



OPEN

# Plasma proteomics reveals markers of metabolic stress in HIV infected children with severe acute malnutrition

Gerard Bryan Gonzales<sup>1,2</sup>✉, James M. Njunge<sup>3,4</sup>, Bonface M. Gichuki<sup>3,4</sup>, Bijun Wen<sup>5,6</sup>, Isabel Potani<sup>3</sup>, Wieger Voskuil<sup>3,7,8</sup>, Robert H. J. Bandsma<sup>3,5,6</sup> & James A. Berkley<sup>3,4,9</sup>

HIV infection affects up to 30% of children presenting with severe acute malnutrition (SAM) in Africa and is associated with increased mortality. Children with SAM are treated similarly regardless of HIV status, although mechanisms of nutritional recovery in HIV and/or SAM are not well understood. We performed a secondary analysis of a clinical trial and plasma proteomics data among children with complicated SAM in Kenya and Malawi. Compared to children with SAM without HIV ( $n = 113$ ), HIV-infected children ( $n = 54$ ) had evidence (false discovery rate (FDR) corrected  $p < 0.05$ ) of metabolic stress, including enriched pathways related to inflammation and lipid metabolism. Moreover, we observed reduced plasma levels of zinc- $\alpha$ -2-glycoprotein, butyrylcholinesterase, and increased levels of complement C2 resembling findings in metabolic syndrome, diabetes and other non-communicable diseases. HIV was also associated (FDR corrected  $p < 0.05$ ) with higher plasma levels of inflammatory chemokines. Considering evidence of biomarkers of metabolic stress, it is of potential concern that our current treatment strategy for SAM regardless of HIV status involves a high-fat therapeutic diet. The results of this study suggest a need for clinical trials of therapeutic foods that meet the specific metabolic needs of children with HIV and SAM.

Malnutrition, specifically undernutrition in all its forms, remains a global public health burden that accounts for 45% of all death among children under 5 years old<sup>1</sup>. Despite careful monitoring and adherence to guidelines set by the World Health Organization, whilst in general, uncomplicated SAM cases treated in the community do well, up to 25% of children with complicated severe acute malnutrition (SAM) treated in a hospital environment do not survive<sup>2-5</sup>. Furthermore, about one in five children treated for complicated SAM and discharged alive, die in the first year after discharge in low-resource settings<sup>6-8</sup>. However, our understanding of the pathophysiology underlying the poor prognosis for these children is surprisingly limited.

Infection with the human immunodeficiency virus (HIV) is a common co-morbidity of SAM in sub-Saharan Africa affecting up to 30% of admissions among SAM cases<sup>9</sup>. HIV-infected or exposed children are significantly more likely to be stunted, wasted, and underweight<sup>10</sup>. They also more often present with other clinical complications and greater susceptibility to infections, thus further complicating their clinical management, which may include providing more aggressive antimicrobial therapy and higher caloric nutritional intervention<sup>11</sup>. Moreover, response to clinical management is also less predictable and less well-understood in HIV-infected children compared to their uninfected counterparts<sup>12</sup>. Although acute opportunistic infections play a key role in the outcome of these children, intestinal pathology including inflammation and malabsorption, and metabolic perturbations

<sup>1</sup>Department of Gastroenterology, Faculty of Medicine and Health Sciences, Ghent University, Ghent, Belgium. <sup>2</sup>VIB Inflammation Research Centre, Ghent, Belgium. <sup>3</sup>The Childhood Acute Illness & Nutrition (CHAIN) Network, Nairobi, Kenya. <sup>4</sup>KEMRI/Wellcome Trust Research Programme, Kilifi, Kenya. <sup>5</sup>Centre for Global Child Health, The Hospital for Sick Children, Toronto, ON, Canada. <sup>6</sup>Department of Nutritional Sciences, Faculty of Medicine, University of Toronto, Toronto, Canada. <sup>7</sup>Global Child Health Group, Emma Children's Hospital, Amsterdam University Medical Centres, Amsterdam, The Netherlands. <sup>8</sup>Department of Global Health, Amsterdam Institute for Global Health and Development, Amsterdam University Medical Centres, Amsterdam, The Netherlands. <sup>9</sup>Nuffield Department of Medicine, Centre for Tropical Medicine & Global Health, University of Oxford, Oxford, UK. ✉email: Gerard.gonzales@ugent.be

	All	HIV (+)	HIV (-)	Unknown HIV status	p*
n (%)	843	179 (21%)	618 (73%)	46 (5%)	
Median age in months [IQR]	16 [10–25]	21 [12–31]	16 [10–25]	10 [8–17]	< 0.001
% girls (n)	45% (359)	45% (81)	45% (278)	56% (26)	0.95
Mean MUAC in cm [95% CI]	11.2 [11.1–11.3]	10.5 [10.36–10.7]	11.4 [11.3–11.5]	11.2 [10.9–11.5]	< 0.001
Mean weight-for-age z-score [95% CI]	-4.01 [-4.11 to -3.92]	-4.51 [-4.72 to -4.31]	-3.92 [-4.03 to -3.80]	-3.56 [-3.94 to -3.92]	< 0.001
% mortality (n)	15% (127)	26% (47)	10% (64)	34% (16)	< 0.001°
% oedematous (n)	31% (264)	30% (54)	33% (203)	15% (7)	0.50
<b>Site</b>					
Coast Provincial General Hospital, Kenya	39% (329)	25% (45)	40% (247)	80% (37)	Reference
Kilifi County Hospital, Kenya	22% (187)	22% (40)	23% (145)	4% (2)	0.08
Queen Elizabeth Central Hospital, Malawi	39% (327)	52% (94)	36% (226)	15% (7)	< 0.001

**Table 1.** Descriptive characteristics of the study participants. \*Comparison between HIV(+) and HIV(-). °Adjusted for age, sex and site of recruitment.

may also be present. However, mechanisms driving poor nutritional recovery of children with HIV even when detected co-morbidities are treated remain poorly understood<sup>12</sup>.

We hypothesised that inflammatory, metabolic and other pathways which are likely to be involved in the response to infection, survival and nutritional recovery differ between children with SAM with and without HIV. We conducted a secondary analysis of clinical data and biological samples from a randomised clinical trial in Kenya and Malawi<sup>13</sup>.

## Results

**Patient characteristics.** Table 1 presents the baseline characteristics of the children in the randomised trial. A total of 843 complicated SAM children were recruited for the randomised trial, of which 179 (22%) patients were HIV(+). Age was higher and MUAC was lower in HIV(+) children than HIV(-) counterparts. Most HIV cases were found in Malawi. Sex and the presence of oedema were not associated with HIV status. Mortality was more than two times higher among in HIV(+) compared to HIV(-) ( $p < 0.001$ ). Children whose HIV status were unknown had the highest mortality of 34%, which indicates bias due to frequent death before testing could be undertaken or refusal of testing when a child was more severely ill.

