



OPEN Early inflammation as a footprint of increased mortality risk in infants living with HIV from three African countries

Elena Morrocchi^{1,2,4,5}, Giuseppe R. Pascucci^{1,2,4,5}, Nicola Cotugno^{1,3}, Chiara Pighi¹, Sara Dominguez-Rodriguez^{4,7}, Maria Raffaella Petrara⁵, Alfredo Tagarro^{4,6,7}, Louise Kuhn^{8,9}, Mark F. Cotton¹⁰, Kennedy Otwombe^{11,12}, Maria G. Lain¹³, Paula Vaz¹³, Shaun L. Barnabas¹⁰, Moira J. Spyer¹⁴, Elisa Lopez¹⁵, Sheila Fernández-Luis¹⁵, Tacilta Nhampossa¹⁶, Almoustapha I. Maiga¹⁷, Oumar Dolo¹⁷, Anita De Rossi^{5,18}, Pablo Rojo¹⁹, Carlo Giaquinto^{20,21}, Mathias Lichterfeld^{22,23}, Avy Violari¹¹, Theresa Smit²⁴, Osee Behuhuma²⁴, Nigel Klein^{24,25}, Lesley De Armas²⁶, Savita Pahwa²⁶, Paolo Rossi^{1,3}, Paolo Palma^{1,3}✉ on behalf of EPIICAL consortium*

In this work our aim was to identify early biomarkers in plasma samples associated with mortality in children with perinatal HIV treated early in life, to potentially inform early intervention targeting this vulnerable group. 20/215 children (9.3%) with perinatal HIV, enrolled within 3 months of age died prematurely within the first year of the study, despite early ART initiation. Using a propensity score, we selected 40 alive study participants having similar clinical and virological records compared to the deceased group. 13 HIV unexposed (HU) healthy children were additionally used as controls. Baseline plasma samples were analyzed using a targeted proteomic approach, and to assess pathogen-associated and damage-associated molecular patterns (PAMPs, DAMPs) levels. Data from deceased participants were compared to both control groups, with multivariate logistic regression models used to evaluate the association between mortality and plasma proteins. We developed a machine learning model to predict mortality risk, finding that IL-6 and CXCL11 not only were higher in deceased children than Matched-children with HIV ($p < 0.001$ and $p = 0.0034$) but also predictive of mortality (accuracy of 77%); levels of PAMPs were higher in deceased children ($p = 0.0016$). Thus, measuring early inflammatory biomarkers, particularly IL-6, could help mortality risk prediction and potentially guide targeted interventions.

Keywords Predictive model, Inflammatory biomarkers, IL-6, Pediatric population, HIV infection, PAMPs and DAMPs

¹Research Unit of Clinical Immunology and Vaccinology, Bambino Gesù Children's Hospital, IRCCS, Rome, Italy. ²Precision Vaccines Program, Boston Children's Hospital, Boston, MA, USA. ³Department of Systems Medicine, Molecular Medicine, and Applied Biotechnology, University of Rome Tor Vergata, Rome, Italy. ⁴Instituto de Investigación Sanitaria Hospital 12 Octubre, Madrid, Spain. ⁵Section of Oncology and Immunology, DiSCOG, University of Padova, Padova, Italy. ⁶Pediatrics Department, Hospital Universitario Infanta Sofía, Infanta Sofía University Hospital and Henares University Hospital Foundation for Biomedical Research and Innovation (FIIB HUIS HHEN), San Sebastián de los Reyes, Madrid, Spain. ⁷Pediatrics Research Group, Universidad Europea de Madrid, Madrid, Spain. ⁸Gertrude H. Sergievsky Center, Vagelos College of Physicians and Surgeons, Columbia University Irving Medical Center, New York, NY, USA. ⁹Department of Epidemiology, Mailman School of Public Health, Columbia University Irving Medical Center, New York, NY, USA. ¹⁰Family Centre for Research With Ubuntu, Department of Paediatrics and Child Health, Faculty of Medicine and Health Sciences, Stellenbosch University, Cape Town, South Africa. ¹¹Perinatal HIV Research Unit, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa. ¹²School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa. ¹³Fundação Ariel Glaser contra o SIDA Pediátrico, Maputo, Mozambique. ¹⁴University College London, Great Ormond Street Institute for Child Health, Infection, Immunity and Inflammation, London, UK. ¹⁵ISGlobal, Barcelona Institute for Global Health, Hospital Clínic - Universitat de Barcelona, Barcelona, Spain. ¹⁶Centro de Investigação em Saúde de Manhiça (CISM), Manhiça, Mozambique. ¹⁷University Clinical Research

Center (UCRC), University of Sciences, Techniques and Technologies of Bamako (USTTB), Bamako, Mali. ¹⁸Istituto Oncologico Veneto (IOV-IRCCS), Padua, Italy. ¹⁹Universidad Computense de Madrid, Instituto de Investigación Hospital 12 de Octubre, Madrid, Spain. ²⁰Department for Women's and Children's Health, University of Padova, Via Giustiniani, 3 - 35128 Padua, Italy. ²¹Penta - Child Health Research, Corso Stati Uniti, 4 - 35127 Padua, Italy. ²²Infectious Disease Division, Brigham and Women's Hospital, Boston, MA, USA. ²³Ragon Institute of MGH, MIT and Harvard, Cambridge, MA, USA. ²⁴Africa Health Research Institute (AHRI), Durban, South Africa. ²⁵Great Ormond Street Institute for Child Health (GOS ICH), University College London (UCL), London, UK. ²⁶Department of Microbiology and Immunology, Miami Center for AIDS Research, Miller School of Medicine, University of Miami, Miami, USA. ⁴⁵Elena Morrocchi and Giuseppe R. Pascucci: These authors contributed equally to this work. *A list of authors and their affiliations appears at the end of the paper. ✉email: paolo.palma@opbg.net

Child deaths have rapidly decreased worldwide, from over 17 million per year in the 1970's to around 5.3 million in 2019^{1,2}. However, these reductions are relatively modest in low- and middle-income (LMIC) countries, which account for up to 99% of all child deaths^{3,4}. Infectious diseases such as HIV are leading contributors to death among children under the age of 5⁵, in sub-Saharan Africa and Southern Asia¹. Of note, infectious causes of child mortality are often preventable through simple interventions if symptoms are recognized early and rapidly treated. Therefore, the identification of early markers informative of HIV-related mortality in children, currently absent, would be crucial.

Despite improved availability of antiretroviral agents, HIV-related mortality and vertical transmission still persist with children currently accounting for 4% of the total people living with HIV⁶. Through a targeted proteomic approach and quantification of plasma levels of circulating levels of pathogen-associated molecular patterns (PAMPs), estimated by measuring 16S ribosomal (r)DNA and damage-associated molecular patterns (DAMPs) by evaluating mitochondrial (mt) DNA molecules, we searched for early biomarkers of mortality in a cohort of infants with perinatal HIV.

Results

Mortality incidence in the first 12 months of life

As shown in the study design in Fig. 1A, 20 deceased children (deceased-HIV +), 9.3% of the cohort, 40 survivors (Matched-HIV +) matched through the propensity score and 13 HU were included in the study. Causes of deaths related to HIV progression in 8/20 cases (40%) are listed in Fig. 1B. Clinical, virological and immunological characteristics of the cohort are listed in Table 1. Only age at enrollment was significantly higher in the HU group compared to Deceased-HIV + and Matched-HIV +. Data collected from the mothers are listed in Table 2.

