Plasma calprotectin as a biomarker of mortality at antiretroviral treatment initiation in advanced HIV – pilot study [version 1; peer review: 2 approved with reservations]

Faith W. Kamau1,2, Agnes Gwela1,3, Andrew K. Nyerere4, Victor Riitho5, James M. Njunge1,3, Moses M. Ngari1,3, Andrew J. Prendergast5, James A. Berkley1,3,6

1Clinical Research, KEMRI/Wellcome Trust Research Programme, Kilifi, Kilifi County, 320-80108, Kenya
2Department of Molecular Biology and Biotechnology, Pan African University Institute for Basic Sciences, Technology and Innovation, Juja, Nairobi, 62000-00200, Kenya
3Childhood Acute Illness & Nutrition (CHAIN) Network, Nairobi, Nairobi, 43640-00100, Kenya
4Department of Medical Microbiology, Jomo Kenyatta University of Agriculture and Technology, Juja, Nairobi, 62000–00200, Kenya
5Blizard Institute, Queen Mary University of London, London, London, E1 2AT, UK
6Centre for Tropical Medicine & Global Health, University of Oxford, Oxford, Oxfordshire, OX3 7FZ, UK

V1 First published: 13 Mar 2020, 5:46
https://doi.org/10.12688/wellcomeopenres.15563.1
Latest published: 13 Mar 2020, 5:46
https://doi.org/10.12688/wellcomeopenres.15563.1

Abstract

Background: In advanced HIV, significant mortality occurs soon after starting antiretroviral treatment (ART) in low- and middle-incomes countries. Calprotectin is a biomarker of innate response to infection and inflammatory conditions. We examined the association between plasma calprotectin at initiation of ART and mortality among individuals with advanced HIV.

Methods: We conducted a pilot case-cohort study among HIV infected adults and children over 5 years old with CD4+ <100/mm³ at ART initiation at two Kenyan sites. Participants received three factorial randomised interventions in addition to ART within the REALITY trial (ISRCTN43622374). Calprotectin was measured by ELISA in archived plasma of those who died within 24 weeks (cases) and randomly selected participants who survived for 48 weeks (non-cases) for whom samples were available. Factors associated with baseline plasma calprotectin were investigated using linear regression. To test association with mortality, Cox proportional hazards models with inverse sampling probability weights and adjusted for age, sex, site, BMI, viral load, randomised treatments, and clustered by CD4 count were fitted.

Results: Baseline median (IQR) plasma calprotectin was 6.82 (2.65–12.5) µg/ml in cases (n=39) and 5.01 (1.92–11.5) µg/ml in non-cases (n=58). Baseline calprotectin was associated with age, neutrophil count and the presence of cough, but not other measured indicators of infection. In adjusted multivariable models, baseline calprotectin was associated with subsequent mortality: HR 1.64 (95% CI 1.11 - 2.42)
and HR 2.77 (95% CI 1.58 - 4.88) for deaths during the first twenty-four and four weeks respectively. Calprotectin levels fell between baseline and 4 weeks among both cases and non-cases irrespective of randomised interventions.

**Conclusion:** Among individuals with advanced HIV starting ART in Kenya, plasma calprotectin may have potential as a biomarker of early mortality. Validation in larger studies, comparison with other biomarkers and investigation of the sources of infection and inflammation are warranted.

**Keywords**
HIV, CD4, Mortality; Adult, Neutrophil, Antiretroviral, Biomarker, Prognostic
Introduction
Approximately one-quarter of individuals newly diagnosed with HIV in sub-Saharan Africa have advanced disease at presentation (IeDEA and ART Cohort Collaborations et al., 2014). Advanced HIV is characterized by immunosuppression, infection and immune activation which may independently drive mortality despite antiretroviral therapy (ART). Measurements of soluble biomarkers such as soluble CD14 (sCD14), C-reactive protein (CRP) and IL-18 have highlighted that inflammation and innate immune responses predict all-cause mortality, cardiovascular events, and other morbidities in HIV infected individuals, even after ART initiation (Duprez et al., 2012; Kuller et al., 2008; Sandler & Douek, 2012). Overall, innate immune activation seems more important than T-cell activation for disease progression in sub-Saharan Africa (Hunt et al., 2016; Serrano-Villar et al., 2014). Inflammation and co-infection can also arise from disruption of intestinal tight junctions leading to increased mucosal permeability. Translocation from the intestine of bacteria and their products including lipopolysaccharides has been demonstrated in some (Brenchley et al., 2006; Marchetti et al., 2013; Nazli et al., 2010; Sandler & Douek, 2012), but not all studies (Fitzgerald et al., 2019).

