Author's response to reviews

Title: A predictive tool for an effective use of 18F-FDG PET in assessing activity of sarcoidosis

Authors:

Rémy LM Mostard (r.mostard@atriummc.nl)
Sander MJ van Kuijk (sander.vankuijk@maastrichtuniversity.nl)
Johny A Verschakelen (johny.verschakelen@uzleuven.be)
Marinus JPG van Kroonenburgh (m.van.kroonenburgh@mumc.nl)
Patty J Nelemans (patty.nelemans@maastrichtuniversity.nl)
Petal AHM Wijnen (petal.wijnen@mumc.nl)
Marjolein Drent (m.drent@hetnet.nl)

Version: 3 Date: 27 June 2012

Author's response to reviews: see over
Dear Editor,

First of all we would like to thank you for giving us the opportunity to resubmit a revised version of our manuscript entitled: 'Clinical prediction of 18F-FDG PET activity in sarcoidosis' (MS number 2132005557097815; New title: A predictive tool for an effective use of 18F-FDG PET in assessing activity of sarcoidosis). Before we go in detail on the reviewer’s comments, the authors first want to thank both reviewers kindly for their critical look, constructive comments and suggestions. We addressed all points raised by the reviewers and changed the manuscript accordingly, if appropriate. Enclosed you will find our revised manuscript. In the revised version we incorporated the comments and suggestions of reviewer 2 dr. Francesco Bonella.

We hope that the revised manuscript fulfils your requirements now and that you will consider this manuscript for publication in BMC Pulmonary Medicine.

Awaiting your reply,
Yours sincerely,

Prof. dr. M. Drent, MD, PhD, professor of interstitial lung diseases
Faculty of Health, Medicine and Life Science, University Maastricht
PO Box 3100, 6202 NC Maastricht, NL
Phone: +31 43 3882087
email: m.drent@maastrichtuniversity.nl
Reviewer: Giorgio Treglia

Reviewer's report:
The Authors performed a valuable study about the use of FDG PET in patients with sarcoidosis. The article is original, interesting, well-written and suitable for publication in the journal.

We appreciate the reviewer’s positive comments.
MS number: MS number 2132005557097815
MS title: Clinical prediction of 18F-FDG PET activity in sarcoidosis. New title: A predictive tool for an effective use of 18F-FDG PET in assessing activity of sarcoidosis.
MS authors: Mostard et al.

Reviewer: Francesco Bonella

First of all we would like to thank you for your constructive comments. We have addressed all the points raised in your comments to improve our manuscript.

Reviewer’s report:
General comments
The manuscript by Mostard et al reports on their experience with 18F-FDG PET in patients affected by sarcoidosis with persistent disabling symptoms. The authors found that applying a prediction model based on serum IL-2R and HRCT, it was possible to identify patients that need 18F-FDG PET to assess residual or occult inflammatory activity in sarcoidosis. The manuscript is well written and of certain interests, above all considering that the role of 18F-FDG PET is still under debate.

There are some concerns that need to be addressed.
Major concerns
C1 A recent study of Mostard et al (Ref 8 in the paper) demonstrated that the combination of sIL-2R and Neopterin yielded a sensitivity of 80 % and a specificity of 100% for detecting inflammatory activity as shown by PET. The reported PPV and NPV for the combined biomarkers (CMI) were 100% and 65% respectively. Neopterin or serum IL-2R alone showed both a weaker association. Why did the authors exclude Neopterin, or the combination of these 2 biomarkers from the prediction tool’s calculation? Moreover, did the authors perform ROC analysis with the combination of the two biomarkers (Neopterin and sIL-2R) and HRCT score to predict PET positivity? In my opinion it is an interesting point. If they performed that, which Se, Sp, PPV and NPV for PET positivity have been found?

