Author's response to reviews

Title: Clinical and Immunological outcomes according to adherence to first-line HAART in a urban and rural cohort of HIV-infected patients in Burkina Faso, West Africa

Authors:

Emanuele Focà (emanuelefoca@gmail.com)
Silvia Odolini (Silvia.odolini@gmail.com)
Giorgia Sulis (giorgia.sulis@gmail.com)
Stefano Calza (stecalza@gmail.com)
Virginio Pietra (lvia_nanoro@yahoo.fr)
Paola Rodari (rodaripaola@libero.it)
Pierfrancesco Giorgetti (pifagiorgetti@hotmail.it)
Alice Noris (khenet@tin.it)
Paul Ouedraogo (st.camille@fasonet.bf)
Jacques Simpore (jacques.simpore@yahoo.fr)
Salvatore Pignatelli (pignatellis@hotmail.com)
Francesco Castelli (castelli@med.unibs.it)

Version: 3 Date: 13 December 2013

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

enclosed for your evaluation our revised manuscript entitled “Clinical and Immunological outcomes according to adherence to first-line HAART in a urban and rural cohort of HIV-infected patients in Burkina Faso, West Africa” for possible publication as research Article on your estimated BMC Infectious Diseases Journal.

First of all we would like to thank the Editor and Referees for the interesting comments that were very helpful to improve the quality of our work. We were able to address all of their comments as you can see below:

Editor’s comment:

'This manuscript is interesting and I commend Authors for this effort undertaking, However a extensive revision is necessary before it can be accepted for publication. In making this revision, please be patient and respond point-by-point to the very constructive comments and criticisms raised by the referees. This is critical for paper acceptance for publication. Such revision must address conceptual (mainly statistical), style and language problems. In addition to the referees' comments (which I agree upon), please consider the following:

1) Survival (and inferential) analyses need to be better explained. My concern is that adherence of patients who presented later for care (e.g., those with AIDS) was lower than adherence of the fitter patients. If this was the case, increased risk of death was not due to lower adherence levels but to disease severity at baseline. Therefore, correlation between adherence and outcome need to be corrected for stage of the disease at baseline, at least. Moreover, I'm not sure that adherence as a summary measure is the right way to stratify patients in the survival analysis. I understand that patients were ranked by number of visits with optimal adherence among 10. Why 10? Did each patient receive 10 visits? If not (as likely is) it is obvious that patients who died or were lost to follow-up received less visits (probably someone less than 10). So the probability of having optimal adherence should be corrected for the number of visits.

Probably, a stronger (and more convincing) model may be built up by considering adherence as a time-dependent variable together with others either at baseline (including stage of the disease) or during the follow-up.

Response: We agree with the concern raised by the editor. Indeed we evaluated the effect of adherence score as continuous (for 1 point increasing) as a time-dependent variable in a Cox model. A multivariate model was fitted including relevant clinicals significant in univariate analysis (Center, Age, HAART regimen including FDC and WHO HIV clinical stage and Calendar Year included as stratification variables).

2) Conclusions should be somehow less strong. Indeed, the Authors have not validated their methods against others and did not assess whether a different combination of parameters (or additional ones) were better predictors of the outcomes.

R: Thanks for this suggestion: we redrafted our conclusions looking for a less strong and more appropriate results’ interpretation.

3) As underlined by one additional referee who returned his comments later, international literature is ignored in many sections. This seems to make the authors to ‘reinvent the wheel’, many times with weird
consequences, such as validating their own assumptions backwards. Please make a more extensive and updated review of the literature.

R: Thanks for this suggestion. Actually we revised the cited literature looking for a critical update as well as a direct comparison between our results and available data. We would like to remark that our objective was not to validate our score, so we apologize if in the first draft this aspect was not well illustrated.

Referee 1, Calcagno:

Major Compulsory revisions
1) Ethics: I find acceptable what is discussed in the Mat\&Met section provided the authors state (if true, obviously) that the administration of the adherence questionnaire was not an intervention but it was part of routine clinical practice (since year xxxx).

Response: thanks for your comment. We added the following sentence within the Ethics section: “The Study was conducted in compliance with the Helsinki declaration. The Joint Centre Medicale Saint Camille and Centre de Recherche Biomoleculaire Pietro Annigoni Ethics Committee have been duly informed about the ongoing research and have given permission. In fact, the CNERS (Comité National d’éthique pour la Recherche en Santé) is questioned only in case of interventional clinical trials as well as in international or national pharmacological studies, therefore in this case was not informed. Moreover, being a retrospective non-interventional study founded on post hoc analysis of data present in all the patients’ files, and collected only for clinical indications, no written informed consent was asked to patients.” (Methods section, Ethics paragraph)

2) abstract: I would add the prevalence of adherence class instead of mortality rates.

R: thanks for the comment. We revised the manuscript, following the Editor’s comments and considered adherence score as time-dependent variable. Therefore there is no overall definition of adherence and then removed any reference to it in the abstract.

3) I would suggest to expand the methods section further explaining how the test was administrated

R: Thanks for your suggestion: we expanded the methods section, particularly the paragraph “adherence score” in order to better clarify the questionnaire administration (lines 122-133).

4) I think some information/analysis are missing and they can improve the completeness of the paper:
a) numbers of patients lost to follow/up (global/per year/per adherence class) b) number of deaths (global/per year/per adherence class)
c) Maybe I missed it but is the multivariate analysis corrected by calendar year and prevalence of lost to follow up and type of HAART?
d) I know missing data can be many but are the causes of death available?
e) Perhaps listing the third drugs used in the regimens (EFV vs NVP, LPV vs ATV) and the prevalence of tenofovir use may be useful (given the fact that different drugs have different forgiveness and therefore they may be associated with different selection of RAMs):
R: Thanks for your comment. We really agree with the referee regarding all cited points:

a) We do have these data: the cumulative prevalence of patients lost to follow-up is showed in Table 1. Regarding lost to follow-up per year we found the following distribution: 3 in 2003; 9 in 2004; 4 in 2005; 8 in 2006; 11 in 2007; 5 in 2008. (lines 184)

b) We added the global number of deaths in Table 1 (60). Regarding deaths per year we found: 1 in 2003; 26 in 2004; 11 in 2005; 6 in 2006; 12 in 2007; 4 2008 (lines 185).

c) We updated the survival model accounting for calendar year (as stratifying variable) and taking FDCs as answered to the second referee.

d) Unfortunately, the causes of death could not be recorded, as commonly happens in resource-limited settings.

e) LPV was the only PI available in our setting (though IDV was seldom used) and once-daily regimens were not available (Atripla was actually received at the end of 2008). For this reason, we deemed appropriate describing therapeutic regimens (table 1) as 2NRTi+NNRTI and 2NRTi+PI/r (only few patients received unboosted IDV and none was treated with 3NRTIs). Moreover we included “taking a fixed dose combination” among variables, and it resulted significantly related to survival (table 3), demonstrating that patients receiving a FDC-containing regimen were less at risk of dying. Unfortunately, we were not able to evaluate RAMs since we did not have neither viral load nor genotype test.

Minor essential revisions
I would suggest the article to be checked by an English native speaker

R: Thank you for this suggestion. A native English speaker reviewed the article, we hope that now the English may be acceptable.

Discretionary revisions
I have no idea of what F2660

R: “F_{2,260}” represented the F distribution value with 2 and 260 degrees of freedom used in the ANOVA tests for comparing quantitative variables among the three groups. Since we prefer to show our population as only one cohort, in the revised manuscript any such test was removed.

Referee 2, Gebremeskel

This is an important study on ART adherence from a resource-limited country in Sub Saharan Africa where the toll of HIV is high and would add to the limited knowledge base in the region. The novelty of the research is the use of a multi-domain approach to measuring the multi-faceted adherence behavior. The approach seems to potentially be a feasible one. However the major weakness of the paper is the lack of detail on the adherence instrument that was used in the study. Although purported to be a novel approach, the authors have not clearly formulated how the instrument was developed, on what conceptual framework it was built upon, and if it has been pilot-tested or validated before use in this study.

Response: Thanks for your comment. We agree with the referee. Our score has not been validated yet and our aim was not to validate this score as an adherence marker, but only to explore the role of this score in predicting survival and/or CD4+ T-cell count recovery. We therefore redrafted our conclusion looking for a less strong and more appropriate results’ interpretation, also taking into account the editor’s comments. Notably, we expanded the limitations paragraph.
The paper lacks a lot of detail on statistical analysis, and should address in detail how the multivariable models were built and what initial variables were considered for inclusion. Repeated measures analysis would benefit a lot from more detail.

R: Thanks for this comment. We updated the methods section accordingly.

Additionally, the language used sometimes detracts from full understanding and a bit of proof reading might be important.

R: Thank you for this suggestion. A native English speaker reviewed the article, we hope that now the English may be acceptable.

Background:
The background is fairly comprehensive and make a good rationale for why we need to have a novel adherence measurement instrument that captures different facets. The objectives of the study are explicitly spelled out.

Minor: In line 76, ‘multi-method’ approach is not a clear phrase
R: we changed the sentence: (line 77 “Since no measurement strategy can be considered optimal especially when used alone, an holistic approach is currently suggested”)

The sentence starting on line 84 does not clearly capture the conclusion from the reference (11).
R: Thanks for this suggestion: we changed the sentence accordingly (line 87) “Previously conducted studies concerning adherence to antiretrovirals found that a 95% minimum rate of drug assumption is required in order to guarantee a viro-immunological effectiveness of the therapy [16]”

The flow from the background to the aims section could be better structured
R: thank you for the suggestion: we restructured the section accordingly.

Methods. Describes the study population fairly well
Major: However the major weakness of the study is the lack of substantial detail about the adherence instrument (the ‘adherence score’). Even if the components of the score are mentioned, it is not apparent what each component constitutes. There is no information if the score was pilot-tested and/or validated before use in this study.
R: we agree with the referee’s comment. However, our score has not been validated yet and our aim was not to validate this score as an adherence marker, but only to explore the role of this score in predicting survival and/or CD4+ T-cell count recovery. However, following the first referee’s suggestions, we redraft the methods section, paragraph “adherence score” and limitations (lines 122 and 248).

It is important to have the instrument (with the questionnaires) as a supplement.
R: Thank you for the suggestion: unfortunately, in the patients’ files we have a single page for each visit. Therefore, we reported the total score alone, since a specific form to evaluate our questionnaire was not provided. We are aware that this suggestion (as well as some others) cannot be properly answered since the retrospective nature of our study makes it difficult to further investigate available data concerning both clinical and socio-demographic aspects.

One would question if all the five components should be weighed equally and summed up as a composite score.
R: we agree with the referee’s question but also in this aspect we do not have the value of each component of the score. Therefore, the relative weight of each component could not be evaluated. We included this aspect in the “limitation” section of the manuscript (lines 248)

On what grounds were the 0-2 score given for each component?
R: The score assignment was given discretionarily by a medical doctor, As explicated in the Methods section, paragraph “adherence score” (lines 122)

What is meant by ‘proper knowledge of medical prescriptions’?
R: This means “knowledge of drug posology”, and we changed the text accordingly (line 123)

How was lost to follow up being defined?
R: we added this definition in the methods section: a patient was considered lost to follow up if at the end of the study he/she misses 2 consecutive visits without a further contact. For the survival analysis, we assumed that a patient lost to follow up is a non-adherent patient; therefore we assigned 0 to the adherence score (lines 132)

The duration of treatment interruption (‘gap’) is an important aspect of adherence that is associated with virologic outcomes, but does not seem to be incorporated in the instrument. Is there a length of time for a gap to be considered as interruption, or are all ‘interruptions’ equal?
R: we agree with this comment. However we are not able to evaluate treatment discontinuations properly. Following our clinical practice, if patient stopped the treatment, was considered as “0” for the adherence score.

The duration of adherence measurement is variable [15 days, 30 days, and 3 months]. But it seems all the visits are being weighed equally even if the length of assessment is different. For all the components, evaluations that have variable time periods does not make sense as this might impact on ‘estimation’ by the physician. Another important drawback is that the whole instrument relies on the estimation by the medical doctor
R: thank you for this comment: we agree with this remark but we would like to point out that adherence measurement reflects the visit timing according to WHO current guidelines (15 days, 1 month, 3 month etc). Regarding the estimation, it is purely at the physician’s discretion. We are aware that this may be a limitation so we added this aspect in the devoted section accordingly (lines 253).

How were the three adherence categories created?
R: thank you for the comment but we deleted the three classes to avoid some interpretation bias. We considered the score as a time-varying variable so there is no more a global definition of adherence for a patient but rather a time related status. In addition, we considered two different groups (0-7 vs. 8-10 score points) in the linear mixed model and as predictor of death in order to distinguish patients with an optimal adherence versus those with a sub-optimal adherence (results section).

Major: The statistical analysis section is lacks considerable detail. This has to be substantially expanded with regards to how the multivariable model was built and what variables were considered initially for inclusion. Survival analysis descriptions are not adequately addressed. The repeated measures analysis needs further expounding.
R: more details on statistical analysis were added to the methods section, paragraph “statistical
analysis”.

Minor :
Explicitly state that it is a retrospective study
R: following this suggestion we clarify that our study is retrospective (lines 39)

Line 101 could go the background section
R: thank you for this suggestion: we moved the sentence accordingly (lines 96-98)

Results:
The section reports the findings fairly well.
Minor: Line 152 : give proportions of those in stage 2 or 3 (instead of the F-test results)
R: With respect to baseline WHO clinical stage, we considered an early stage (1 and 2) compared to an
advanced stage (3 and 4). We added proportions in table 1.

With regards to CD4 count increase, was baseline taken into consideration for analyses looking at changes
from 6month to year?
R: We confirm that baseline was taken into consideration.

Major:
It is not clear why chi-square test was used to test associations between factors and adherence Apart from
Figure 2 and table 3 , there is not report at all about the survival analysis. needs detail.
R: thanks for the suggestion: as said we deleted the global three adherence classes and so the association
tests. More details on statistical analysis were added to the methods section.

Discussion: Discusses the results in light of findings from other studies, and warns against direct
comparisons with other studies

Minor
Line 215 : these results are not reported in the results section.
R: More details were provided in the results section and all data we discussed were also reported there.

The reference for the HR is not provided in Table 3. Table and Figures
R: we provided the reference for HR in the table 3

Minor . Figure 1This is a confusing graph. It could be better scaled.
R: thank you for this suggestion, we better scaled the figure 1

Minor : Figure 2: the statistical significance of the Kaplan-Meier curve should be indicated
R: KM figure is now removed from the test as no established modification is available for time-varying
covariates.

Major .Table 3 should be expanded to include the other predictors considered.
R: thank you for this suggestion: we improved table 3 by adding the univariate model and considering all
included variables.
Referee 3
Comments:
Title: The subtitle (a longitudinal Cohort study) should be excluded! It is redundant. The word “cohort” is already part of the title and cohort studies must be longitudinal (either prospective or retrospective).
Response: We changed the text accordingly (see title).

Abstract:
“Is known that in developing Countries adherence evaluation is a very challenging task.”
Adherence is a challenge everywhere, in consequence of many different reasons, many times complementary and sometimes mutually reinforcing. There is no a priori reason patients from Burkina Faso should be more or less adherent than, for instance, uninsured Americans or drug dependent people anywhere. Optimal or less than optimal adherence is a function of a complex combination of individual and contextual variables.
R: Thank you for this comment: we added the word “also” before the word “in”, accordingly (line 36)

Developing countries is nowadays a very confusing category. “Low/Middle/High Income countries” should be used here. For instance, in the context of the current European crisis, Greece has a lower per capita GDP than most of the so-called “developing countries”.
R: we thank for this comment but we do not agree. There is no established convention for the designation of “developed” and “developing” Countries or areas within the United Nation system. However, GDP per capita is not a measure of individual income. According to HDI we believe that life expectancy, education and income indicators must be considered in order to define a “low/middle/high income Country”.
However, we changed all the “disputed” definition into “Low-income Countries” as suggested.

“Literacy” is a key sociodemographic variable, which has been fully incorporated into key indicators such as HDI, and should be described as such!
Thank you for this comment: unfortunately we are not able to calculate the education level as in HDI. We considered as “illiterates” those who were unable to read and write (analphabets). Besides, we considered as literates those who had gone to school for at least one year (then able to read and write).

I have no idea what does a “custom” questionnaire mean. Please, use instead a regular concept people could easily understand when they read the abstract.
R: we used the word “custom” to define a “home-made” and non-standardized questionnaire. According to the referee’s suggestion, we deleted the word “custom” from the manuscript.

I must confess I could not follow such system of points that seems to refer to visits instead of measuring adherence to ARVs. Please, rephrase it, clarifying what does such original (?) process of ranking points (per visit?) actually mean.
R: Thank you for this comment. However, as we mentioned in the background section: “WHO usually defines adherence to any treatment as “the extent to which a person’s behavior-taking medication, following a diet, and/or executing lifestyle changes, corresponds with agreed recommendations from a healthcare provider”. Since no measurement strategy can be considered optimal especially when used
alone, an integrated multi-method approach is currently suggested. Therefore, not only punctuality to consultation but also “standard” adherence ARVs measurement such as pill count were considered.

CD4 absolute counts are quite inaccurate biomarkers of disease progress. I could not understand the reason the authors used it. Of course, they have results respecting overall lymphocyte counts, so proportions would be much more useful and accurate, and would not mean to spend a single extra cent.

R: thank you for this comment: we agree that CD4 T-cell count alone is not an accurate marker of disease progression. However, in some settings (such as our setting in Burkina Faso) CD4+ T-cell count may be the ONLY immunological marker available. Therefore we analyzed it.

Please, rephrase:
“patients with higher adherence 51 (Group B p<0.001 and Group C p=0.014) and followed in a urban Center (p<0.001) were associated with lower”. Patients (i.e. concrete individuals) are NOT associated with something. Some of their characteristics may or may not be associated with a given outcome (such as optimal versus less than optimal adherence), so they be more or less likely to adhere to something. Variables and concrete individuals should not confused. Variables are amenable to statistical analysis, whereas individuals are something much more complex than what can be summarized by a small set of variables (in this specific case, a metrics defined by a ranking process, the place where patients have been followed, and a single biomarker!).

R: thank you for this suggestion: we apologize for the mistake. We redrafted our conclusion accordingly.

Do the authors call absolute CD4 counts “Immunological outcomes” or do they include other data not mentioned before in their abstract?

R: We thank for the referee’s comment: unfortunately, no immunological markers (other than absolute CD4+ T cell count) were available in our setting. Consequently, we used CD4+ T-cell count recovery as “immunological outcome”.

“suggesting that our score might be a feasible 56 and suitable tool to easily and precisely monitoring HAART adherence.” This conclusion does not make sense. Something can NOT be based on a given criterion and then back-validates itself. Of course, statistical findings must be based on something, but cannot at the same time validate findings AND basic assumptions. This would be a circular way of defining ways to measure things and doing analysis. Actually, wrong assumptions may generate consistent findings, but cannot tell anything about their validity, so consistent associations between A and B may perfectly match, despite the fact both can be right (or totally wrong). So, in case I use dark green glasses to look at a horse, I may say he or she is green, what is perfectly consistent considering my glasses, but would means I would be finding green horses in nature!

R: Thank you for this comment: following the editor’s comments as well, we redrafted our conclusion looking for a less strong and more appropriate results’ interpretation.

Main Text:
Background
at the end of 2011 – Please update it!
Were WHAT? Reported? “These dramatic data suggest that HIV infection in Africa is a major public health problem whose burden is very difficult to assess and control.” This has been repeated again and again by UNAIDS, the PEPFAR reports, the UNGASS sessions over three decades! Please, delete it or replace it with some concrete/innovative statement.
R: we changed the text in the light of these suggestions by updating all the epidemiological data.

“even in resource-limited settings” Why EVEN? Is there any intrinsic characteristic of people living in such countries that should make them “a priori” impervious to the benefits of medical science?  
R: In accordance with the referee’s comment we eliminated the word “even” (see text)

“adherence monitoring represents a useful marker of HAART effectiveness” Monitoring is NOT a marker!, but rather an active process of tracking and evaluating things, providing feedback for policymakers and health professionals.  
R: thank you for this comment. We reformulated this sentence, as suggested (line 69)

“Despite this, only few studies have been carried on to better assess adherence patterns in poor resource settings” This is by no means true. There are hundreds of papers assessing adherence in such settings. See, for instance, PEPFAR and/or CDC reports, available at: http://www.pepfar.gov/reports/progress/index.htm Or some peer-reviewed papers, available at:  
http://www.ncbi.nlm.nih.gov/pubmed/?term=lowincome+countries+and+ARVs  
R: We fully agree with this comment. Therefore, we better updated the state of art by adding some references. In particular this sentence was modified (line 72)

“suitable methods should be implemented”  
They HAVE been implemented!, See, for instance, the creative papers by W. El-Sadr and her team as follows:  
http://www.ncbi.nlm.nih.gov/pubmed/?term=el-sadr+w  
R: thank you for this comment: we agree with the referee’s comment and we updated the cited references as suggested.

“non absolute adherence is actually a normal behavior among HIV-infected individuals.” What does mean NORMAL here?  
R: thanks for your comment: we eliminated this sentence to avoid some conceptual misunderstandings that would need another paper to be adequately explicated,.  

“Aims of this study are: (1) to assess HAART adherence through a custom score”  
Whatever such “custom” score can mean, it is not an AIM, but rather a tool or means to achieve the actual aim, i.e. to accurately measure adherence in this concrete setting, among these concrete patients.  
R: thank you for this comment, we rephrased the sentence in order to better focalize the objectives’ presentation (lines 93).