Calprotectin is a soluble 24 kDa dimer of calcium-binding proteins S100A8 and S100A9 (Brophy & Nolan, 2015) produced by neutrophils and other cells following activation in response to infection and inflammation. Calprotectin, measured in either stool or plasma, is a recognised biomarker of inflammation and bacterial infections including sepsis (Banerjee et al., 2015; Bjarnason, 2017; Huang et al., 2016; Jonsson et al., 2017; Simm et al., 2016; Walsham & Sherwood, 2016). The S100A9 sub-unit of plasma calprotectin has been associated with reduced immune reconstitution after ART (Drozd et al., 2016), enhanced antimicrobial defence transiently induced by antiviral treatment (Muller et al., 1994) and HIV-associated neurocognitive disorders (Colon et al., 2016). Njunge et al. recently demonstrated an association between increased plasma calprotectin and early post-discharge mortality among HIV-uninfected children hospitalized for severe acute malnutrition (Njunge et al., 2019).

We considered that calprotectin may be of value as a prognostic biomarker in advanced HIV and conducted a pilot study to evaluate associations between plasma calprotectin and mortality in individuals with advanced HIV infection at ART initiation who participated in the Reduction of Early Mortality in HIV Infected Adults and Children Starting Antiretroviral Therapy (REALITY) trial (Hakim et al., 2017) (Kityo et al., 2018) (Mallewa et al., 2018).

Methods
Study population
This was a nested case-cohort study within the REALITY trial (ISRCTN43622374) that enrolled HIV-infected adults and children aged five years or more with a CD4+ T cell count <100 cells/mm³ and without previous ART treatment at 8 sites in Kenya, Uganda, Malawi, and Zimbabwe. Participants in the REALITY trial were enrolled between August, 2013 and April, 2015 and randomised to three factorial treatments compared to standard of care: enhanced antimicrobial prophylaxis (single-dose albendazole, 5 days of azithromycin, 12 weeks of fluconazole (100 mg), and 12 weeks of fixed-dose combination of cotrimoxazole (800/160 mg)/isoniazid (300 mg)/pyridoxine (25 mg) once daily) (Hakim et al., 2017); additional raltegravir (Kityo et al., 2018); and ready to use supplementary food (RUSF) (Mallewa et al., 2018).

Study design
This pilot case-cohort study utilised archived plasma samples from the two Kenyan sites participating in the REALITY trial: Kilifi County Hospital and the Academic Model for the Prevention and Treatment of HIV/AIDS Centre at Moi Teaching Referral Hospital, Eldoret. REALITY enrolled a total of 139 participants in Kilifi, of whom 29 (20%) died, and 195 in Eldoret, of whom 14 (7.2%) died. This pilot study capitalised on a broader ongoing immunology case-cohort sub-study within REALITY that included study participants with a sample set of plasma, baseline stool, PBMCs and data on CD8+ T cell counts which would allow investigation of biomarkers.

The case cohort study aimed to randomly select 45% of all Kilifi participants and 10% of Eldoret study participants to reflect mortality, stratified by CD4+ count (0–24, 25–49, and 50–99 cells/mm³) plus any remaining unselected deaths by week 24 (the trial primary endpoint; Figure 1). Random selection was performed in the trial database using the uniform random number function in STATA. However, in Kilifi baseline CD8 measurements were missing on approximately one-third of participants due to reagent unavailability at specific time periods (i.e. missing at random) and one-quarter were also missing stored specimens for similar reasons. Therefore, the required number of participants in Kilifi were selected from those with complete samples and baseline CD8 available to ensure that data could be obtained, with weights (see below) reflecting the original population.

Deaths were weighted as 1 and non-deaths were weighted as 106/42 in Kilifi and 174/16 in Eldoret. These values were chosen in order that the sample represented the full trial population in terms of deaths and survivors at these sites using the inverse probability of selection from all REALITY patients ≥13 years of age alive at week 48 (regardless of immunology sub-study membership and available samples). Demographic, clinical and laboratory data were collected during the REALITY trial using standardised case report forms (Hakim et al., 2017). Complete blood counts, including neutrophils and CD4+ counts were done at local laboratories.