As stated in the ‘potential predictors’ section, because of the limited number of patients with a negative PET scan (the least frequent outcome in this study) and the usual recommendation to include one predictor for at least ten events, we had to select the two predictors with the strongest associations with PET-positivity. Based on the results of recent studies, the following clinical characteristics were selected in view of their association with PET-positivity: elevated serological inflammatory markers (sIL-2R and neopterin), HRCT abnormalities as assessed by the HRCT scoring system, and lung function tests. Of the serological inflammatory parameters, positive sIL-2R had shown the strongest association with PET-positivity in the previous study (Ref 8 in the paper). This made us decide to include sIL-2R instead of neopterin. Obviously, this decision was also made as neopterin values were missing in 45 (47.4%) of the studied patients. Moreover, we would like to stress that in general, measurement of sIL-2R levels is more accessible in clinical practice compared with neopterin.

The inclusion of only a limited number of tests increases the compliance, the willingness to apply it and the cost-effectiveness of a clinical prediction rule. Furthermore, we preferred to include a non-serological parameter as second predictor in the clinical prediction rule. Lung function tests had no superior predictive value compared to the HRCT scoring results, so it was decided to include the total HRCT score. Thus, two potential predictor variables (sIL-2R and total HRCT scoring results) were included in the predefined model and the model according to the needs of the methodological standards for clinical prediction rules.

To clarify this point, we have added this sentence to the end of the ‘potential predictors’ section: ‘Neopterin was not added as predictor since neopterin values were missing in almost half of the studied patients. Moreover, from a practical point of view sIL-2R also is preferable considering that in clinical practice accessibility to neopterin measurement is less compared to sIL-2R.’
C2 A recent study of Mostard et al (Ref 8 in the paper) showed that 75 % of the patients with persistent disease related symptoms had signs of inflammatory activity as detected by PET. Using the currently proposed internal predicting rule (calculated retrospectively) for PET positivity, PET would be needed in up to 46% of these complicated patients, on the basis of the best possible cut-off. Considering that the 95 patients selected for this retrospective study belong to the original cohort of 122 patients (Ref 8), is this discrepancy in the percentage (75 % vs 46 %) due to the adjustment for overfitting by using the shrinkage factor? Does it mean that about 30 % of the patients that have inflammatory activity (based on biomarkers but not on HRCT score) would not have additional benefit from receiving a PET? If my interpretation is correct, this percentage of patients has pulmonary inflammatory activity that can be assessed by HRCT.

Can the authors elucidate this point?

It should be mentioned that the aim of our recent study by Mostard et al (Ref 8 in the paper) was to assess the presence of inflammatory activity using PET in sarcoidosis patients with unexplained persistent disabling symptoms and the association between PET findings and serological inflammatory markers. Inflammatory activity was considered to be present in case PET positive findings were demonstrated. Using the strategy of combining serological marker testing provided enough information to determine the presence of inflammatory activity in 58% of the studied patients (52/89) in that study. However, since the negative predictive value of serological marker testing was moderate (65%), negative serological test results did not exclude the presence of inflammatory activity and therefore PET appeared to offer added value in assessing inflammatory activity in the subgroup of symptomatic sarcoidosis patients without serological signs of inflammatory activity. The aim of the current study was to develop a prediction rule that can be used in clinical practice to identify symptomatic sarcoidosis patients for whom there is a high probability that PET will show the presence of inflammatory activity. Although there is a considerable overlap in the patient population of both studies, the design of both studies is completely different, since in the current study a multivariable logistic regression model was used. Therefore, the results can unfortunately not be compared in the way that is suggested by the reviewer. In the recently published study of Mostard et al (Ref 8 in the paper), the positive predictive value of both the separate and the combined results of the serological inflammatory markers for detecting inflammatory activity as shown by PET was 100% in all cases. This is in accordance with the high predictive value for PET positivity of a positive sIL-2R result alone in the current study (92.2%). Therefore, this means that patients with inflammatory activity based on positive biomarkers (sIL-2R in the current study and ACE, neopterin and sIL-2R in the published study, ref 8) in general not have additional benefit from receiving a PET for assessment of inflammatory activity. Based on the results of the current study, HRCT appeared to offer added value for assessment of inflammatory activity in the patients with negative biomarkers. Of these patients, only those with a low HRCT score (the cut-off level depends on the positive predictive value that is desired for clinical decision making) would have to be referred for PET.

