Check for updates

RESEARCH ARTICLE

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

Faith W. Kamau[1,2], Agnes Gwela[1,3], Andrew K. Nyerere[4], Victor Riitho[5], James M. Njunge[1,3], Moses M. Ngari[1,3], Andrew J. Prendergast[5], James A. Berkley[1,3,6]

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

v2 **First published:** 13 Mar 2020, **5**:46
https://doi.org/10.12688/wellcomeopenres.15563.1
Latest published: 27 Nov 2020, **5**:46
https://doi.org/10.12688/wellcomeopenres.15563.2

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 collected before ART treatment and mortality among individuals with advanced HIV. **Methods:** We conducted a pilot case-cohort study among HIV infected adults and adolescents over 13 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 collected at baseline (before ART) and after 4 weeks of treatment was measured in archived plasma of those who died within 24 weeks (cases) and randomly selected participants who survived (non-cases). Association with mortality was assessed using 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 (0-24, 25-49, and 50-99 cells/mm3). **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

Open Peer Review

Reviewer Status ✓ ✓

	Invited Reviewers	
	1	**2**
version 2 (revision) 27 Nov 2020	✓ report	✓ report
version 1 13 Mar 2020	? report	? report

1. **Rachel A. Silverman**, Virginia Polytechnic Institute and State University, Blacksburg, USA

2. **Rupak Shivakoti**, Columbia University Mailman School of Public Health, New York, USA

Any reports and responses or comments on the article can be found at the end of the article.

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. **Conclusions:** 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

KEMRI wellcome

This article is included in the KEMRI | Wellcome Trust gateway.

Corresponding author: James A. Berkley (jberkley@kemri-wellcome.org)

Author roles: Kamau FW: Conceptualization, Formal Analysis, Investigation, Methodology, Writing – Original Draft Preparation; **Gwela A** : Conceptualization, Formal Analysis, Investigation, Methodology, Supervision, Writing – Review & Editing; **Nyerere AK**: Methodology, Project Administration, Supervision, Writing – Review & Editing; **Riitho V**: Conceptualization, Investigation, Methodology, Project Administration, Writing – Review & Editing; **Njunge JM**: Conceptualization, Formal Analysis, Investigation, Methodology, Project Administration, Resources, Validation, Writing – Review & Editing; **Ngari MM**: Data Curation, Formal Analysis, Writing – Review & Editing; **Prendergast AJ**: Conceptualization, Funding Acquisition, Methodology, Project Administration, Resources, Supervision, Writing – Review & Editing; **Berkley JA**: Conceptualization, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Supervision, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This work was supported by the Wellcome Trust through funding to the Medical Research Council (MRC)/ UK Department for International Development (DFID)/Wellcome Global Health Trials Scheme which supports AJP [108065] and JAB [MR/M007367/1]. FK was supported through IDeAL, a programme of the DELTAS Africa Initiative [DEL-15-003]. The DELTAS Africa Initiative is an independent funding scheme of the African Academy of Sciences (AAS)'s Alliance for Accelerating Excellence in Science in Africa (AESA) and supported by the New Partnership for Africa's Development Planning and Coordinating Agency (NEPAD Agency) with funding from Wellcome [107769] and the UK government. JMN, MN, AG, JAB were supported by the Bill & Melinda Gates Foundation [OPP1131320]. The REALITY trial was funded by the Joint Global Health Trials Scheme (JGHTS) of the UK DFID, the Wellcome Trust, and the UK MRC [G1100693]. Additional funding support was provided by the PENTA foundation and core support to the MRC Clinical Trials Unit at University College London (London, UK) [MC_UU_12023/23, MC_UU_12023/26].
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2020 Kamau FW *et al*. This is an open access article 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.

How to cite this article: Kamau FW, Gwela A, Nyerere AK *et al.* **Plasma calprotectin as a biomarker of mortality at antiretroviral treatment initiation in advanced HIV – pilot study [version 2; peer review: 2 approved]** Wellcome Open Research 2020, **5**:46 https://doi.org/10.12688/wellcomeopenres.15563.2

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

REVISED **Amendments from Version 1**

Abstract, in methods, clarified sub-study minimum age 13 years (R2); that calprotectin was measured before ART initiation; that mortality to 24 weeks was used (as per primary endpoint of trial) rather than 48 weeks extended follow up (R1&2); the analysis is more succinctly described; in results, effect sizes 'for death' changed to 'for subsequent mortality' and that 4-week changes involved cases who survived to 4 weeks and controls.

Introduction, clarified: paragraph 1, outcomes are despite starting ART; paragraph 3, we took samples before ART (R1&2).