Enzyme-linked immunosorbent assay (ELISA)
Plasma calprotectin was measured in duplicate at recommended dilutions using a solid-phase enzyme-linked immunosorbent assay (ELISA) human calprotectin kit (Hycult Biotech HK379-02) according to the manufacturer’s instructions and the absorbance read at 450 nm using a Synergy 4 (BioTek) plate reader. Calprotectin was measured using samples...
collected at baseline and four weeks after treatment initiation. The four-week time period allowed an appropriate time frame for the detection of immunological changes following treatment initiation.

**Statistical analysis**

Data were analysed using STATA version 13.1 (STATA corp. TX, USA). The baseline characteristics of the selected study participants were presented as mean (SD) for normally distributed variables, as median with interquartile range (IQR) for non-normally distributed variables, and as numbers (percentage) for categorical data. To compare differences in characteristics between cases and non-cases, Student’s t-test and Mann-Whitney rank-sum test (of non-normally distributed variables) were used. Chi-square or Fisher’s exact tests were used to compare proportions and Spearman’s correlation to assess correlation.

Factors associated with plasma calprotectin levels at baseline were assessed using a linear regression model including age, sex, site, CD4+, viral load and neutrophil count, but excluding the randomised interventions. Individual clinical and laboratory features of infection: the presence of fever, cough, diarrhoea, known tuberculosis, measured body temperature, CD8+ T cell count, and cryptococcal antigen test were then added one by one, retaining only those which were statistically significant. The beta coefficient indicated the strength of the effect on plasma calprotectin.

The association between baseline plasma calprotectin and mortality was assessed using a Cox proportional hazards regression model with inverse probability weights (Buchanan et al., 2014) and stratified by CD4 count to reflect the proportions selected in the case-cohort design to help address bias, as described above. Time at risk was defined from enrolment to 24 weeks which was the primary endpoint of the parent trial. The hazard ratio per unit increase in calprotectin was estimated, after adjustment for age, sex, BMI, log viral load, site and each of the randomised treatment arms. Akaike (AIC) and Bayesian (BIC) information criteria were calculated to assess model performance.

The Wilcoxon paired sign-rank test was used to assess the change in plasma calprotectin between baseline and after four weeks following treatment. Linear regression adjusted for age, sex and site was used to examine whether the subsequent death and the three randomised interventions were associated with the change in plasma calprotectin between baseline and 4 weeks.

**Ethical considerations**

The study protocol was approved by the Kenya Medical Research Institute (KEMRI) Scientific and Ethics Review Committee approval number SSC 2231. Written informed consent was obtained from all participants using local languages, which included permission for storage and testing for this work.
Results
Characteristics of study participants
Of the selected 97 participants, 39 were cases and were 58 non-cases (Figure 1). The baseline characteristics of the participants are shown in Table 1. A total of 16/39 (41%) cases died between enrolment and four weeks, and a further 23/39 (59%) cases died between four and 24 weeks. Death occurred at a median of 31 (Interquartile range, IQR 18–72) days after enrolment. Cases had a lower BMI, lower haemoglobin and CD8+ T cell counts than non-cases, and fewer cases were randomized to receive RUSF in univariate analyses.

Baseline plasma calprotectin
At enrolment, crude calprotectin levels were median (IQR) 6.82 (2.65 to 12.5) µg/mL in cases (n=39) and 5.01 (1.92 to 11.5) µg/ml in non-cases (n=58). Higher age was associated with lower plasma calprotectin at baseline and there was positive association with female sex, neutrophil count and reporting a cough (Table 2). Other putative markers of infection (CD8+ T cells, fever, tuberculosis, body temperature, diarrhoea, and cryptococcal antigen test) did not emerge as factors affecting plasma calprotectin levels.

Baseline calprotectin was significantly associated with subsequent mortality to 24 weeks ($P=0.03$) and mortality within the first 4 weeks ($P<0.001$) in multivariable models adjusted for potential confounders (Table 3).