IL-6 as the first plasmatic molecule to discriminate mortality risk at the entry visit among HIV + grouEPIICALPs

The 92 plasmatic proteins analyzed could partially cluster the participants by study group (Fig. 2A), with the top 15 proteins contributing to the clusterization shown with lollipop graphs (Fig. 2A).

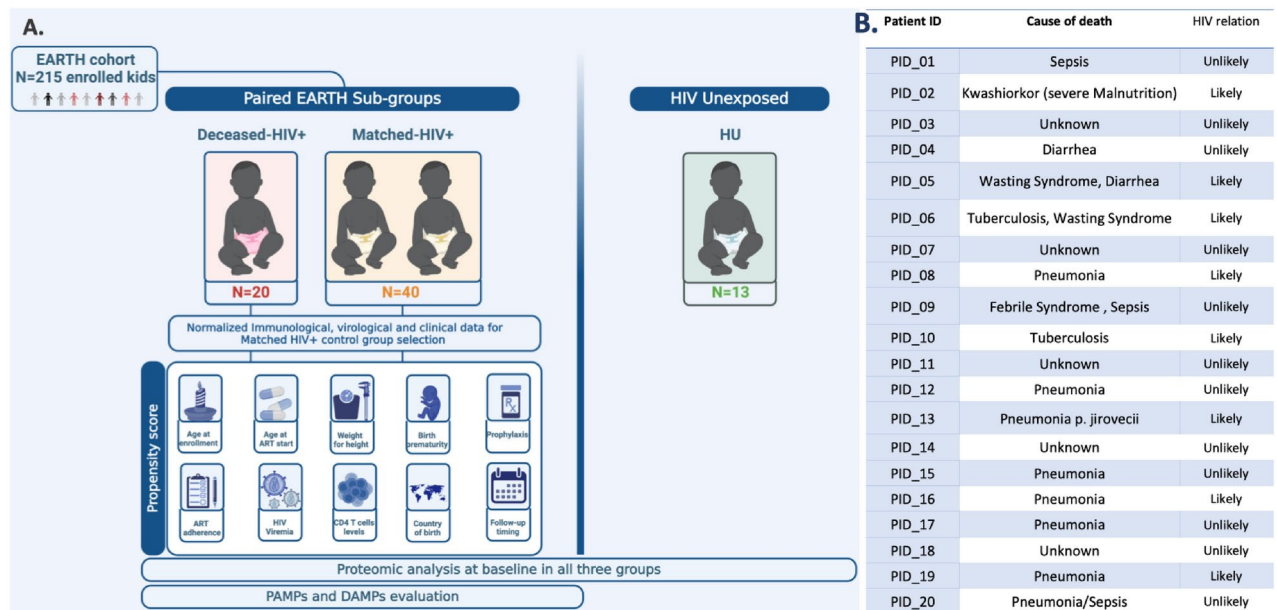


Fig. 1. **A** Study design depicting the numbers of study participants enrolled within the EARTH study and of the HU group, with a list of Deceased-HIV + data used to generate the propensity score used to select the Matched-HIV +. **B** Mortality reasons and their HIV progression relationship in the Deceased-HIV group. Created with BioRender.com.

	Deceased-HIV + (n = 20)	Matched-HIV + (n = 40)	HU (n = 13)	Significance
Age at sampling (months, min–max)	1.34 (0.26–3.67)	1.18 (0.1–4.23)	2.03 (2–12.66)	p < 0.001 (Matched-HIV + vs HU), p < 0.001 (deceased-HIV + vs HU)
Gender (F%)	40%	40%	30%	
Prematurity < 37 weeks (Yes%)	20%	22%	NA	
Mode of delivery (Vaginal%)	95%	82.5%	NA	
Breast milk (Yes%)	90%	92.5%	NA	
Gestational Age (weeks, min–max)	39 (32–42)	38 (26–41)	NA	
Birth Weight (gr, min–max)	2780 (1830–4000)	2800 (1900–3900)	NA	
Length at birth (cm, min–max)	48 (42–52)	49 (36–54)	NA	
Head circumference at birth (cm, min–max)	33 (29–39)	33 (29–39)	NA	
WHO classification_Clinical status (1:2:3:4)	16:3:0:1	38:2:0:0	NA	
Prophylaxis (Yes%)	85%	85%	NA	
Serious adverse events after birth (Yes%)	20%	10%	NA	
Site of Birth (AHRI:ARIEL:CISM:PHRU:TCH)	3:4:10:2:1	7:7:12:8:6	NA	
Weight at visit (gr, min–max)	3900 (2500–6000)	3900 (2015–6700)	NA	
Length at visit (cm, min–max)	52.5 (47–59)	51 (41.5–61)	NA	
Head circumference at visit (cm, min–max)	38 (33–42)	38 (32–42.7)	NA	
Prophylaxis start from birth (h, min–max)	4 (0–336)	6 (0–192)	NA	
VL-RNA (cp/ml, min–max)	1,205,095 (30,401–1e + 07)	298,604.5 (20–1e + 07)	NA	
CD4%	33.5 (1–57)	34 (13–61.04)	NA	
CD4 count (cells/mm ³ , min–max)	1436 (14–3163)	1728 (360–4418)	NA	
Time On ART (months, min–max)	0 (0–0.46)	0 (0–0.46)	NA	
Age At ART start (months, min–max)	1.34 (0.26–3.67)	1.11 (0.1–4.23)	NA	
ART_Regimen: LAMIDIVUDINE, LOPINAVIR, ABACAVIR, COTRIMOXAZOL; LAMIDIVUDINE, LOPINAVIR, ABACAVIR; LAMIDIVUDINE, ZIDOVUDINE, LOPINAVIR; LAMIDIVUDINE, NEVIRAPINE, ABACAVIR; LAMIDIVUDINE, NEVIRAPINE, ZIDOVUDINE	0:10:8:0:2	1:28:7:1:3	Not Applicable	

Table 1. Medical records of all participants presenting data as median and range (min–max). AHRI: Africa Health Research Institute; ARIEL: Ariel Foundation in Mozambique; CISM: Centro de Investigação em Saúde de Manhica; PHRU: Perinatal HIV Research Uni; TCH: Tygerberg Children's Hospital.

Variable	Deceased-HIV +	Matched-HIV +	Significance
WHO classification_Mother Clinical status (1:2:3:4)	15:1:3:1	29:3:6:1	
Paid Job (Yes%)	10%	12%	
Marital status (Yes%)	45%	47.5%	
Educational grade (No school:Primary:Secondary:University)	0:11:9:0	4:13:22:1	
HIV diagnosis (Before pregnancy:During pregnancy:During admission for delivery:After discharge)	9:9:0:2	22:12:2:4	
Tuberculosis infection (No:Disease:Latent)	15:0:5	32:2:5	p = 0.0025
Tuberculosis prophylaxis (Yes%)	35%	27.5%	
ART treatment (No:Before pregnancy: during pregnancy:at delivery:After delivery)	2:6:11:1	1:17:16:6	
ART adherence (No:Low:High:Good)	4:2:7:7	2:9:11:18	
Weight, kg (min–max)	58.2 (47–96.7)	62.3 (40–116)	
Height, cm (min–max)	155.75 (148–168)	159.75 (146–180)	
BMI (min–max)	24.41 (19.05–34.8)	24.27 (15.06–36.63)	

Table 2. Medical records collected from mothers of both HIV groups. Data are listed as median and relative ranges (min–max).