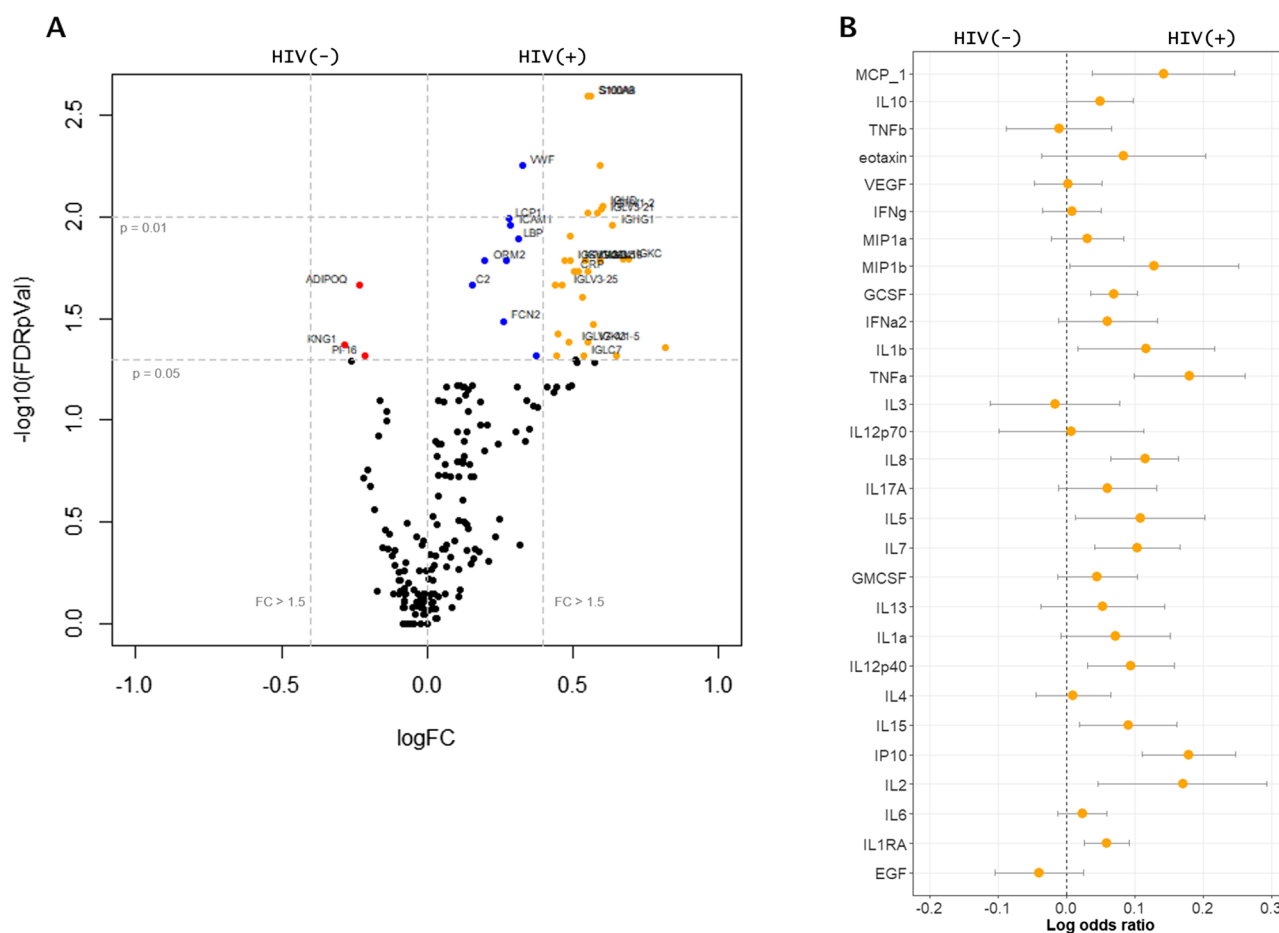
Among HIV(+), 33% were already receiving an anti-retroviral treatment (ART) regime: 53/179 (30%) on highly active antiretroviral therapy (HAART), and 7/179 (4%) on Nevirapine only. About half of the children (90/179) were naïve for ART whereas HIV treatment status was unknown for 16% (29/179). Mortality was not significantly different among children on HAART, ART naïve and children with unknown HIV treatment status (Supplementary Table 1).

**HIV is associated with increased inflammation, immune activation, dysregulated lipid metabolism, and increased proteolysis in children with SAM.** Among the children included in the proteomics study, 54 were HIV (+) and 113 were HIV(-) (Table 2). In this sub-population, age, sex and the presence of oedema were not significantly associated with HIV. HIV(+) children also had significantly lower MUAC and higher mortality than HIV(-) children.

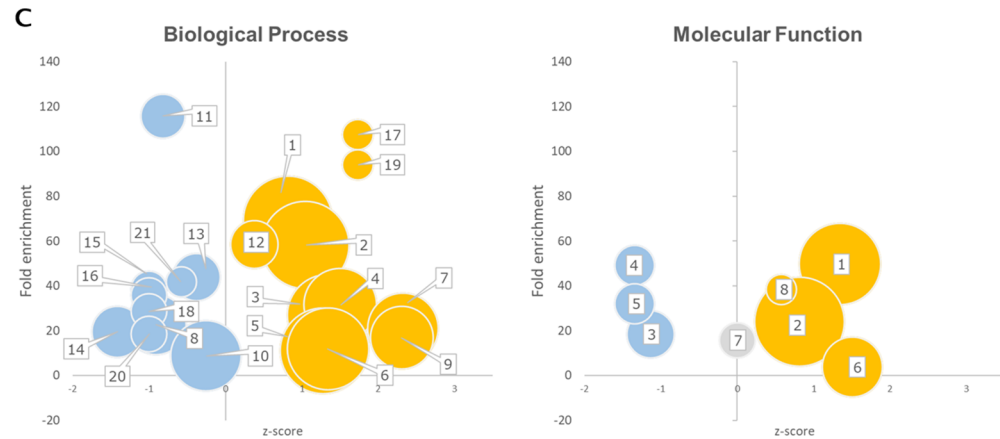
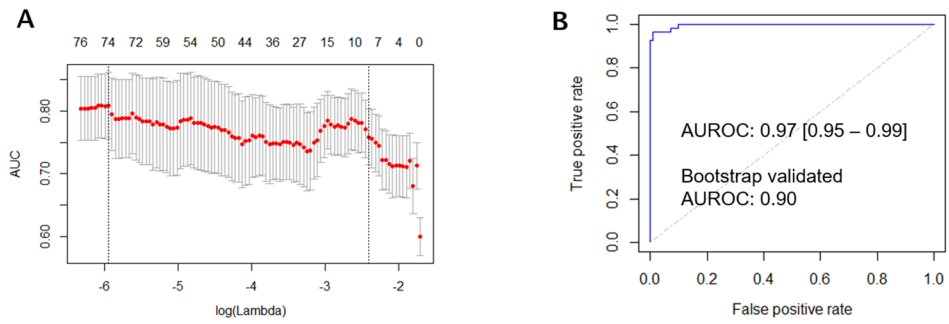
A total of 204 circulating proteins were annotated and compared between children with and without HIV infection. Of these, levels of 42 proteins were found to be significantly associated with HIV status in the initial univariate analysis (Fig. 1A) (Supplementary Table 2). Specifically, HIV(+) was associated with higher circulating levels of immunoglobulins, inflammatory proteins such as calprotectin (S100 calcium binding protein A8 and S100 calcium binding protein A9), complement proteins, and proteins related to host response to infection (i.e. lipopolysaccharide binding protein, galectin 3 binding protein and CD5 molecule-like protein). Enrichment analysis suggested that HIV(+) children had higher levels of proteins associated with classical complement pathway activation, immune activation and inflammation than HIV(-) children. Neutrophil aggregation and chemokine production appeared to be the pathways most highly enriched in HIV(+) compared to HIV(-) SAM children. To substantiate these results, we quantified chemokine and cytokine levels in plasma. As shown, most chemokines had the tendency to be associated with HIV infection, where elevated plasma concentration of 12 were significantly associated with HIV status in SAM children (Fig. 1B), namely: monocyte chemoattractant protein 1 (MCP1), macrophage inflammatory protein 1 beta (MIP1b, CCL4), granulocyte colony-stimulating factor (GCSF), interleukin 1 beta (IL1b), tumour necrosis factor alpha (TNFa), interleukins 2,5,7, 8 and 15 (IL2, 5, 7, 8, 15), interleukin 12 subunit beta (IL12p40), interferon gamma-induced protein 10 (IP-10), and interleukin-1 receptor antagonist (IL-1RA).

