Reviewer’s report

Title: The predictive value of mental health for long-term sickness absence: The Major Depression Inventory (MDI) and the Mental Health Inventory (MHI-5) compared

Version: 3 Date: 3 May 2013

Reviewer: Niels Smits

Reviewer’s report:

This paper compares the predictive utility of two mental health inventories. One criterion is long-term sickness and the other true mental health status.

Major Compulsory Revisions

I have a major problem with this paper. Whereas they do their best to collect data on the first criterion, long-term sickness, there is no data on the true mental health status. Although in many applied studies researchers do so (as cited by the authors), inventories cannot be used to estimate prevalence as clearly shown by statisticians, for example, Hand (1987) and McNamee (2004) and also Spee et al. (2012). To estimate prevalence the gold standard must be used. The gold standard is the best diagnostic test available, in case of mental health this would be something like the MINI. Screeners are quick and dirty proxies of the gold standard, which may be very valuable for detecting patients at risk, but which must not be used in place of the gold standard. Moreover, screeners may be used to make an efficient guess of prevalence when used together with the gold standard in two-phase-sampling: the screener is used to split a sample up into a low and a high risk group, and a subsample is taken from each for administering the expensive gold standard. In each group the proportion of positive outcomes is calculated and are reweighted to obtain an overall prevalence estimate. The message is that the gold standard must be used for at least a subsample.

As can be read in many textbooks on clinical measurement such as Kraemer (1992), the proportion of positive outcomes on a screener/test is called the level of the test. The proportion of positive outcomes on the gold standard is prevalence, and the level cannot be used to estimate prevalence! In addition, the level varies with the threshold used on the screener, whereas prevalence is constant. One may argue that using positive and negative predictive values and the level (Q) and 1-Q, prevalence P can be estimated because

\[ P = \text{PPV} \times Q + (1 - \text{NPV}) \times (1 - Q) \]

For example, in North Denmark P, PPV, NPV and Q are known, and in South Denmark only Q is known using the screener; it may seem as if prevalence in South Denmark can be calculated by just entering Q in the previous equation and assuming the relation between test and screener in North and South are equal. But for it to work, NPV and PPV need to be the same in both parts of the country, and this is only true if prevalence in South, the thing your tried to estimate, is equal to prevalence in North! In short prevalence
estimation cannot be performed using a screener only. Moreover, it is trivial that different screeners with different cut-offs give different levels of the test.

For the paper to be published, the whole prevalence estimation part should be deleted, if no gold standard is used as criterion. Alternatively, it should be just a subanalysis in which `prevalence estimation' is replaced by the term `level of the test'. This level shows how often the tests would flag respondents as being at risk.

Minor Essential Revisions

In the introduction, the authors talk of just `prevalence'. Be more specific about the disorders the screeners are used for to detect.

In the paper it says `Error! Reference source not found' at several places.

The smoking groups are not exhaustively categorized. What about an occasional smoker with one or two cigarettes a week?

References


Spee, H and Smits, N and De Koning H. The usefulness of the Kessler Psychological Distress Scale (K10) for estimating the prevalence of depression and anxiety disorders. TSG Tijdschrift voor Gezondheidswetenschappen 2012; 90: 145-149.

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