PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Body mass index, proteinuria and total lymphocyte counts in predicting treatment responses among ART naïve individuals with HIV initiated on antiretroviral treatment in Dar es Salaam, Tanzania, 2019: a cohort study
AUTHORS	Munseri, Patricia; Jassely, Lazaro; Tumaini, Basil; Hertzmark, Ellen

VERSION 1 – REVIEW

REVIEWER	Kalluvya, Samuel Bugando Medical Centre, Internal Medicine
REVIEW RETURNED	14-Feb-2022

GENERAL COMMENTS	.In sub Saharan Africa there is transitioning from efavirenz to dolutegravir based regimens which can cause weight gain even in patients without optimal HIV viral suppression.Therefore weight gain and hen is not a good surrogate marker of viral suppression when dolutegravir is used. If an AZT based regimen is used and HIV lipodystrophy develops with weight loss, causing a drop in BMI even in fully suppressed patients. In this situation cART failure will be diagnosed using the BMI drop criterion. Protenuria may also arise or recur from UTI in a virally supressed individual after six months of cART, making proteinuria to be an unreliable marker of cART failure. When will Enhaced Adherence counselling(EAC) be introduced, and how will re suppression be assessed after EAC sessions if the proposed markers of cART failure are used? Immunological non responders will be wrongly diagnosed as having cART failure if total lymphocyte counts are used as determinants of cART failure. Considering that dried blood spots can be used to get HIV viral loads even from remote rural areas to ensure early detection of virological failure with its associated benefits of limiting accumulation of HIV drug resistance and preserving future treatment options,do we need probably late markers in sub Saharan Africa with a very limited ARV formulary.? Authors have to address those questions/concerns.

REVIEWER	Sabin, Caroline UCL Medical School
REVIEW RETURNED	24-Feb-2022

GENERAL COMMENTS	The authors of this manuscript have described associations between three measures (weight change, change in proteinuria and change in lymphopenia status) assessed three months after

	ART initiation and virological suppression rates at 6 months with the aim of developing a prognostic model that can identify individuals at risk of non-suppressed viraemia at an earlier point after ART initiation. The manuscript is clearly written and the results have some clinical value. I do, however, feel that the final model may not be optimal and would suggest some further analyses that might help (see below). 1. As a general point, the authors could use more person-centred language throughout - so use of 'people with HIV' rather than 'HIV-infected', and people/individuals rather than patients. 2. Introduction, paragraph 2 - the authors make a statement about the frequency of monitoring after ART initiation. Could they clarify which country this relates to (I assume Tanzania, but this isn't stated)? 3. Methods - a clear statement of the inclusion/exclusion criteria for participants in the study would be helpful. I assume the study included all ART-naive participants starting ART for the first time? What is 'baseline'? 4. Sample size calculation - this is very confusing and the minimum detectable risk ratios were actually very high meaning that important predictors might be missed (for example, older age and female sex might both be predictors of non-virological response but these were not significant, possibly due to lack of power). Some discussion of the implications of this choice of sample size would be valuable. 5. Statistical methods - the authors pre-define what constitutes an increase/decrease in BMI or change in proteinuria/lymphopenia status, converting what are essentially continuous measurements (change in BMI, say) into categorical measurements (decrease, stable, increased BMI). Whilst this may be the most obvious categorisation from a clinical perspective, it may not result in an optimal model - different categorisations of the variables may result in models that have better sensitivity/specificities, say. It would be interesting, therefore, to see further analyses that attempt to identify the optimal categorisation of these variables, before including these together into the final model. 6. A brief rationale for the use of modified Poisson rather than logistic regression would be helpful. 7. To generate the final prognostic model, the authors have rounded each of the parameter estimates to 1 - the final score is therefore simply the number of negative characteristics that were present and can only range from 0-3; the approach also weights each of the characteristics equally although this may not be optimal. Why not generate a score that actually uses the estimates themselves rather than rounding - given that most people now have access to mobile phones, it shouldn't be that difficult to find a way to implement the model in a clinic setting. As it is, the ROC curve isn't that helpful (as the score can only take limited discrete values). 8. Results. How many people in these clinics started ART over the period of recruitment, and what proportion were recruited into the

	study? How did the characteristics of those included compare to the rest of the people attending the clinics? 9. Discussion - this really lacks any discussion of the limitations of the study. For example, any prognostic model is generally overly optimistic in terms of sensitivity and specificity, and thus models should always be validated in external populations, Are there other markers that might have greater prognostic value in such a model that might be available?