	All	HIV (+)	HIV (-)	<i>p</i> *
n	167	54	113	
Median age in months [IQR]	15 [10–26]	15 [10–26]	15 [10–24]	0.433
n girls (%)	76 (45%)	27 (50%)	49 (43%)	0.506
Mean MUAC at admission (cm) [95% CI]	10.9 [10.7–11.1]	10.2 [9.8–10.5]	11.3 [11.0–11.5]	<0.001
n oedematous (%)	49 (29%)	15 (28%)	34 (30%)	0.856
n mortality (%)	79 (47%)	36 (67%)	43 (38%)	<0.001°
<b>Use of antiretroviral medication</b>				
Naïve		27 (50%)		
Highly active antiretroviral therapy (HAART)		14 (26%)		
Nevirapine only		3 (6%)		
Unknown		10 (18%)		

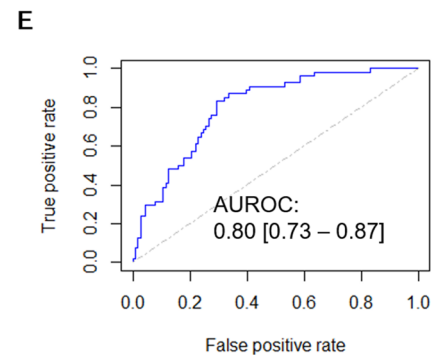
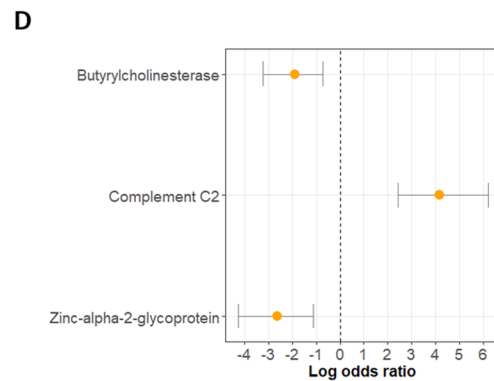
**Table 2.** Patient characteristics of those subjected to proteomics analysis. \*Comparison between HIV(+) and HIV(-). °Adjusted for age, sex, site of recruitment, oedema.



**Figure 1.** Univariate analysis of plasma proteome and individual plasma cytokines associated with HIV. **(A)** Volcano plot showing several significantly different (FDR adjusted  $p$  value < 0.05) proteins and their  $\log_2$  HIV(+) versus HIV(-) fold change. Red points represent those significantly higher in plasma of HIV(-), blue points significantly enriched in plasma of HIV(+) and orange points significantly higher than 1.5 folds in HIV(+) compared to HIV(-) SAM children. Vertical lines indicate significance level at  $p = 0.05$  and  $0.01$ ; horizontal lines indicate more than 1.5 folds enrichment. **(B)** Log odds plots showing association of chemokine markers analysed using Luminex platform and HIV status. Points indicate log odds ratio for every log increase in plasma protein concentration; bars indicate 95% confidence interval.



Code	Pathway	Code	Pathway
1	complement activation	1	antigen binding
2	complement activation, classical pathway	2	serine-type endopeptidase activity
3	receptor-mediated endocytosis	3	serine-type endopeptidase inhibitor activity
4	Fc-gamma receptor signalling pathway involved in phagocytosis	4	immunoglobulin receptor binding
5	Proteolysis	5	endopeptidase inhibitor activity
6	immune response	6	identical protein binding
7	regulation of immune response	7	glycoprotein binding
8	negative regulation of endopeptidase activity	8	lipid transporter activity
9	Fc-epsilon receptor signalling pathway		
10	innate immune response		
11	complement activation, alternative pathway		
12	regulation of complement activation		
13	retina homeostasis		
14	platelet degranulation		
15	positive regulation of B cell activation		
16	phagocytosis, recognition		
17	complement activation, lectin pathway		
18	phagocytosis, engulfment		
19	very-low-density lipoprotein particle assembly		
20	B cell receptor signaling pathway		
21	blood coagulation, intrinsic pathway		



◀ **Figure 2.** Multivariate analysis of plasma proteome associated with HIV. (A) Elastic net (EN) regularized regression lambda parameter optimization curve, optimal lambda parameter was chosen based on the highest area under the receiver operator curve (AUROC); (B) AUROC (0.97 [95% CI 0.95–0.99]) of the EN model generated using the lambda parameter, alpha parameter was set to 0.75, final model extracted 34 protein features, optimism-adjusted bootstrap validation of the generated EN model, validated AUROC = 0.90 [95% CI 0.90–0.90] using 2000 iterations; (C) Gene ontology (GO-terms) enrichment analysis of proteins extracted by the EN model. X-axis represents z-scores; y-axis, fold enrichment, and size of the spheres represent the number of proteins involved in the particular pathway. Gold circles represent pathways enriched in HIV(+) whereas blue circles are pathways more associated with HIV(–). The grey circle indicate that there are as much proteins in this pathway that are significantly upregulated and downregulated in HIV. Only significantly enriched pathways ( $p < 0.05$  after FDR adjustment) are plotted. See main text for explanation of the plots. Pathways enriched are identified in the table. (D) Log odds ratio plot of the three proteins extracted after bootstrap validation with log odds on the x-axis and bars indicating 95% confidence interval obtained using weighted logistic regression with HIV as outcome variable and the three proteins as covariates. Weights used were obtained by inverse probability of treatment weights; (E) predictive ability of the weighted logistic regression model using the three bootstrap validated proteins with HIV as outcome variable, AUROC = 0.80 [95% CI 0.73–0.87].

Out of the 43 differentially expressed proteins, three proteins were found to be negatively associated with HIV status on initial univariate analysis, namely: adiponectin, kininogen-1 and peptidase inhibitor 16. Among HIV(+) children, there were no statistically significant associations with receiving HAART ( $n = 14$ ) compared to ART naïve ( $n = 27$ ) children (Supplementary Fig. 1), recognising our study was not powered for this comparison. Furthermore, sensitivity analysis to address the possibility of HIV maternal antibodies in younger children, showed no significant interaction of age above or below 18 months and individual proteins plasma levels towards HIV status, although power to detect was limited.

The weighted EN model extracted 73 circulating proteins (Fig. 2A) that are associated with HIV status with AUROC = 0.97 [95% CI 0.95–0.99] (Fig. 2B) and misclassification error rate of 2.4%. Optimism-adjusted validated AUROC after bootstrapping was 0.90 [95% CI 0.90–0.902], indicating a robust model. Pathway enrichment analysis highlighted that apart from immune activation, HIV(+) children with SAM had increased levels of proteins involved in proteolysis and lipid mobilisation pathways, specifically increased very low-density lipoprotein assembly, indicating metabolic dysregulation related to cholesterol and triglyceride metabolism among HIV(+) patients (Fig. 2D).

After 2000 bootstrap iterations during bootstrap validation, 3 proteins were consistently extracted by the EN model >80% of the time (Fig. 2D), namely: butyrylcholinesterase (BChE), complement C2 and zinc- $\alpha$ -2-glycoprotein (ZAG), indicating that these three proteins are likely to be the most important features associated with HIV in children with complicated SAM. Weighted logistic regression model of these 3 proteins showed good discrimination of HIV status (AUROC = 0.80 [95% CI 0.74–0.87]) (Fig. 2E).

