blood global hematology

American Society of Hematology
2021 L Street NW, Suite 900,
Washington, DC 20036
Phone: 202-776-0544 | Fax 202-776-0545
bloodglobal@hematology.org

Circulating plasma EBV DNA: a potential tool to facilitate diagnosis of pediatric Burkitt lymphoma in sub-Saharan Africa

Tracking no: BGH-2025-000216R2

Karena Volesky-Avellaneda (National Cancer Institute, United States) Rena Xian (Johns Hopkins School of Medicine, United States) Todd Young (National Cancer Institute, United States) Ryan Baugher (Frederick National Laboratory for Cancer Research, United States) David Ogwang (ST. MARY'S HOSPITAL LACOR, Uganda) Patrick Kerchan (African Field Epidemiology Network, Uganda) Steven Reynolds (NIAID/NIH, United States) Constance Tenge (Moi University, Kenya) Pamela Were (Moi Teaching and Referral Hospital, Kenya) Robert Kuremu (Moi University, Kenya) Walter Wekesa (Moi University, Kenya) Nestory Masalu (Bugando Medical Center,) Esther Kawira (Shirati Health, Education, and Development Foundation, Tanzania-United Republic) stephanie mellot (National Cancer Instiute, United States) Heidi Lawhorn (National Cancer Instiute, United States) Kristen Pike (Frederick National Laboratory for Cancer Research, operated by Leidos Biomedical Research, Inc. for the National Cancer Institute, United States) Jennifer Stone (Johns Hopkins School of Medicine, United States) laura walsh (Johns Hopkins School of Medicine, United States) Tobias Kinyera (ST. MARY'S HOSPITAL LACOR, Uganda) Isaac Otim (ST. MARY'S HOSPITAL LACOR, Uganda) Hadijah Nabalende (ST. MARY'S HOSPITAL LACOR, Uganda) Herry Dhudha (Bugando Medical Center, Tanzania-United Republic) Leona Ayers (Ohio State University, United States) Kishor Bhatia (Lantern Pharma, United States) James Goedert (Anchor Mental Health, United States) Dan Soppet (National Cancer Institute, United States) Stephen Hewitt (National Cancer Institute, United States) Hyokyoung Hong (National Cancer Institute, United States) Richard Ambinder (Johns Hopkins University School of Medicine, United States) Sam Mbulaiteye (National Cancer Institute, United States)

Abstract:
Due to diagnostic delays in sub-Saharan Africa, many pediatric patients with Burkitt lymphoma (BL) perish. A subtype of BL, accounting for 95% of pediatric BL cases in sub-Saharan Africa, is characterized by the presence of Epstein-Barr Virus (EBV). We assessed plasma EBV DNA in 400 BL cases and 400 controls frequency-matched on sex, age, and country enrolled in the Epidemiology of Burkitt lymphoma in East African Children and Minors (EMBLEM) study in Uganda, Kenya, and Tanzania (2010-2016). EBV was measured using digital droplet PCR assay targeting EBV BamH1-W internal repeats, duplexed with RPP30 human housekeeping gene. The study population was predominately male (63% of cases and 64% of controls), with an average age of approximately 7.5 years. EBV DNA was detected in 309 (77.3%) BL cases and 62 (15.5%) controls. The mean plasma EBV DNA levels were 5.00 (standard deviation [SD] 1.63) log10 copies/mL for BL cases versus 1.94 (SD 1.35) for controls ($p<0.0001$). Excluding 50 BL cases and 61 controls with indeterminate (RPP30 and EBV negative) samples, the maximum sensitivity of plasma EBV DNA detection to discriminate BL cases from controls was 88.3% (95% CI 84.4-91.5%), with 81.7% (77.2-85.7%) specificity, and area under the curve 91.9% (89.7-93.9%). A specificity of 100% was achieved at a threshold of 4.19 log10 EBV copies/mL, which reduced sensitivity to 66.6%. Assay accuracy varied from 83%-87.4% at different thresholds. Testing for EBV DNA could facilitate the recognition of BL earlier in sub-Saharan Africa, a critical step in improving BL cure rates in the region.

Conflict of interest: No COI declared

COI notes:

Preprint server: No;

Author contributions and disclosures: R.R.X., S.M.M., K.B., and R.F.A. conceived the idea and designed the study. K.V.A. performed the statistical analysis; H.G.H. advised on statistical methods. S.M.M, M.D.O, S.J.R., P.K., C.N.T., W.N.W., N.M., and E.K. supervised the fieldwork; K.V.A. and S.M.M wrote the first draft of the manuscript. I.O., P.A.W., T.K., H.N., H.D., L.W. A., K.B., S.J.R., and J.J.G. conducted and monitored fieldwork. T.B.Y., R.N.B., S.D.M., H.E.L., H.H., J.S., L.W. performed laboratory testing of samples. K.M. P. coordinated laboratory work. All authors contributed to manuscript revision, read, and approved the submitted version.

Non-author contributions and disclosures: No;

Agreement to Share Publication-Related Data and Data Sharing Statement: The datasets used in this study may be requested from the corresponding author, Sam M. Mbulaiteye. Additional details on the EMBLEM study design and implementation manuals, including data collection forms used for data collection, can be downloaded from the EMBLEM website: https://emblem.cancer.gov/.

Clinical trial registration information (if any):

Circulating plasma EBV DNA: a potential tool to facilitate diagnosis of pediatric Burkitt lymphoma in sub-Saharan Africa

Running header: Plasma EBV DNA for pediatric BL diagnosis

Karena Volesky-Avellaneda[1*], Rena R. Xian[2*], Todd B. Young[3], Ryan N. Baugher[3], Martin D. Ogwang[4], Patrick Kerchan[5], Steven J. Reynolds[6], Constance N. Tenge[7], Pamela A. Were[8], Robert T. Kuremu[7], Walter N. Wekesa[7], Nestory Masalu[9], Esther Kawira[10], Stephanie D. Mellott[3], Heidi E. Lawhorn[3], Kristen M. Pike[3], Jennifer Stone[11], Laura Walsh[11], Tobias Kinyera[12], Isaac Otim[12], Hadijah Nabalende[12], Herry Dhudha[10], Leona W. Ayers[13], Kishor Bhatia[1], James J. Goedert[1], Daniel R. Soppet[3], Stephen M. Hewitt[3], Hyokyoung G. Hong[1], Richard F. Ambinder[14†], and Sam M. Mbulaiteye[1†‡]

[1]Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD, USA.[2]Department of Pathology, School of Medicine, Johns Hopkins, Baltimore, MD, USA.[3]Frederick National Laboratory for Cancer Research, National Cancer Institute, Frederick, MD, USA.[4]St. Mary's Hospital, Lacor, Gulu, Uganda.[5]Kuluva Hospital, Arua, Uganda.[6]Division of Intramural Research, National Institute of Allergy and Infectious Diseases, Bethesda, MD, USA.[7]Moi University, Eldoret, Kenya.[8]Academic Model Providing Access To Healthcare (AMPATH), Moi Teaching and Referral Hospital, Eldoret, Kenya.[9]Bugando Medical Center, Mwanza, Tanzania.[10]Shirati Health, Education, and Development Foundation, Shirati, Tanzania.[11]School of Medicine, Johns Hopkins University, Baltimore, MD, USA.[12]African Field Epidemiology Network, Gulu, Uganda.[13]Ohio State University, Columbus, OH, USA.[14]Department of Oncology, School of Medicine, Johns Hopkins University, Baltimore, MD, USA.

* Shared first co-authorship
† Shared senior co-authorship
‡ Correspondence: Sam M. Mbulaiteye
Division of Cancer Epidemiology & Genetics
National Cancer Institute, NIH, HHS
9609 Medical Center Dr, Rm. 6E118 MSC 3330
Bethesda, MD, 20892
Office Tel: +1 (240) 276-7108/ Office Fax: +1 (240) 276-7836
Email: mbulaits@mail.nih.gov

Data Sharing Statement

For original data, please contact Sam M. Mbulaiteye (mbulaits@mail.nih.gov). Study details on the EMBLEM study design and implementation manuals, including data collection forms can be downloaded from the EMBLEM website (https://emblem.cancer.gov/)

Word count: Text: 3,582; Abstract: 250; Figures: 3; Tables: 3; Reference count: 60.

Key points

- Plasma EBV DNA copy number discriminated BL cases from healthy children in East Africa with sensitivity, specificity, accuracy above 80%.
- Circulating EBV DNA copy number could facilitate diagnosis of BL earlier, reduce diagnostic delays, and increase BL cure in Africa.

