Reviewer's report

**Title:** Outcome Measures of a 6 Minute Walk Test: Relationships with Physiologic and Computed Tomography Findings in Patients with Sarcoidosis

**Version:** 4  **Date:** 12 May 2010

**Reviewer:** Giorgio Bedogni

**Reviewer's report:**

I was asked to review this paper as BMC statistical referee.

I read version 4 of the manuscript together with a reply to Dr Holland.

**Major Compulsory Revisions**

Depending on the outcome (DSP or distance), you modeled from 10 to 11 variables per patient. As a *truly best case* scenario you would need at least 10 continuous events per patient to do this modeling with adequate power. This means from 100 to 110 patients. However, because you are modeling 2 outcomes, theoretically the number of patients should double and become 200 to 220.

While I *fully* understand the difficulty or recruiting patients with a rare disease such as sarcoidosis even at highly specialized tertiary centers as yours, you must consider this limitation of your study.

Ideally, one should reduce the list of predictors to a clinically relevant and statistically manageable one. In a *truly best case* scenario that would mean 6 predictors. However, you have already performed the analysis and this decision cannot be taken post-hoc. Thus, my suggestion is to perform stepwise regression using backward elimination and bootstrap selection of predictors. Bootstrap will also allow to calculate 95% confidence intervals for the determination coefficient and for the root mean squared error of the estimate.

Although this will *not* increase power, I expect that it will protect somewhat from spurious associations. You should clearly write that the low sample size you studied (relative to the number of modeled variables) makes it mandatory that your prediction models be validated in external populations (besides the obvious fact that external validation is the only gold-standard for any prediction system).

Please, report regression models for ALL variables, even if they are not significant. The tables should report: regression coefficient, standard error of regression coefficient, true p-value (NOT just < 0.05 or > 0.05), determination coefficient and root mean square error of the estimate for both univariable and multivariable analysis. Using bootstrap will allow to calculate confidence intervals for these metrics.

The standard errors of the regression coefficients that you give in the text are
*surprisingly* high. They are about 20-30 times higher than the regression coefficients. Of course, I suppose that with “beta” you mean “regression coefficient” and with “SE” you mean the standard error of the regression coefficient and there is no typing error. If this is the case, this estimate is very unstable. Please, comment on this very important point.

Minor Essential Revisions

Please, consider that the effects that you attribute to treatment may be due to unknown confounders.

Discretionary Revisions

Please add page numbers.

**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