Methods: *study population*: paragraph 1, description of the parent trial moved up and in paragraph 2, moved up the names of the sites for this sub-study; *study design*: paragraph 1, reworded participant selection percentages (R1&2); *study design*, paragraph 2 and *ELISA tests*, paragraph 1: stated baseline samples taken prior to ART initiation.

Methods, *statistical analysis*: clarified regression analysis indicating it was stepwise, including variables with biological plausibility, retaining variables with P<0.1 (R1); better explaining stratification and methods for examining randomised interventions in the last paragraph.

Results, under *baseline plasma calprotectin*, added regression coefficients to the text, reworded infection markers (R1), and in Table 2 added interpretations (R1); under *association with mortality*, added hazards ratios to the text (R2); Table 3, added CD4 strata to the footnote (R1); under *plasma calprotectin after 4 weeks*, added regression coefficients to the text (R2); Table 4, added row for all participants and clarified cases must have survived 4 weeks to be included here (R1), indicating N=78 below Table 4 and Figure 2.

Discussion, paragraph 1, clarification of 24-week mortality and moved up comments on antimicrobial prophylaxis; paragraphs 3 & 4, expanded discussion of sample size and other limitations (R1&2); under *conclusions*, clarified sample timing and added potential practical usage (R1).

Any further responses from the reviewers can be found at the end of the article

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, C-reactive protein 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 starting ART treatment (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 subunit 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 prior to 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

The REALITY trial (ISRCTN43622374) enrolled HIV-infected adults and children aged five years or more with a $CD4^+$ T cell count <100 cells/mm^3 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).

This pilot study capitalised on a broader ongoing immunology case-cohort sub-study that included study participants aged 13 years or more with a sample set of plasma, baseline stool, PBMCs and data at two Kenyan sites: Kilifi County Hospital and the Academic Model for the Prevention and Treatment of HIV/AIDS Centre at Moi Teaching Referral Hospital, Eldoret.

Study design

This pilot was a case-cohort study. The REALITY trial had enrolled a total of 139 participants in Kilifi, of whom 29 (20%)

died, and 195 in Eldoret, of whom 14 (7.2%) died (Figure 1). 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, and weighted (see below) to reflect the original trial population. The immunology case-cohort sub-study randomly selected 45% of all participants at Kilifi and 10% of study participants at Eldoret as a sub-cohort in order to reflect enrolment and mortality in the full REALITY trial, stratified by $CD4^+$ count (0–24, 25–49, and 50–99 cells/mm^3) to avoid imbalance in this exposure, plus any remaining unselected deaths by week 24 (the trial primary endpoint). Non-cases were randomly selected from those who survived through 24 weeks from the trial database using the uniform random number function in STATA.

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 24 (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. Samples analysed were those prior to receiving ART at baseline, then 4 weeks after starting ART.

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, before ART treatment was initiated, and at four weeks after ART initiation. The four-week time period allowed an