Abstract

Due to diagnostic delays in sub-Saharan Africa, many pediatric patients with Burkitt lymphoma (BL) perish. A subtype of BL, accounting for 95% of pediatric BL cases in sub-Saharan Africa, is characterized by the presence of Epstein-Barr Virus (EBV). We assessed plasma EBV DNA in 400 BL cases and 400 controls frequency-matched on sex, age, and country enrolled in the Epidemiology of Burkitt lymphoma in East African Children and Minors (EMBLEM) study in Uganda, Kenya, and Tanzania (2010-2016). EBV was measured using digital droplet PCR assay targeting EBV BamH1-W internal repeats, duplexed with RPP30 human housekeeping gene. The study population was predominately male (63% of cases and 64% of controls), with an average age of approximately 7.5 years. EBV DNA was detected in 309 (77.3%) BL cases and 62 (15.5%) controls. The mean plasma EBV DNA levels were 5.00 (standard deviation [SD] 1.63) $\log_{10}$ copies/mL for BL cases versus 1.94 (SD 1.35) for controls ($p<0.0001$). Excluding 50 BL cases and 61 controls with indeterminate (RPP30 and EBV negative) samples, the maximum sensitivity of plasma EBV DNA detection to discriminate BL cases from controls was 88.3% (95% CI 84.4-91.5%), with 81.7% (77.2-85.7%) specificity, and area under the curve 91.9% (89.7-93.9%). A specificity of 100% was achieved at a threshold of 4.19 $\log_{10}$ EBV copies/mL, which reduced sensitivity to 66.6%. Assay accuracy varied from 83%-87.4% at different thresholds. Testing for EBV DNA could facilitate the recognition of BL earlier in sub-Saharan Africa, a critical step in improving BL cure rates in the region.

Keywords: Non-Hodgkin Lymphoma, epidemiology, Burkitt lymphoma, Epstein-Barr virus, *Plasmodium falciparum* malaria, HIV/AIDS.

Introduction

Burkitt lymphoma (BL) is an aggressive B-cell lymphoma characterized by translocation of *MYC* into the vicinity of immunoglobulin gene enhancers, either heavy chain- or light chain-genes, leading to constitutive expression of *MYC* and high cellular proliferation.[1] In fact, BL is one of the fastest growing tumors with a doubling time 24-48 hours.[2] Historically grouped into endemic, sporadic, and immunodeficiency-associated subtypes,[3-5] recent studies suggest that BL is better defined by Epstein-Barr Virus (EBV) infection status in the tumor.[6] The presence or absence of EBV correlates with distinct somatic molecular and epigenetic patterns.[7]

BL incidence is 20 times higher in sub-Saharan Africa (SSA) than elsewhere, and is responsible for most childhood cancers in many SSA countries.[8] Although BL is highly curable when treated promptly,[9,10] only 38-55% of cases in SSA are cured.[11-13] This is largely due to treatment delays from symptom onset to the start of cancer treatment.[14] The median treatment delay is 13 weeks in Uganda and Tanzania,[14] including seven weeks during which the patient was within the health care system. These delays increase the likelihood that cases will present with advanced stage disease[15] and with a compromised performance status,[16] both indicators of poor prognosis.

Since 95% of BL cases in SSA are EBV positive,[1] and tumor cells typically carry a median of 50 EBV genome copies per cell,[17,18] dying tumor cells could release sufficient tumor EBV DNA in plasma that could be measured to facilitate earlier recognition of BL and potentially reduce treatment delays. EBV DNA tests have been extensively

evaluated in other EBV-associated malignancies,[19-21] but not in BL. An EBV DNA test could also be used as an intermediate disease-predictive biomarker to screen for BL or map BL hotspots.[22,23] The few studies that evaluated plasma EBV DNA in pediatric BL cases in SSA (described in Supplementary Table 1) [24-28] reported EBV DNA detection in 60-100% of BL cases, but also in 5-84% of children without BL, raising concerns about insufficient specificity. Due to varying study designs, specimens tested, detection technologies, and EBV genes targeted (EBNA3C, BALF51, EBNA2, or LMP-1), it is challenging to draw conclusions on the clinical utility of EBV detection in pediatric BL cases. We initially conducted a pilot study of 25 BL cases and 25 controls from Uganda[24] in the Epidemiology of Burkitt lymphoma in East African Children and Minors (EMBLEM) study[29,30] using quantitative polymerase chain reaction (qPCR) targeting EBV's BamH1-W repeats.[31] Targeting BamH1-W may be more sensitive than EBNA3C, BALF51, EBNA2, or LMP-1 because there are 9-12 repeats per viral genome versus one copy for the other EBV genes.[32] The assay discriminated BL cases from controls with 88% sensitivity, 100% specificity, and area under the curve (AUC) of 94% for circulating plasma EBV DNA using a threshold of 2.57 $\log_{10}$ copies/mL. Herein, we report expanded results from 400 BL cases and 400 controls from Uganda, Tanzania, and Kenya EMBLEM study participants using a custom digital droplet PCR (ddPCR) BamH1-W EBV assay.

Materials and methods

Study population

The EMBLEM study methods have been reported previously.[29,30] Briefly, BL cases and controls aged 0-15 years were enrolled from six neighboring regions in Uganda,

Tanzania, and Kenya study from November 2010 to September 2016. In 74% of BL cases the diagnosis was made by local histology or cytology, with only clinical diagnoses for the remainder.[30] Controls were healthy-appearing children from 295 randomly selected villages.[33-35] The controls were frequency-matched (hence, matched population controls) on age, sex, and area to the historical BL cases reported in their region (in the 10 years prior to EMBLEM study enrollment).[36,37] For this study, we randomly selected 400 BL cases from 697 cases and 400 controls from 2,934 matched controls, with frequency matching to the selected cases on sex, age, and country. Ethical approval for the study was given by the Uganda Virus Research Institute Research and Ethics Committee (UVRI, GC/127), Uganda National Council for Science and Technology (HS 816), Tanzania National Institute for Medical Research (NIMR/HQ/R.8c/Vol. IX/1023), Moi University/Moi Teaching and Referral Hospital (000536), and the US National Cancer Institute (10-C-N133). Written informed consent was obtained from the participants' parent or guardian and assent from children aged seven years or older.

Venous blood samples were collected (before treatment in BL cases) in EDTA tubes and separated into buffy coat, plasma, and red cell blood cell layers by centrifugation (1300g for 15 minutes) and stored at -80°C. Frozen samples were shipped to the NCI Frederick National Laboratory for Cancer Research under liquid nitrogen vapor for long-term storage.

Plasma DNA extraction

DNA was extracted from 250 µL aliquots of plasma from previously unthawed plasma sample vials. Plasma aliquots were thawed via heat block at 37°C for five minutes, then transferred to a KingFisher™ Deep-Well 96 Plate (Life Technologies Corporation Carlsbad, CA, USA) on the automated Hamilton Microlab® Prep™ Liquid Handling System (Hamilton Company Reno, NV, USA) where DNA extraction was performed using the Apostle MiniMax™ High Efficiency cfDNA Isolation Kit (Beckman Coulter Inc., Sykesville, MD, USA) with a modified 55°C incubation and reagents and volume scaled down for 250 µL plasma input. Elution incubation was modified from the manufacturer's protocol to be performed at 55°C.

EBV DNA and RPP30 ddPCR

The samples (with their case status masked) were tested at the CLIA Molecular Diagnostics Laboratory (CMDL), Frederick National Laboratory for Cancer Research using a novel ddPCR assay targeting a 71 bp sequence within the *Bam*H1-W repeat region of the EBV genome (probe sets EBV_BamH1-W_ddPCR_FAM: *5'-/6-FAM/ CACACACTA/ZEN/CACACACCCACCCGTCTC/IABkFQ/-3'* ; EBV_BamH1-W_ddPCR_AF: CCAACACTCCACCACACC*;* EBV_BamH1-W_ddPCR_ZR: TCTTAGGAGCTGTCCGAGG, Primers: Life Technologies Corporation Carlsbad, CA, USA; Probes: Integrated DNA Technologies, Coralville, IA, USA). The ddPCR technology was selected because it eliminates the need for standard curves for quantification of EBV DNA copies while maintaining precision, making it more feasible

for small volume samples and potentially less expensive and thus more suitable for use in SSA. The assay was optimized and performed.

