Author's response to reviews

Title: Length of sick leave - Why not ask the sick-listed? Sick-listed individuals predict their length of sick leave more accurately than professionals

Authors:

Nils Fleten (Nils.Fleten@ism.uit.no)
Roar Johnsen (Roar.Johnsen@medisin.ntnu.no)
Olav Helge Forde (Olav.Helge.Forde@ism.uit.no)

Version: 3 Date: 21 September 2004

Author's response to reviews: see over
Dear Editors

MS: 1758495353377884  
Length of sick leave - Why not ask the sick-listed? Sick-listed individuals predict their length of sick leave more accurately than professionals.

We have revised the manuscript according to the comments of the reviewer on the following pages. Most of the suggestions are accepted, but there are a few proposals where we still have some reluctance making the changes.

We hope that the changes made, increases the clarity of the manuscript.

All authors have read and accepted the changes.

Yours sincerely

Nils Fleten
Comments and revisions according to review’s by Harald Reiso.

We think some of the confusion on the terms diagnostic accuracy and diagnostic validity might be a result of different interpretation of diagnosis. According to our understanding the identification of a group of individuals shearing a common diagnosis (here long term sick-listed) or prognosis, here the length of ongoing sick leave, is a diagnostic procedure, and the ability to identify persons that will have a defined length of sick leave is a matter of correctness of diagnoses or prognosis, or diagnostic accuracy.

The length of sick leave predicted by National Insurance collaborators and the sick-listed themselves are the diagnostic tests and the accuracy of these tests might be described with ROC area, sensitivity, specificity, and likelihood ratio according to the STARD statement[1].

The prognostic value, as the diagnostic value, of the tests (predicted lengths) is presented by sensitivity, specificity, likelihood ratios and positive predicted value (PPV) of different thresholds, (cut-offs) of predicted length. The terms diagnostic and predictive validity seems not to be clearly differentiated in the literature, and are synonymously used. We accept that predictive validity might be easier to interpret, as the question is how well the length of sick leave can be predicted.

To clarify the use of terms, we have adapted analyses of the characteristics of the tests to the terms in the STARD check list [1], and changed diagnostic validity to predictive validity discussing the practical usefulness of the tests.

We have tried to stress in the text if cut-offs refers to thresholds in observed or predicted length, and hope this make the use less confusing.

The first paragraph on page 9 presents the material, and it might be discussed if it is more properly placed under Methods, and were moved to the Results according to reviews by Kurt Svardsudd. We find 95% CI here no more relevant than in presenting the number of males and females in the material with 95% CI.

We have followed the suggestion to use medical consultants throughout the manuscript. However, we think the use of officers clearly refers to National Insurance Office (NIO) officers. NIO officers would be preferable to insurance officers, as the later might probably introduce confusion.

We still find that the table 4 is important, giving predictive validity of identifying those sick-listed for at least half a year, as this is relevant for daily work in the National Insurance Offices.
The manuscript is accordingly revised:

- Objective in the abstract: …, the objectives of this study were to inquire the diagnostic accuracy of length of sick leaves as predicted in Norwegian National Insurance offices, and to compare their prediction with the self-predictions of the sick-listed

- Last sentence in Abstract Methods to: Diagnostic accuracy of the predictions was analysed by ROC area, sensitivity, specificity, likelihood ratio, and positive predictive value was included in analyses of predictive validity.

- Last sentence in Background p 5 to: …, the objectives of this study were to inquire diagnostic accuracy of prediction within the NIOs, and to compare their prediction with the self-predictions of the sick-listed.

- Observed length of sick leaves p 7: The reference standard lengths of individual sick leaves were……

- Statistics p 7: The diagnostic accuracy of predicted lengths was compared on the basis of sensitivity, specificity, likelihood ratio and the area under the receiver operating characteristics curves (ROC area) [6,7]. The non-parametric standard error and 95% CI for the ROC area were calculated in SPSS-11. The ROC curve represents plots of the true-positive rate (sensitivity) and the false positive rate (1 – specificity) at the average of two consecutive categories of the assessments (>= 0 weeks, >= 4 weeks, >= 8 weeks etc). The ROC curves of the mean assessment by NIO officers and medical consultants include even intermediate points representing half categories. The predictive validity is presented as sensitivity, specificity, positive predictive value (PPV) and likelihood ratio at different thresholds, cut-offs, in predicted length [8]. Reliability of predicted length was analysed with agreement between assessors, the kappa value [9,10].

- Second paragraph p 10: Changing the observed length to be identified from 12 weeks to 8 or 26 weeks did not significantly change the diagnostic accuracy as assessed by the ROC area.

- The figures on p 11: Out of every 1000 sick-listed persons, 333 will be sick-listed for more than 12 weeks according to the prevalence in this study.

Reference List