## Discussion

In this study, we report plasma proteomic differences associated with HIV status, suggesting that HIV imposes additional metabolic and inflammatory insults among HIV(+) children with SAM. Our results show that pathways involved in inflammatory response, complement cascade activation and lipid metabolism dysregulation are associated with HIV status. Circulating levels of several plasma chemokines were also found to be higher in HIV(+) among children with SAM. Greater inflammatory responses in these children could be related to the higher inpatient mortality of HIV(+) compared to HIV(–) children with SAM.

An earlier metabolomics study in Uganda reported reduced serum levels of adiponectin and leptin, whereas serum triglycerides, ketones and even-chain acylcarnitines were higher in HIV(+) children with SAM indicating perturbed lipid metabolism<sup>14</sup>. Our current study therefore concurs with this finding, as we also found reduced plasma levels of adiponectin in HIV(+) SAM children compared to HIV(–) SAM children, along with upregulation of pathways involved in lipid transport and metabolism, specifically very low-density lipoprotein assembly.

Using optimism-adjusted bootstrap validation of the EN model, we found three proteins: complement c2, BChE and ZAG robustly distinguished HIV(+) from HIV(–) in children with SAM, demonstrating the ability of multivariate analysis techniques, such as EN, to uncover underlying relationships between protein markers which would be difficult to identify when analysed individually. The activation of the complement system during HIV infection has been previously discussed at length, which is associated with the increased cellular invasion of HIV in cells<sup>15–17</sup>.

On the other hand, BChE is a protein synthesized in the liver and abundant in plasma, which hydrolyses acetylcholine. Although very similar to its sister protein, acetylcholinesterase, biological functions of BChE appear to be more varied but less understood<sup>18</sup>. In a recent study in China, low circulating BChE was found to be highly associated with HIV severity, was predictive of mortality in adults, and was proposed as a plausible strategy for severity classification among adults with HIV<sup>19</sup>. BChE is also reported to be reduced in SAM, stress and inflammation<sup>20</sup>. In animal studies, BChE deficiency was found to strongly affect fat metabolism and promotes hepatic lipid accumulation<sup>21</sup>. Serum BChE levels have been found to have a significant negative correlation with serum total cholesterol and serum low-density-lipoprotein cholesterol among people with diabetes mellitus<sup>22</sup>.

ZAG is a newly characterized adipokine that is involved in lipolysis, body weight regulation and may also be involved in the development of insulin resistance<sup>23</sup>. Reduction in plasma levels of ZAG was previously reported to be implicated in dyslipidaemia in HIV(+) adults under ART treatment<sup>23</sup>. Reduced circulating levels of ZAG has also been found among adults with clinically diagnosed metabolic syndrome, based on guidelines of the

	Stabilization phase	In-patient rehabilitation phase	Out-patient rehabilitation phase
	Days 1–7	Weeks 2–6	Lengths vary depending on site
Complicated SAM	F75	F100	RUTF
Uncomplicated SAM	–	–	RUTF
<b>Composition</b>			
Energy (kcal per 100 mL F75/F100 or 100 g RUTF)	75	100	5.2–5.5
Protein (% total energy)	5	12	10–12
Fat (% total energy)	32	53	45–60

**Table 3.** Nutritional management protocol for children with severe acute malnutrition<sup>39</sup>.

United States National Cholesterol Education Program (NCEP) Expert Panel Adult Treatment Panel (ATP) III criteria<sup>24</sup>. Serum ZAG levels have been reported lower among adults with impaired glucose tolerance and type 2 diabetes mellitus<sup>25</sup>. Taken together, our results therefore suggest that children with both HIV and SAM manifest hallmarks of metabolic stress similar to those occurring in metabolic syndrome and other non-communicable diseases (NCD).

This study is the first proteomics investigation on the interaction between HIV and SAM. In summary, our results, which together with the previously published metabolomics study<sup>14</sup>, strengthens evidence on the increased metabolic stress and altered metabolic response among children living with both HIV and SAM. Our results also concur with previous studies that reported elevated metabolic stress among non-malnourished adults living with HIV leading to increased prevalence or risk for metabolic syndrome, cardiovascular diseases, diabetes and other non-communicable diseases<sup>26–32</sup>.

Metabolic abnormalities have previously been reported to be attributed HAART use among HIV(+) patients<sup>33</sup>. In a recent systematic review, use of two classes of HAART, protease inhibitors and nonnucleoside reverse transcriptase inhibitors, has been found to be associated with abnormalities in plasma lipid profiles<sup>34</sup>. However, dysregulation in lipid metabolism has also been reported in HAART-naïve patients, which indicates that HIV infection alone cause lipid metabolism perturbations. An earlier longitudinal study of 50 men in the USA reported notable declines in serum total cholesterol after HIV infection compared to results of blood analysis from last seronegative visit. Large increases in total cholesterol and low-density lipoproteins (LDL) were detected after HAART initiation<sup>35</sup>. However, many other studies reported increases in total cholesterol among HIV-infected patients naïve to HAART. For instance, in a study of ART-naïve HIV-infected adults in Ethiopia, malnutrition and lipid abnormalities (specifically total cholesterol) were associated with CD4 + T cell counts<sup>36</sup>. In *in vitro* studies, transfection of a T-cell (RH9) with HIV led to the enhanced production of free fatty acids and LDL<sup>37</sup>. Furthermore, monocytes isolated from HIV-infected patients both taking HAART and HAART-naïve, were found to have altered expression patterns of receptors linked with lipid metabolism (i.e. FXR, PXR, PPAR $\alpha$ , GR, RAR $\alpha$  and RXR) compared to monocytes of HIV-uninfected controls<sup>38</sup>. For our study however, we are unable to ascertain whether the lipid metabolism dysregulation we observed is due primarily on the viral load itself or the use of HAART due to lack of power for this sub-analysis. Majority of the participants subjected to proteomics analysis were HAART-naïve (50%), where 26% were on HAART, 6% were on Nevirapine alone and we had no data on treatment of 18% of the patients (Table 2). In all these studies cited, authors argue to need for monitoring of lipid profiles in HIV-infected populations. Hence, lipid monitoring may also inform nutritional and clinical recovery of children with SAM and HIV and could be implemented to improve clinical care for these children.

However, despite our knowledge that HIV-infected populations have altered metabolic requirements compared to HIV-uninfected counterparts, WHO guidelines for the nutritional management for SAM are globally the same regardless of HIV status, which is summarized in Table 3<sup>39</sup>. Nutritional management for in-patient children with SAM involves provision of a low-protein, low-fat milk-based food, F75, every three hours. F75 is used during clinical stabilization occurring during the first few days after admission and is not intended for weight gain. Once the children are clinically stabilized and are able to tolerate the milk/solute load, children are transitioned to F100, a higher-calorie, high-fat milk intended to boost weight gain or to Ready-to-Use Therapeutic Food (RUTF), a peanut-based calorie-dense diet. Upon discharge from in-patient care, children are referred to community based nutritional therapeutic centres where they are provided with RUTF on a 2 weekly basis.