The analytical sensitivity of the assay was 1-2 EBV copies/ng of DNA based on spiking experiments with BL-derived cell lines NAMALWA (1-2 EBV copies per cell) and RAJI (50-60 copies per cell). The assay was duplexed with a human housekeeping gene, ribonuclease P/MRP subunit p30 (RPP30; Assay ID: dHsaCPE5038241; Bio-Rad, Hercules, CA, USA), to allow for normalization on a per cell basis and to confirm successful extraction of DNA if EBV DNA was not detected in a sample. Samples underwent droplet generation in 20 µL reactions per manufacturer's protocol on a QX200 Manual Droplet Generator (Bio-Rad) using ddPCR™ 2X Supermix for Probes (No dUTP; Bio-Rad), BamH1-W primer/probe mix (900 nM/250 nM final concentrations in reaction), and RPP30 gene expression assay. Cycling was done on the C1000 TOUCH CYCLER w/96W DP RM (Bio-Rad) under the following conditions with a ramp rate 2°C per second: 1 cycle at 95°C for 10 minutes, 40 cycles at 94°C for 30 seconds and 60°C for 1 minute, 1 cycle at 98°C for 10 minutes, and an infinite hold step at 4°C. Sample results were read using the Bio-Rad QX200 Droplet Reader (Bio-Rad) and analyzed using Bio-Rad QX Manager v1.2.345 software. BamH1-W and RPP30 concentrations in copies/µL were determined by applying manual thresholds to individual 2-D amplitude scatter plots in the software and back calculated to copies/mL. The reliability of the ddPCR assay was assessed in replicate samples from five BL cases and five controls (six embedded within-batch and four between-batches). The laboratory staff were blinded to the case/control or replicate status of the samples.

Statistical analysis

Results of samples with positive droplets of EBV BamH1-W were log-transformed to base 10. The reliability of the results was evaluated by calculating the coefficient of variation (CV) as the standard deviation [SD]/mean $\log_{10}$ EBV DNA copies/mL of the study replicate samples and expressed as a percent. Seven replicate samples (from two BL cases and five controls) were concordantly EBV negative, and three replicate samples (all from BL cases) were concordantly EBV positive with an average CV of 1.8%. The mean $\log_{10}$ EBV copies/mL in BL cases and controls were compared using the unpaired t-test. The diagnostic value of plasma EBV DNA for BL was assessed using non-parametric receiver operator characteristic (ROC) curves with bootstrapping (1000 replicates) to calculate the area under the curve (AUC) and associated 95% confidence intervals (CIs).[38,39] The threshold that maximized discrimination of BL cases from controls was evaluated using EBV copies/mL in increments of 0.5 $\log_{10}$, thresholds determined subjectively by visual inspection of the data, and 2.57 $\log_{10}$/mL based on our pilot study.[24] Additionally, the R command “cutpointr” was used to implement a bootstrapping method (1000 resamples) to select an optimal Youden’s Index (sensitivity + specificity – 1), defined as the median of the 1000 thresholds. For each threshold, the sensitivity (true positive/true positives + false negatives), specificity (true negatives/true negatives + false positives), accuracy (true positives + true negatives/total number), and Youden’s Index are reported.[40,41] The positive and negative predictive values were not calculated because these values are affected by the prevalence of disease, therefore our 1:1 matched study design with 50% prevalence of BL would give misleading results. The associations between plasma EBV DNA copies and demographic, geographic, and

laboratory characteristics were assessed separately in the cases and controls using the two-sample Wilcoxon rank-sum test or Kruskal-Wallis quality-of-populations rank test. Additionally, Pearson's correlations and corresponding *P*-values were calculated to assess the associations between case status and each of the following variables: $\log_{10}$ RPP30 copies/mL, white blood cell count (10^9/L), and $\log_{10}$ thick parasite count. Data were analyzed in Stata (StataCorp. 2023. Stata Statistical Software: Release 17. College Station, TX), ROC curves and threshold selection based on maximizing Youden's Index produced in R,[42] and dot plots in JMP (v18.0.0, SAS Institute Inc., Cary, NC,1989–2024).

Results

Most (93%) BL cases in this study were diagnosed histologically or cytologically. As shown in Table 1, the BL cases and controls had similar mean age of ~7.5. Half of the participants were from Uganda, one-third from Kenya, and the remaining from Tanzania. The geographic characteristics of the BL cases and their frequency-matched controls fit the expected pattern of residing in areas suitable for intense *P. falciparum* transmission i.e., over-represented in rural villages, villages near surface water, and enrolled during the wet season.[29]

Plasma EBV DNA was detected in 309 (77.3%) BL cases and in 62 (15.6%) controls, including 21 cases and 3 controls that were negative for RPP30 (Table 1). Fourteen percent of the participants (50 cases and 61 controls) were negative for both EBV and RPP30, so their samples were considered indeterminate due to insufficient DNA. *P.*

falciparum-negative infection status was observed in 36 (72%) of the EBV-indeterminate cases and in 36 (59%) of the EBV-indeterminate controls (Supplementary Table 2). Otherwise, the characteristics of the indeterminate samples were unremarkable. Excluding the indeterminate samples, plasma EBV was detected in 88.3% of the BL cases and 18.3% of the controls.

Among those with detectable EBV, mean plasma EBV load was about 3.1 logs higher in BL cases (5.00 $\log_{10}$ EBV copies/mL, standard deviation: [SD] 1.63) than controls (1.94 $\log_{10}$ EBV copies/mL [SD 1.35], P<0.001, Figure 1). Supplementary Figure 1 shows these results inclusive of the EBV negative participants. No controls had EBV detected above the 25th percentile (4.2 $\log_{10}$ copies/mL) of the case values. Approximately 17% of cases (all RPP30 positive) had EBV DNA copies below the controls' lowest quartile (1.47 $\log_{10}$ copies/mL); most of these cases were from Kenya (34 of 53) (Supplementary Table 3).

Overall, detectable plasma EBV DNA correctly classified BL cases from controls with an AUC of 86.0%, which increased to 91.9% when excluding indeterminate samples (Fig. 2). Assay sensitivity was highest when all EBV positive samples were included (equivalent to a viral load threshold of 1.35 $\log_{10}$/mL), corresponding to 77.7% in analyses without- versus 88.3% in analyses excluding indeterminate samples (Table 2). The corresponding specificities were 84.5% and 81.7%, respectively. Viral load thresholds based on visual inspection and bootstrapping were similar (2.21 versus 2.27

$\log_{10}$/mL) and had similar sensitivities, specificities, accuracy, and Youden's Index (Table 2). A viral load threshold of 2.57 $\log_{10}$/mL based on our pilot study[24] yielded comparable results. For EBV plasma DNA copy number $\geq$4.19 $\log_{10}$/mL (a value slightly higher than the highest value in the controls), the specificity was 100%, but sensitivity fell to 58.3% without exclusion of indeterminate samples, and the accuracy was 79.1%. The sensitivity improved to 66.6% when indeterminate samples were excluded with specificity of 100.0% and accuracy of 83.0%.

Among the controls, no differences in plasma EBV loads were noted by demographic, geographic, or laboratory characteristics (Table 3). Among BL cases, plasma EBV loads were higher in females than males (*P*=0.0118) and increased with increasing age (*P*=0.0004). Plasma EBV copies were highest in Uganda BL cases (5.73 $\log_{10}$ copies/mL), intermediate in those from Tanzania (5.14 $\log_{10}$ copies/mL), and lowest in those from Kenya (4.61 $\log_{10}$ copies/mL, *P*=0.0001). EBV loads did not differ by the location of the home, season of enrollment, or RPP30 detection status (Table 3). The proportion positive for EBV positivity was similar for BL cases presenting in the head only (78.0%) versus head and abdomen/abdomen only (75.3%) (Figure 3), but EBV loads were lowest in BL presenting only in the head (mean: 4.56 $\log_{10}$ copies/mL) compared to abdominal (5.21 $\log_{10}$ copies/mL) or disseminated/other (5.51 $\log_{10}$ copies/mL).

Mean $\log_{10}$ RPP30 levels were similar in cases and controls (3.84 $\log_{10}$ copies/mL in cases and 3.28 $\log_{10}$ copies/mL in the controls). Mean RPP30 level was significantly lower in EBV-positive versus EBV-negative cases (3.77 versus 4.33, *P*=0.0001), but

levels in the positive versus negative controls were comparable (3.34 versus 3.27, *P*=0.206) (Supplementary Figure 2). RPP30 levels in samples from the three countries were also similar (Supplementary Figure 3). EBV levels did not correlate with RPP30 copies/mL (Supplementary Figure 4), were not different in those who were RPP30 positive versus negative (5.38 versus 5.36), or with white blood cell counts (Supplementary Figure 5).

The lower prevalence of recent and current *P. falciparum* infection observed in BL cases compared to controls has consistently been observed in our previous studies.[29,43] Among the controls, EBV levels were weakly correlated with $\log_{10}$ *P. falciparum* parasitemia (ρ=0.1908, *P*=0.034), but showed no correlation in the cases (Supplementary Figure 6). EBV loads did not differ by the *P. falciparum* infection status (Table 3). Many of the BL cases with EBV levels below the lower quartile observed in the controls were *P. falciparum* negative (46 of 53 cases) (Supplementary Table 3).