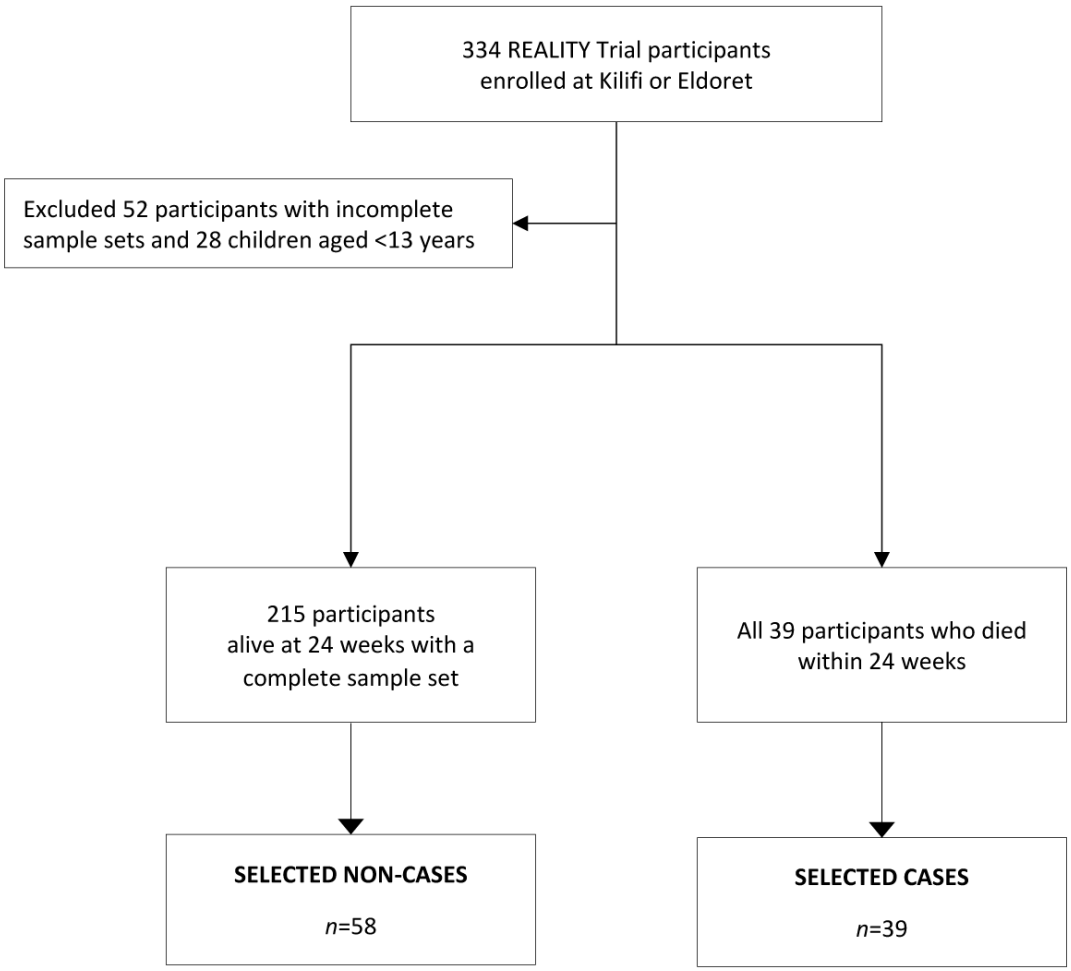

Figure 1. Participant selection.

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 stepwise 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 with biological plausibility: 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 at P<0.1. The beta coefficient indicated the strength of the effect on plasma calprotectin.

The association between baseline plasma calprotectin and mortality was assessed using a stratified Cox proportional hazards regression model with inverse probability weights (Buchanan *et al.*, 2014) with three strata by CD4 count (0–24, 25–49, and 50–99 cells/mm^3) 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. Mortality within the first four weeks, when most deaths occurred, was also assessed. 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. Generalised linear modelling, adjusted for age, sex and site and the three randomised interventions and stratified by $CD4^+$ group as above was used to examine relative risk for subsequent mortality.

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, 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, unadjusted 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 (β=-0.02; 95% CI -0.40 to -0.01; P=<0.01) and there was positive association with female sex (β=-0.35; 95% CI 0.04 to 0.66; P=0.03), neutrophil count (β=-0.19; 95% CI 0.12 to 0.27; P=<0.01) and reporting a cough (β=-0.51; 95% CI -0.13 to 0.88; P=0.01) (Table 2). Other putative markers of infection ($CD8^+$ T cells, fever, tuberculosis, body temperature, diarrhoea, and cryptococcal antigen test) were not statistically significant when their effect on plasma calprotectin levels was tested.

Association with mortality

Baseline calprotectin was significantly associated with subsequent mortality to 24 weeks (HR 1.82 (95% CI 1.08 to 3.08), *P*=0.03) and mortality within the first 4 weeks (HR 2.77 (95% CI 1.58 to 4.88), *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 and non-cases (Table 4). The three randomised interventions did not have statistically significant effects on the change in calprotectin between baseline and 4 weeks: additional raltegravir ß1.02 (95% CI -2.26 to 4.29); enhanced antimicrobial prophylaxis ß -0.55 (95% CI -3.86 to 2.77); ready-to-use supplementary food ß 2.18 (95% CI -1.153 to 5.52). Two cases who died after 4 weeks (10%) and four non-cases (7%) had a >2-fold rise in plasma calprotectin between baseline and 4 weeks (Figure 2). Overall, change in plasma calprotectin between baseline and 4 weeks was not associated with subsequent mortality: RR 1.02 (95% CI 0.93 - 0.11) per µg/ml (P=0.664).

Discussion

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 (<100cells/mm^3). Plasma calprotectin at the time of ART initiation was associated with subsequent mortality within 24 weeks as well as 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. Interestingly, although none of the three interventions significantly affected the change in plasma calprotectin in this pilot study, the point estimate for enhanced antimicrobial prophylaxis, which was associated with

Table 1. Participant characteristics.