Considering evidence of biomarkers of metabolic syndrome and NCD in HIV(+) children with SAM, it is of potential concern that our current treatment strategy involves a high-fat therapeutic diet. About 50% of much needed calories during the growth catch-up phase are supplied as lipids, which HIV(+) children may not be able to efficiently assimilate. Alterations in lipid metabolism in HIV(+) children with SAM may also mean that the high amounts of dietary lipids could be deposited as ectopic fat in the liver and muscle, predisposing to insulin resistance, diabetes, cardiovascular problems and other NCDs later in life. Although long-term metabolic follow-up studies could be done for HIV(+) children previously treated for either complicated and uncomplicated SAM, significant barriers are the high mortality rate in earlier studies of HIV(+) children with SAM, cost and difficulty tracing them years later. The results of this study indicate a need for clinical trials of F100 or RUTF modified to meet the expected metabolic needs of HIV(+) children with SAM. This could initially be done in relatively small groups with outcomes that include measuring metabolic stress.

Several studies on nutritional intervention strategies among HIV-infected adults have been reported. For instance, a study in the USA showed that dietary fat intake, specifically saturated fats, was significantly associated with hypertriglyceridemia among HIV-infected adults (18–60 years)<sup>40</sup>. Moreover, in a preclinical model, high

saturated fat consumption was found to accelerate immunodeficiency virus disease progression in macaques, specifically increased mortality hazard and circulating levels of pro-inflammatory cytokines, especially IL8<sup>41</sup>, which has been previously reported to be associated with lipodystrophy among HIV patients<sup>42</sup>. In our study, we also found a significant association between high plasma IL8 concentration and HIV in SAM children. Hence, modifying the saturated fat composition of the milk-based F75 and F100 could potentially lower metabolic stress.

The European Society for Parenteral and Enteral Nutrition (ESPEN) have given a grade A recommendation for the use of medium-chain triglyceride (MCT)-based diet on HIV(+) patients with diarrhoea and severe under-nutrition in its 2006 ESPEN Guidelines on Enteral Nutrition<sup>43</sup>. Grade A recommendations are given to strategies based on meta-analysis or at least one randomised control trial. In this case, the recommendation was based on a prospective, randomized double-blind comparative trial on 24 adult patients with HIV and diarrhoea of more than 4-week duration, fat malabsorption, and loss of 10–20% of ideal body weight<sup>44</sup>. In this study, the authors found improved outcomes from diarrhoea and fat malabsorption from MCT than long-chain triglyceride-based diet among HIV(+) adults.

HIV infection has been reported to be accompanied by substantial damage to gut integrity and changes in gut microbiome composition<sup>45</sup>. In this study, we observed increased circulating levels of LPS binding protein, which is a marker of bacterial translocation from the gut into the bloodstream. Therefore, understanding the interaction between HIV and gut microbiota could provide insights into aetiology and interventional points of view. As more evidence on the role of gut microbiota and gut integrity on health outcomes emerge, we must also be aware of the potential impact of antibiotics and nutritional therapeutic strategies on the microbiome. Markers of gut health and microbiome restoration among children with HIV and SAM therefore need to be studied in parallel with improved/modified RUTF formulations to fully elucidate the mechanisms of their efficacy.

Lastly, the long-term metabolic effect of nutritional intervention strategies for SAM still remains unresolved. Most specifically, the potential metabolic stress associated with the rapid weight gain during the nutritional rehabilitation phase after SAM and its implications on nutritional outcomes during adulthood demands urgent research attention, especially for HIV(+) children with SAM.

Limitations of this study include absence of data on viral load and CD4+ counts of the patients, which could provide a deeper understanding of the results. Furthermore, in this study, we did not find association between oedematous malnutrition and HIV status, although several studies have found higher HIV prevalence among non-oedematous children with SAM<sup>46–48</sup>. In our study however, we found high in-patient mortality rate (16/46, 34%) among children with unknown HIV status, where 39/46 (85%) had non-oedematous SAM. Considering the high rate of mortality, these children may have been HIV(+). This highlights the need for earlier HIV screening among children with SAM. Finally, a deeper understanding of the comorbidity of HIV and SAM would require studies also involving non-malnourished HIV+ and HIV– children preferably in various geographical and social contexts. Hence, further studies are needed fully characterize the interplay between HIV infection and malnutrition.

## Conclusion

Plasma proteomics reveals that HIV(+) children with SAM manifest hallmarks of metabolic stress similar to those observed in non-communicable diseases. This could be related to the poor nutritional recovery and high mortality of HIV(+) children with SAM despite clinical and nutritional intervention. The results of this study indicate a need for clinical trials modifying the composition of F100 or RUTF to meet the specific metabolic needs of HIV(+) children with SAM during rehabilitation phase. This could initially be done in relatively small groups with outcomes that include measuring metabolic stress.

## Methods

**Patient recruitment and study design.** This is a secondary analysis of a nested case control study from a randomised controlled trial (NCT02246296), which tested the effect of a lactose-free, low-carbohydrate F75 milk to limit carbohydrate malabsorption, diarrhoea and refeeding syndrome among children hospitalized for complicated SAM at Queen Elizabeth Central Hospital in Blantyre, Malawi, Kilifi County Hospital and Coast General Hospital, Mombasa, Kenya<sup>13</sup>. Children aged 6 months to 13 years were eligible for enrolment into the trial at admission to hospital if they had SAM, defined as: mid-upper arm circumference (MUAC) < 11.5 cm or weight-for-height Z score < – 3 if younger than 5 years of age, BMI Z score < – 3 if older than 5 years, or oedematous malnutrition at any age and had medical complications or failing an appetite test, as defined by WHO guidelines<sup>49</sup>. Children were excluded if they had a known allergy to milk products and did not provide consent. Biological samples were obtained before the children received the randomised treatment irrespective of HIV status. Unless a child's HIV positive status was documented, HIV status was assessed by offering an antibody test at admission plus appropriate counselling. For this analysis, patients that tested positive on an HIV antibody test were considered HIV(+) and children with missing or declined HIV test were excluded.

To compare the proteomic profiles between HIV infected and non-infected children with SAM, we used data from a nested case–control study to investigate inpatient mortality. Of 127 children who died, 92 had sufficient samples available for proteomics analysis. Since the main outcome of our current study is HIV, we excluded deaths with unknown HIV status (n = 13), resulting to 79 cases included in this analysis. Among children who survived, 92 had been randomly selected in the nested case–control study matched on site of recruitment. After excluding children with unknown HIV status (n = 4), 88 controls from the nested case–control study were used for this analysis. Proteomic, cytokine, and chemokine data was generated using plasma samples collected at admission during enrolment to the trial. A weighted analysis was designed to help overcome selection bias, as described in the data analysis section below.

**Proteomics, cytokine and chemokine analysis.** Untargeted proteomics and targeted cytokines and chemokines analysis of plasma samples were performed following methods described previously<sup>50</sup>. The targeted protein panel included: epidermal growth factor (EGF); eotaxin; granulocyte-colony stimulating factor (GCSF); granulocyte-macrophage colony-stimulating factor (GMCSF); interferon alpha-2 (IFN $\alpha$ 2); interferon gamma (IFN $\gamma$ ); interleukins 10, 12p40, 12p70, 13, 15, 17A, 1a, 1b, 1RA, 2 to 8; interferon gamma-induced protein 10 (IP10); monocyte chemoattractant protein 1 (MCP1), macrophage inflammatory protein 1 alpha and beta (MIP1a & b); tumour necrosis factor alpha (TNF $\alpha$ ) and beta (TNF $\beta$ ); and vascular endothelial growth factor (VEGF).

**Data analysis.** Data analyses were performed using R v3.5<sup>51</sup>. Analysis of the prevalence of HIV(+), nutritional status and their associations with inpatient mortality utilised the entire trial dataset (N = 843). Analysis of categorical data was performed using Fisher's test and generalised linear models for continuous outcomes. Logistic regression was used to analyse binary outcomes adjusting for age, sex, presence of oedema, and site of recruitment. These associations were also adjusted for MUAC. As a sensitivity analysis to address the possibility of confounding due to HIV maternal antibodies in younger children, a test of interaction between age above or below 18 months and individual proteins towards HIV status was performed.