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Dr. Samuel Kalluvya, Bugando Medical Centre, Bugando Medical Centre
Comments to the Author:
In sub Saharan Africa there is transitioning from efavirenz to dolutegravir based regimens which can cause weight gain even in patients without optimal HIV viral suppression.Therefore weight gain and hen is not a good surrogate marker of viral suppression when dolutegravir is used.
Response: We thank the reviewer for the comment. We acknowledge that most people living with HIV in many sub-Saharan countries are now transitioning from efavirenz based regimens to dolutegravir based regimens.
The concept that individuals gain weight when on integrase strand inhibitors and lose weight on efavirenz based regimens is interesting. However, there is limited research in this area. We take note of a study done in South Africa by Griesel et al that indicated that the effects of weight loss from efavirenz are due to increase in efavirenz concentrations over time (48 weeks) with an explanation that weight loss resulted from mitochondrial toxicity and neuropsychiatric effects. The issue of mitochondrial toxicity and neuropsychiatric manifestations among patients on efavirenz was addressd by modifying the dose to 600mg. In the South African study, patients switched from efavirenz to dolutegravir did not have an increase in weight due to dolutegravir but rather due to failure of the patients to gain weight while on efavirenz due to toxicity.

1. Griesel R, Kawuma AN, Wasmann R, Sokhela S, Akpomiemie G, Venter WF, et al. Concentration-response relationships of dolutegravir and efavirenz with weight change after starting antiretroviral therapy. British Journal of Clinical Pharmacology. 2022; 88(3): 883- 893. https://doi.org/10.1111/bcp.15177

Our study indicates that a decrease in BMI is associated with HIV viral non suppression at six months, while individuals with HIV whose BMI remained normal or increased were virally suppressed.
We suggest that a clinician who cares for people living with HIV on treatment should be on the lookout rather than assume that this could be a side effect of efavirenz.
Although this study was conducted on participants who were on efavirenz based regimens, we believe that our findings may apply to any regimen and should raise an alarm for clinicians caring for patients on ARV if their patients have an unintentional weight loss with subsequent drop in BMI. We have included the point in the discussion section on page 20 from line 318 - 321

If an AZT based regimen is used and HIV lipodystrophy develops with weight loss, causing a drop in BMI even in fully suppressed patients. In this situation cART failure will be diagnosed using the BMI drop criterion.

Response: This is an interesting issue. Fortunately, none of the participants in this study were on an AZT based regimen and all participants were ART naïve. A drop in weight and subsequent BMI should raise an alarm to a clinician of an early warning sign for treatment failure first before assuming that the reason is lipodystrophy.

Protenuria may also arise or recur from UTI in a virally supressed individual after six months of cART, making proteinuria to be an unreliable marker of cART failure.
Response: We thank the reviewer for this observation. Proteinuria is an indicator of glomerular injury. Renal biopsy findings among patients with HIV showed that the presence of proteinuria or microalbuminuria is mainly indicative of HIV associated nephropathy (HIVAN) especially in among Africans, with prevalence of HIVAN ranging from 53-79%. Our study has indicated that proteinuria was associated with HIV non-suppression, and therefore, the presence of proteinuria should alert a clinician to exclude treatment failure first, and if excluded, look for alternative explanations.
See:

1. Carter JL, Tomson CR, Stevens PE, Lamb EJ. Does urinary tract infection cause proteinuria or microalbuminuria? A systematic review. Nephrology Dialysis Transplantation. 2006; 21(11):3031-7. https://doi.org/10.1093/ndt/gfl373
2. Han TM, Naicker S, Ramdial PK, Assounga AG. A cross-sectional study of HIV-seropositive patients with varying degrees of proteinuria in South Africa. Kidney International. 2006; 69(12):2243-50. https://doi.org/10.1038/sj.ki.5000339

When will Enhaced Adherence counselling(EAC) be introduced, and how will re suppression be assessed after EAC sessions if the proposed markers of cART failure are used?
Response: We thank the reviewer for the comment. Our study findings are not intended to replace enhanced adherence counseling but to facilitate EAC after a clinician observes that the person living with HIV has a drop in BMI, presence of proteinuria and reduced lymphocyte counts. A clinician will review the patient's adherence while exploring further for treatment failure.