Characteristic	Cases (n=39)	Non-Cases (n=58)	P Value
Site			0.55
Kilifi (%)	26 (67)	42 (72)	
Eldoret (%)	13 (33)	16 (28)	
Sex			0.84
Male (%)	18 (46)	28 (48)	
Female (%)	21 (54)	30 (52)	
Age (years)	41 (7.5)	39 (10.5)	0.35
BMI (kg/m^2)	16.9 (3.5)	18.3 (2.7)	0.04
Full blood count			
Haemoglobin (g/dL)	9.3 (7.3 - 10.5)	9.9 (9.0 - 11.8)	0.03
CD4$^+$ count (cells/mm^3)	22 (8 - 44)	21 (11 - 64)	0.33
CD8$^+$ count (cells/mm^3)	377 (242 - 707)	646 (410 - 931)	0.02
Viral load (×10^3/mL) (copies/mL)	255 (132 - 760)	254 (115 - 521)	0.80
Plasma Calprotectin (µg/mL)	6.82 (2.65 – 12.5)	5.01 (1.92 – 11.5)	0.23
Neutrophil count (×10^9/L)	2.26 (1.29 - 3.66)	1.77 (1.26 - 2.71)	
Antimicrobial prophylaxis			0.43
Standard (%)	24 (62)	31 (53)	
Enhanced (%)	15 (38)	27 (47)	
Additional raltegravir			0.30
No Raltegravir (%)	18 (46)	33 (57)	
Raltegravir (%)	21 (54)	25 (43)	
Nutritional supplement			0.02
No RUSF (%)	12 (31)	32 (55)	
RUSF (%)	27 (69)	26 (45)	

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.

reduced mortality in the parent trial (Hakim *et al.*, 2017) was in a negative direction. 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.*, 2014; Sorensen *et al.*, 2015;

Table 2. Multivariable stepwise linear regression analysis of non-randomised variables with log plasma calprotectin at baseline.

Variable	Beta with interpretation	95% CI	P value
Age per year	-0.02 (calprotectin is lower in older participants)	-0.40 to -0.01	<0.01
Sex (female)	0.35 (females had higher calprotectin than males)	0.04 to 0.66	0.03
Site (Eldoret)	-0.22 (calprotectin was lower in Eldoret than Kilifi)	-0.57 to 0.14	0.23
Log viral load	0.01 (viral load was not associated with calprotectin)	-0.13 to 0.10	0.86
CD4+ 50-99/mm^3	Reference		
CD4+ 25-49/mm^3	-0.76 (calprotectin did not vary by CD4$^+$ strata)	-0.54 to 0.39	0.75
CD4+ 0-24/mm^3	0.06 (calprotectin did not vary by CD4$^+$ strata)	-0.29 to 0.41	0.75
Neutrophils ×109/L	0.19 (calprotectin is positively associated with neutrophil count)	0.12 to 0.27	<0.01
Cough	0.51 (calprotectin was higher in patients with cough)	0.13 to 0.88	0.01

N=95. CD4$^+$ T cell count was stratified into 0-24, 25-49, and 50-99 cells/mm^3. 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.

	Deaths within 24 weeks		Deaths within 4 weeks	
Variable	**HR [95% CI]**	**P**	**HR (95% CI)**	**P**
Age per year	1.04 (0.99 - 1.09)	0.06	1.04 (0.98 – 1.11)	0.04
Sex (female)	0.84 (0.33 – 2.18)	0.73	0.29 (0.09 – 0.97)	0.19
BMI per kg/m^2	0.83 (0.67 - 1.01)	0.07	0.94 (0.76 - 1.18)	0.60
Log viral load	0.97 (0.57 - 1.63)	0.91	0.91(0.48 - 1.73)	0.77
Site (Eldoret)	0.70 (0.23 – 2.17)	0.29	0.35 (0.06 – 2.14)	0.25
Calprotectin per µg/ml	1.82 (1.08 – 3.08)	0.03	2.77 (1.58 - 4.88)	<0.001
Information criteria	AIC 87.9; BIC 111		AIC 41.9; BIC 65.0	

Cox proportional hazards model stratified by three CD4 count groups (0-24, 25-49, and 50-99 cells/mm^3) and adjusted for the three randomised interventions. BMI: Body Mass Index.

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 statistically significantly associated with baseline plasma calprotectin. This could result due to the small number of samples tested, hence insufficient power.

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. Moreover, the use of ELISA is feasible in these setups, and already in application as a confirmatory test for HIV positive diagnosis. Besides the sample size of this pilot study, limitations included that despite weighting there was potential for bias during sample selection as included participants were required to have a full set of samples, which varied by site, and children under thirteen years old were excluded. The parent trial only enrolled patients with advanced disease, therefore we recommend that future studies be carried on patients with less advanced HIV to elucidate whether plasma calprotectin has similar predictive value, along with other proteins, metabolites and cytokines.

Conclusions

Findings from this pilot study suggest that plasma calprotectin at the time of ART initiation has value in predicting early mortality among HIV patients with advanced disease. There is at least one quantitative serum calprotectin lateral-flow test available, and ELISA- or lateral flow-based tests may be useful in patient care. However, further 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.

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

	Median Calprotectin µg/ml at baseline	Median Calprotectin µg/ml at 4 weeks	Median* change from baseline to 4 weeks	IQR
All participants alive at 4 weeks	6.70	5.52	-0.58*	-5.20 to +0.96
Cases that died after 4 weeks	8.03	5.78	-0.02	-5.40 to +0.68
Non-cases	5.03	4.94	-0.72	-5.00 to +1.27

Only participants alive at 4 weeks are shown (N=78).