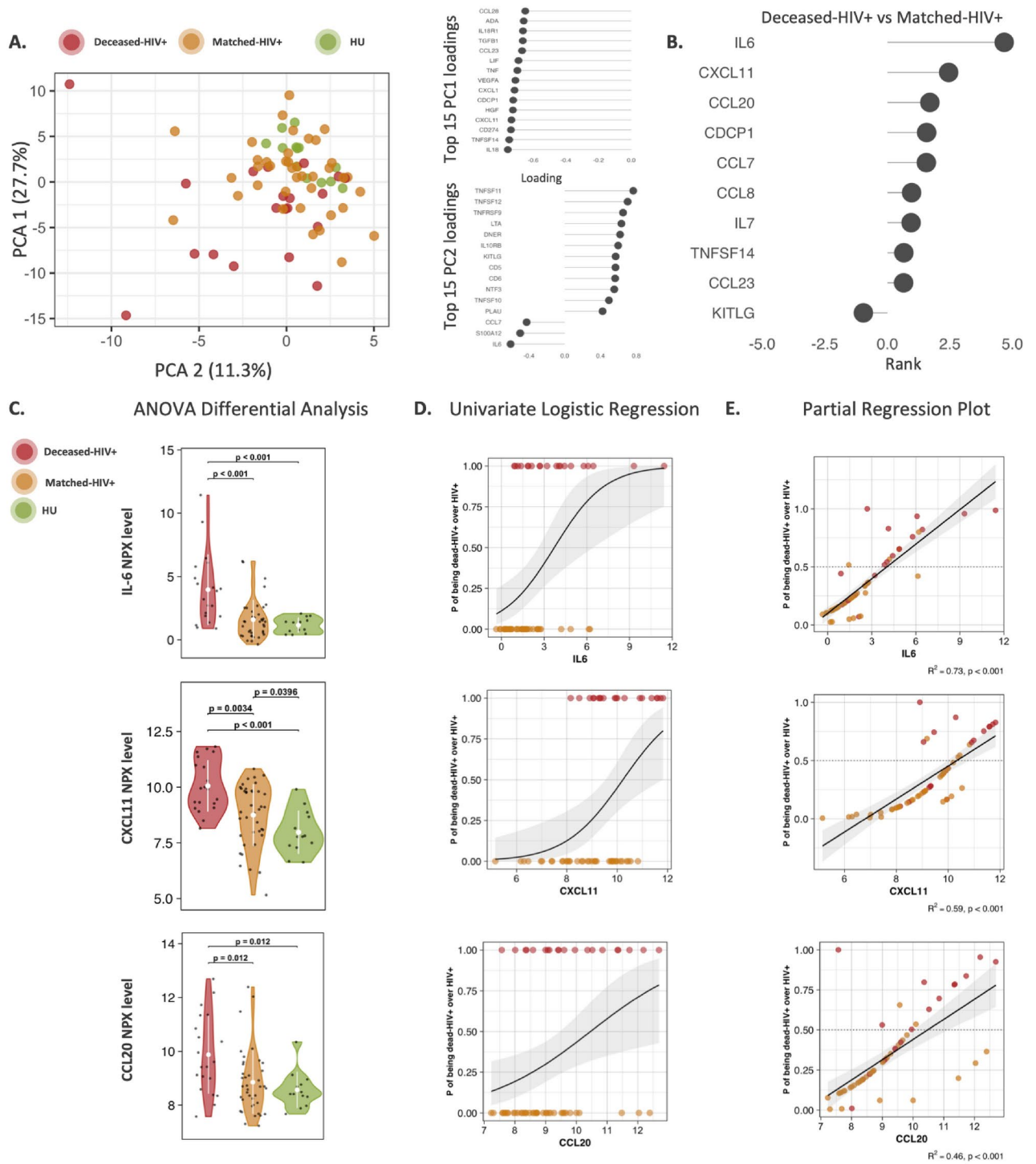


Fig. 2. **A** Principal Component Analysis (PCA) was performed using the prcomp R function to evaluate the distribution of participants according to the data variance of inflammatory plasma markers (Olink T96) data variance. The top 15 proteins contributing to clusters both in PC1 and PC2 components are listed on the side of PCA. **B** Proteins differentially produced between Deceased-HIV + and Matched-HIV + (ANOVA) and ordered by the rank values calculated as $-\log_{10}(\text{Adjusted } p\text{-value}) \times \log_2(\text{Fold Change})$. Positive values indicate up-regulated proteins. **C** Violin plots showing the NPX distribution and significance of the 3 proteins with the highest rank values. Statistical analysis was performed using ANOVA. **D** Univariate logistic regression analysis of the 3 proteins with the highest rank values. The curve shows how the probability of classification between the two groups (Deceased-HIV + in red and Matched-HIV + in orange) varies as protein levels change. **E** Partial regression plot showing the correlation between the multivariate regression analysis classification probability and the protein levels. Linear regression analysis was used to quantify this relationship.

ANOVA among all three groups identified 23 differentially produced proteins (adjusted p -value < 0.05), of which 10 differed between Deceased-HIV + and Matched-HIV + groups (Supplementary Fig. 1), ranked according to p -values and fold changes (Fig. 2B). IL-6, CXCL11 and CCL20 occupied the first three positions indicating their highest impact. IL-6 ranked highest and was significantly higher in the Deceased-HIV + group than Matched-HIV + and HU ($p < 0.001$) (Fig. 2C). CXCL11 (C-X-C Motif Chemokine Ligand 11) a chemokine involved in the chemotactic response especially for activated T lymphocytes, was higher in Deceased-HIV + group than Matched-HIV + and HU ($p = 0.0034$ and $p < 0.001$ respectively) (Fig. 2C). Also, CCL20, a chemotaxin that attracts lymphocytes to the sites of inflammation, was significantly higher in Deceased-HIV + than both control groups ($p = 0.012$) (Fig. 2C).

To better evaluate the association between the mortality outcome and the plasma proteins, we built a multivariate logistic regression model for each protein including potentially confounding variables as covariates. In fact, after testing multiple clinical variables (see materials and methods section) we selected the type of delivery at birth, and HIV WHO status classification. We found 27 proteins significantly associated with mortality, 9 of which were included in the 10 proteins identified through ANOVA analyses between the Deceased-HIV + and Matched-HIV + groups (Supplementary Fig. 2). To visualize these results in two dimensions, IL6, CXCL11 and CCL20 were plotted by univariate logistic regression (Fig. 2D) and partial regression (Fig. 2E) plots. A strong linear association ($p < 0.001$) between the probability informed by the multivariate model and the proteins NPX values is shown in Fig. 2E.