Plasma calprotectin after 4 weeks
Between baseline and four weeks, calprotectin declined both in cases who were still alive at 4 weeks (n=21) and in non-cases (n=57), without evidence of difference between cases

<table>
<thead>
<tr>
<th>Table 1. Participant characteristics.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristic</td>
</tr>
<tr>
<td>Site</td>
</tr>
<tr>
<td>Kilifi (%)</td>
</tr>
<tr>
<td>Eldoret (%)</td>
</tr>
<tr>
<td>Sex</td>
</tr>
<tr>
<td>Male (%)</td>
</tr>
<tr>
<td>Female (%)</td>
</tr>
<tr>
<td>Age (years)</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
</tr>
<tr>
<td>Full blood count</td>
</tr>
<tr>
<td>Haemoglobin (g/dL)</td>
</tr>
<tr>
<td>CD4+ count (cells/mm³)</td>
</tr>
<tr>
<td>CD8+ count (cells/mm³)</td>
</tr>
<tr>
<td>Viral load (×10³/mL) (copies/mL)</td>
</tr>
<tr>
<td>Plasma Calprotectin (µg/mL)</td>
</tr>
<tr>
<td>Neutrophil count (×10⁹/L)</td>
</tr>
<tr>
<td>Antimicrobial prophylaxis</td>
</tr>
<tr>
<td>Standard (%)</td>
</tr>
<tr>
<td>Enhanced (%)</td>
</tr>
<tr>
<td>Additional raltegravir</td>
</tr>
<tr>
<td>No Raltegravir (%)</td>
</tr>
<tr>
<td>Raltegravir (%)</td>
</tr>
<tr>
<td>Nutritional supplement</td>
</tr>
<tr>
<td>No RUSF (%)</td>
</tr>
<tr>
<td>RUSF (%)</td>
</tr>
</tbody>
</table>

Categorical data represented as number (percentage) and continuous data as mean (SD) for the normally distributed and as median (IQR) for the non-normally distributed. Abbreviations: BMI, body mass index; RUSF, ready to use supplementary food; CD, cluster of differentiation; HIV, human immunodeficiency virus.
Table 2. Association of non-randomised variables with log plasma calprotectin at baseline.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Beta</th>
<th>95% CI</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age per year</td>
<td>-0.02</td>
<td>-0.40 to -0.01</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Sex (female)</td>
<td>0.35</td>
<td>0.04 to 0.66</td>
<td>0.03</td>
</tr>
<tr>
<td>Site (Eldoret)</td>
<td>-0.22</td>
<td>-0.57 to -0.14</td>
<td>0.23</td>
</tr>
<tr>
<td>Log viral load</td>
<td>-0.01</td>
<td>-0.13 to 0.10</td>
<td>0.86</td>
</tr>
<tr>
<td>CD4+ 50-99/mm</td>
<td>Reference</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD4+ 25-49/mm</td>
<td>-0.76</td>
<td>-0.54 to 0.39</td>
<td>0.75</td>
</tr>
<tr>
<td>CD4+ 0-24/mm</td>
<td>0.06</td>
<td>-0.29 to 0.41</td>
<td>0.75</td>
</tr>
<tr>
<td>Neutrophils ×10³/L</td>
<td>0.19</td>
<td>0.12 to 0.27</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Cough</td>
<td>0.51</td>
<td>0.13 to 0.88</td>
<td>0.01</td>
</tr>
</tbody>
</table>

CD4+ T cell count was stratified into 0-24, 25-49, and 50-99 cells/mm³. The following variables were tested but excluded: CD8+ T cells, fever, tuberculosis, body temperature, diarrhoea, and cryptococcal antigen test.

Table 3. Hazards ratios for mortality in the first 4 and 24 weeks.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Deaths within 24 weeks</th>
<th>Deaths within 4 weeks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age per year</td>
<td>HR [95% CI] P</td>
<td>HR [95% CI] P</td>
</tr>
<tr>
<td>Sex (female)</td>
<td>1.04 (0.99 - 1.09) 0.06</td>
<td>1.04 (0.98 - 1.11) 0.04</td>
</tr>
<tr>
<td>BMI per kg/m²</td>
<td>0.84 (0.33 - 2.18) 0.73</td>
<td>0.29 (0.09 - 0.97) 0.19</td>
</tr>
<tr>
<td>Log viral load</td>
<td>0.83 (0.67 - 1.01) 0.07</td>
<td>0.94 (0.76 - 1.18) 0.60</td>
</tr>
<tr>
<td>Site (Eldoret)</td>
<td>0.97 (0.57 - 1.63) 0.91</td>
<td>0.91 (0.48 - 1.73) 0.77</td>
</tr>
<tr>
<td>Calprotectin per µg/ml</td>
<td>0.70 (0.23 - 2.17) 0.29</td>
<td>0.35 (0.06 - 2.14) 0.25</td>
</tr>
<tr>
<td>Information criteria</td>
<td>1.82 (1.08 - 3.08) 0.03</td>
<td>2.77 (1.58 - 4.88) &lt;0.001</td>
</tr>
</tbody>
</table>

Cox proportional hazards model stratified by three CD4 count groups and adjusted for the three randomised interventions. BMI: Body Mass Index.