Discussion

Improving cure rates of pediatric cancers to >60% in SSA is an important goal of the Global Initiative for Childhood Cancer (GICC).[44] However, limited access to reliable pathology diagnosis[45,46] is an obstacle towards achieving that goal for BL. Delayed diagnosis of BL, which is characterized by high BL tumor kinetics,[2] contributes to presentation when BL is advanced, with patients often *in extremis* and therefore unable to tolerate chemotherapy related to other factors. For example, 162 of 562 BL patients

identified could not be treated in the 3rd Groupe Franco-Africain d'Oncologie Pediatrique (GFAOP) Lymphomes Malins B (GFA-LMB) study in West Africa because they presented with advanced stage disease, poor nutrition/performance status, or died before shortly after presentation.[47] Moreover, less than one-tenth of these cases were diagnosed by pathology, underscoring the need for other diagnostic approaches to facilitate recognition of BL in SSA.

Our results demonstrate that plasma EBV DNA has good to excellent discrimination of BL cases from controls in a large sample from SSA. Previous studies of plasma EBV DNA in pediatric BL cases in SSA[24-28] reported high EBV DNA detection (60-100%) in BL cases and variable detection (5.3-84%) in children without BL. However, each study used a different EBV gene target and methods with variable analytic sensitivities. A study in Malawi also supported the utility of plasma EBV DNA for pediatric BL diagnosis.[25] Findings from that study showed that plasma EBV DNA declined during treatment of pediatric BL and levels spiked in the cases that relapsed,[25] suggesting that plasma EBV DNA also may be a useful treatment monitoring test. Measuring plasma EBV DNA is minimally invasive and could decrease diagnostic delays and increase the confidence in treating pediatric BL in settings where pathology cannot be obtained in a timely manner. While high instrument and reagents costs preclude introducing ddPCR EBV testing at peripheral centers in SSA now, the results pave the way for innovation to away from the multi-instrument non-self-contained ddPCR technology to a sealed/self-contained, multiplex-capable single instrument technology, with a plus/minus or semi-quantitative result read out. Further research to standardize sample collection and

processing and using controls attending the same facility as the BL cases with symptoms resembling BL (jaw or abdominal swellings)[1] would help to validate our results and further demonstrate the clinical utility of EBV testing may be used as a tool in conjunction with other tools (i.e., assessment of clinical features) for earlier detection of BL in SSA. Since plasma EBV in BL cases is likely to consist of tumor-derived virus, other research directions may include evaluating methylation status of EBV DNA as biomarker for tumor-derived EBV DNA.[21] Insofar as EBV-associated tumor DNA is known to be CpG methylated while EBV virion DNA is never methylated, methods to exclude measurement of unmethylated viral DNA might enhance the specificity EBV DNA measurements in plasma with regard to tumor diagnosis.[32]

In our controls, we found a weak correlation between asymptomatic *P. falciparum* parasitemia and plasma EBV load. This is consistent with evidence that *P. falciparum* proteins reactivate EBV in people with asymptomatic infection[48] and those with acute malaria.[27,49] However, in one study,[27] plasma EBV DNA levels became undetectable in 85% of those children after malaria treatment, suggesting that EBV kinetics related to malaria are transient and malaria can be ruled out by repeat testing of suspected patients presenting with concurrent acute malaria. Since current National Malaria Control policy does not recommend treating asymptomatic *P. falciparum* parasitemia,[50] investigating the kinetics of EBV viremia in asymptomatic patients with or without treatment is an open question. As none of the controls had EBV levels above the lower 25th percentile of BL cases, BL should still be considered in asymptomatically infected suspected cases with significantly elevated EBV levels.

Among the BL cases, EBV DNA copy number was higher in BL cases where there was also abdominal disease versus head-only disease. This is consistent with the hypothesis that EBV DNA detected may reflect tumor DNA bulk. Previous studies have correlated anatomic site with disease stage,[51] i.e., head-only disease with limited stage disease and abdominal involved disease with advanced stage disease.[52,53] The higher EBV DNA copy number in females versus males, across countries and age group categories in the cases may reflect the higher proportion of BL cases with abdominal involved disease correlated with those characteristics and the correlation of abdominal involved disease with higher plasma EBV DNA copy number. The variation of plasma EBV levels in BL cases across countries is intriguing and unexplained. Possibly, the differences could be a clue into genotypic differences in EBV circulating in different countries. For example, the prevalence of EBV type 2, which has impaired ability to immortalize B cells and is associated with lower EBV levels than EBV type 1 in some studies,[54] is more prevalent in Kenya,[55] but big differences have not been reported in Uganda.[56]

The strengths of our study include having a relatively large sample size enrolled from three countries with detailed covariate data, which enabled us to increase the generalizability of our findings. Also, we used previously unthawed samples with detailed sample pre-analytic collection and processing information. Our assay used 250 µL input plasma volume, which is feasible in pediatric patients and suitable for packaging into point-of-care technology. We note that cell-stabilizing tubes rather than

EDTA tubes used in this study may improve the detection of circulating tumor DNA. Further, dual-centrifugation processing of whole blood, which is increasingly used for plasma cell-free DNA analyses,[57] may also improve test performance for BL diagnosis. While our use of healthy controls is ideal for assessing a screening assay, our results should be replicated using controls with BL-like symptoms attending the same health facilities where BL cases seek care. Our finding of significantly lower RPP30 levels in EBV positive cases suggests that there could be interference with EBV DNA reducing detection of RPP30 above certain thresholds. However, this interference apparently does not affect EBV levels, which were comparable in those who were RPP30 positive versus negative.

Our research fits in the aspiration goals for cancer research in low- and middle-income countries to leverage technology to improve cancer control.[58] The plasma EBV assay used here leveraged ddPCR technology to use small volume samples to detect EBV. The BamH1-W region is advantageous because it is present as multiple copies in the EBV genome,[32] potentially increasing the analytic sensitivity of the assay compared to targets present as a single copy per EBV genome.[59] However, variability in the number of BamH1-W copies (2-9 per genome) in circulating isolates may introduce variability in real-world results. Our results enable new translational research questions about optimal sample types, the methods to collect, prepare, and process specimens, standardization of assays (gene target, volume, and technology), and assay utility, including screening or prognostic applications in SSA. Other research may focus on

methylation status of EBV to characterize the origin, tumor or non-tumor,[21] or BL-associated EBV sequence variants.[60]

In conclusion, we showed that circulating plasma EBV DNA discriminates BL cases from controls with high specificity in a large study conducted in three African countries. Our work opens new research directions to replicate results and to address methodological and translational questions that will guide the application of circulating plasma EBV assays in SSA.

ACKNOWLEDGEMENTS

The EMBLEM study was funded by the National Cancer Institute, National Institutes of Health, under Contract Nos. HHSN261200800001E, HHSN261201100063C, and HHSN261201100007I (Division of Cancer Epidemiology and Genetics), and in part (SJR) by the Division of Intramural Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health. The EBV work was supported by the Office of Clinical Research Bench to Bedside Program Funds-award # 883923, NIH and grants R21CA232891, R01 CA250069, U01 CA271252, P30CA06973 to Johns Hopkins University. The authors acknowledge the research contributions of the Cancer Genomics Research Laboratory, NCI, for DNA extraction, quantification, aliquoting and shipping under NCI Contract No. 75N910D00024.The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products or organizations imply endorsement by the U.S. Government. The authors also acknowledge the Information Management Systems (Silver Spring, Maryland), Westat, Inc (Rockville, Maryland) and the African Field Epidemiology Network (Kampala, Uganda) for coordinating EMBLEM fieldwork in Uganda.

AUTHORSHIP CONTRIBUTIONS

R.R.X., S.M.M., K.B., and R.F.A. conceived the idea and designed the study. K.V.A. performed the statistical analysis; H.G.H. advised on statistical methods. S.M.M, M.D.O, S.J.R., P.K., C.N.T., W.N.W., N.M., and E.K. supervised the fieldwork; K.V.A. and S.M.M wrote the first draft of the manuscript. I.O., P.A.W., T.K., H.N., H.D., L.W. A., K.B., S.J.R.,

and J.J.G. conducted and monitored fieldwork. T.B.Y., R.N.B., S.D.M., H.E.L., H.H., J.S., L.W. performed laboratory testing of samples. K.M. P. coordinated laboratory work. All authors contributed to manuscript revision, read, and approved the submitted version.

Disclosure of conflicts of interest

None to declare

References

1. López C, Burkhardt B, Chan JKC, et al. Burkitt lymphoma. *Nat Rev Dis Primers*. 2022;8(1):78.

2. Iversen U, Iversen OH, Bluming AZ, Ziegler JL, Kyalwasi S. Cell kinetics of African cases of Burkitt lymphoma. A preliminary report. *Eur J Cancer*. 1972;8(3):305-308.

