Reviewer's report

**Title:** Reading and writing difficulties in adolescence and risk of health related work exclusion: a cohort study

**Version:** 1  **Date:** 26 April 2011

**Reviewer:** Torbjorn Torsheim

**Reviewer's report:**

This is a well-written manuscript on a central topic. Identifying modifiable risk factors of long term sick leave is a central objective in many countries.

The main strength of study is the design, with record linkage of a large epidemiological study community cohort with health service register data.

**Major compulsory revisions**

1. The main weakness of the study is the measurement and conceptualisation of reading and writing difficulties. While the current measurement likely captures some true variance, there might be substantial measurement error. The clinical validity is questionable, and misclassification is probable. While it is easy to understand that clinical classification is non-feasable in a large epidemiological study, clearly some assessment of measurement accuracy is needed. On p.12 the authors discuss the measurement, asserting that the current measure obtained a base rate that is similar to other studies. Obtaining a plausible base rate is valuable, but clearly insufficient to demonstrate the quality of the indicators on the level of individual respondents. The authors should also pay attention to the problems of combining reading difficulties and writing difficulties together, as these might occur in isolation.

2. The fact that reading and writing difficulties were assessed in students belonging to different age cohorts warrants more attention, as strategies for dealing with RW might differ substantially across school grades and thus threaten the comparability of RWD across age groups. This is underscored by the observation that 70% of the RWD group was middle school attendees, whereas only 50% of the non-RWD group was middle school attendees. Controlling for age might help to remove selection effects, but does not remove lack of comparability.

3. Potentially mediating factors should receive a more explicit treatment throughout the text. In the introduction section, as well as in the methods section, the authors do not differentiate between confounders and mediating factors, while the causal status of these two classes of variables clearly differs. The discussion section includes a treatment of mediating factors, and rival explanations for the association, including working ability, mental health and completed education. The analysis plan seems to focus on isolating the effects of mental health, since mental health is added in a separate block. However, the
rationale for testing mediation procedure was not clearly stated.

4. The choice of analysis needs further explanation. In figure 1 an age by RWD effect is depicted, with increasing differences across age. The figure implies a longitudinal trend, yet the modelling approach did seem to include a longitudinal dependent measures. Rather, a cross-sectional logistic model on any (?) long term sickness leave during ages 24 to 28 was specified. The author should provide arguments to support their chosen analytical approach, and why the window of ages 24 to 28 was selected. Further, it is unclear why siblings were analysed separately.

5. The concluding arguments should be more balanced. Based on relatively weak associations, the importance of targeting reading and writing difficulties seems overstated, especially since the study does not identify plausible pathways that mediate the impact of RWD.

**Level of interest:** An article whose findings are important to those with closely related research interests

**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.