Discussion

This pilot study focused on assessing plasma calprotectin mortality among patients with advanced HIV disease determined by a very low CD4+ cell count. Plasma calprotectin at the time of ART initiation was associated with subsequent mortality. It appeared that calprotectin levels at ART initiation are likely to be predictive of deaths occurring in the first 4 weeks. There was no evidence that the enhanced opportunistic infection intervention in the REALITY trial, which was associated with a 27% reduction in mortality to 24 weeks, (Hakim et al., 2017) affected changes in calprotectin during the first 4 weeks. However, the sample size for this exploratory analysis was limited. This pilot study requires validation in a larger sample and in other clinical settings before being used to guide further investigations or specific interventions.

To the best of our knowledge, this is the first study looking at plasma calprotectin in the context of mortality in advanced HIV. Previous studies have focused on faecal calprotectin as a biomarker of enteropathy, which is typically elevated in HIV-positive compared to HIV-negative individuals and progressively increases with a reduction in CD4+ T cell count (Hestvik et al., 2012). Faecal calprotectin is elevated in both early and chronic HIV infection and the elevated levels of faecal calprotectin have been associated with microbial translocation and enteropathy (Pastor et al., 2019).

Our results indicated a significant positive correlation between plasma calprotectin and neutrophil counts at baseline which has been observed previously (Cotoi et al., 2013; Sorensen et al., 2015; Sun et al., 2014) and likely reflects neutrophil expansion and activation as a main source of calprotectin (Chatzikonstantinou et al., 2016). Plasma calprotectin may be a marker of systemic inflammation as a result of microbial translocation (Deeks et al., 2013; Jonsson et al., 2017) or could indicate the presence of opportunistic infections among advanced HIV-positive patients, rather than only inflammation due to infection with HIV. However, markers of infection markers apart from cough were not associated with baseline plasma calprotectin.

The main strength of our study is that it was carried out in typical African hospital-based HIV clinics in which all eligible patients from both urban and peri-urban areas were recruited and is thus a reasonable reflection of advanced HIV patients in sub-Saharan Africa. The limitations of our study included the fact that this pilot study included a small number of subjects, young children were excluded from the parent trial, and despite weighting there was potential for bias during sample selection as included participants were required to have a full set of samples. Our study examined the potential of plasma calprotectin alone; future studies should compare this biomarker with other proteins, metabolites or cytokines to evaluate performance and help elucidate mechanisms underlying early mortality in advanced HIV disease.

Conclusions

Findings from this study suggest that plasma calprotectin may have value in predicting early mortality among HIV patients with advanced disease at the time of ART.
Initiation. Validation of plasma calprotectin as a clinical tool is needed before incorporating the biomarker to guide enhanced investigation for infections, more frequent follow up or specific interventions.

**Data availability**

**Underlying data**

De-identified REALITY trial data are available from MRC CTU at UCL, which encourages optimal use of data by employing a controlled access approach to data sharing, incorporating a transparent and robust system to review requests and provide secure data access consistent with the relevant ethics committee approvals. All requests for data are considered and can be initiated by contacting mrcctu.ctuenquiries@ucl.ac.uk Quoting “REALITY Trial Immunology”.

**Acknowledgements**

We acknowledge Diana Gibb, Sarah Walker and Alex Szubert for their leadership of the REALITY trial and assistance in case-cohort design, and the coordination, site staff and participants of the REALITY trial. We also thank the Training Department at KWTRP, and the Department of Molecular Biology and Biotechnology, PAUSTI, Juja, Kenya. This manuscript is published with the permission of the Director, Kenya Medical Research Institute (KEMRI).