3. Mbulaiteye SM, Clarke CA, Morton LM, et al. Burkitt lymphoma risk in U.S. solid organ transplant recipients. *Am J Hematol*. 2013;88(4):245-250.

4. Guech-Ongey M, Simard EP, Anderson WF, Engels EA, Bhatia K, Devesa SS, Mbulaiteye SM. AIDS-related Burkitt lymphoma in the United States: what do age and CD4 lymphocyte patterns tell us about etiology and/or biology? *Blood*. 2010;116(25):5600-5604.

5. Mbulaiteye SM, Parkin DM, Rabkin CS. Epidemiology of AIDS-related malignancies an international perspective. *Hematol Oncol Clin North Am*. 2003;17(3):673-696, v.

6. Epstein MA, Achong BG, Barr YM. Virus Particles in Cultured Lymphoblasts from Burkitt's Lymphoma. *Lancet*. 1964;1(7335):702-703.

7. Alaggio R, Amador C, Anagnostopoulos I, et al. The 5th edition of the World Health Organization Classification of Haematolymphoid Tumours: Lymphoid Neoplasms. *Leukemia*. 2022;36(7):1720-1748.

8. Mbulaiteye SM, Devesa SS. Burkitt Lymphoma Incidence in Five Continents. *Hemato*. 2022;3(3):434-453.

9. Gopal S, Gross TG. How I treat Burkitt lymphoma in children, adolescents, and young adults in sub-Saharan Africa. *Blood*. 2018;132(3):254-263.

10. Ozuah NW, Lubega J, Allen CE, El-Mallawany NK. Five decades of low intensity and low survival: adapting intensified regimens to cure pediatric Burkitt lymphoma in Africa. *Blood advances*. 2020;4(16):4007-4019.

11. Chapman H, Ntemi PS, Gisiri M, Vasudevan L, Kashaigili HJ, Schroeder K. Retrospective analysis of pediatric patients with Burkitt lymphoma treated in Tanzania following the implementation of the 2016 National Treatment Guidelines: Poor outcomes to current standard-of-care therapy. *Pediatr Blood Cancer*. 2024;71(9):e31145.

12. Buckle G, Maranda L, Skiles J, et al. Factors influencing survival among Kenyan children diagnosed with endemic Burkitt lymphoma between 2003 and 2011: A historical cohort study. *Int J Cancer*. 2016;139(6):1231-1240.

13. Mutyaba I, Wabinga HR, Orem J, Casper C, Phipps W. Presentation and Outcomes of Childhood Cancer Patients at Uganda Cancer Institute. *Global Pediatric Health*. 2019;6:2333794X19849749.

14. Mawalla WF, Morrell L, Chirande LF, et al. Treatment Delays in Children and Young Adults with Lymphoma: report from an East Africa Lymphoma Cohort Study. *Blood Adv*. 2023.

15. Ozuah NW, Lubega J, Allen CE, El-Mallawany NK. Five decades of low intensity and low survival: adapting intensified regimens to cure pediatric Burkitt lymphoma in Africa. *Blood Adv*. 2020;4(16):4007-4019.

16. Okongo F, Ogwang DM, Liu B, Maxwell Parkin D. Cancer incidence in Northern Uganda (2013–2016). *International journal of cancer*. 2019;144(12):2985-2991.

17. Nonoyama M, Huang CH, Pagano JS, Klein G, Singh S. DNA of Epstein-Barr virus detected in tissue of Burkitt's lymphoma and nasopharyngeal carcinoma. *Proc Natl Acad Sci U S A*. 1973;70(11):3265-3268.

18. Lei H, Li T, Li B, et al. Epstein-Barr virus from Burkitt Lymphoma biopsies from Africa and South America share novel LMP-1 promoter and gene variations. *Sci Rep*. 2015;5:16706.

19. Chan KCA, Woo JKS, King A, et al. Analysis of Plasma Epstein-Barr Virus DNA to Screen for Nasopharyngeal Cancer. *N Engl J Med*. 2017;377(6):513-522.

20. Lo YM. Quantitative analysis of Epstein-Barr virus DNA in plasma and serum: applications to tumor detection and monitoring. *Ann N Y Acad Sci*. 2001;945:68-72.

21. Shamay M, Kanakry JA, Low JSW, et al. CpG methylation in cell-free Epstein-Barr virus DNA in patients with EBV-Hodgkin lymphoma. *Blood Adv*. 2020;4(8):1624-1627.

22. Gondwe Y, Salima A, Manda A, et al. Spatial distribution of incident pediatric Burkitt lymphoma in central and northern Malawi and association with malaria prevalence. *Pediatr Blood Cancer*. 2022;69(10):e29867.

23. Broen K, Dickens J, Trangucci R, et al. Burkitt lymphoma risk shows geographic and temporal associations with Plasmodium falciparum infections in Uganda, Tanzania, and Kenya. *Proc Natl Acad Sci U S A*. 2023;120(2):e2211055120.

24. Xian RR, Kinyera T, Otim I, et al. Plasma EBV DNA: A Promising Diagnostic Marker for Endemic Burkitt Lymphoma. *Front Oncol*. 2021;11:804083.

25. Westmoreland KD, Montgomery ND, Stanley CC, et al. Plasma Epstein-Barr virus DNA for pediatric Burkitt lymphoma diagnosis, prognosis and response assessment in Malawi. *Int J Cancer*. 2017;140(11):2509-2516.

26. Mulama DH, Bailey JA, Foley J, et al. Sickle cell trait is not associated with endemic Burkitt lymphoma: an ethnicity and malaria endemicity-matched case-control study suggests factors controlling EBV may serve as a predictive biomarker for this pediatric cancer. *Int J Cancer*. 2014;134(3):645-653.

27. Donati D, Espmark E, Kironde F, et al. Clearance of circulating Epstein-Barr virus DNA in children with acute malaria after antimalaria treatment. *J Infect Dis*. 2006;193(7):971-977.

28. Kabyemera R, Masalu N, Rambau P, et al. Relationship between non-Hodgkin's lymphoma and blood levels of Epstein-Barr virus in children in north-western Tanzania: a case control study. *BMC Pediatr*. 2013;13:4.

29. Peprah S, Ogwang MD, Kerchan P, et al. Risk factors for Burkitt lymphoma in East African children and minors: A case-control study in malaria-endemic regions in Uganda, Tanzania and Kenya. *Int J Cancer*. 2020;146(4):953-969.

30. Hong HG, Gouveia MH, Ogwang MD, et al. Sickle cell allele HBB-rs334(T) is associated with decreased risk of childhood Burkitt lymphoma in East Africa. *Am J Hematol*. 2024;99(1):113-123.

31. Kanakry JA, Li H, Gellert LL, et al. Plasma Epstein-Barr virus DNA predicts outcome in advanced Hodgkin lymphoma: correlative analysis from a large North American cooperative group trial. *Blood*. 2013;121(18):3547-3553.

32. Ambinder RF. Plasma Epstein-Barr Virus DNA for Screening. *N Engl J Med*. 2017;377(6):584-585.

33. Maziarz M, Kinyera T, Otim I, et al. Age and geographic patterns of Plasmodium falciparum malaria infection in a representative sample of children living in Burkitt lymphoma-endemic areas of northern Uganda. *Malar J*. 2017;16(1):124.

34. Peprah S, Dhudha H, Ally H, et al. A population-based study of the prevalence and risk factors of low-grade Plasmodium falciparum malaria infection in children aged 0-15 years old in northern Tanzania. *Trop Med Int Health*. 2019;24(5):571-585.

35. Peprah S, Tenge C, Genga IO, et al. A Cross-Sectional Population Study of Geographic, Age-Specific, and Household Risk Factors for Asymptomatic Plasmodium falciparum Malaria Infection in Western Kenya. *Am J Trop Med Hyg*. 2019;100(1):54-65.

36. Aka P, Kawira E, Masalu N, Emmanuel B, Brubaker G, Magatti J, Mbulaiteye SM. Incidence and trends in Burkitt lymphoma in northern Tanzania from 2000 to 2009. *Pediatr Blood Cancer*. 2012;59(7):1234-1238.

37. Ogwang MD, Bhatia K, Biggar RJ, Mbulaiteye SM. Incidence and geographic distribution of endemic Burkitt lymphoma in northern Uganda revisited. *Int J Cancer*. 2008;123(11):2658-2663.

38. Jin H, Lu Y. The ROC region of a regression tree. *Statistics & Probability Letters*. 2009;79(7):936-942.

39. Linnet K. Comparison of quantitative diagnostic tests: type I error, power, and sample size. *Stat Med*. 1987;6(2):147-158.

40. Ruopp MD, Perkins NJ, Whitcomb BW, Schisterman EF. Youden Index and optimal cut-point estimated from observations affected by a lower limit of detection. *Biom J*. 2008;50(3):419-430.

