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

Authors:

Helena Posova (hmare@lf1.cuni.cz)
Dana Horakova (dana.horakova@lf1.cuni.cz)
Vaclav Capek (venca@cicconia.cz)
Tomas Uher (tomas.uher@vfn.cz)
Zdenka Hruskova (zdenka.hruskova@vfn.cz)
Eva Havrdova (ehavr@lf1.cuni.cz)

Version: 2 Date: 24 Oct 2016

Author’s response to reviews:

Reviewer #2: 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.

The ROC curves are constructed from all time points at once. As a predicted outcome we consider an EDSS resp. relapse that appears in (and only in) one year after the corresponding measurement. This detail is important as we are looking at the data backwards – we are explaining what is happening one year before the EDSS resp. relapse.
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.

No, this threshold is taken from the ROC curve as explained above. Then the threshold is applied to measurements before baseline, at baseline, at 6M, etc. Every time it divides patients into two groups. For those groups the survival probability curves are constructed and compared.

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' CD19 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.

This is a misunderstanding, see the explanation above. The threshold is taken from all time points.

The survival curve is drawn, in all cases, from the baseline. From the picture it can be seen that the value at 48M nicely predicts our outcome, but unfortunately too late.

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?

The principle of construction of the ROC curve and the corresponding threshold is the same as described above. It includes all time points at once.

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?

The reason is that we do not have baseline measurements for all patients and so we are not able to compute their ratio to baseline.

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?

The explanation is again the same as above

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

Added:

The results obtained from patients after therapy escalation (natalizumab, fingolimod) and also after changing for copaxone were excluded from this part of our study.

> 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?

Both. 114 patients experienced the second clinical attack, but we have data only from 104 patients at 48 months. The data close to the time of relapse were much more interesting for us, and were used for statistical analysis.

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

I agree with this, but some authors prefer to use absolute number of lymphocyte subsets.

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

Yes, corrected.

> Thanks for the author’s explanations.
Reviewer #3: I reviewed revised version of the above article and I also read other reviewer's comments with interest.

Authors addressed some of my criticisms and they somehow handled NK contradiction in discussion section. Authors emphasized CD56bright NK cells and its functional link to T lymphocytes an chronic inflammation as explanation of NK result inconsistency pointed out in my original review but still mention NK result in abstract results as reliable conclusion of the study.

Excluded from the abstract.

I believe that percentages of patients switched to other treatments (however not included in statistical analysis) should be mentioned in the text and results and not only as a footnote to tables.

Added in the text

In clinical data section authors should be consistent and show numbers but also percentages (after 48 month 114 patients (59.6%)....and (69 the third one).... what was a percentage and mention third relapse (not the "third one"); possibly graph with percentages would be suitable option here.

Added in the text

My original proposition that authors should put accent on naive cells

should be re-considered.

Changes in naïve CD4+ lymphocytes were the most convincing ones in statistical analyses, but during patients’ follow-up (without statistics) we mostly noted changes in NK cells and B lymphocytes. The same experience was obtained in various European centres as we had the opportunity to discuss while presenting the poster (ECTRIMS 2015). These subpopulations are routinely examined and their significance seems to be higher when their changes are assessed together, in each patient separately.

From my point of view article is still not "friendly" to general medical professional and I absolutely agree with the last sentence of revised article.