The proteomics, cytokines and chemokines analyses were secondary analyses of data collected from a nested case-control study with inpatient mortality as its primary outcome, hence with strong selection bias. The analysis for the association between HIV status and individual proteins was therefore performed using logistic regression analysis with inverse probability weighting (IPW) to correct for selection bias<sup>52-55</sup>. Weights ( $w$ ) were calculated as suggested by Samuelsen<sup>53</sup> wherein the weight for each observation selected into the nested case-control study was computed as the inverse of the probability of being selected for the nested study from the main clinical trial. The probability of inclusion was therefore calculated as:

$$p(i) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 \times 1 + \beta_2 \times 2 \dots + \beta_n \times n)}};$$

where  $p(i)$  is the probability of inclusion in the nested case-control study and  $\times 1, \times 2, \dots, \times n$  are HIV status, sex, age, presence of oedema, mid-upper arm circumference, and site of recruitment of the  $i$ th observation (child) based on the entire trial population. Inverse probability weight is therefore:

$$w(i) = \frac{1}{p(i)}$$

Differences in individual proteins abundances were considered statistically significant when  $p < 0.05$  after adjustment for multiple comparisons using Benjamini-Hochberg false discovery rate (FDR)<sup>56</sup>.

Multivariate analysis was undertaken in order to determine several proteins that are collectively associated with HIV status, some of which may not be significantly associated to HIV independently. This was performed using a weighted elastic net (EN) model implemented using the "glmnet" package in R<sup>57</sup>. EN is a penalized regression approach that was developed to help overcome problems caused by high dimensional data. It is an integration of two regularized approaches, ridge regression and least absolute shrinkage and selection operator (LASSO), wherein the contribution of each of these models to the final EN model is controlled by the  $\alpha$  parameter<sup>57,58</sup>. The strong penalization imposed by LASSO draws coefficients to zero thereby eliminating non-predictive proteins features, whereas ridge regression addresses potential multi-collinearity problems in high-dimensional data<sup>57,58</sup>.

Weighted EN model generation was performed with HIV status as outcome, protein profile as predictors, and  $w$  as observation weights. The penalization parameter lambda, which influences the shrinkage of variable coefficients to zero thus eliminating some non-contributing variables, was determined by estimating the area under the receiver operator curve (ROC) of the population using ten-fold cross validation. Several alpha parameter values were assessed and a final value of 0.85 was taken to achieve a compromise between predictive ability and fewer number of features extracted. The final lambda parameter was based on the value which gave the highest area under the ROC (AUROC) value.

Proteins with significant association with HIV status after correction for false discovery and those extracted by the EN model were then uploaded to The Database for Annotation, Visualization and Integrated Discovery (DAVID) v6.8 Bioinformatics Resource<sup>59</sup> to assess the Gene ontology (GO) enriched pathways of the differentially expressed proteins.

EN model validity was judged based on the AUROC and misclassification error rate. The fitted EN model performance measured as optimism-corrected AUC was validated using bootstrap, following the procedure of Smith et al.<sup>60</sup>. Bootstrapping was performed on 2000 iterations using the "BootValidation" package in R. Protein features extracted at least 80% of all iterations by the bootstrap EN model were then considered to be the most relevant protein biomarkers. To test how well these proteins can discriminate HIV status, they were then fitted on a weighted logistic regression with HIV as outcome.

**Visualisation of significantly enriched GO terms.** Bubble plots were used to visualise the significantly enriched pathways ( $p < 0.05$  after adjustment for FDR) obtained from DAVID. The p-values in DAVID were obtained using a modified Fisher's exact test<sup>61</sup>. The y-axis represents the fold enrichment which indicates the magnitude of the enrichment, as calculated in DAVID. Fold enrichment is defined as:



$$\text{fold enrichment} = \frac{(m/n)}{M/N},$$

where  $m$  is the number of proteins significantly associated with HIV status or proteins extracted by the EN model that belong to a particular pathway, while  $M$  is the total number of proteins belonging to the same pathway. Variable  $n$  is the number of all proteins significantly associated with HIV status or extracted by the EN model and  $N$  is the total number of all proteins in the human background. Therefore, a fold enrichment of ten indicates that 10% of the proteins significantly associated with HIV status belong to a particular pathway, and 1% of all annotated proteins in the human background belongs to the same pathway<sup>61</sup>. However, the proponents of this metric warn that big fold enrichments could be obtained from a small number of proteins, which could be due to small  $n$  or pathways with fewer members.

The x-axis on the hand represents the enrichment z-score for a particular pathway<sup>62</sup>, which is calculated as follows:

$$z - \text{score} = \frac{(up - down)}{\sqrt{\text{count}}};$$

where  $up$  is the total number of proteins upregulated,  $down$  is the total number of proteins downregulated, and  $count$  is the total number of proteins in the input which belongs to a particular pathway. Variables  $up$  and  $down$  were based on the weighted logistic regression for each individual protein. Hence, if five proteins belonging to pathway  $x$  were upregulated and two were downregulated, the z-score for pathway  $x$  would be:  $(5-2)/\sqrt{7} = 1.13$ . A positive z-score indicates that the particular pathway is overall upregulated in HIV(+), whereas a negative z-score indicates an overall downregulation<sup>62</sup>.

**Ethics approval.** The secondary analyses of the trial were approved by the Kenyan National Ethics Committee, KEMRI-SERU (KEMRI/RES/7/3/1). The trial was registered at clinicaltrials.gov (NCT02246296).