C3 A comment concerns the title of the paper. I find that the title in the current form is quite ambiguous. The internally validated rule developed by the authors is based on biomarkers and HRCT and predicts PET positivity, and not PET activity. I suggest to re-formulate the title as following: “A predictive tool for an effective use of 18F-FDG PET in assessing activity of sarcoidosis”.

This is a very good point that we agree with. We thank the reviewer for his suggestion and changed the title accordingly: “A predictive tool for an effective use of 18F-FDG PET in assessing activity of sarcoidosis”.
C4 The authors define as “persistent disabling symptoms” the presence of more than one symptom that had substantial influence on quality of life, and that could not be explained from the results of routine investigations, including the absence of lung functional or chest radiographic deterioration. I cannot find either in the text or in the tables, a temporal indication on the persistency of these symptoms. Did the authors use a time range of 6 months or 2 years to define symptoms as persisting? Or can the authors indicate a mean observational time?

In the sentence just before the definition of ‘persistent disabling symptoms’ is stated that the indication for performing PET was the presence of unexplained disease-related disabling symptoms persisting for at least one year. To avoid the potential confusion which the reviewer noted, we have added an additional sentence in the materials and methods section after ‘Persistent disabling symptoms were defined as ...deterioration’: ‘The symptoms had to be present for at least one year.’

C5 Moreover I have some perplexities about the use of the term “unexplained” for the symptoms, as in materials and methods on page 5. The term “non organ specific symptoms” may fit better the definition given by the authors on page 4. I suggest to use the same term all over in the text.

We agree with the reviewer and changed the term all over in the text.

Minor concerns

C6 On page 13, the authors write that “Among patients with a normal sIL-2R level, PET appeared ...”. Can the authors report somewhere in the text or in the table 1 an upper limit of normal for sIL-2R in their lab? In the current manuscript and in the previous study of the same group (Ref 8) only the range of sIL-2R in 40 healthy controls has been reported (240-3154 IU/ml). However, Grutters et al (Chest, 2003) reported a value of 710 IU/ml as upper limit of normal. Do the authors intend, on page 13, patients with a sIL-2R level < 3154 IU/ml? If yes, please clarify this point in the methods.

As mentioned in the materials and methods section, subsection ‘Laboratory and lung function results’: Serum levels of sIL-2R were analyzed using commercially available Diaclone ELISA kits (Sanquin, Amsterdam, The Netherlands) and considered elevated if >3154 pg·mL\(^{-1}\). The different upper limit of normal in the mentioned study of Grutters et al (Chest 2003) was due to the use of another ELISA (DPC, Breda, The Netherlands) for the quantitatively determination of sIL-2R in that study.

C7 In the table 2 the comparison between the original model and the model after internal validation is shown. The p-value for the model after internal validation is not reported. Will the original p value also be influenced by the application of the shrinkage factor (0.93)? If this calculation is possible, please insert the p-value for the new regression.

We have refrained from computing p-values after the internal validation step because no formal testing procedure is performed after this point. The coefficients of the internally validated model are not direct estimates based on the data, but corrected for the likely magnitude of overfitting based on the bootstrap samples. However, it is possible to compute confidence intervals around these estimates, and we have added them to table 2. These can be used to assess significance, if appropriate, and reflect accuracy in these estimates.

C8 The title of the table 3 should be changed, for clarity reasons, as following: “Sensitivity, specificity, positive and negative predictive values for PET activity at consecutive cut-off points of the prediction rule score”.

We thank the reviewer for this suggestion and amended the title of table 3 accordingly. However, in view of the comments that the reviewer made in C3, we replaced ‘PET activity’ by ‘PET positivity’.