Reviewer’s report

Title: Baseline Hospital Performance and the Impact of Medical Emergency Teams: Modelling vs. Conventional Subgroup Analysis

Version: 1 Date: 3 August 2009

Reviewer: Sara Brookes

Reviewer’s report:

This paper considers the scenario where the effectiveness of an intervention is modified by another factor, in this instance the baseline measure of the primary outcome. As the authors suggest, this is a relatively common occurrence. Such potential effect-modifiers are often measured and analysed as categorical variables. I believe that the authors are attempting to consider an analytical approach when the potential effect-modifier is a continuous measure as it is here, but the overall aims of the study are not set out clearly which makes the paper a little hard to follow.

Major Compulsory Revisions

I do have some major concerns with the paper. Firstly, the content is fairly theoretical and may be better suited in a statistics journal. The authors provide little detail of the theory, which may be because of the intended audience, however, this makes it quite hard to follow in places.

They give some more detailed information in the discussion, for example about quadratic and linear interactions – this requires more information in lay terms as to how each can be interpreted and should be earlier in the paper to aid interpretation of both the methods and results.

“Conventional” subgroup analyses

My main concern with the paper is that it is not clear what comparison they are making. If I understand the paper correctly the authors are comparing a formal test of interaction with a continuous variable with subgroup-specific tests using a dichotomised version of the potential effect-modifier. I disagree that subgroup-specific tests as the authors describe remain the “conventional approach to examining for potential effect modifiers. Such tests are widely acknowledged to be misleading in terms of producing false-positive and false-negative results and most importantly involve no direct comparison of differences between the subgroups. I would say that “conventional” subgroup analysis would be to perform a formal test of interaction, whether that be with a continuous or categorical effect-modifier.

In any case, the inherent problems of such tests mean that it should not be a comparison method to use here as it is not testing the same thing. If the paper simply sets out to demonstrate that a formal interaction test using a continuous measure is better than subgroup-specific test then this has already been done on
a theoretical and applied level many times.

I do not understand why the authors have used change in incidence rather than incidence at follow up as their primary outcome. They state in the discussion that this is for ease of interpretation but I do not see how it is any easier to interpret than a score at follow up - performing an ANCOVA as they mention would be a superior test. A model containing change and baseline incorporates the measure at baseline twice.

If this is a misinterpretation of what the authors have set out to demonstrate then the paper needs far more clarification and detail of the overall objectives and methods used.

**Level of interest:** An article of limited interest

**Quality of written English:** Acceptable

**Statistical review:** Yes, and I have assessed the statistics in my report.

**Declaration of competing interests:**

I declare that I have no competing interests.