Reviewer's report

Title: Early identification of adult offenders in high risk child psychiatric in-patients: A 30 years follow up study using official crime records.

Version: 2 Date: 14 December 2010

Reviewer: Morten Hesse

Reviewer's report:

The manuscript is a register based follow-up of children treated for psychiatric disorders at a single centralized unit in Norway. The study addresses important questions, such as the relative contribution of conduct disorder, attention and hyperactivity symptoms, and intelligence.

The methodology appears to be sound, except for the two issues described below.

Major essential revision:

Please provide substantial more data to give us an understanding of the process of giving a diagnosis based on the file review. For instance, what types of information was considered critical for each diagnosis? What kinds of information were used to exclude a diagnosis?

The single diagnosis that is best described, mental retardation, could benefit from having some examples given for types of tests, and whether other information was available, such as details of school performance.

For other diagnoses, very little information is given. Please note that a BMC article has no word limit.

The authors have chosen to give diagnoses as mutually exclusive categories. For conduct disorder and ADHD/hyperkinetic disorder, this may give face value because it relates to the question of an interaction between the two. However, it does give some curious results, such as when the authors analyze the impact of having mental retardation only, controlling for mental retardation. It would be more informative if the authors made 7 different dummy-coded variables, one for each diagnosis into the regression analysis. In a second step, interactions between variables of interest could be included (e.g., ADHD and conduct disorder, which may or may not interact). The exploratory nature of forward stepwise regression means that it should not be used for a study that is hypothesis testing, such as this. Forward stepwise, as well as backwise, capitalizes heavily on chance, and produces wrong coefficients, standard errors, and estimates of overall prediction (see http://www.stata.com/support/faqs/stat/stepwise.html, for number of pertinent points, with numerous references).

In contrast with stepwise regression, forcing an interaction into the model after first assessing the model without the interaction does not have these associated
problems. If this second step was significant, it would indicate that the authors found that the evidence from their study supported that it is the combination of CD and ADHD that sets individuals at risk for crime. If not, they would reject it in their sample. This would add an important point to their discussion.

Minor essential revisions:
Please give a reference to the ICD-10.

Please provide one or more references to the validity of the use of file-based diagnoses (I could think of one: Journal of Personality Assessment, 70(3), 1998, 416-426, Grann and colleagues, but there must be more than that out there).

Discretionary revisions
Personally, I would also like to see if entering a CD*family disturbance interaction had anything interesting to say. Should the authors choose I suggest that each interaction is entered separately, because if multiple interactions with the same variable are entered in several interactions, it is likely to cause spurious associations and cloud the overall impact of that variable taken alone.

Level of interest: An article of outstanding merit and interest in its field

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.