41. Nahm FS. Receiver operating characteristic curve: overview and practical use for clinicians. *Korean J Anesthesiol*. 2022;75(1):25-36.

42. R Development Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2018.

43. Peprah S, Ogwang MD, Kerchan P, et al. Inverse association of falciparum positivity with endemic Burkitt lymphoma is robust in analyses adjusting for pre-enrollment malaria in the EMBLEM case-control study. *Infect Agent Cancer*. 2021;16(1):40.

44. Smith ER, Cotache-Condor C, Leraas H, et al. Towards attainment of the 2030 goal for childhood cancer survival for the World Health Organization Global Initiative for Childhood Cancer: An ecological, cross-sectional study. *PLOS Glob Public Health*. 2024;4(8):e0002530.

45. Ogwang MD, Zhao W, Ayers LW, Mbulaiteye SM. Accuracy of Burkitt lymphoma diagnosis in constrained pathology settings: importance to epidemiology. *Arch Pathol Lab Med*. 2011;135(4):445-450.

46. Orem J, Sandin S, Weibull CE, et al. Agreement between diagnoses of childhood lymphoma assigned in Uganda and by an international reference laboratory. *Clin Epidemiol*. 2012;4:339-347.

47. Bouda GC, Traore F, Couitchere L, et al. Advanced Burkitt Lymphoma in Sub-Saharan Africa Pediatric Units: Results of the Third Prospective Multicenter Study of the Groupe Franco-Africain d'Oncologie Pediatrique. *J Glob Oncol*. 2019;5:1-9.

48. Chene A, Donati D, Guerreiro-Cacais AO, et al. A molecular link between malaria and Epstein-Barr virus reactivation. *PLoS Pathog*. 2007;3(6):e80.

49. Lam KM, Syed N, Whittle H, Crawford DH. Circulating Epstein-Barr virus-carrying B cells in acute malaria. *Lancet*. 1991;337(8746):876-878.

50. Nyeko R, Otim F, Obiya EM, Abala C. Anti-malarial drug use, appropriateness and associated factors among children under-five with febrile illnesses presenting to a tertiary health facility: a cross sectional study. *Malaria Journal*. 2023;22(1):103.

51. Ziegler JL. Burkitt's Lymphoma. *New England Journal of Medicine*. 1981;305(13):735-745.

52. Nkrumah FK, Perkins IV. Burkitt's lymphoma. A clinical study of 110 patients. *Cancer*. 1976;37(2):671-676.

53. Olweny CL, Katongole-Mbidde E, Otim D, Lwanga SK, Magrath IT, Ziegler JL. Long-term experience with Burkitt's lymphoma in Uganda. *Int J Cancer*. 1980;26(3):261-266.

54. Coleman CB, Daud II, Ogolla SO, et al. Epstein-Barr Virus Type 2 Infects T Cells in Healthy Kenyan Children. *The Journal of Infectious Diseases*. 2017;216(6):670-677.

55. Daud II, Coleman CB, Smith NA, et al. Breast Milk as a Potential Source of Epstein-Barr Virus Transmission Among Infants Living in a Malaria-Endemic Region of Kenya. *The Journal of Infectious Diseases*. 2015;212(11):1735-1742.

56. Petrara MR, Penazzato M, Massavon W, et al. Epstein-Barr virus load in children infected with human immunodeficiency virus type 1 in Uganda. *J Infect Dis*. 2014;210(3):392-399.

57. Lockwood CM, Borsu L, Cankovic M, et al. Recommendations for Cell-Free DNA Assay Validations: A Joint Consensus Recommendation of the Association for Molecular Pathology and College of American Pathologists. *J Mol Diagn*. 2023;25(12):876-897.

58. Pramesh CS, Badwe RA, Bhoo-Pathy N, et al. Priorities for cancer research in low- and middle-income countries: a global perspective. *Nat Med*. 2022;28(4):649-657.

59. Lee AWM, Lee VHF, Ng WT, et al. A systematic review and recommendations on the use of plasma EBV DNA for nasopharyngeal carcinoma. *Eur J Cancer*. 2021;153:109-122.

60. Liao HM, Liu H, Chin PJ, et al. Epstein-Barr Virus in Burkitt Lymphoma in Africa Reveals a Limited Set of Whole Genome and LMP-1 Sequence Patterns: Analysis of Archival Datasets and Field Samples From Uganda, Tanzania, and Kenya. *Front Oncol*. 2022;12:812224.

TABLE 1. Characteristics of BL cases and controls investigated for circulating plasma EBV DNA

Characteristics	BL cases		Controls	
	n	%	n	%
All participants	400[a]	100.0	400	100.0
Demographics				
Sex[b]				
Male	252	63.0	256	64.0
Female	148	37.0	144	36.0
Age, years				
Mean ± SD	7.5 ± 3.5	…	7.6 ± 3.4	…
Age group, years				
0–3	25	6.3	20	5.0
3–5	104	26.0	103	25.8
6–8	123	30.8	122	30.5
9–11	88	22.0	93	23.3
12–15	60	15.0	62	15.5
Geographic				
Country				
Kenya	148	37.0	133	33.3
Tanzania	48	12.0	50	12.5
Uganda	204	51.0	217	54.3
Urban/rural status of home[c]				
Urban	142	35.5	144	36.0
Rural	198	49.5	256	64.0
Missing	60	15.0	0	0.0
Proximity of home to surface water[d]				
Far from water	65	16.3	176	44.0
Near water	274	68.5	224	56.0
Missing	61	15.3	0	0.0
Season of enrollment[e]				
Wet	232	58.0	216	54.0
Dry	168	42.0	184	46.0
Laboratory				
P. falciparum infection status[f]				
Negative	298	74.5	222	55.5
Recent	26	6.5	54	13.5
Current	71	17.8	123	30.8
Missing	5	1.3	1	0.3
EBV/RPP30 status				
EBV and RPP30 positive	288	72.0	59	14.8
EBV positive only	21	5.3	3	0.8
RPP30 positive only	41	10.3	277	69.3
EBV and RPP30 negative[g]	50	12.5	61	15.3

BL Burkitt lymphoma, EBV Epstein Barr virus, EMBLEM Epidemiology of Burkitt Lymphoma in East African Children and Minors, *Plasmodium falciparum P. falciparum*, RPP30 Ribonuclease P/MRP protein subunit p30, SD standard deviation

a. BL cases were diagnosed by histology or cytology, except for 29 (7.3%) cases whose diagnosis was based only on clinical information.

b. Sex was defined genetically in all individuals, except 37 who lack genetic data but had self-reported sex.

c. Participants' residential village was classified as urban based on whether the population count of children aged 0–15 years in their village was equal or greater than the average population count for all the villages in the study region; otherwise, the village was classified as "rural".

d. Participants' residential village was classified as "near water" if any part of the village boundary was within 500 meters of all season surface water (lake, swamp, or river); otherwise, the village was classified as "far from water".

[e] Seasons were defined by calendar months using data from the National Weather Bureaus, with April to June and September to December defined as “wet” season months and January to March and July to August defined as the “dry” season months.

[f] *P. falciparum* infection status was grouped into three categories: “current” in those in whom asexual parasite forms could be visualized in thick films under the microscope, “recent” in those without visible asexual parasites forms but having detectable parasite antigenemia using commercial rapid antibody capture assays rapid diagnostic tests for *P. falciparum*-specific antigen histidine rich protein 2 (HRP2) and pan-lactate dehydrogenase (PanLDH).

[g] Participants negative for both EBV and RPP30 are included in the main analyses but excluded as a sensitivity analyses.

