Reviewer’s report

Title: Peripheral blood lymphocytes immunophenotyping predicts disease activity in clinically isolated syndrome patients

Version: 1 Date: 20 Aug 2016

Reviewer: Ming-Feng Liao

Reviewer's report:

I think the manuscript is improved and better understood after revision. The results may apply to clinical practice to predict relapse. The author also had intact data with long term follow up and is worth to be published. However, I remain curious about the details of statistical analysis and suggest to clarify the methodology of statistical analysis more clearly before publish. Bellows are my understandings and questions of the statistical analysis methods. If I misunderstand author's study designs, please clarify in the manuscripts.

The author used the survival curves (Kaplan-Meier estimators) to predict disease progression as in Villar's study (Villar, 2011). The cut-off value in the Villar's survival curve is "> 3.5%" (percentage of CD5+ B lymphocytes), which is based on mean +/- 2 standard deviations (SD) of the percentage of blood B lymphocyte (CD5+) of the control group. In this study, the cut-off value in survival curve is chosen by the results of different ROC curves (at base line or different time-points?). The author used three indicators in ROC curve. (a) the population relative value, (b) the population relative value change compared to baseline and (c) the population relative value change compared to a measurement one year before.

1. Is "population relative value" means percentage of different lymphocyte at base line? For example, in Figure 1, threshold of value: 9.05% is gotten by percentage of CD19 lymphocyte (CD19 lymphocyte/total lymphocyte) at base line? Please clarity in the manuscript. As I understood, the author used this threshold or cut-off value (gotten from the data at base line) to check "every time-point a cohort of patients" (ex: in Figure 1, in 12 months, 59 patients' CD 19 lymphocyte/total lymphocyte > 9.05%, in 24 months, 38 patients' CD 19 lymphocyte/total lymphocyte > 9.05%). Is this method reasonable? Can we use threshold gotten by ROC curve from the data at base line to check different time-point? (ex: 6 months, 12 months from base line, et al.) or we need to use different threshold chosen from data from different time-point (ex: in 12M, in 24M, in 36 M or in 48M) in the survival curve of different time-points? Please explain in discussion. What is the meaning of survival curve checked in 48 M? 48 month seems is the end of study.

2. In the indicator of "ratio vs BL", is the thresholds in ROC curve gotten from "every time-point" or only the "ratio (data in 12M) vs BL"? For example, is threshold 1.638 in Figure 1 gotten from the data of ratio (at 12M) vs BL? Did the survival curve of ratio vs BL in 24 M,
36M, 48 M use the same threshold? In Figure 1, total patients number of "value in 48 M" is 181 (21+160), but the total patients number of "ratio vs BL in 48 M" is 165 (80+85). What is the reason of this discrepancy?

3. In the indicator of "ratio vs LY", is the threshold of survival curve different in different time points (in 12M, in 24M, in 36M, in 48M)? How to get the threshold 0.808 in Figure 1? Is the threshold 0.808 gotten from the data of "ration vs LY (12M vs base line)"? In my understanding, the threshold or the data of "ratio vs LY in 12M", "ratio vs LY in 24M", "ratio vs LY in 36M", and "ratio vs LY in 48 M" should be different? Again, if this is reasonable to use the same cuff-off value or threshold in the survival curve in different time-points?

4. The patients received natalizumab, fingolimod, and copaxone were excluded from this study as author mentioned. Please emphasize in the manuscript.

There are some minor questions:

1. In table 1a, 104 patients with relapse at 48 months. However, in result (page 6, line 9), 114 patients experienced the second clinical attack at 48 months. Which data is correct?

2. I think table 2a, b provide similar information as table 1a,b and maybe redundant.

3. Page 9, line 40: 9.5) should be 9.5%?

Thanks for the author's explanations.

Are the methods appropriate and well described?
If not, please specify what is required in your comments to the authors.

No

Does the work include the necessary controls?
If not, please specify which controls are required in your comments to the authors.

Unable to assess

Are the conclusions drawn adequately supported by the data shown?
If not, please explain in your comments to the authors.

Unable to assess
Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I recommend additional statistical review

Quality of written English
Please indicate the quality of language in the manuscript:

Needs some language corrections before being published

Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests except I had a few multiple sclerosis patients.

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal