Author’s response to reviews

**Title:** Transient detectable viremia and the risk of viral rebound in patients from the Swiss HIV Cohort Stud

**Authors:**

Jim Young (james.young@usb.ch)
Martin Rickenbach (Martin.Rickenbach@chuv.ch)
Alexandra Calmy (Alexandra.Calmy@hcuge.ch)
Enos Bernasconi (Enos.Bernasconi@eoc.ch)
Cornelia Staehelin (Cornelia.Staehelin@insel.ch)
Patrick Schmid (Patrick.Schmid@kssg.ch)
Matthias Cavassini (Matthias.Cavassini@chuv.ch)
Manuel Battegay (Manuel.Battegay@usb.ch)
Guldrych Günthard (guldrych.guenthard@usz.ch)
Heiner Bucher (heiner.bucher@usb.ch)

**Version:** 3  
**Date:** 21 July 2015

**Author’s response to reviews:** see over
We thank the reviewers and editor for their constructive comments and for their encouragement. The changes I've made in response are described below.

Regards Jim Young 22/07/15

Reviewer: David Hanna

Minor Essential Revisions:

1. Abstract: The thresholds used to define “blip” and “viral rebound” should be included in the Methods for clarification. [I assume the Reviewer would like to see these thresholds in the Abstract, in addition to in the Methods].

It is difficult to do this clearly yet concisely in an abstract. These definitions have to be precise because they are absolutely fundamental to the analysis (being the exposure and outcome respectively). Therefore on balance I think it is best to rely on the reader's intuition of these concepts when reading the abstract and then to define these fully and precisely in the manuscript. Most readers will have a good intuitive understanding of a "blip" – and we then provide a precise definition in the second sentence of the Introduction – and all readers should have a good intuitive understanding of "viral rebound".

2. Abstract: The time frame for the study is not presented but should be included. Otherwise, the reader does not know what "newer" is in relation to. This should also be included in either the Introduction or Methods of the main text, as it is never clearly stated.

Again this is hard to do because selection of data for analysis was by assay rather than by time (see Methods, Patient selection). The Reviewer makes another related suggestion: "Abstract: The authors may want to consider describing the assays considered in this analysis as “more sensitive” instead of “newer”, since the former is the actual characteristic affecting measurement."

I think the best way to address the problem is to make two minor changes. First I have amended the description of assays from "newer" to "more sensitive" because I agree with Reviewer: increased sensitivity is both the characteristic affecting measurement and the characteristic that determined the selection of data for analysis. Second I have added a footnote to Table 1 that gives (1) the year each assay was first used in these data and (2) the period each assay was in common use in these data. Note that Table 1 already summarises the year in which the suppression episode started, for both first and subsequent episodes. Most (90%) episodes started after 2005.

3. Introduction, line 61: The authors note early studies but the actual studies are not cited. Please include references for these studies.

I did not include the many references because these are nicely reviewed in the Grennan paper (which I then cited). I've made a minor change to alert the reader to the fact that this information has been summarised in the Grennan paper and the relevant sentence now reads:

"Early studies typically found no association between blips and subsequent viral rebound but were limited by their small sample size (see Table 1 in [2])."

4. Introduction, line 63-64: It would be useful to note the lower detection limit for the Roche Amplicor assays here, in order to understand the justification for the current study.
Thanks – yes, that ought to be mentioned. I have added this information.

5. It would be helpful to include a clearer justification for why there are separate analyses for first and subsequent episodes of viral suppression, perhaps in the Methods section. Also, it should be more clearly described that the “subsequent” episodes are episodes occurring after viral rebound.

Good point. These are not separate analyses; rather a single analysis but with separate baseline hazard functions for first and subsequent episodes of viral suppression. This ought to be explained. However the model used in the Grennan paper also has this property so that needs to be acknowledged too. Hence the 'Statistical methods' section now reads:

"We fitted a variety of proportional hazard models to data from first episodes and used these results to select a suitable model for our analyses of both first and subsequent episodes (Appendix A, Additional file 1). A suitable model should have separate baseline hazard functions for both first and subsequent episodes because although the effect of covariates may be the same in both first and subsequent episodes, the rate of viral rebound is likely to be higher in subsequent episodes (see Figure 1). The selected model was a generalised linear model for interval censored time to event data [20] with strata for first and subsequent episodes but we also fitted the gap-time Cox model used in an earlier study [2]. The gap-time model is a standard Cox model stratified by suppression episode and with time reset to zero at the beginning of each new episode (see {Kelly 2000}). However viral rebound is interval censored because it is only known to have occurred at some point between one measurement and the next. The standard Cox model is known to underestimate hazard ratios when measurement error is added to event times {Meier 2003}.

I have also tried to clarify what I mean by "subsequent episodes" and the relevant text reads:

"Patients could contribute more than one suppression episode to our analyses if they again achieved viral suppression after a viral rebound. Some patients contributed a subsequent suppression episode (after a first viral rebound) but did not have a first suppression episode measured using acceptable assays."  

6. Results, lines 178-179: Description of alternate definition for viral rebound should first be introduced in the Methods section. I have now done this (Methods, Suppression episodes).

7. Results, line 189: Authors should clarify what comparisons are being made with the two hazard ratios. I have re-written these comparisons to make them clearer:

"There was some evidence that viral rebound was associated with blips measured using the TaqMan version 2 assay relative to the Amplicor ultrasensitive assay (HR 1.31, 95% CI 0.93 to 1.86, Table 2). However blip magnitude was not appreciably more predictive of viral rebound with this assay than with other assays (HR 1.11, 95% CI 0.97 to 1.26, per 100 copies/mL using the TaqMan version assay, and HR 1.08, 95%CI 1.02 to 1.15, per 100 copies/mL using other assays). And low magnitude blips were not appreciably less predictive of viral rebound with this assay than with other assays, although this comparison lacks power (HR 1.11, 95% CI 0.60 to 2.04, using the TaqMan version 2 assay, and HR 1.21, 95%CI 0.90 to 1.63, using other assays).

8. I would be curious to know what the HR is for continuous blip threshold using the standard Cox model. Perhaps this could be included as part of Table 2 (along with the Weibull result).
I have added this information to the text – I am reluctant to make Table 2 more complicated in case I confuse the reader:

"In the gap-time model, the relative risk of viral rebound with increasing blip magnitude was estimated to be HR 1.08 (95% CI 1.03 to 1.14) per 100 copies/mL."

There is no real difference between the two models. As I note in Appendix A "Relative to Model 4, Models 1 and 2 appear to under and over estimate, respectively, although all estimates lead to the same clinical conclusions and each model shows an increase in the relative risk of viral rebound with increasing blip magnitude." With a gradual increase in risk (rather than a threshold effect), all models are likely to give essentially the same estimate when blip magnitude is represented by a continuous variable.

9. Table 2: Why is CD4 cell count not part of the standard Cox model, and similarly why is # RNA tests per year not part of the Weibull model? A footnote should be helpful to explain this.

I have reproduced the model used in an earlier study so our results for this model can be directly compared to the results in this study (as noted in the first paragraph of the Discussion).

As suggested, I have added a footnote to Table 2 that reads:

"The gap-time Cox model in [2] has the number of RNA tests per year as a covariate but not CD4 cell count. The number of RNA tests per year is not an appropriate covariate in models for interval censored data – see Appendix A, Additional file 1. Current (time updated) CD4 cell count was added to the model for interval censored data because it is a strong predictor of HIV progression even in patients with a suppressed viral load [21]."

10. Appendix A: The last sentence of the Methods section states that “we cannot fit [Model 4, Weibull model] to data from both first and subsequent suppression episodes”. Since this is assumed to be the primary model used for the paper, this sentence should be clarified.

It’s fair to say I did a fairly average job of explaining this issue in Appendix A. I would have used Model 4 for data from both first and subsequent suppression episodes except that it is just not possible with the software available. It’s theoretically possible, but programming the necessary likelihood would be challenging – and very time consuming. So I did the next best thing. I fit two other models that have been proposed for interval censored data – models that could be extended to analyse data from both first and subsequent episodes (which of course I wanted to do to increase the power of the analysis). I then selected the one that gave results closest to those from Model 4 when fit to data from just first episodes.

I do apologise for the confusion. I have made the changes to the Statistical methods section already described. And I have made many changes to Appendix A. I won’t reproduce all those changes here but I think that cumulatively this will eliminate the confusion.

Discretionary Revisions:

1. Abstract: The authors may want to consider describing the assays considered in this analysis as “more sensitive” instead of “newer”, since the former is the actual characteristic affecting measurement.

Agree – change made, and changes made throughout the manuscript for consistency.
2. Abstract: If there is room, it would be helpful to include the findings (or at least the P-value for trend) from the model with continuous blip magnitude, in order to be convinced that there is not a threshold effect at 500 copies/mL but rather a continuous/linear effect of blip magnitude on viral rebound (as suggested in the Conclusions). The Conclusion suggests that blips above 200 copies/mL may result in viral rebound, but the data in the Results section do not exactly support this.

I have added this result to the Abstract – the word count is under 250 so perhaps I can get away with it. I agree that both results – categorical and continuous – are needed to support the conclusion that a clinician ought to intervene at 200 copies/mL rather than at 500 copies/mL.

3. Introduction, line 74: Should define what “blip magnitude” here means, as it is unclear. It could refer either to HIV RNA level used to define the blip, or the number of blips.

I have amended this sentence so it reads:

"In this study, we consider whether blip magnitude – the level of transient HIV RNA in blood plasma – is predictive of subsequent viral rebound using data from the Swiss HIV Cohort Study (SHCS)."

4. Methods, line 103: Here, “size” is used instead of “magnitude”. Should use consistent terminology.

Indeed I should – change made.

5. Discussion, line 242: The word “these” might be changed to “our” to clarify that the authors are referring to the present study, not the cited studies.

Yes, that helps; change made.

6. Figures 1 and 2: What is the justification for use of 125, 350, and 750 copies/mL for the survival curves of blip magnitude per 100 copies? It seems that survival curves for 100, 200, 300, 400, etc. would be more intuitive for the reader.

I wanted to show curves for a continuous measure of blip magnitude at the midpoints of the categories that this continuous measure replaced – to illustrate the equivalence between the two. So I provided a curve at the midpoint between 50 and 200 copies/mL and so on.

7. Figure 2: For curves among non-adherent patients, why not show individual curves per copies/mL to provide more graphical evidence that the probability of continued suppression is independent of blip magnitude?

I was planning to do just that but among non-adherent patients, blip magnitude is so weakly associated with outcome that the estimate is a hazard ratio slightly below 1.0 (rather than above 1.0). The hazard ratio is reported in the manuscript: HR 0.96, 95% CI 0.80 to 1.17, per 100 copies/mL. This reverses the order of the curves, although they are all very close. I thought this reversal of curves would just confuse the reader and so provided a single curve that represents the clinical implications of the estimate – that is, that blip magnitude can be ignored in non-adherent patients.

8. Tables A2 and A3: I would recommend listing all of the covariates in the model in the footnote instead of referring the reader to Table 2.

I have listed the covariates in the footnotes, as suggested.
Reviewer: Ravindra Kumar Gupta

Minor essential revisions

1. Is NNRTI the standard first line in the SHCS? I assume so given the greater predictive value of blips in those on non NNRTI based regimens. I think this should be clarified in the paper.

There really isn’t a standard first line therapy in Switzerland. Health insurance is compulsory and health insurers will fund a variety of regimens and clinicians (and patients) can chose between regimens. As Table 1 shows, slightly more patients achieved a first suppression episode on a boosted PI than on an NNRTI.

I have given one reference with background information on the Swiss HIV Cohort Study. I’ve checked this reference but it says very little about the regimens that have been used and of course these have changed over time. I have added two new references: one compares the regimens used in the Cohort over time to International AIDS Society guidelines; the other describes the most common regimens and differences in their use in between Cohort sites.

I’ve added a new sentence to the Methods and this reads:

"Most patients (95%) in the SHCS received regimens recommended in clinical guidelines (Wandeler 2011). During 2005 to 2009, the most common first regimens were efavirenz with either tenofovir and emtricitabine or abacavir and lamivudine; lopinavir (boosted with ritonavir) with either tenofovir and emtricitabine or zidovudine and lamivudine; or atazanavir (boosted with ritonavir), tenofovir and emtricitabine (Elzi 2012)."

2. Could the authors comment on the lower limit of quantification of 20 copies/ml? would they say it is not necessary? I think this point is important as clinicians might otherwise think an assay with LLQ of 20 is better than one with LLQ of 40 copies/ml.

I think we address this point in the third paragraph of the Discussion:

"The increased sensitivity of the TaqMan version 2 assay, however, does not seem material in these data, in contrast to studies suggesting such effects might be important [8,9], because blip magnitudes measured with this assay were not dramatically more or less predictive of viral rebound than blip magnitudes measured with other assays."

"Most patients (95%) in the SHCS received regimens recommended in clinical guidelines (Wandeler 2011). During 2005 to 2009, the most common first regimens were efavirenz with either tenofovir and emtricitabine or abacavir and lamivudine; lopinavir (boosted with ritonavir) with either tenofovir and emtricitabine or zidovudine and lamivudine; or atazanavir (boosted with ritonavir), tenofovir and emtricitabine (Elzi 2012)."

"Most patients (95%) in the SHCS received regimens recommended in clinical guidelines (Wandeler 2011). During 2005 to 2009, the most common first regimens were efavirenz with either tenofovir and emtricitabine or abacavir and lamivudine; lopinavir (boosted with ritonavir) with either tenofovir and emtricitabine or zidovudine and lamivudine; or atazanavir (boosted with ritonavir), tenofovir and emtricitabine (Elzi 2012)."

"Most patients (95%) in the SHCS received regimens recommended in clinical guidelines (Wandeler 2011). During 2005 to 2009, the most common first regimens were efavirenz with either tenofovir and emtricitabine or abacavir and lamivudine; lopinavir (boosted with ritonavir) with either tenofovir and emtricitabine or zidovudine and lamivudine; or atazanavir (boosted with ritonavir), tenofovir and emtricitabine (Elzi 2012)."