Deceased-HIV + shows enrichment in pathways involved in pathogens fight

Gene ontology (GO) analyses were used to outline the pathways enriched in the proteins identified by ANOVA and logistic regression analyses (Fig. 3). We observed that pathways involved in activating human immune responses such as cytokine-cytokine receptor interactions and signaling pathways mediated by chemokine, P13K-Akt, IL-17, TLR, and TNF were significantly enriched between Deceased-HIV vs. HU. Such differences, albeit at a lower level, also emerged between Deceased-HIV + and Matched-HIV +. Conversely, MAPK, Ras, and Ras-associated protein-1 (Rap1) pathways were altered only between the HIV groups vs HU.

The analysis of the biological roles of the 27 proteins associated with mortality risk through the LR analysis (Fig. 3B), confirmed similar enriched pathways as in the GO results except for the MAPK pathway which instead showed a weaker association with mortality.

Furthermore, additional pathways involved in responding to pathogens such as malaria, amoebiasis, legionellosis, pertussis, salmonella and African trypanosomiasis were observed in the Deceased-HIV + group at enrollment compared to matched group (Fig. 3B). The network plot of the proteins found to be associated with mortality (by both ANOVA and LR) indicated that are all effector proteins and up-regulated in Deceased-HIV + compared to Matched-HIV (Fig. 3C).

PAMPs and DAMPs levels at enrollment

Plasma aliquots at baseline were also investigated for PAMPs, measured by quantification of 16S rRNA and DAMPs measured by quantifying mtDNA. While DAMPs showed similar levels in all study participants, higher levels of PAMPs ($p = 0.0016$) were detected for Deceased-HIV + (Fig. 3D). A positive association between DAMPs and proteins that inform mortality (by both ANOVA and LR) was found in Deceased-HIV + (Fig. 3E). Indeed, IL6, CXCL11, CCL20 and CCL8 were positively correlated with DAMPs in Deceased-HIV +. Conversely, Matched-HIV + were positively correlated only to the extent of viremia and negatively to CD4 percentage and PAMPs levels (Data not shown).

IL-6 as a molecule to be tested at entry to predict mortality risk

Predictive models informed by the targeted plasma proteomic levels and participants' clinical data collected through the REDCap application were tested. The best model mispredicted 5 subjects (3 Deceased-HIV + children and 2 Matched-HIV + children) giving an accuracy of 77%, sensitivity of 57% specificity of 87% (Fig. 4). This model shown in Fig. 4 was built using CXCL11 and IL-6 as variables. Of note, 2 out of 3 incorrectly identified Deceased-HIV + children died of pneumonia. One Matched-HIV + survivor who was mispredicted in this model later died of tuberculosis at 22 months of age.

Discussion

HIV-associated mortality in infants remains relatively high in sub-Saharan Africa despite the significant advances in ART. During the first year of the project, 20 children (9.3%)⁷ died prematurely. Early mortality predictors in this population would facilitate optimized patient's follow-up schedule aiming to reduce the risk of death with closer visits. Using plasma to evaluate proteomic inflammatory pathways, together with PAMPs and DAMPs levels, we hypothesized that a differential concentration of inflammatory molecules at systemic levels at study enrollment could identify those children with higher risk of mortality within the first year of the study. ANOVA analysis highlighted among all differentially expressed molecules, IL-6, CXCL11 and CCL20 as the most impactful, all higher in deceased children than survivors and unexposed controls. IL-6 is a classic example of a cytokine with multifaceted and overlapping functions⁸, and as in our study, was linked to mortality risk prediction in cohorts of young adults and adults^{9,10}. For example in the adult Multicenter AIDS Cohort Study (MACS) in the US, higher IL-6 levels showed more than a threefold hazard for long-term mortality risk, suggesting how even despite ART-induced HIV suppression, underlying inflammation significantly contributed to an increased mortality risk¹⁰. Additionally, in a recent trial, the anti-IL-6 monoclonal antibody Tocilizumab, used with the goal of reducing inflammation and thus serious clinical consequences, was safe and able to reduce inflammation in people living with HIV¹¹.