Immunological non responders will be wrongly diagnosed as having cART failure if total lymphocyte counts are used as determinants of cART failure.
Response: We thank the reviewer for the comment. The aim of our research was to aid a clinician or the ART team to triage people living with HIV who could be at risk of treatment failure. The three parameters: drop in BMI, presence of proteinuria and low lymphocyte counts will raise an alarm so that the individual with HIV is further investigated to exclude treatment failure. Immunological non responders are the persons who will need primary and secondary prophylaxis for opportunistic infection. Therefore, if picked by low lymphocyte counts further action is needed to exclude treatment failure and subsequent provision of prophylaxis. We have included the reviewer's point in the discussion section on page 20 line 307-309

Considering that dried blood spots can be used to get HIV viral loads even from remote rural areas to ensure early detection of virological failure with its associated benefits of limiting accumulation of HIV drug resistance and preserving future treatment options,do we need probably late markers in sub Saharan Africa with a very limited ARV formulary.?
Response: We thank the reviewer. We agree with the reviewer that dried blood spots (DBS) are a good alternative in remote settings. However dry blood spots are usually reserved for children (infants) for HIV detection and the test is not routine for HIV viral load testing in adults. Although this would be a good alternative in resource constrained settings, testing for DBS will need resources to ship the DBS to a laboratory with a capacity for analysis, this might still pose a challenge in such settings with difficulties in transportation and turnaround time for results to be utilized by the clinician at the primary care and treatment center.

We believe the limitation in ARV formulary should not deprive a clinician from assessing a patient and offering the best available care, as currently most resource-constrained settings have up to three treatment options when treatment failure arises.
The option of using late markers of viral non suppression could
encourage secondary transmission of resistant viral strains in the population with ultimate loss of the first line regimen.

Authors have to address those questions/concerns.

.

Reviewer: 2
Prof. Caroline Sabin, UCL Medical School
Comments to the Author:
The authors of this manuscript have described associations between three measures (weight change, change in proteinuria and change in lymphopenia status) assessed three months after ART initiation and virological suppression rates at 6 months with the aim of developing a prognostic model that can identify individuals at risk of non-suppressed viraemia at an earlier point after ART initiation. The manuscript is clearly written and the results have some clinical value. I do, however, feel that the final model may not be optimal and would suggest some further analyses that might help (see below).

1. As a general point, the authors could use more person-centred language throughout - so use of 'people with HIV' rather than 'HIV-infected', and people/individuals rather than patients.
Response: We thank Dr. Sabin. We have used a more person-centered language as proposed by the reviewer. We have now changed “HIV-infected individuals” to “individuals with HIV”. The changes have been marked in track change.

2. Introduction, paragraph 2 - the authors make a statement about the frequency of monitoring after ART initiation. Could they clarify which country this relates to (I assume Tanzania, but this isn't stated)?
Response: We thank the reviewer and we have improved the introduction section on page 6 line 82 by clarifying the country. The sentence now reads” Individuals with HIV are routinely assessed for weight, height, renal function, and complete blood count before initiation of combined antiretroviral treatment (ART) in resource constrained settings including Tanzania”

3. Methods - a clear statement of the inclusion/exclusion criteria for participants in the study would be helpful. I assume the study included all ART-naive participants starting ART for the first time? What is 'baseline'?
Response: We thank the reviewer for helping us improve the clarity of our manuscript. We have included inclusion criteria under the Methods section on page 9 from lines 122-124 that now reads Participants were included in the study if they were newly diagnosed with HIV and were ART naïve aged 18 years or older.
Baseline means at treatment initiation changed in Data collection

