Reviewer's report

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

Version: 1 Date: 16 July 2009

Reviewer: Thomas Lumley

Reviewer's report:

This paper describes modeling of interactions as an alternative to subgroup analysis in situations where the treatment effect may be different depending on baseline variables. In their example, the MERIT trial of the introduction of Medical Emergency Teams, the linear regression approach and the associated graphs show a fairly strong dependence of treatment effect on the baseline incidence of various outcomes. That is, hospitals with a high baseline incidence are the only ones to benefit from treatment.

Major compulsory revisions

1. The main conclusion in the abstract, that linear regression analysis is more sensitive, is probably true but a single case study, while interesting, cannot provide much evidence for this contention. Placing this as the main conclusion of the paper is seriously overstating the evidence. Similarly, there appears to be no evidence at all presented for the qualification “in particularly when the number of units included in the analysis is small” to the conclusion.

2. The Discussion claims that “This approach preserves the rigour of a trial design and does not represent a subgroup analysis, as it uses all data to draw conclusions”. I do not see that there is any qualitative distinction between this approach and a subgroup analysis. Both approaches can be represented as fitting a regression model with an interaction: a quadratic interaction in the proposed approach and a step-function interaction in for a subgroup analysis. The subgroup analysis has the advantage of being easier to pre-specify – the proposed regression approach in this paper involved using a test for interaction in order to decide what model to fit, a sequential approach that makes formal inference more difficult (for example, the shaded confidence intervals in Figures 1 and 2 do not incorporate uncertainty due to model selection). The real advantage of this approach over subgroup analysis is that it allows for smooth linear or nonlinear relationships with baseline variables, and these smooth relationships are often more plausible.

3. The Discussion also claims that the straightforward comparison of randomized groups in the MERIT study “implied an untested assumption: the treatment effect was not influenced by the baseline incidence of the study outcomes.” Comparing
the outcomes in the two treatment arms of a randomized trial does not make any untested assumptions; it is an unconditionally valid analysis. In the case of the MERIT study, the main-effects analysis shows that any benefit of Medical Emergency Teams is small on average. The interaction analysis augments this finding by suggesting that the benefit may still be large in a few hospitals, those with high rates of the outcomes. This does not make the original conclusion invalid. There might well be other important interactions still to be found, including interactions involving variables that were not measured. If it were necessary to assume that no such interactions existed, it would be impossible to analysis any randomized trial. It is entirely reasonable to argue that using an interaction with baseline incidence is a more useful analysis of the MERIT trial, but it is incorrect to claim that omitting this interaction from the analysis requires the assumption that no interaction exists.

**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