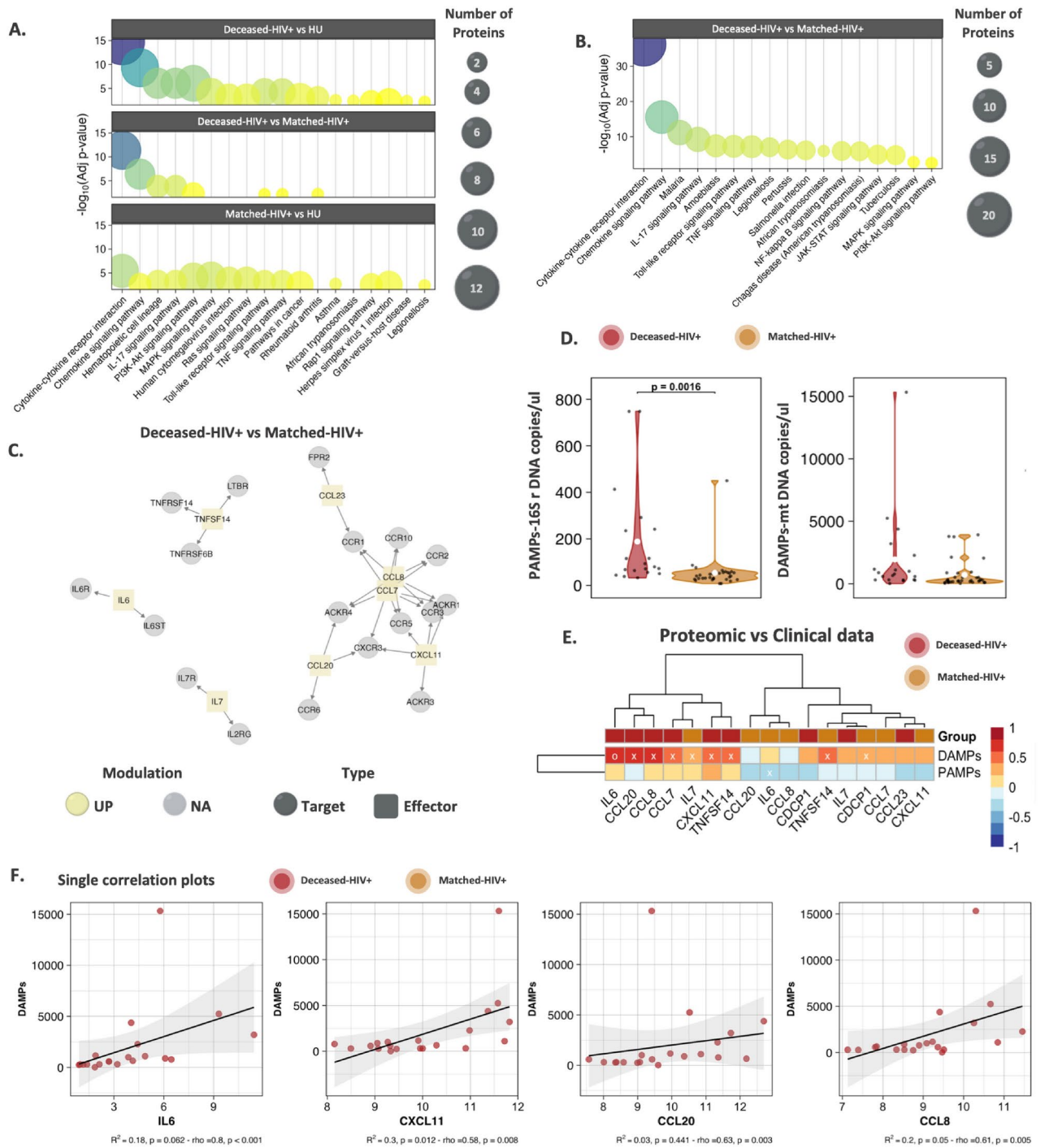


Fig. 3. Pathway enrichment analysis in the KEGG databases³¹ using the R package “enrich” v3.2 to investigate the biological role of the **A** proteins differentially produced between the 3 groups (ANOVA analysis) and **B** proteins associated with mortality through multivariate logistic regression analysis. Only pathways with an adjusted p-value < 0.01 were shown. The size of each bubble corresponds to the number of differentially produced proteins that enrich that pathway. The color gradient is proportional to the \log_{10} adjusted p-value. **C** Analysis of protein–protein interaction networks showing up-regulated (yellow) proteins and their function [target (circle) or effector (square)] in Deceased-HIV + versus Matched-HIV +. The relationships between proteins are sourced from the Omnipath database using the OmnipathR v3.2.8 R package³². **D** Violin plots illustrating PAMPs and DAMPs distributions and differences. **E** Spearman correlation heatmap between PAMPs and DAMPs values and the levels of the 9 proteins in common between the ANOVA and the logistic regression analyses. The red color gradient indicates positive correlations, while the “x” highlights significant results ($p < 0.05$). **F** Scatter plots to focus on the strongest relationships between DAMPs and proteins observed in the heatmap in panel **B**. These associations were measured by linear regression (R^2) and Spearman correlation (ρ).

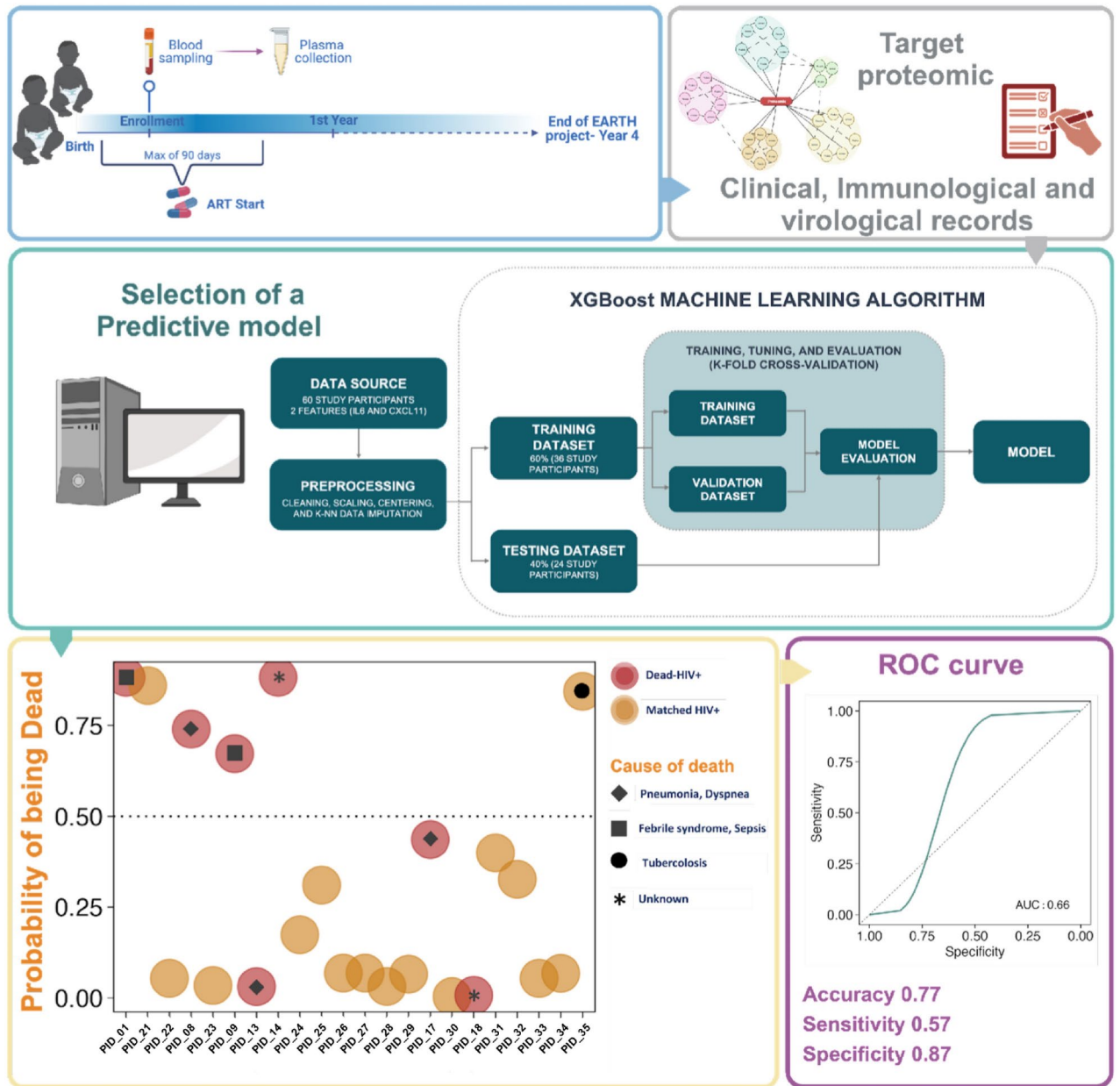


Fig. 4. Cartoon that illustrates the stepwise workflow in constructing the mortality predictive model. Created with BioRender.com.

Similarly, cytokine CXCL11 with chemoattractant properties has an important role in inducing immune cell activation¹². For example, the combination of CXCL11 levels with CXCL9 and CXCL10 in primary HIV infection was correlated to HIV disease progression¹³, or polymorphisms in the CXCL11 gene were associated to sustained virologic responses in HCV/HIV co-infected patients¹⁴. Again, blood-increased levels of chemokine (C-C motif) ligand 20, also known as Macrophage Inflammatory Protein-3 Alpha (MIP-3α) was recently negatively correlated to CD4+ T cells depletion in people living with HIV¹⁵.

Gene ontology analysis of proteins associated with mortality risk and of proteins differentially produced between the deceased-HIV + and survivors revealed the enrichment in pathways primarily revolved around stimulating human immune responses (encompassing interactions between cytokines and cytokine-receptors, and the modification of several signaling pathways such as chemokine, IL-17, TNF, and P13K-AKT). Moreover, gene ontology also revealed enrichment in pathways involved in human immune defense against bacteria. Since the number of co-infections was similar between the groups (measured as serious events rate), the higher PAMP levels found in deceased children, could support either the presence of microbial translocation at the time of sampling, but also an impairment of the immune system to counteract the inflammation. In particular, HIV-induced damage to the gastrointestinal tract results in the translocation of microbial products, known as PAMPs, into the systemic circulation^{16,17} along with the release of endogenous damage-associated molecular

patterns (DAMPs) due to cellular injury^{18–20}. The presence of both PAMPs and DAMPs contributes to chronic immune system activation, thereby triggering a pro-inflammatory cascade characterized by elevated expression of inflammatory markers, as observed in the deceased children who vertically acquired HIV. These results suggest that the measurement of such inflammatory molecules might have significant repercussions in the early detection of mortality cases.