* 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.

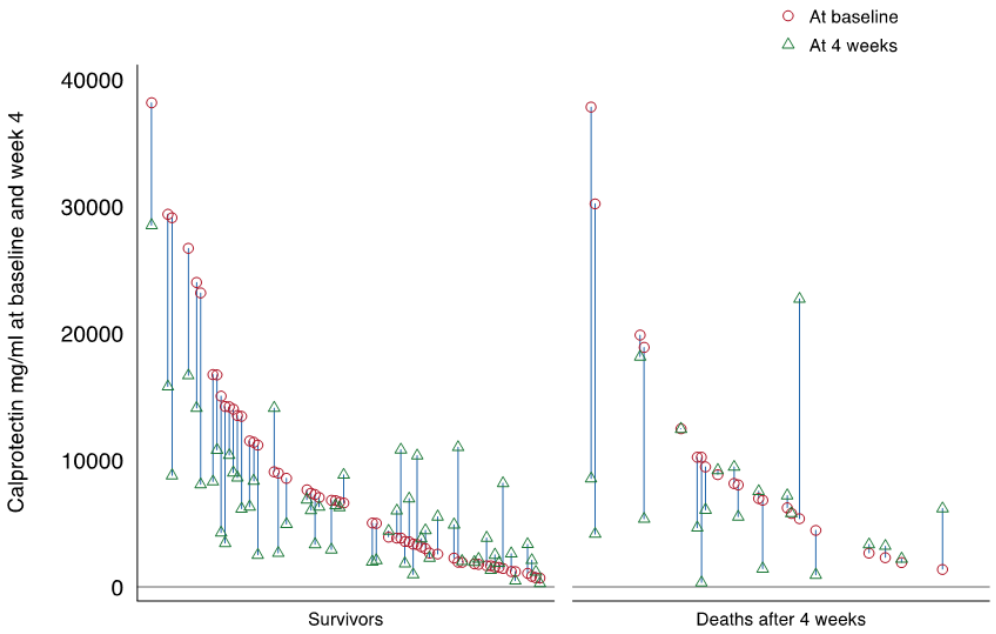

Figure 2. Plasma calprotectin at baseline and at four weeks. Participants sorted by baseline calprotectin values (high to low), N=78.

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).

References

Banerjee A, Srinivas M, Eyre R, *et al.*: **Faecal calprotectin for differentiating between irritable bowel syndrome and inflammatory bowel disease: a useful screen in daily gastroenterology practice.** *Frontline Gastroenterol.* 2015; **6**(1): 20–26.
PubMed Abstract | Publisher Full Text | Free Full Text

Bjarnason I: **The Use of Fecal Calprotectin in Inflammatory Bowel Disease.** *Gastroenterol Hepatol (N Y).* 2017; **13**(1): 53–56.
PubMed Abstract | Free Full Text

Brenchley JM, Price DA, Schacker TW, *et al.*: **Microbial translocation is a cause of systemic immune activation in chronic HIV infection.** *Nat Med.* 2006; **12**(12): 1365–71.
PubMed Abstract | Publisher Full Text

Brophy MB, Nolan EM: **Manganese and microbial pathogenesis: sequestration by the Mammalian immune system and utilization by microorganisms.** *ACS Chem Biol.* 2015; **10**(3): 641–651.
PubMed Abstract | Publisher Full Text | Free Full Text

Buchanan AL, Hudgens MG, Cole SR, *et al.*: **Worth the weight: using inverse probability weighted Cox models in AIDS research.** *AIDS Res Hum Retroviruses.* 2014; **30**(12): 1170–1177.
PubMed Abstract | Publisher Full Text | Free Full Text

Chatzikonstantinou M, Konstantopoulos P, Stergiopoulos S, *et al.*: **Calprotectin as a diagnostic tool for inflammatory bowel diseases.** *Biomed Rep.* 2016; **5**(4): 403–407.
PubMed Abstract | Publisher Full Text | Free Full Text

Colon K, Perez-Laspiur J, Quiles R, *et al.*: **Macrophage secretome from women with HIV-associated neurocognitive disorders.** *Proteomics Clin Appl.* 2016; **10**(2): 136–143.
PubMed Abstract | Publisher Full Text | Free Full Text

Cotoi OS, Dunér P, Ko N, *et al.*: **Plasma S100A8/A9 correlates with blood neutrophil counts, traditional risk factors, and cardiovascular disease in middle-aged healthy individuals.** *Arterioscler Thromb Vasc Biol.* 2014; **34**(1): 202–210.
PubMed Abstract | Publisher Full Text