4. Sample size calculation - this is very confusing and the minimum detectable risk ratios were actually very high meaning that important predictors might be missed (for example, older age and female sex might both be predictors of non-virological response but these were not significant, possibly due to lack of power). Some discussion of the implications of this choice of sample size would be valuable.
We thank the reviewer for this comment. We understand that with our small sample size we cannot uncover all the potential risk factors in our data. On the other hand, the small sample size means that

only the strongest risk factors will be “statistically significant,” and these are likely to be true over a wide range of locations and times.
We have modified this section to read:

Response: We thank the reviewer, we have modified this section on page 9 line 128-136 that now reads: To determine the minimal detectable relative risks for the study variables, we considered two-sample tests of the expected highest risk category versus the expected lowest risk category. For the dichotomous potential risk factors, we assumed a total number of 215 subjects, split roughly as our actual data are split (with numbers rounded to the nearest 5 to mimic a pre-study power calculation). For BMI change we used 80 for the reference group (gain), 125 for stable, and 20 for the loss group. The minimum detectable risk ratios were 3.77 for decreased BMI, 2.56 for stable BMI, 2.94 for lymphopenia and for proteinuria, 2.47 for stage greater than 1, 2.47 for age of 40, years and above 2.59 (or < 0.11) for female sex, 2.74 for secondary or higher education, 2.44 for unemployment and for never married.

5. Statistical methods - the authors pre-define what constitutes an increase/decrease in BMI or change in proteinuria/lymphopenia status, converting what are essentially continuous measurements (change in BMI, say) into categorical measurements (decrease, stable, increased BMI). Whilst this may be the most obvious categorisation from a clinical perspective, it may not result in an optimal model - different categorisations of the variables may result in models that have better sensitivity/specificities, say. It would be interesting, therefore, to see further analyses that attempt to identify the optimal categorisation of these variables, before including these together into the final model.
Response: We thank the reviewer for the insightful comment. We employed BMI change by 10% whose results of univariable analysis are presented below:

BMI changes in 10 percentage points

a. Gain >10%
b. Stable
c. Loss >10%

BMI change from baseline to six months	Total	HIV non suppression at six months n (%)	Univariable RR
Loss >10%	10	9 (90.0)	9.23 (3.55-23.99)
Stable	164	33 (20.1%)	2.06 (0.77-5.51)
Gain >10%	41	4 (9.8%)	1

This analysis indicates a statistically significant association between a 10% BMI loss and HIV non-suppression even though the numbers tend to get smaller. We think this categorization is not optimal since BMI loss is then able to pick only 9 out of the 46 (20%) with viral non-suppression. We therefore maintain the categorization initially presented (see Table 2) as it would be of value for a clinician to pick failing patients earlier rather than wait for a 10% loss in BMI. We have modified the discussion section to reflect the 10% categorization on page 17 line 245-247

6. A brief rationale for the use of modified Poisson rather than logistic regression would be helpful.
Response: We thank the reviewer for the suggestion we have improved clarity to describe the reason for use of modified Poisson regression. Poisson regression analysis is used for non-rare outcomes while logistic regression analysis is for rare outcomes. HIV non-suppression is a non-rare outcome with a prevalence of 10% or more and in this study the prevalence was 21%. To improve clarity of the manuscript, we have added a sentence in the methods section under Statistical Methods on page 12

from line 125-126 that reads “To determine the relationships between individual predictors and viral non-suppression at six months, we first used modified Poisson regression for univariable analysis with an assumption that viral non suppression is a non-rare outcome with an estimate of more than 10%, to determine which variables to include in the multivariable model”
See:

1. Spiegelman D, Hertzmark E. Easy SAS calculations for risk or prevalence ratios and differences. American Journal of Epidemiology. 2005; 162(3):199-200. https://doi.org/10.1093/aje/kwi188
2. Dwivedi AK, Mallawaarachchi I, Lee S, Tarwater P. Methods for estimating relative risk in studies of common binary outcomes. Journal of Applied Statistics. 2014; 41(3):484-500. https://doi.org/10.1080/02664763.2013.840772

7. To generate the final prognostic model, the authors have rounded each of the parameter estimates to 1 - the final score is therefore simply the number of negative characteristics that were present and can only range from 0-3; the approach also weights each of the characteristics equally although this may not be optimal. Why not generate a score that actually uses the estimates themselves rather than rounding - given that most people now have access to mobile phones, it shouldn't be that difficult to find a way to implement the model in a clinic setting. As it is, the ROC curve isn't that helpful (as the score can only take limited discrete values).