TABLE 2. Discrimination of BL cases from controls based on different $\log_{10}$ EBV copies/mL thresholds

Threshold selection method[d]	$\log_{10}$ EBV copies/mL[a]	Correctly classified		Misclassified		Sensitivity	Specificity	Accuracy[b]	Youden's index[c]
		BL cases	Controls	BL cases	Controls				
All samples, N=800									
Maximum sensitivity	1.35	309	338	91	62	77.3%	84.5%	80.9%	61.8%
	1.50	297	360	103	40	74.3%	90.0%	82.1%	64.3%
	2.00	281	377	119	23	70.3%	94.3%	82.3%	64.6%
	2.50	271	391	129	9	67.8%	97.8%	82.8%	65.6%
Based on Xian et al	2.57	270	393	130	7	67.5%	98.3%	82.9%	65.8%
	3.00	259	396	141	4	64.8%	99.0%	81.9%	63.8%
	3.50	249	398	151	2	62.3%	99.5%	80.9%	61.8%
	4.00	238	399	162	1	59.5%	99.8%	79.6%	59.3%
Maximum specificity	4.19	233	400	167	0	58.3%	100.0%	79.1%	58.3%
Visual inspection	2.21	278	385	122	15	69.5%	96.3%	82.9%	65.8%
Bootstrap method	2.27	275	387	125	13	68.8%	96.8%	82.8%	65.6%
Excluding indeterminate samples[e], N=689									
Maximum sensitivity	1.35	309	277	41	62	88.3%	81.7%	85.1%	70.0%
	1.50	297	299	53	40	84.9%	88.2%	86.5%	73.1%
	2.00	281	316	69	23	80.3%	93.2%	86.6%	73.5%
	2.50	271	330	79	9	77.4%	97.3%	87.2%	74.7%
Based on Xian et al	2.57	270	332	80	7	77.1%	97.9%	87.4%	75.0%
	3.00	259	335	91	4	74.0%	98.8%	86.2%	72.8%
	3.50	249	337	101	2	71.1%	99.4%	85.1%	70.5%
	4.00	238	338	112	1	68.0%	99.7%	83.6%	67.7%
Maximum specificity	4.19	233	339	117	0	66.6%	100.0%	83.0%	66.6%
Visual inspection	2.21	278	324	72	15	79.1%	95.6%	87.4%	75.0%
Bootstrap method	2.27	275	326	75	13	78.6%	96.2%	87.2%	74.8%

BL Burkitt lymphoma, EBV Epstein-Barr virus, mL milliliter, RPP30 Ribonuclease P/MRP protein subunit p30

[a.] The thresholds displayed start with the $\log_{10}$ copies/mL value (1.35) that would classify all those with detectible EBV as meeting the threshold, then possible thresholds are presented in increments of 0.50 $\log_{10}$ copies/mL to the threshold (4.19) that would produce a specificity of 100.0%.

[b.] Proportion of correct classifications calculated as the number of BL participants and controls correctly classified divided by the total.

[c.] Composite measure of sensitivity and specificity, calculated by adding the sensitivity and specificity and subtracting 1.
[d.] The threshold selected in the prior study by Xian et al., 2021, visual inspection, or bootstrap method to optimize Youden's Index.
[e.] Indeterminate samples were those testing both RPP30 and EBV negative (RPP30 negative, but EBV positive samples were included).

TABLE 3. Frequency of detection of circulating EBV DNA in plasma and median $\log_{10}$ EBV copies/mL among BL cases and controls

Characteristics	BL cases				Controls			
	No. EBV+ / No. tested	EBV+ (row %)	Median	*P*-value for median[a]	No. EBV+ / No. tested	EBV+ (row %)	Median	*P*-value for median[a]
All participants	309/400	77.3	5.38	…	62/400	15.5	1.74	…
Sex								
Male	203/252	80.6	5.20	**0.0118**	38/256	14.8	1.75	0.2905
Female	106/148	71.6	5.62		24/144	16.7	1.73	
Age group, years								
0–3	15/25	60.0	2.22	**0.0004**	3/20	15.0	2.23	0.4722
3–5	84/104	80.8	5.25		20/103	19.4	1.76	
6–8	99/123	80.5	5.41		23/122	18.9	1.75	
9–11	69/88	78.4	5.73		12/93	12.9	1.72	
12–15	42/60	70.0	5.40		4/62	6.5	1.91	
Country								
Kenya	100/148	67.6	4.64	**0.0001**	21/133	15.8	1.53	0.7611
Tanzania	39/48	81.3	5.14		4/50	8.0	1.91	
Uganda	170/204	83.3	5.73		37/217	17.1	1.74	
Urban/rural status of home								
Urban	106/142	74.7	5.51	0.9041	15/144	10.4	1.71	0.9611
Rural	160/198	80.8	5.37		47/256	18.4	1.74	
Proximity of home to surface water								
Far from water	54/65	83.1	5.32	0.1980	23/176	13.1	1.75	0.8851
Near water	211/274	77.0	5.41		39/224	17.4	1.74	
Season of enrollment								
Wet	176/232	75.9	5.47	0.3763	31/216	14.4	1.72	0.2067
Dry	133/168	79.2	5.21		31/184	16.9	1.99	
***P. falciparum* infection status**								
Negative	227/298	76.2	5.39	0.5041	24/222	10.8	1.48	0.0713
Recent	20/26	76.9	5.68		4/54	7.4	1.90	

Current	58/71	81.7	4.98		34/123	27.6	1.84	
RPP30 detection								
Positive	288/329	87.5	5.38	0.8891	59/336	17.6	1.74	0.7808
Negative	21/71	29.6	5.36		3/64	4.7	1.50	
Tumor anatomic site involvement								
Head only	99/127	78.0	4.88	**0.0001**	NA		NA	NA
Head and abdomen and abdomen only	177/235	75.3	5.73					
Other and disseminated	11/13	84.6	5.91					

BL Burkitt lymphoma, EBV Epstein Barr Virus, mL milliliter, NA not applicable, *Plasmodium falciparum P. falciparum*, RPP30 Ribonuclease P/MRP protein subunit p30

[a.] The *P*-value is from either the two-sample Wilcoxon rank-sum test or Kruskal-Wallis quality-of-populations rank test, as appropriate.

Figure legends

Figure 1. Dot plot of quantified circulating log_{10} EBV copies/mL in plasma for 309 BL cases and 62 controls with detectable EBV

BL Burkitt lymphoma, EBV Epstein-Barr virus, IQR interquartile range, mL milliliter, SD standard deviation

Figure legend

The box plots show the distribution of quantified EBV as log_{10} copies/mL, where each dot represents one individual. The boxes cover the interquartile range (IQR) and the horizontal line in the box represents the medians, the whiskers are 1.5 times the IQR, and the dots beyond the whiskers are outliers. The *P*-value tested the hypothesis of non-equality of means of EBV values in BL cases and controls (unpaired t-test).

Figure 2. Receiver operating characteristic (ROC) curves to summarize the performance of detection EBV DNA in plasma to discriminate BL cases from controls (a) all BL cases and controls, N=800; and (b) BL cases and controls, excluding those lacking detectible RPP30 and EBV, N=689

AUC area under the curve, BL Burkitt lymphoma, CI confidence interval, EBV Epstein-Barr virus, ROC receiver operating characteristic

Figure legend: These ROC curves plot the false positive rate (1 – specificity) against the true positive rate (sensitivity) for the detection of EBV in BL participants and controls, separately for all 800 BL participants and controls (panel a) and for 698 BL participants and controls that had either EBV or RPP30 detected. The shaded purple region represents the 95% confidence intervals based on the bootstrap method with 1000 replicates. The gray diagonal line is the line of no discrimination (AUC=50%). Since the majority (85%) of controls had no detectable EBV, there is a vertical incline at a false positive rate of 0 because controls without detectable EBV are correctly classified. There is a plateau at 77.3% (panel a) and 88.3% (panel b) because that is the proportion of BL cases with detectable EBV.

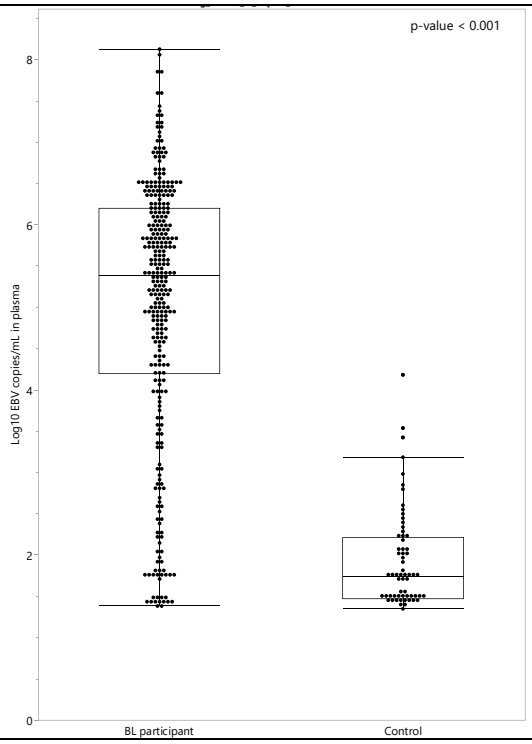
Figure 3. Dot plot of log_{10} EBV copies/mL for 287 BL cases with detectable EBV and anatomical site information, by anatomical site

BL Burkitt lymphoma, EBV Epstein-Barr virus, IQR interquartile range, mL milliliter, SD standard deviation

Figure legend: The box plots show the distribution of quantified EBV as log_{10} copies/mL in 287 BL participants, where each dot represents one BL participant. The boxes cover the interquartile range (IQR) and the horizontal line in the box represents the medians, the whiskers are 1.5 times the IQR, and the dots beyond the whiskers are outliers. The ANOVA test *P*-value was 0.003; post-hoc testing (Tukey honestly Significant Difference) *P*-values between groups were 0.003 (head only vs head/abdomen and abdomen-only), 0.142 (head-only vs disseminated/other), and 0.820 (head/abdomen-only vs disseminated and other).