**Table 4. Change in plasma calprotectin between baseline and four weeks.**

<table>
<thead>
<tr>
<th></th>
<th>Median Calprotectin µg/ml at baseline</th>
<th>Median Calprotectin µg/ml at 4 weeks</th>
<th>Median* change from baseline to 4 weeks</th>
<th>IQR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cases</strong></td>
<td>8.03</td>
<td>5.78</td>
<td>-0.02</td>
<td>-5.40 to +0.68</td>
</tr>
<tr>
<td><strong>Non-cases</strong></td>
<td>5.03</td>
<td>4.94</td>
<td>-0.72</td>
<td>-5.00 to +1.27</td>
</tr>
</tbody>
</table>

Only participants alive at 4 weeks are shown.

* For the change between baseline and 4 weeks in all participants P=0.005; change in cases vs. non-cases P=0.38 adjusted for age, sex, site and randomised interventions.

**Figure 2. Plasma calprotectin at baseline and at four weeks.** Participants sorted by baseline calprotectin values (high to low)
References


Published Abstract | Publisher Full Text | Free Full Text


Published Abstract | Publisher Full Text


Published Abstract | Publisher Full Text


Published Abstract | Publisher Full Text | Free Full Text


Published Abstract | Publisher Full Text | Free Full Text


Published Abstract | Publisher Full Text | Free Full Text


Published Abstract | Publisher Full Text | Free Full Text


Published Abstract | Publisher Full Text | Free Full Text


Published Abstract | Publisher Full Text | Free Full Text


Published Abstract | Publisher Full Text | Free Full Text


Published Abstract | Publisher Full Text | Free Full Text


Published Abstract | Publisher Full Text | Free Full Text


Published Abstract | Publisher Full Text | Free Full Text


Published Abstract | Publisher Full Text | Free Full Text


Published Abstract | Publisher Full Text | Free Full Text


Published Abstract | Publisher Full Text | Free Full Text


Published Abstract | Publisher Full Text | Free Full Text


Published Abstract | Publisher Full Text | Free Full Text


Published Abstract | Publisher Full Text | Free Full Text


Published Abstract | Publisher Full Text | Free Full Text


Published Abstract


Published Abstract | Publisher Full Text | Free Full Text


Published Abstract | Publisher Full Text | Free Full Text


Published Abstract | Publisher Full Text | Free Full Text


Published Abstract | Publisher Full Text


Published Abstract | Publisher Full Text | Free Full Text
Open Peer Review

Current Peer Review Status:  ?  ?

Version 1

Reviewer Report 26 June 2020

https://doi.org/10.21956/wellcomeopenres.17044.r39054

© 2020 Shivakoti R. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Rupak Shivakoti
Columbia University Mailman School of Public Health, New York, NY, USA

This is a pilot study assessing the relationship of plasma calprotectin, a marker of inflammation, with mortality among individuals with advanced HIV initiating treatment. The authors use a nested case-cohort design to measure inflammation at time of ART initiation and use data from the parent study to assess mortality during treatment. Their results show that plasma calprotectin could be a useful marker for subsequent mortality, especially within the next 4 weeks, in individuals with advanced HIV. This is a clearly written manuscript and the design seems appropriate. There are, however, a few major questions that should be addressed.

Comments:

○ Given differences in immunity for children compared to adults, the percentage of children in this study should be specified and it would be helpful to know whether the baseline levels of calprotectin were different between children and adults. It is likely that the study is underpowered to look at association with mortality separately for children and adults but the authors should discuss whether combining them (including under 5 as suggested in discussion) in one analysis might or might not make sense for this specific marker.

○ It is not clear why 45% of Kilifi participants vs. only 10% of Eldoret were selected for the random subcohort. It is also not clear how this case-cohort was selected from the larger immunology case-cohort.

○ The case-cohort design has a random subcohort. How many cases out of 39 were part of the random subcohort? It is not clear from reading this what the sample size of the random subcohort is.

○ In the methods, the authors say: “The linear regression to examine whether subsequent death and the 3 randomized interventions were associated with the change in plasma calprotectin between baseline and 4 weeks”

  ○ The mortality analysis is confusing as written. I don't think the results were presented in the manuscript? For example, it is not clear what the exposure and outcome
variable for this analysis is. If the exposure is change in plasma calprotectin, then the outcome is presumably death after 4 weeks (which would exclude many deaths) and a linear regression does not seem appropriate. But if the exposure is death (i.e. before 4 weeks), then a 4 week sample would not be available.