Response: We thank the reviewer for thinking about this. A score will simplify implementation for a clinician.
We think that use of a score based on the estimates, which were, of course, optimized to our data, would seem to be more precise, but would be more likely less accurate on data from other times and places. Regardless of the small number of discrete values, the ROC curve shows that even our model had good predictive ability.

8. Results. How many people in these clinics started ART over the period of recruitment, and what proportion were recruited into the study? How did the characteristics of those included compare to the rest of the people attending the clinics?
Response: To improve clarity based on the reviewer’s comment we have updated the results section on page 13 line 203 the sentence now reads: There were a total of 220 ARV-naive individuals who were initiated on ART during the study period and all were recruited in the study.

9. Discussion - this really lacks any discussion of the limitations of the study. For example, any prognostic model is generally overly optimistic in terms of sensitivity and specificity, and thus models should always be validated in external populations,
Response: Thank you, Dr. Sabin, for encouraging us to improve our Discussion section. We have added a description to the discussion section on page 22 from line 343-that reads “Our findings require validation in a study with a larger sample size. A small sample may have constrained some predictors of non-virological response. Similar studies conducted in different locations are also needed since local conditions and treatment standards may influence some observed patterns, both in prevalence and effect”.

Are there other markers that might have greater prognostic value in such a model that might be available?
Response: In an attempt to look for other predictors, we have added a recommendation on page 22 line 348 that reads “We recommend further studies to examine the relationship between virological response and anemia as well as opportunistic infections and AIDS associated malignancies.”
Reviewer: 1

Competing interests of Reviewer: none

Reviewer: 2
Competing interests of Reviewer: I have no competing interests

VERSION 2 – REVIEW

REVIEWER	Kalluvya, Samuel Bugando Medical Centre, Internal Medicine
REVIEW RETURNED	20-Apr-2022

GENERAL COMMENTS	According to WHO reccommendations which are used in sub Saharan Africa and other resource limited settings,virologic failure is is only confirmed if two HIV viral loads results are above 1000copies/ml,that is there is no resuppression after three months of Enhanced Adherence Counselling and support(EAC).That is a single high HIV viral load is not sufficient to make a diagnosis of HIV virologic failure and switch from the initial cART regimen to a subsequentmore expensive second line cART regimen.With the proposed predictors in this study-BMI,proteinuria,and total lymhocyte count, how will resuppression in case of previously poor adherence which led to an initialbe ass spiking of viral load to above 1000copies/ml be defined/assessed?According to WHO resuppression is defined as a drop of viral load to below 1000copies/ml.This is amajor limitation of this study which needs to be discussed adequately,The proposed predictors of virologic failure do not seem to be reversible after optimisation of adherence through EAC whilistHIV viral load would resuppress tobelow 1000copies/ml.Thus the diagnosis of HIV virologic failure would not be upheld averting unnecessary switch to a more expensive, difficult to adhere to second line cART regimen. The issue of INSTI associated weight gain should not simply be discarded,it is areal clinical practice problem and so it should be listed as a limitation.