In our final predictive model of mortality risk in this cohort in which IL-6 had a critical impact, it was reached an accuracy of 77%, a sensitivity of 57% and specificity of 87%. Five patients were misclassified, especially those who died of pneumonia. While the model had good predictive capability, other markers would undoubtedly be important to be included to foster model predictive performance¹⁴. Interestingly, a participant, misclassified by the model as a death, died at 22 months of age of tuberculosis. Given the strong predictive ability of IL-6, measuring its levels in plasma as a routine marker may provide useful information on the general health status in vulnerable infants, favoring closer follow-up visits and possibly avoiding severe irreversible outcomes²¹. In summary, although we are conscious that the limited number of subjects in our study may have impacted the robustness of our results and that larger studies could help us to corroborate our findings, we still identified that early inflammation might predict mortality in the pediatric population. Therefore, implementing future interventions that target the reduction of early inflammation could be crucial in decreasing early mortality rates in African infants, potentially improving long-term health outcomes and survival in this vulnerable population.

Materials and methods

Ethics statement

The independent Ethics and Institutional Review Boards of all the participating institutions and Countries approved the study. Study participants' management and biological sampling were performed in accordance with guidelines of the Declaration of Helsinki guidelines.

Informed consent

Informed consent to participate to the study was obtained from parents or legal guardians of all enrolled participants. The participants' numbers inserted in the manuscript, table, and figures are for readability purposes only and not related to the actual study patient IDs, so that the de-identification of study participants is maintained.

Study participants

This sub-study is from a prospective, observational cohort, from the multicenter transnational Early Anti-Retroviral Treatment in HIV- infected Infants (EARTH) study (NCT05784584) through the Early treated Perinatally HIV Infected individuals: Improving Children's Actual Life (EPIICAL) Project (epiical.com). Study participants who tested positive for both HIV-DNA and RNA within 3 months from birth, either due to perinatal or breastfeeding infection and also started ART therapy within the following 90 days were enrolled. Recruitment was performed in 4 urban and 2 rural areas: Johannesburg, Cape Town, uMkhanyakude District in Kwa-Zulu Natal (South Africa), Maputo and Manhica (Mozambique), and Bamako (Mali)²². Of 215 children enrolled, 20 died in the first year of the project (Fig. 1A) and were included in our study sample (mean age at sample collection = 1.34 months). A comparison cohort of 40 survivors from the same study (mean age at sample collection = 1.18 months) was selected through propensity score matching, discussed later in methods. With the exception of one child that could only be matched to one surviving child, due to sample unavailability, the others were matched to two surviving children.

The date of blood sampling at baseline, corresponding to age at sampling in Table 1, reports an overlap of 58.3% (35/60) with ART start.

An additional group of healthy HIV unexposed children (HU, mean age = 2.03 months) enrolled from Maputo (N = 13) within the TARA project (Towards AIDS Remission Approaches, 5R01AI127347), was chosen for comparison. All the available clinical information for the three groups is listed in Table 1. Serious adverse events at each study visit were systematically collected. These included the onset of any new infectious disease, including malaria events, any hospitalization, any new antimicrobial treatment initiation or congenital disease diagnosis.

Plasma isolation

EDTA whole blood samples were processed according to EARTH study Standard Operating Procedures (SOP) at local laboratories. EDTA blood tubes were first centrifuged at 400 g for 10 min (18–25 °C), plasma was then recovered and transferred to a 15 ml conical centrifuge tube for second centrifugation at 800 g for 10 min (18–25 °C) to remove any cellular debris. Thereafter plasma was aliquoted (0.5–1 ml/cryovial) and stored at – 80 °C until shipment on dry ice to the Cryolab facility in Rome for storage at – 80 °C until further analysis.

Propensity score calculation and HIV group selection

A propensity score was used to select EARTH survivors with similar baseline clinical, virological, and immunological characteristics to the deceased group to limit potential confounding factors for downstream analysis (Fig. 1). The variables used to match the participants were selected according to factors associated with mortality in previous publications^{22–24}. We performed a 2:1 nearest propensity score matching without replacement with a propensity estimated using a random forest on the covariates. After matching, all standardized mean differences for the covariates were below 0.1 and all standardized mean differences for squares and two-way interactions between covariates were below 0.15, indicating adequate balance. The propensity score matching was performed using MatchIt R library²⁵.

Targeted proteomics

Plasma samples collected at baseline were randomized on a 96-well plate and the Multiplex Proximity Extension Assay (PEA technology) was performed using a Target 96 Inflammation Olink kit. 1ul of plasma samples was used to investigate for the presence of inflammatory proteins. Data were reported in an arbitrary unit as normalized protein expression (NPX), provided by Olink, that enables individual protein analysis across a sample set analyzed in log₂ scale, wherein a higher NPX correlates with higher protein expression. Data were analyzed as previously described^{26,27}.

DAMPs and PAMPS measurement

Aliquots of 200ul of plasma were used to extract DNA with QIAamp DNA Mini Kit (QIAGEN, Hilden, Germany) as described by manufacturer's instructions, and eluted in 50 µl of AE Buffer. Real-time PCR assays were performed to quantify circulating levels of 16S ribosomal (r)DNA as PAMPS, and mitochondrial (mt)DNA as DAMPs. Specifically, forward primer: 5'-AGTTTGATCCTGGCTCAG-3', reverse primer: 5'-GWATTACCGCGGCKGCTG-3' and probe 5'-FAM-GCTGCCTCCCGTAGGAGT-BHQ-3' were used to amplify 16S rDNA; and forward: 5'-AGGACAAGAGAAATAAGGCC-3', reverse: 5'-TAAGAAGAGGAATTGAACCTCTGACTGTAA-3' and probe 5'-FAM-TTCACAAAGCGCCTTCCCCCGTAAATGA-BHQ-3' for mtDNA^{28,29}. Results were expressed as 16S rDNA copies/µl plasma and mtDNA copies/µl plasma, respectively.

Analysis

For each continuous variable, normality was assessed using the Shapiro test. Variance homogeneity was evaluated using Bartlett's test for normally distributed variables, otherwise using Fligner's test. Statistical comparisons of proteomics and continuous clinical variables between more than two groups (Deceased-HIV +, Matched-HIV +, and HU) were performed using One-way ANOVA followed by Tukey–Kramer's post hoc test if the distributions were normally distributed and homogeneous. If variances were unequal, Welch ANOVA followed by Games-Howell's post hoc test were used, provided that distributions were normally distributed. For not normally distributed distributions, Kruskal–Wallis's test was used followed by Dunn's post hoc test. Post hoc tests were applied only to variables with a p-value adjusted for the false discovery rate (FDR) < 0.05. Differences between the two groups Deceased-HIV + and Matched-HIV + were evaluated using the t-test whether both the distributions were normally distributed and homogeneous. If variances were unequal but the distributions were normally distributed, the Welch t-test was used. The Mann–Whitney test was used for distributions not normally distributed. Correction for multiple comparisons was addressed by adjusting the p-values with the FDR method. Results were considered statistically significant for those with a p-value adj < 0.05. Categorical clinical data were analyzed using Fisher's exact test. To highlight the weight of differentially produced proteins, we plotted ranking values calculated as: $-\log_{10}(\text{Adjusted p-value}) * \log_2(\text{Fold Change})$. The rank is proportional to both the significance and the magnitude of the difference. Spearman's correlation and linear regression analysis were used to examine the association between variables. Statistical analyses were performed with R (version 4.1.1).

