice Research Database

**Version:** 1  **Date:** 22 April 2013

**Reviewer:** David Goldsby

**Reviewer's report:**

I think this is an interesting article that covers an important topic. While there are a number of revisions suggested they shouldn't be difficult to implement and will make the article suitable for publication.

**Major Compulsory Revisions**

1. Is there a look-back or black-out period used to help prevent the inclusion of cancers that have already been diagnosed? If not, this could contribute to the higher incidence in the early years of the study and explain part of the reduction over the study period. As a sensitivity analysis, the authors could compare the reported rates in say 2004-2008 (all have at least 5 prior years without a cancer diagnosis) with the rates that would have been calculated if only the data for 2004-2008 were available.

2. The authors need to explain the modelling results more clearly and use more stable reference categories. For example, the abstract states that patients aged 85+ years have RR of 9.21. No reference category for the comparison is given (people aged 0-14 years) and the actual reference category used has a relatively small sample size. The latter also applies to the reference category for ethnicity. Reference categories would be better with more events/people to give more precise RR estimates.

3. Figure 1 shows the overall incidence rate is lower than both the male and female rates in 2002-2004. Is this possible? Was some form of standardisation used that hasn’t been described?

4. The results list ethnicity and continuity of care as significant predictors for diagnosis by emergency admission but with “no difference between patient groups”. This is not true, as the group of people with unknown values for both variables have lower rates and more generally there needs to be a difference in order to be significantly associated.

5. From the discussion: “Although the adjusted results are inconsistent, crude analyses showed national variation in diagnoses by emergency admission, supporting previous results [11].” The term “inconsistent” seems inappropriate here and might underplay the importance of the results. If the previous study reported crude national variation then the lack of variation after adjusting for other factors means that emergency diagnosis rates don’t vary by region.
Minor Essential Revisions

1. Table 2 is essentially a repeat of Table 1 with two extra columns and a couple of variables removed. These could be combined into one table, with blank values in the “Adjusted” fields where the variables were not included in the final model.

2. In the table(s), do not list the p values for every intra-variable comparison, just give the overall p value for each variable.

3. Leave the results for referrals before diagnosis out of the table(s) and just describe it as being <5 cases in the text. The “<5” in the table does not mask the actual numbers as they can be figured out via the column totals.

4. Explain why skin cancer cases were excluded. Also, why were they then listed in the results as being among the most frequently recorded?

5. The service use results list the mean and SD for number of consultations, but are these the means among people who had at least one consultation? If 0.46% of people have a consultation then a mean of 1.2 per person seems unlikely.

6. The service use results say that among emergency diagnoses 3.3% had a prior admission and gives n=817, but this is 100% of those who had an emergency diagnosis. The same applies to the non-emergency group. Showing the range from 0 seems unnecessary given that so few had it, perhaps it better to just list the maximum value here.

7. Is the number of prior emergency admissions treated as a continuous variable in modelling? If so, the RR should be listed as 0.31 per emergency admission prior to the first diagnosis.

8. The results list the cancer types diagnosed by emergency admission as being most commonly breast, colorectal or “Other”, but these three groups do not comprise 100% and the “Other” group is not defined anywhere.

9. The third paragraph of the discussion says there were crude differences by the number of consultations, and also differences based on having a consultation. Is this repetition or is there a difference that needs to be clarified?

10. Is there a gold standard for identifying cancer diagnoses? Is it the cancer registry? How do the sources of cancer diagnoses used in this study compare?

11. Give some indication of the representativeness of this sample compared to all cancer cases in the region over a similar period.

Discretionary Revisions

1. The sample are described as “random patients” – perhaps make this “randomly selected patients”

2. A brief description of Read codes would be informative.

3. The conclusion talks about “poor quality care” – can this be made more specific?

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.