Received: 29 November 2019; Accepted: 17 June 2020

Published online: 08 July 2020

## References

- Black, R. E. *et al.* Maternal and child undernutrition and overweight in low-income and middle-income countries. *Lancet* **382**, 427–451. [https://doi.org/10.1016/S0140-6736\(13\)60937-X](https://doi.org/10.1016/S0140-6736(13)60937-X) (2013).
- Heikens, G. T. *et al.* Case management of HIV-infected severely malnourished children: Challenges in the area of highest prevalence. *Lancet* **371**, 1305–1307. [https://doi.org/10.1016/s0140-6736\(08\)60565-6](https://doi.org/10.1016/s0140-6736(08)60565-6) (2008).
- Irena, A. H., Mwambazi, M. & Mulenga, V. Diarrhea is a major killer of children with severe acute malnutrition admitted to inpatient set-up in Lusaka, Zambia. *Nutr. J.* **10**, 110–110. <https://doi.org/10.1186/1475-2891-10-110> (2011).
- Maitland, K. *et al.* Children with severe malnutrition: Can those at highest risk of death be identified with the WHO protocol?. *PLoS Med.* **3**, e500. <https://doi.org/10.1371/journal.pmed.0030500> (2006).
- Hossain, M. *et al.* Efficacy of World Health Organization guideline in facility-based reduction of mortality in severely malnourished children from low and middle income countries: A systematic review and meta-analysis. *J. Paediatr. Child Health* **53**, 474–479. <https://doi.org/10.1111/jpc.13443> (2017).
- Kerac, M. *et al.* Follow-up of post-discharge growth and mortality after treatment for severe acute malnutrition (FuSAM Study): A prospective cohort study. *PLoS ONE* **9**, e96030. <https://doi.org/10.1371/journal.pone.0096030> (2014).
- Berkley, J. A. *et al.* Daily co-trimoxazole prophylaxis to prevent mortality in children with complicated severe acute malnutrition: A multicentre, double-blind, randomised placebo-controlled trial. *Lancet. Glob. Health* **4**, e464–473. [https://doi.org/10.1016/s2214-109x\(16\)30096-1](https://doi.org/10.1016/s2214-109x(16)30096-1) (2016).
- Nemetchek, B. *et al.* Paediatric postdischarge mortality in developing countries: A systematic review. *BMJ Open* **8**, e023445. <https://doi.org/10.1136/bmjopen-2018-023445> (2018).
- Fergusson, P. & Tomkins, A. HIV prevalence and mortality among children undergoing treatment for severe acute malnutrition in sub-Saharan Africa: A systematic review and meta-analysis. *Trans. R. Soc. Trop. Med. Hyg.* **103**, 541–548. <https://doi.org/10.1016/j.trstmh.2008.10.029> (2009).
- Ndirangu, M., Wariero, J. O., Sachs, S. E., Masibo, P. & Deckelbaum, R. J. Nutritional status of under-five children in HIV-affected households in western Kenya. *Food Nutr. Bull.* **32**, 159–167. <https://doi.org/10.1177/156482651103200208> (2011).
- Rose, A. M., Hall, C. S. & Martinez-Alier, N. Aetiology and management of malnutrition in HIV-positive children. *Arch. Dis. Child.* **99**, 546–551. <https://doi.org/10.1136/archdischild-2012-303348> (2014).
- Trehan, I., O'Hare, B. A., Phiri, A. & Heikens, G. T. Challenges in the management of HIV-infected malnourished children in Sub-Saharan Africa. *AIDS Res. Treat.* **2012**, 8. <https://doi.org/10.1155/2012/790786> (2012).
- Bandsma, R. H. J. *et al.* A reduced-carbohydrate and lactose-free formulation for stabilization among hospitalized children with severe acute malnutrition: A double-blind, randomized controlled trial. *PLoS Med.* **16**, e1002747. <https://doi.org/10.1371/journal.pmed.1002747> (2019).
- Mody, A. *et al.* Effects of HIV infection on the metabolic and hormonal status of children with severe acute malnutrition. *PLoS ONE* **9**, e102233. <https://doi.org/10.1371/journal.pone.0102233> (2014).
- Tacnet-Delorme, P. *et al.* In vitro analysis of complement-dependent HIV-1 cell infection using a model system. *J. Immunol.* **162**, 4088 (1999).
- Stoiber, H., Kacani, L., Speth, C., Würzner, R. & Dierich, M. P. The supportive role of complement in HIV pathogenesis. *Immunol. Rev.* **180**, 168–176. <https://doi.org/10.1034/j.1600-065X.2001.1800115.x> (2001).
- Stoiber, H., Speth, C. & Dierich, M. P. Role of complement in the control of HIV dynamics and pathogenesis. *Vaccine* **21**, S77–S82. [https://doi.org/10.1016/S0264-410X\(03\)00203-2](https://doi.org/10.1016/S0264-410X(03)00203-2) (2003).
- Oliveira, J. D. *et al.* Effect of BCHE single nucleotide polymorphisms on lipid metabolism markers in women. *Genet. Mol. Biol.* **40**, 408–414. <https://doi.org/10.1590/1678-4685-GMB-2016-0123> (2017).
- Xu, L. *et al.* Butyrylcholinesterase levels on admission predict severity and 12-month mortality in hospitalized AIDS patients. *Mediators Inflamm.* **2018**, 5201652. <https://doi.org/10.1155/2018/5201652> (2018).
- Santarpia, L., Grandone, I., Contaldo, F. & Pasanisi, F. Butyrylcholinesterase as a prognostic marker: A review of the literature. *J. Cachexia Sarcopenia Muscle* **4**, 31–39. <https://doi.org/10.1007/s13539-012-0083-5> (2013).