Deeks SG, Tracy R, Douek DC: **Systemic effects of inflammation on health during chronic HIV infection.** *Immunity.* 2013; **39**(4): 633–645.
PubMed Abstract | Publisher Full Text | Free Full Text

Drozd DR, Graham SM, Crane HM, *et al.*: **Short Communication: Effect of Antiretroviral Therapy on Circulating Damage-Associated Molecular Pattern Molecules and CD4 Immune Reconstitution in HIV-Infected Individuals.** *AIDS Res Hum Retroviruses.* 2016; **32**(9): 876–878.
PubMed Abstract | Publisher Full Text | Free Full Text

Duprez DA, Neuhaus J, Kuller LH, *et al.*: **Inflammation, coagulation and cardiovascular disease in HIV-infected individuals.** *PLoS One.* 2012; **7**(9): e44454.
PubMed Abstract | Publisher Full Text | Free Full Text

Fitzgerald FC, Lhomme E, Harris K, *et al.*: **Microbial Translocation Does Not Drive Immune Activation in Ugandan Children Infected With HIV.** *J Infect Dis.* 2019; **219**(1): 89–100.
PubMed Abstract | Publisher Full Text | Free Full Text

Hakim J, Musiime V, Szubert AJ, *et al.*: **Enhanced Prophylaxis plus Antiretroviral Therapy for Advanced HIV Infection in Africa.** *N Engl J Med.* 2017; **377**(3): 233–245.
PubMed Abstract | Publisher Full Text | Free Full Text

Hestvik E, Olafsdottir E, Tylleskar T, *et al.*: **Faecal calprotectin in HIV-infected, HAART-naïve Ugandan children.** *J Pediatr Gastroenterol Nutr.* 2012; **54**(6): 785–90.
PubMed Abstract | Publisher Full Text

Huang L, Li J, Han Y, *et al.*: **Serum Calprotectin Expression as a Diagnostic Marker for Sepsis in Postoperative Intensive Care Unit Patients.** *J Interferon Cytokine Res.* 2016; **36**(10): 607–616.
PubMed Abstract | Publisher Full Text

Hunt PW, Lee SA, Siedner MJ: **Immunologic Biomarkers, Morbidity, and Mortality in Treated HIV Infection.** *J Infect Dis.* 2016; **214**(Suppl 2): S44–50.
PubMed Abstract | Publisher Full Text | Free Full Text

IeDEA and ART Cohort Collaborations, Avila D, Althoff KN, *et al.*: **Immunodeficiency at the start of combination antiretroviral therapy in low-, middle-, and high-income countries.** *J Acquir Immune Defic Syndr.* 2014; **65**(1): e8–e16.
PubMed Abstract | Publisher Full Text | Free Full Text

Jonsson MK, Sundlisaeter NP, Nordal HH, *et al.*: **Calprotectin as a marker of inflammation in patients with early rheumatoid arthritis.** *Ann Rheum Dis.* 2017; **76**(12): 2031–2037.
PubMed Abstract | Publisher Full Text

Kityo C, Szubert AJ, Siika A, *et al.*: **Raltegravir-intensified initial antiretroviral therapy in advanced HIV disease in Africa: A randomised controlled trial.** *PLoS Med.* 2018; **15**(12): e1002706.
PubMed Abstract | Publisher Full Text | Free Full Text

Kuller LH, Tracy R, Belloso W, *et al.*: **Inflammatory and coagulation biomarkers and mortality in patients with HIV infection.** *PLoS Med.* 2008; **5**(10): e203.
PubMed Abstract | Publisher Full Text | Free Full Text

Mallewa J, Szubert AJ, Mugyenyi P, *et al.*: **Effect of ready-to-use supplementary food on mortality in severely immunocompromised HIV-infected individuals in Africa initiating antiretroviral therapy (REALITY): an open-label, parallel-group, randomised controlled trial.** *Lancet HIV.* 2018; **5**(5): e231–e240.
PubMed Abstract | Publisher Full Text | Free Full Text

Marchetti G, Tincati C, Silvestri G: **Microbial translocation in the pathogenesis of HIV infection and AIDS.** *Clin Microbiol Rev.* 2013; **26**(1): 2–18.
PubMed Abstract | Publisher Full Text | Free Full Text

Muller F, Frøland SS, Aukrust P, *et al.*: **Elevated serum calprotectin levels in HIV-infected patients: the calprotectin response during ZDV treatment is associated with clinical events.** *J Acquir Immune Defic Syndr (1988).* 1994; **7**(9): 931–939.
PubMed Abstract

Nazli A, Chan O, Dobson-Belaire WN, *et al.*: **Exposure to HIV-1 directly impairs mucosal epithelial barrier integrity allowing microbial translocation.** *PLoS Pathog.* 2010; **6**(4): e1000852.
PubMed Abstract | Publisher Full Text | Free Full Text