Multivariate logistic regression models

Multivariate logistic regression models were used to evaluate the association between mortality and plasma proteins. To exclude the effect of potentially confounding variables, we selected the following clinical variables as potential covariates: gestational birth, weight, age, viral load (VL) RNA, % of CD4, white blood cell count (WBC), type of birth delivery, birth site, prophylaxis, gender, WHO classification, mother's body mass index (BMI) and mother's WHO HIV status. Some of these variables were already used to produce the propensity-matching score. We decided to evaluate them also as covariates to exclude potential residual imbalance between the Deceased-HIV + and the Matched-HIV + groups. Multicollinearity was evaluated using the findLinearCombs function of the caret R package³⁰ and data imputation of variables with not available (NA) values $\leq 15\%$ applying the k-nearest neighbor (k-NN) method provided by the caret R package. Therefore, a stepAIC backward selection was performed in the logistic regression model built with these clinical variables to select the best-fitting model. Finally, we build a logistic regression model for each protein using the type of birth delivery and the WHO classification variables as covariates. We considered protein-related adjusted p-values < 0.05 and an odds ratio (OR) > 2 to be associated with mortality. We visualized the association of these proteins with the outcome by univariate logistic regression and partial regression plots.

Machine learning

We built a machine-learning model to predict the mortality risk including just IL-6 and CXCL11 to reduce the overfitting risk. Data were scaled and centered, and missing values (< 15%) were imputed using the k-nearest neighbor (k-NN) method provided by the R caret package³⁰. The dataset of 60 patients was divided into training (60%) and testing (40%) data sets. This ratio between the data sets was chosen to allow enough observations in the testing dataset for good accuracy evaluations. Four ML algorithms were tested, including Logistic Regression, Random Forest, eXtreme Gradient Boosting (XGBoost) and k-nearest neighbor (k-NN). In the training step of each algorithm, the optimal hyperparameters were selected by the random search method and the repeated 5-time tenfold cross-validation was used as a resampling method. The accuracy, sensitivity, and specificity computed in the testing data set were used to evaluate the models by the R caret package³⁰. According to these parameters, the XGBoost model shows the best performance with an accuracy of 77%, a specificity of 87%, and a sensitivity of 57%.

Data availability

Data will be made available upon reasonable request to the corresponding author Dr. Paolo Palma, after the paper is accepted.

Received: 21 May 2024; Accepted: 23 September 2024

Published online: 28 October 2024

References

- Bassat, Q. *et al.* Causes of death among infants and children in the child health and mortality prevention surveillance (CHAMPS) network. *JAMA Netw. Open* **6**, e2322494 (2023).
- Perin, J. *et al.* Global, regional, and national causes of under-5 mortality in 2000–19: an updated systematic analysis with implications for the Sustainable Development Goals. *Lancet Child Adolesc. Health* **6**, 106–115 (2022).
- Wang, Y., Zhao, Y., Bollas, A., Wang, Y. & Au, K. F. Nanopore sequencing technology, bioinformatics and applications. *Nat. Biotechnol.* **39**, 1348–1365 (2021).
- Burstein, R. *et al.* Mapping 123 million neonatal, infant and child deaths between 2000 and 2017. *Nature* **574**, 353–358 (2019).
- Agathis, N. T. *et al.* Mortality Among Children Aged <5 Years Living with HIV Who Are Receiving Antiretroviral Treatment - U.S. President's Emergency Plan for AIDS Relief, 28 Supported Countries and Regions, October 2020–September 2022. *MMWR Morb. Mortal. Wkly. Rep.* **72**, 1293–1299 (2023).
- Teasdale, C. A., Marais, B. J. & Abrams, E. J. HIV: prevention of mother-to-child transmission. *BMJ Clin Evid* **2011**, 0909 (2011).
- Domínguez-Rodríguez, S. *et al.* Machine learning outperformed logistic regression classification even with limit sample size: A model to predict pediatric HIV mortality and clinical progression to AIDS. *PLoS One* **17**, e0276116 (2022).
- Tanaka, T. & Kishimoto, T. The biology and medical implications of interleukin-6. *Cancer Immunol. Res.* **2**, 288–294 (2014).
- Wada, N. I. *et al.* Inflammatory biomarkers and mortality risk among HIV-suppressed men: A multisite prospective cohort study. *Clin. Infect. Dis.* **63**, 984–990 (2016).
- Borges, Á. H. *et al.* Interleukin 6 is a stronger predictor of clinical events than high-sensitivity C-reactive protein or D-dimer during HIV infection. *J. Infect. Dis.* **214**, 408–416 (2016).
- Funderburg, N. T. *et al.* Interleukin 6 blockade with tocilizumab diminishes indices of inflammation that are linked to mortality in treated human immunodeficiency virus infection. *Clin. Infect. Dis.* **77**, 272–279 (2023).
- Karin, N. CXCR3 ligands in cancer and autoimmunity, chemoattraction of effector T cells, and beyond. *Front. Immunol.* **11**, 976 (2020).
- Yin, X. *et al.* The combination of CXCL9, CXCL10 and CXCL11 levels during primary HIV infection predicts HIV disease progression. *J. Transl. Med.* **17**, 417 (2019).
- Pineda-Tenor, D. *et al.* CXCL9, CXCL10 and CXCL11 polymorphisms are associated with sustained virologic response in HIV/HCV-coinfected patients. *J. Clin. Virol.* **61**, 423–429 (2014).
- Aziz, N., Detels, R., Chang, L. C. & Butch, A. W. Macrophage inflammatory protein-3 alpha (MIP-3α)/CCL20 in HIV-1-infected individuals. *J. AIDS Clin. Res.* **7**, 587 (2016).
- Dalchini, A. *et al.* Biological aging and immune senescence in children with perinatally acquired HIV. *J. Immunol. Res.* **2020**, 8041616 (2020).
- Brenchley, J. M. *et al.* Microbial translocation is a cause of systemic immune activation in chronic HIV infection. *Nat. Med.* **12**, 1365–1371 (2006).
- Zhang, Q. *et al.* Circulating mitochondrial DAMPs cause inflammatory responses to injury. *Nature* **464**, 104–107 (2010).
- Vénéreau, E., Ceriotti, C. & Bianchi, M. E. DAMPs from cell death to new life. *Front. Immunol.* <https://doi.org/10.3389/fimmu.2015.00422> (2015).
- Roh, J. S. & Sohn, D. H. Damage-associated molecular patterns in inflammatory diseases. *Immune Netw.* **18**, e27 (2018).
- Hakim, J. *et al.* Enhanced prophylaxis plus antiretroviral therapy for advanced HIV infection in Africa. *N Engl. J. Med.* **377**, 233–245 (2017).
- Tagarro, A. *et al.* High mortality following early initiation of antiretroviral therapy in infants living with HIV from three African countries. *eClinicalMedicine* **73**, 102648 (2024).
- Venkatesh, K. K. *et al.* Morbidity and mortality among infants born to HIV-infected women in South Africa: implications for child health in resource-limited settings. *J. Trop. Pediatr.* **57**, 109–119 (2011).
- Oumer, A., Kubsu, M. E. & Mekonnen, B. A. Malnutrition as predictor of survival from anti-retroviral treatment among children living with HIV/AIDS in Southwest Ethiopia: survival analysis. *BMC Pediatr.* **19**, 474 (2019).
- Ho, D. E., Imai, K., King, G. & Stuart, E. A. MatchIt: nonparametric preprocessing for parametric causal inference. *J. Stat. Soft.* <https://doi.org/10.18637/jss.v042.i08> (2011).
- Morrocchi, E. *et al.* Perinatally human immunodeficiency virus-infected adolescents and young adults demonstrate distinct BNT162b2 messenger RNA coronavirus disease 2019 vaccine immunogenicity. *Clin. Infect. Dis.* **75**, S51–S60 (2022).
- Amodio, D. *et al.* Similarities and differences between myocarditis following COVID-19 mRNA vaccine and multiple inflammatory syndrome with cardiac involvement in children. *Clin. Immunol.* **255**, 109751 (2023).
- Petrara, M. R. *et al.* Immune activation, immune senescence and levels of Epstein Barr Virus in kidney transplant patients: Impact of mTOR inhibitors. *Cancer Lett.* **469**, 323–331 (2020).
- Petrara, M. R. *et al.* Impact of monotherapy on HIV-1 reservoir, immune activation, and co-infection with Epstein-Barr virus. *PLoS ONE* **12**, e0185128 (2017).
- Kuhn, M. Building predictive models in R using the **caret** package. *J. Stat. Soft.* <https://doi.org/10.18637/jss.v028.i05> (2008).
- Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).
- Türei, D. *et al.* Integrated intra- and intercellular signaling knowledge for multicellular omics analysis. *Mol. Syst. Biol.* **17**, e9923 (2021).