21. Chen, V. P. *et al.* Butyrylcholinesterase deficiency promotes adipose tissue growth and hepatic lipid accumulation in male mice on high-fat diet. *Endocrinology* **157**, 3086–3095. <https://doi.org/10.1210/en.2016-1166> (2016).
22. Sridhar, G. *et al.* Serum butyrylcholinesterase in type 2 diabetes mellitus: A biochemical and bioinformatics approach. *Lipids Health Dis.* **4**, 18. <https://doi.org/10.1186/1476-511x-4-18> (2005).
23. Ceperuelo-Mallafre, V. *et al.* Zinc alpha-2 glycoprotein is implicated in dyslipidaemia in HIV-1-infected patients treated with antiretroviral drugs. *HIV Med.* **13**, 297–303. <https://doi.org/10.1111/j.1468-1293.2011.00976.x> (2012).
24. Lei, L. *et al.* Circulating zinc- $\alpha$ 2-glycoprotein levels are low in newly diagnosed patients with metabolic syndrome and correlate with adiponectin. *Nutr. Metab. Lond.* **14**, 53. <https://doi.org/10.1186/s12986-017-0210-6> (2017).
25. Yang, M. *et al.* Zinc- $\alpha$ 2-glycoprotein is associated with insulin resistance in humans and is regulated by hyperglycemia, hyperinsulinemia, or liraglutide administration: Cross-sectional and interventional studies in normal subjects, insulin-resistant subjects, and subjects with newly diagnosed diabetes. *Diabetes Care* **36**, 1074–1082. <https://doi.org/10.2337/dc12-0940> (2013).
26. Bastard, J. P. *et al.* Diabetes and dyslipidaemia are associated with oxidative stress independently of inflammation in long-term antiretroviral-treated HIV-infected patients. *Diabetes Metab.* <https://doi.org/10.1016/j.diabet.2019.02.008> (2019).
27. Nix, L. M. & Tien, P. C. Metabolic syndrome, diabetes, and cardiovascular risk in HIV. *Curr. HIV/AIDS Rep.* **11**, 271–278. <https://doi.org/10.1007/s11904-014-0219-7> (2014).
28. Sobieszczyk, M. E. *et al.* Metabolic syndrome after HIV acquisition in South African Women. *J. Acquir. Immune Defic. Syndr.* **73**, 438–445. <https://doi.org/10.1097/qai.0000000000001123> (2016).
29. Policarpo, S., Rodrigues, T., Moreira, A. C. & Valadas, E. Cardiovascular risk in HIV-infected individuals: A comparison of three risk prediction algorithms. *Rev. Port. Cardiol.* **38**, 463–470. <https://doi.org/10.1016/j.repc.2019.08.002> (2019).
30. Lagathu, C. *et al.* Metabolic complications affecting adipose tissue, lipid and glucose metabolism associated with HIV antiretroviral treatment. *Expert Opin. Drug Saf.* **18**, 829–840. <https://doi.org/10.1080/14740338.2019.1644317> (2019).
31. Sears, S. *et al.* Metabolic syndrome among people living with HIV receiving medical care in Southern United States: Prevalence and risk factors. *AIDS Behav.* <https://doi.org/10.1007/s10461-019-02487-8> (2019).
32. Seth, A. & Sherman, K. E. Fatty liver disease in persons with HIV infection. *Top. Antivir. Med.* **27**, 75–82 (2019).
33. Palios, J., Kadoglou, N. P. E. & Lampropoulos, S. The pathophysiology of HIV-/HAART-related metabolic syndrome leading to cardiovascular disorders: The emerging role of adipokines. *Exp. Diabetes Res.* **103063–103063**, 2012. <https://doi.org/10.1155/2012/103063> (2012).
34. Silva, B., Peixoto, G., da Luz, S., de Moraes, S. & Peres, S. Adverse effects of chronic treatment with the Main subclasses of highly active antiretroviral therapy: A systematic review. *HIV Med.* **20**, 429–438. <https://doi.org/10.1111/hiv.12733> (2019).
35. Riddler, S. A. *et al.* Impact of HIV infection and HAART on serum lipids in men. *JAMA* **289**, 2978–2982. <https://doi.org/10.1001/jama.289.22.2978> (2003).
36. Adal, M., Howe, R., Kassa, D., Aseffa, A. & Petros, B. Malnutrition and lipid abnormalities in antiretroviral naïve HIV-infected adults in Addis Ababa: A cross-sectional study. *PLoS ONE* **13**, e0195942. <https://doi.org/10.1371/journal.pone.0195942> (2018).
37. Rasheed, S., Yan, J. S., Lau, A. & Chan, A. S. HIV Replication enhances production of free fatty acids, low density lipoproteins and many key proteins involved in lipid metabolism: A proteomics study. *PLoS ONE* **3**, e3003. <https://doi.org/10.1371/journal.pone.0003003> (2008).
38. Renga, B. *et al.* HIV-1 infection is associated with changes in nuclear receptor transcriptome, pro-inflammatory and lipid profile of monocytes. *BMC Infect. Dis.* **12**, 274. <https://doi.org/10.1186/1471-2334-12-274> (2012).
39. Trehan, I. & Manary, M. J. Management of severe acute malnutrition in low-income and middle-income countries. *Arch. Dis. Child.* **100**, 283–287. <https://doi.org/10.1136/archdischild-2014-306026> (2015).
40. Joy, T. *et al.* Dietary fat intake and relationship to serum lipid levels in HIV-infected patients with metabolic abnormalities in the HAART era. *AIDS (London, England)* **21**, 1591–1600. <https://doi.org/10.1097/QAD.0b013e32823644ff> (2007).
41. Mansfield, K. G. *et al.* A diet high in saturated fat and cholesterol accelerates simian immunodeficiency virus disease progression. *J. Infect. Dis.* **196**, 1202–1210. <https://doi.org/10.1086/521680> (2007).
42. Lindegaard, B. *et al.* Adipose tissue expression of IL-18 and HIV-associated lipodystrophy. *AIDS* **18**, 1956–1958. <https://doi.org/10.1097/00002030-200409240-00013> (2004).
43. Ockenga, J. *et al.* ESPEN guidelines on enteral nutrition: Wasting in HIV and other chronic infectious diseases. *Clin. Nutr.* **25**, 319–329. <https://doi.org/10.1016/j.clnu.2006.01.016> (2006).
44. Wanke, C. A. *et al.* A medium chain triglyceride-based diet in patients with HIV and chronic diarrhea reduces diarrhea and malabsorption: A prospective, controlled trial. *Nutrition* **12**, 766–771. [https://doi.org/10.1016/S0899-9007\(96\)00221-3](https://doi.org/10.1016/S0899-9007(96)00221-3) (1996).
45. Dillon, S. M., Frank, D. N. & Wilson, C. C. The gut microbiome and HIV-1 pathogenesis: A two-way street. *AIDS (London, England)* **30**, 2737–2751. <https://doi.org/10.1097/QAD.0000000000001289> (2016).
46. Amadi, B. *et al.* Intestinal and systemic infection, HIV, and mortality in Zambian children with persistent diarrhea and malnutrition. *J. Pediatr. Gastroenterol. Nutr.* **32**, 550–554. <https://doi.org/10.1097/00005176-200105000-00011> (2001).
47. Kessler, L., Daley, H., Malenga, G. & Graham, S. The impact of HIV infection on the clinical presentation of severe malnutrition in children at QECH. *Malawi Med. J. J. Med. Assoc. Malawi* **13**, 30–33 (2001).
48. Musoke, P. M. & Fergusson, P. Severe malnutrition and metabolic complications of HIV-infected children in the antiretroviral era: Clinical care and management in resource-limited settings. *Am. J. Clin. Nutr.* **94**, 1716S–1720S. <https://doi.org/10.3945/ajcn.111.018374> (2011).
49. World Health Organization. Management of severe malnutrition: a manual for physicians and other senior health workers. <https://apps.who.int/iris/handle/10665/41999> (1998).
50. Njunje, J. M. *et al.* Biomarkers of post-discharge mortality among children with complicated severe acute malnutrition. *Sci. Rep.* **9**, 5981. <https://doi.org/10.1038/s41598-019-42436-y> (2019).
51. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. <https://www.R-project.org/> (2017).
52. Kim, R. S. & Kaplan, R. C. Analysis of secondary outcomes in nested case–control study designs. *Stat. Med.* **33**, 4215–4226. <https://doi.org/10.1002/sim.6231> (2014).
53. Samuelsen, S. O. A pseudolikelihood approach to analysis of nested case–control studies. *Biometrika* **84**, 379–394 (1997).
54. Pan, Y., Cai, J., Longnecker, M. P. & Zhou, H. Secondary outcome analysis for data from an outcome-dependent sampling design. *Stat. Med.* **37**, 2321–2337. <https://doi.org/10.1002/sim.7672> (2018).
55. Jiang, Y., Scott, A. J. & Wild, C. J. Secondary analysis of case–control data. *Stat. Med.* **25**, 1323–1339. <https://doi.org/10.1002/sim.2283> (2006).
56. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. Series B (Methodol.)* **57**, 289–300 (1995).
57. Friedman, J. H., Hastie, T. & Tibshirani, R. Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* **33**, 22. <https://doi.org/10.18637/jss.v033.i01> (2010).
58. Zou, H. & Hastie, T. Regularization and variable selection via the elastic net. *J. Roy. Stat. Soc. Ser. B. Stat. Method.* **67**, 301–320. <https://doi.org/10.1111/j.1467-9868.2005.00503.x> (2005).
59. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protocols* **4**, 44–57. <https://doi.org/10.1038/nprot.2008.211> (2008).

60. Smith, G. C. S., Seaman, S. R., Wood, A. M., Royston, P. & White, I. R. Correcting for optimistic prediction in small data sets. *Am. J. Epidemiol.* **180**, 318–324. <https://doi.org/10.1093/aje/kwu140> (2014).
61. da Huang, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57. <https://doi.org/10.1038/nprot.2008.211> (2009).
62. Walter, W., Sanchez-Cabo, F. & Ricote, M. GOpot: An R package for visually combining expression data with functional analysis. *Bioinformatics* **31**, 2912–2914. <https://doi.org/10.1093/bioinformatics/btv300> (2015).

## Acknowledgements

The original clinical trial was supported by the Thrasher Fund (Grant number: 9403), awarded to JAB. GBG is a postdoctoral fellow of the Research Foundation—Flanders (FWO). The Ghent University—VLIR-UOS Global Minds Fund supported the travel of GBG to Kenya. JAB and JMN are currently supported by the Bill & Melinda Gates Foundation within the Childhood Acute Illness and Nutrition (CHAIN) Network (Grant OPP1131320). JAB is currently supported by the MRC/DfID/Wellcome Trust Joint Global Health Trials scheme (Grant MR/M007367/1). The funders had no role in the design, conduct, analysis, or writing of this manuscript.

## Author contributions

G.B.G., J.M.N. and J.A.B. designed the study and data analysis. J.M.N. and B.G. performed the proteomics analysis. G.B.G. performed data analysis. B.W., R.B., W.V., I.P. and J.A.B. were involved in the clinical aspects of the study. G.B.G. wrote the initial drafts of the manuscript and all authors contributed to editing and improving the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-020-68143-7>.

**Correspondence** and requests for materials should be addressed to G.B.G.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020