Njunge JM, Gwela A, Kibinge NK, *et al.*: **Biomarkers of post-discharge mortality among children with complicated severe acute malnutrition.** *Sci Rep.* 2019; **9**(1): 5981.
PubMed Abstract | Publisher Full Text | Free Full Text

Pastor L, Langhorst J, Schröder D, *et al.*: **Different pattern of stool and plasma gastrointestinal damage biomarkers during primary and chronic HIV infection.** *PLoS One.* 2019; **14**(6): e0218000.
PubMed Abstract | Publisher Full Text | Free Full Text

Sandler NG, Douek DC: **Microbial translocation in HIV infection: causes, consequences and treatment opportunities.** *Nat Rev Microbiol.* 2012; **10**(9): 655–666.
PubMed Abstract | Publisher Full Text

Serrano-Villar S, Sainz T, Lee SA, *et al.*: **HIV-infected individuals with low CD4/CD8 ratio despite effective antiretroviral therapy exhibit altered T cell subsets, heightened CD8+ T cell activation, and increased risk of non-AIDS morbidity and mortality.** *PLoS Pathog.* 2014; **10**(5): e1004078.
PubMed Abstract | Publisher Full Text | Free Full Text

Simm M, Söderberg E, Larsson A, *et al.*: **Performance of plasma calprotectin as a biomarker of early sepsis: a pilot study.** *Biomark Med.* 2016; **10**(8): 811–818.
PubMed Abstract | Publisher Full Text

Sorensen AK, Holmgaard DB, Mygind LH, *et al.*: **Neutrophil-to-lymphocyte ratio, calprotectin and YKL-40 in patients with chronic obstructive pulmonary disease: correlations and 5-year mortality - a cohort study.** *J Inflamm (Lond).* 2015; **12**: 20.
PubMed Abstract | Publisher Full Text | Free Full Text

Sun X, Meng H, Shi D, *et al.*: **[The correlation between plasma calprotectin and neutrophil count in patients with aggressive periodontitis].** *Zhonghua Kou Qiang Yi Xue Za Zhi.* 2014; **49**(11): 649–651.
PubMed Abstract | Publisher Full Text

Walsham NE, Sherwood RA: **Fecal calprotectin in inflammatory bowel disease.** *Clin Exp Gastroenterol.* 2016; **9**: 21–29.
PubMed Abstract | Publisher Full Text | Free Full Text

Open Peer Review

Current Peer Review Status: ✔ ✔

Version 2

Reviewer Report 07 December 2020

https://doi.org/10.21956/wellcomeopenres.17976.r41660

© **2020 Silverman 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.

✔ **Rachel A. Silverman**
Center for Public Health Practice and Research, Department of Population Health Sciences, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA

The authors have sufficiently addressed my initial comments. I have no further comments to make.

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.

Reviewer Report 03 December 2020

https://doi.org/10.21956/wellcomeopenres.17976.r41659

© **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

The only optional comment for authors to consider is that Table 2, if there is sufficient power, could be analyzed using only the subcohort data. It would be taking advantage of the randomly

selected subcohort vs. using the full case-cohort which was selected based on mortality (not the relevant outcome for Table 2). Overall, the authors have addressed earlier concerns and the new version of the manuscript is stronger and acceptable for indexing.

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.

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 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 was 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.

Author Response 29 Sep 2020

James Berkley,

We thank the reviewer for their comments providing an opportunity to improve our manuscript. Please find out point-by-point responses below.

Reviewer 2 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 was 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.

In the parent study, children recruited were 5-12 years, adolescents were 13-17 years and adults were 18 years and above. In this pilot sub-study we only included participants $\geq$ 13 years and above. As such, data from children is out of scope and not available for this plot study.

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 broader case cohort study involved two of the sites in the broader ongoing immunology case-cohort sub-study which randomly selected 45% of all Kilifi participants and 10% of Eldoret study participants as a sub-cohort reflecting survival and mortality in the full REALITY trial. This has now been clarified on page 6.

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.

We have revised clarified this in the manuscript in methods on page 9 and results on page 11. In the main analysis, factors associated with plasma calprotectin levels at baseline were assessed using a linear regression model. We have now analysed the associations between change in plasma calprotectin between baseline and 4 weeks and deaths after 4 weeks by calculating relative risk.

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.

The sample size has been added to tables and figures. All analyses include the random sub-cohort and all additional deaths as is usual in a nested case-cohort study.

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.
We agree, and we commented in the discussion that this pilot study should be validated in a larger study.
It is useful to also see the hazards and 95% CI (not just p-values) in text of results instead of just the table.
We have added ß coefficients and hazards ratios to the main text.
Discussion should include whether this marker could have utility for individuals with less advanced HIV.
Thank you, we have added this to the discussion.