- Sample size should be stated for the tables and figures. Was the random subcohort used in any of the presented analysis? E.g. due to random selection, it may be most useful for Table 2 analysis if there are enough cases in the subcohort.

- Given the small sample size, selecting covariates for multivariable model based on significance on univariate models can present some issues. Lack of significance could be due to the limited power for this specific study although a real association may exist.

- Minor: It is useful to also see the hazards and 95% CI (not just p-values) in text of results instead of just the table.

- Discussion should include whether this marker could have utility for individuals with less advanced HIV.

**Is the work clearly and accurately presented and does it cite the current literature?**
Yes

**Is the study design appropriate and is the work technically sound?**
Partly

**Are sufficient details of methods and analysis provided to allow replication by others?**
Partly

**If applicable, is the statistical analysis and its interpretation appropriate?**
Partly

**Are all the source data underlying the results available to ensure full reproducibility?**
Yes

**Are the conclusions drawn adequately supported by the results?**
Yes

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** HIV; Inflammation; Epidemiology; HIV treatment outcomes.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.
This study presents results of a pilot study investigating the association between plasma calprotectin at ART initiation with mortality among individuals with advanced HIV. Investigating the use of calprotectin as a prognostic biomarker for advanced HIV could inform its use as a tool in clinical practice to intervene and better prevent early mortality among individuals with advanced HIV initiating ART. The results suggest that plasma calprotectin was associated with mortality within 4 week and within 24 weeks. The case cohort design allowed the researchers to investigate this question accurately without needing samples analyzed from the entire REALITY Trial. The objectives of this pilot study were clearly stated and their results suggest calprotectin could be useful in addressing early mortality among those with advanced HIV initiating ART, an important need when many individuals continue to be diagnosed with HIV late in disease progression due to an opportunistic infection or other late-stage HIV related complication. However, to ensure this article is scientifically sound, there were several areas that could use more clarity or editing. Specifically, there were several results that were lacking associated methods or could use additional clarity. My comments are outlined in order by section below and include both major comments to be addressed in revisions and minor comments related to formatting/typos.

Abstract:

Background:

Major comments:
- Please specify how “ART initiation” is defined? Is the plasma collected prior to receiving any medication?

Methods:

Minor comment:
- Space in Trial ID number before closing parentheses.

Major comments:
- Please clarify how non-cases were selected. It states non-cases were randomly selected among those who survived for 48 weeks. Should this state for at least 48 weeks? See additional comments below for the main body methods section.
- The final sentence “To test association with mortality…” is confusing. Suggest rewording. Also, how was clustering by CD4 count decided/done?

Results:

Major comments:
- The authors described variables associated with baseline calprotectin, but this analysis is not described in the methods. Please add methods for these results in the methods section and explain why this was done.
- Results mention deaths within 4 weeks as an outcome, but this is not include in the
methods. Please ensure all results have methods described.

**Introduction:**
No comments

**Methods:**
**Study design:**
Major comments:
- Please clarify why there were 45% randomly selected from Kilifi but only 10% randomly selected from Eldoret? How were these percentages determined for selection?

Minor comment:
- Please define “baseline” in the context of calprotectin samples collected at baseline and 4 weeks after treatment initiation. So baseline is prior to ART initiation, yes? Please clarify.

**Statistical Analysis:**
Major comments:
- Please clarify the linear regression analysis. Was the linear regression model was univariable or was multivariable? This seems to be describing stepwise model building, yes? It states that the authors only retained those variables which were statistically significant. How was the order of variable addition determined? Generally, model building should not rely solely on statistical significance. Variables should be determined on the basis of plausibility a priori as well. A variable may be associated with an outcome, but especially in a small pilot, this might not be statistically significant due to insufficient power, so significance should not be the determining factor for its inclusion in the final model. Important to look at point estimates in addition to p-values and known relationships between the variables in identifying confounders to include in subsequent adjusted analyses.

- Please clarify in the Cox proportional hazard regression, how was this stratified by CD4 count. What strata were used and how were these determined?

- Please clarify the following in relation to survival time: the abstract states cases died within 24 weeks and non-cases were those who survived for 48 weeks. Manuscript states time at risk was 24 weeks and there is no mention of 48 weeks. What about those who died between 24 and 48 weeks? And the results present 4 weeks as an outcome end point. Please ensure all results have associated methods and that the Abstract has consistent language with the manuscript main body.