Acknowledgements

We acknowledge all study participants and their parents/guardians for giving their consent to participate to the project. We thank Dr. Chiara Medri and Dr. Ilaria Peponi respectively for their experimental and administrative support. Finally, we acknowledge our collaborators from EChiLiBRiST (Enhancing children's lives with biomarkers for risk stratification and Triage) project, for their support.

Author contributions

All authors read and approved the final version of the manuscript. E.M., G.R.P., N.C., C.P., P.P., S.D.R., A.T., and M.R.P. verified the data underlying the manuscript. Conceptualization; E.M., G.R.P., N.C., P.P., P.R. methodology; E.M., G.R.P., N.C., P.P.; software, G.R.P., S.D.R.; formal analysis: G.R.P., S.D.R.; investigation, all authors; resources, P.P. Writing original draft preparation, E.M., G.R.P.; review and editing: E.M., G.R.P., N.C., C.P., S.D.R., M.R.P., A.T., L.K., M.F.C., K.O., M.G.L., P.V., S.B., M.J.S., E.L., S.F.L., T.N., A.I.M., O.D., A.D.R., P.R., C.G., M.L., A.V., T.S., O.B., N.K., L.D.A., S.P., P.R., P.P.; supervision, N.C., A.D.R., A.T., P.P. P.R. Members of the EPIICAL consortium (website: <https://www.epiical.org>, e-mail: communication@epiical.org) discussed the data over the annual meetings, and read and agreed on the last version of the manuscript.

Funding

The present work is part of EPIICAL, a research consortium funded by ViiV Healthcare and led by Penta Foundation.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-74066-4>.

Correspondence and requests for materials should be addressed to P.P.

Reprints and permissions information is available at 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-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024

EPIICAL consortium

Paolo Rossi^{1,3,27}, Silvia Faggion²⁷, Daniel Gomez Pena²⁷, Inger Lindfors Rossi²⁷, Federica D'Ambrosio²⁷, Andrea Oletto²⁷, Francesca Mazzetto²⁷, Musakanya Ching'andu²⁷, Alessandra Nardone²⁷, William James²⁷, Nicola Cotugno^{1,3}, Paola Zangari²⁸, Carla Paganin²⁸, Paolo Palma^{1,3,28}, Alfredo Tagarro^{4,6,7,29}, Anita Rossi^{5,18}, Sara Dominguez-Rodriguez^{4,7,29}, Louise Kuhn^{8,30}, Andrew Yates^{8,9,30}, Mark F. Cotton^{10,31}, Shaun Barnabas^{10,31}, Avy Violari¹¹, Kennedy Otwombe^{11,12}, Paula Vaz¹³, Maria Grazia Lain¹³, Moira Spyer^{14,32}, Eleni Nastouli^{14,32}, Kathleen Gartner^{14,32}, Elisa López^{15,33}, Sheila Fernandez-Luis¹⁵, Denise Nanche¹⁵, Tacilta Nhampossa^{16,33}, Almoustapha Maiga^{17,34}, Mariam Sylla³⁴, Pablo Rojo^{19,29}, Carlo Giaquinto^{20,21,27}, Mathias Lichterfeld^{22,23,35}, Nigel Klein^{24,25}, Carlota Miranda²⁴, Savita Pahwa^{26,36}, Anne-Genevieve Marcelin³⁷, Vincent Calvez³⁷, Caroline Foster³⁸, Deborah Persaud³⁹, Viviana Giannuzzi⁴⁰, Annalisa Landi⁴⁰, Adriana Ceci⁴⁰, Ofer Levy², Philip Goulder⁴¹, Holly Peay⁴², Thanyawee Puthanakit⁴³ & Cissy Kityo⁴⁴

²⁷Fondazione PENTA ETS, Padua, Italy. ²⁸Bambino Gesù Children's Hospital, IRCCS, Rome, Italy. ²⁹Servizio Madrilenno de Salud SERMAS:Hospital 12 de Octubre, Madrid, Spain. ³⁰Columbia University, New York, USA. ³¹Faculty of Medicine and Health Sciences, Stellenbosch University, Cape Town, South Africa. ³²University College London, London, UK. ³³Fundação Manhiça, Manhiça, Mozambique. ³⁴Centre Hospitalier Universitaire Gabriel Touré, Bamako, Mali. ³⁵The Brigham & Women's Hospital, Inc., Boston, USA. ³⁶Miller School of Medicine, University of Miami, Miami, USA. ³⁷Association de Recherche en Virologie et Dermatologie, Paris, France. ³⁸Imperial College Healthcare NHS Trust, London, UK. ³⁹Johns Hopkins University, Baltimore, MD, USA. ⁴⁰Fondazione per la ricerca farmacologica Gianni Benzi, Bari, Italy. ⁴¹Umkhusele Innovation and Research Management, KwaZulu Natal, Pietermaritzburg, South Africa. ⁴²Research Triangle Institute dba RTI International, Durham, NC, USA. ⁴³Chulalongkorn University, Bangkok, Thailand. ⁴⁴Joint Clinical Research Center, Kampala, Uganda.