Competing Interests: No competing interests were disclosed.

Reviewer Report 01 May 2020

https://doi.org/10.21956/wellcomeopenres.17044.r38178

© 2020 Silverman 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.

?

Rachel A. Silverman
Center for Public Health Practice and Research, Department of Population Health Sciences, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA

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 Kilfi 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 (<100cells/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.

Author Response 29 Sep 2020

James Berkley,

We thank the reviewer for their comments providing an opportunity to improve our manuscript. Please find out point-by-point responses below.
Reviewer 1 Please specify how "ART initiation" is defined? Is the plasma collected prior to receiving any medication?
Yes, plasma was collected before any ART was given. This has been added on page 7.
Space in Trial ID number before closing parentheses.
Thank you, corrected.
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.
We randomly selected participants who were confirmed to have survived to the end of the trial at 48 weeks (non-cases) for whom samples were available. This has been clarified on page 8.
The final sentence "To test association with mortality…" is confusing. Suggest rewording. Also, how was clustering by CD4 count decided/done?
We three strata for the $CD4^{+}$ count groups(0-24, 25-49, and 50-99 cells/mm^{3}) as per the sampling frame. The has been revised for clarity n page 8 and the final text of the section also refers to this.

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.
Associations with plasma calprotectin at baseline were assessed using linear regression, this is described in the second paragraph under Statistical methods.
esults mention deaths within 4 weeks as an outcome, but this is not include in the methods. Please ensure all results have methods described.

Thank you, this has been added to the bottom of page 8.
Please clarify why there were 45% randomly selected from Kilfi but only 10% randomly selected from Eldoret? How were these percentages determined for selection?
The broader case cohort study involved two of the sites in the broader ongoing immunology case-cohort sub-study which randomly selected 45% of all Kilifi participants and 10% of Eldoret study participants as a sub-cohort reflecting survival and mortality in the full REALITY trial. This has now been clarified on page 6.
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.
Calprotectin was measured using samples collected at baseline, before ART treatment was initiated, and four weeks after ART treatment initiation. We have clarified this on page 7.
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.
We have now clarified on page 8 and table 2 that the linear regression model was stepwise, always including age, sex, site, CD4+, viral load and neutrophil count, and selecting other biologically plausible variables on the basis of statistical significance at P<0.1. We regarded that including all of the tested variables in the final model
Please clarify in the Cox proportional hazard regression, how was this stratified by CD4 count. What strata were used and how were these determined?
We used three strata for the $CD4^+$ count groups(0-24, 25-49, and 50-99 cells/mm^3) as per the sampling frame. The has been revised for clarity n page 8 and the final text of the section also refers to this. In STATA, the command was 'stcox … , strata(cd4group)'.
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.
Thank you. This is an error, the parent trial primary endpoint was mortality to 24 weeks, with secondary outcomes including those during extended follow up to 48 weeks, but for this analysis, survival or death during 24 weeks was used. This has been corrected through the manuscript. . We have added the 4-week analysis to the methods.
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.
The Wilcoxon paired sign-rank test was used to assess the overall change in plasma calprotectin between baseline and after four weeks, without covariates.
Is enrollment misspelled (enrolment)?

English (United Kingdom) spelling was used since this is a UK-based journal.
Final sentence is grammatically incorrect. I think deleting the “lower” before “hemoglobin”would correct this.
Thank you. We have revised as suggested.
Please define what “crude” means.
We have replaced ‘crude’ with ‘unadjusted’.
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.
We have revised this to indicate P<0.1 was used for selection. We agree with the issue of reliance on P-values but regarded that having many variables in the final model was undesirable in 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?
Thank you. It has been corrected to -0.57 to 0.14.
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?
We have added the interpretations to table 2.
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.
We used three strata for the $CD4^+$ count groups (0-24, 25-49, and 50-99 cells/mm^3) as participants were selected using these strata to avoid imbalance in a key exposure (as per https://pubmed.ncbi.nlm.nih.gov/10763560/). This is added to page 6. All analyses are stratified.
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.
These estimates have been added on page 11 and discussed on page 13.
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 (<100cells/µl).” *This has been reworded as suggested.*
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?
Thank you. The primary analysis is to 24 weeks and this has been clarified along with a more concrete statement.
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.
Thank you. The term 'exploratory' has been removed and we have revised the limitations as suggested.
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.
Thank you, we have made this change. Insufficient power is quite possible and we have included that.
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.
We have remove 'missing at random' since sites were differentially affected and mentioned this under limitations.
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?
We have added this to the conclusions: 'Although capacity for an ELISA-based test may be limited at many clinics, there is at least one quantitative serum calprotectin lateral-flow test available'.

Competing Interests: No competing interests were disclosed.