REVIEWER	Sabin, Caroline UCL Medical School
REVIEW RETURNED	05-Apr-2022

GENERAL COMMENTS	I thank the authors for responding to my comments - I believe the manuscript is now improved with some of the limitations of the research now being clearly stated.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 2
Prof. Caroline Sabin, UCL Medical School
Comments to the Author:
I thank the authors for responding to my comments - I believe the manuscript is now improved with some of the limitations of the research now being clearly stated.
Response: We thank Prof. Sabin for the insightful comments that have helped to shape the manuscript

Reviewer: 1

Dr. Samuel Kalluvya, Bugando Medical Centre, Bugando Medical Centre
Comments to the Author:
According to WHO reccommendations which are used in sub Saharan Africa and other resource limited settings,virologic failure is is only confirmed if two HIV viral loads results are above 1000copies/ml,that is there is no resuppression after three months of Enhanced Adherence Counselling and support(EAC).That is a single high HIV viral load is not sufficient to make a diagnosis of HIV virologic failure and switch from the initial cART regimen to a subsequentmore expensive second line cART regimen.
Response: We thank Dr. Kalluvya for the comments and concerns.
Perhaps we were not clear, and we would like to improve our manuscript with the clarity of our findings.
We found out that the presence of proteinuria, lymphopenia and a 5% or more drop in BMI at six months should alert the clinician that there is a need to explore further with the person living with HIV on treatment. The first step is to check on the ART adherence and to perform extensive adherence counseling in settings where viral load is available viral load testing at six months and subsequently re-testing for viral load after 3 months. But in settings where there is a challenge in obtaining viral load these parameters would be of help in picking patients who need additional support with adherence or change in regimen.
We have changed the manuscript (lines 328 to 332 on page 22) as follows: The presence of proteinuria, lymphopaenia, and a drop in BMI of 5% are relatively simple parameters to monitor among people living with HIV on ART especially, in a setting where viral load monitoring is a challenge. The presence of any of these parameters should alert a clinician on the possibility of viral non-response and review adherence issues including individualized enhanced adherence counselling and subsequent treatment options.

With the proposed predictors in this study-BMI,proteinuria,and total lymhocyte count, how will resuppression in case of previously poor adherence which led to an initialbe ass spiking of viral load to above 1000copies/ml be defined/assessed?
Response: We do not propose a re-defining of viral suppression criteria, what we propose is the use of simple available parameters that will alert a clinician to the possibility that the person living with HIV on treatment could be having issues with their treatment. We have changed the terminology in the manuscript on page 21 lines 310 and 312 from virological failure or treatment failure to read HIV viral non-suppression from any cause
According to WHO resuppression is defined as a drop of viral load to below 1000copies/ml.This is a major limitation of this study which needs to be discussed adequately,The proposed predictors of virologic failure do not seem to be reversible after optimisation of adherence through EAC whilistHIV viral load would resuppress tobelow 1000copies/ml.Thus the diagnosis of HIV virologic failure would not be upheld averting unnecessary switch to a more expensive, difficult to adhere to second line cART regimen.
Response: We thank the reviewer for this interesting comment. This study was conducted among ART naïve individuals initiating on treatment and ended at six months of follow up. We have added a recommendation to further studies to explore what happens to the BMI, proteinuria and lymphocyte counts after six months following extensive adherence counseling among people living with HIV to ascertain if these parameters are reversible.
Proteinuria is a reversible parameter provided other structural causes have been excluded. We observed reversibility of proteinuria in 78 of 89 (87.6%) participants with proteinuria at baseline and at 6 months did not have proteinuria with viral suppression.
Our manuscript now emphasizes that we are not redefining VL non-suppression, but merely trying to aid clinicians in taking the best care possible of their patients, often in situations when viral load or C4 cell counts are not available.
We have included a recommendation in the manuscript on page 22 lines 338-341 that now reads: We

recommend further studies with extended follow up of patients beyond six months to monitor further change in lymphopaenia, proteinuria and drop in BMI of 5% or more especially for individuals maintained on the same regimen after enhanced adherence counselling.
The issue of INSTI associated weight gain should not simply be discarded,it is areal clinical practice problem and so it should be listed as a limitation.
Response: We thank the reviewer for the concern about INSTI and weight gain. We agree that INSTI causes weight gain however we describe the failure to gain weight or a drop in BMI as a clinical parameter that would alert the clinician on treatment responses among people living with HIV initiated on ART.

Reviewer: 2
Competing interests of Reviewer: I have no competing interests.

Reviewer: 1
Competing interests of Reviewer: NO COMPETING INTERESTS