- Why was the Wilcoxon paired sign-rank test was performed (final paragraph of Statistical Analysis). Please explain the objective/purpose of this analysis and how the covariates included in the model were determined.

**Results**
**Characteristics of study participants:**
Minor comments:
- Is enrollment misspelled (enrolment)?

- Final sentence is grammatically incorrect. I think deleting the “lower” before “hemoglobin” would correct this.
Baseline plasma calprotectin:

Minor comments:
- Please define what “crude” means.

Major comments:
- The following language should be adjusted to reduce reliance on p-values for scientific interpretation: “other putative markers of infection (CD8+ T cells, fever, tuberculosis, body temperature, diarrhoea, and cryptococcal antigen test) did not emerge as factors affecting plasma calprotectin levels.” Please change this language to specify these were not “statistically significant.” This could be a result of low power. The point estimates should also be discussed, especially since this is a small pilot.

- Table 2: 95%CI for site (Eldoret) implies statistical significance (-0.57 to -0.14) but p-value is 0.23. Is this row accurate?

- Table 2: Instead of listing the beta, can you please specifically describe what this beta means instead? This is the difference between mean calprotectin levels for a 1 unit change in the variables listed, right? What is the interpretation of this point estimate?

- Table 3: States model was stratified by CD4 counts. How was this stratified and what were the results of the stratified analysis? What are the strata and how were they selected? This should be discussed in the methods.

Plasma calprotectin after 4 weeks:

- Please show or at least describe the following data: “There was no evidence that any of the three randomized interventions affected the change in calprotectin between baseline and 4 weeks (data not shown).” Rephrase to describe lack of statistical significance and also discuss the point estimates.

Discussion:

Minor comment:
- The following could be clarified: “This pilot study focused on assessing plasma calprotectin mortality among patients with advanced HIV disease determined by a very low CD4+ cell count.” Reword to include more detail: “This pilot study focused on assessing the association with pre-ART plasma calprotectin and mortality within 24 weeks among patients with advanced HIV disease defined by a very low CD4+ cell count (<100 cells/µl).”

Major comments:
- The following is confusing: “It appeared that calprotectin levels at ART initiation are likely to be predictive of deaths occurring in the first 4 weeks.” Please use more concrete language to describe your main finding. Also, it is unclear if 24 weeks or 4 weeks is your survival end point. Calprotectin levels were associated with both, right?

- Please describe your limitations specifically. “However, the sample size for this exploratory analysis was limited. This pilot study requires validation in a larger sample and in other clinical settings before being used to guide further investigations or specific interventions.” Should be explicit about your limitations, how they may have impacted your results, and thus why next steps are needed. You do this at the end of the discussion section, so confusing why you mention this in two places. This is also the first time the term “exploratory” is used in the manuscript. Please include this in the methods so it’s clear what you mean by this term.
Please adjust language in the following: “However, markers of infection markers apart from cough were not associated with baseline plasma calprotectin.” The phrase “not associated” is not ideal. Better to state: “However, markers of infection markers apart from cough were not statistically significantly associated with baseline plasma calprotectin.” And discuss the point estimate. Why do you think cough associated but others not? Is this just an issue of multiple comparisons and a spurious association or do you have insufficient power to detect other infection markers? I think this is worth further discussion.

The authors mention “…potential for bias during sample selection as included participants were required to have a full set of sample” but you state earlier that you are confident this was missing at random, so this is assumed not to have biased your results, right? Please include this information here if you are confident in the MAR assumption and it’s impact on your analysis.

Can you please discuss if it will be feasible for the ELISA test to be used in low/middle income locations where it could be of most benefit due to high rates of ART initiation at an advanced HIV stage? Do you think ART clinics have the infrastructure and resources to perform this test to inform patient care or are costs and lab resources potentially another barrier?

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Partly

Are sufficient details of methods and analysis provided to allow replication by others?
Partly

If applicable, is the statistical analysis and its interpretation appropriate?
Partly

Are all the source data underlying the results available to ensure full reproducibility?
Partly

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Epidemiology, HIV treatment and prevention, mortality following ART initiation. This topic was a focus of my recent dissertation work, though I have since switched to other topics related to other STDs and specific issues related to health in Southwest Virginia in the United States, so my knowledge on this specific topic may be a couple of years out of date.

I confirm that I have read this submission and believe that I have an appropriate level of
expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.