W

Figure 1



EBV detected	309/400 (77.3%)	62/400 (15.5%)
EBV not detected	91/400 (22.7%)	338/400 (84.5%)
Among participants with detectable EBV		
Mean ± SD	5.00 ± 1.63	1.94 ± 1.35
Median (IQR)	5.38 (4.20–6.19)	1.74 (1.47–2.21)
Range	1.38–8.13	1.35–4.18

Figure 2

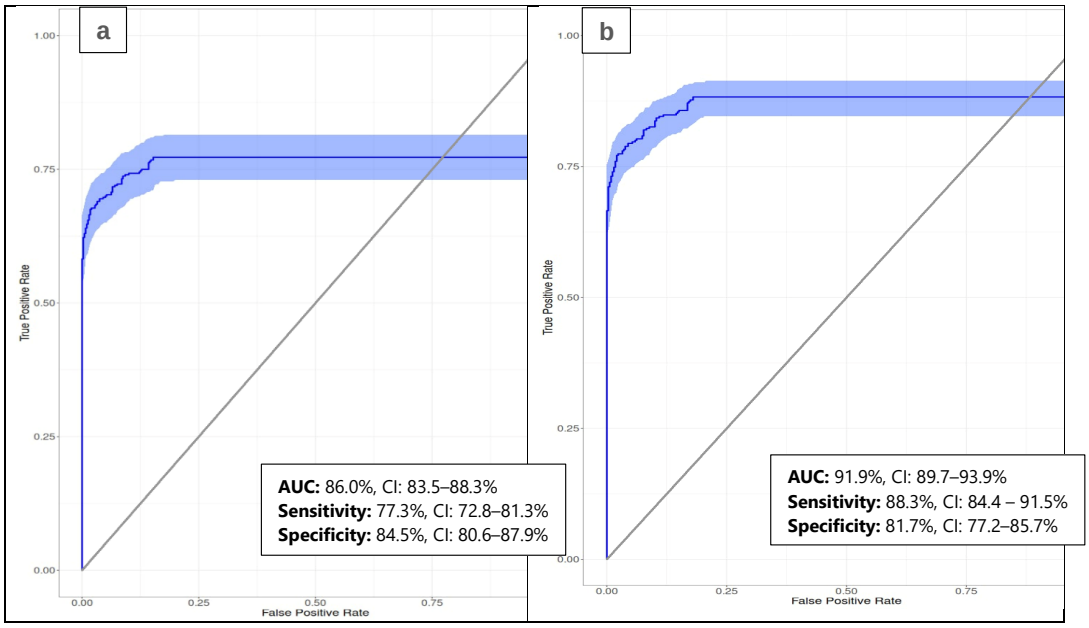
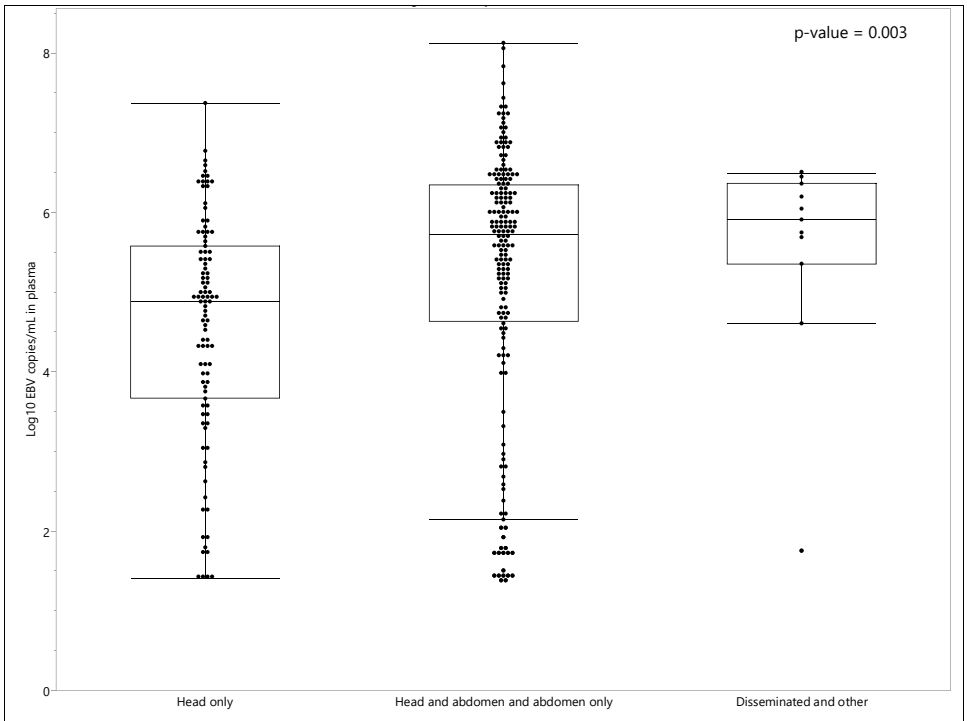
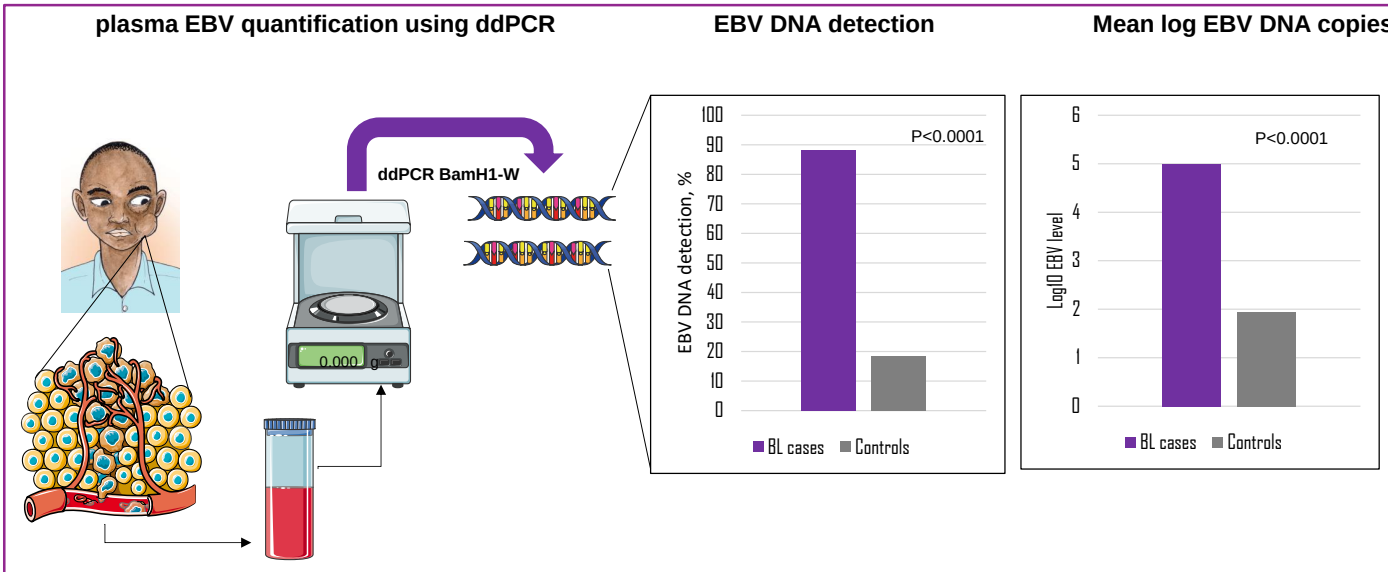

a
b
True Positive Rate
False Positive Rate
1.00
0.75
0.50
0.25
0.00
0.00
0.25
0.50
0.75
AUC: 86.0%, CI: 83.5–88.3%
Sensitivity: 77.3%, CI: 72.8–81.3%
Specificity: 84.5%, CI: 80.6–87.9%
AUC: 91.9%, CI: 89.7–93.9%
Sensitivity: 88.3%, CI: 84.4 – 91.5%
Specificity: 81.7%, CI: 77.2–85.7%

Figure 3



EBV detected	99/127 (78.0%)	177/235 (75.3%)	11/13 (84.6%)
Mean (SD)	4.56 ± 1.43	5.21 ± 1.67	5.51 ± 1.36
Range	1.41-7.37	1.38-8.13	1.76-6.49

Plasma EBV DNA: a potential tool for pediatric Burkitt lymphoma diagnosis in sub-Saharan Africa



Testing for EBV DNA could facilitate the recognition of BL earlier in sub-Saharan Africa, a critical step in improving BL cure rates in the region.

Volesky-Avellaneda *et al. doi:XXXXXXXXXXXXXXXXXXXXX.*