**UNIVERSIDAD COMPLUTENSE DE MADRID**
FACULTAD DE MEDICINA




# TESIS DOCTORAL

## Caracterización de pacientes perinatalmente infectados por VIH-1 con un inicio precoz del tratamiento antirretroviral

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

**Sara Domínguez Rodríguez**

Directores

**Alfredo Tagarro García**
**Pablo Rojo Conejo**

Madrid

**© Sara Domínguez Rodríguez, 2021**

# UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE MEDICINA




# TESIS DOCTORAL

***Caracterización de pacientes perinatalmente infectados por VIH-1 con un inicio precoz del tratamiento antirretroviral***

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

SARA DOMÍNGUEZ RODRÍGUEZ

DIRECTOR

Alfredo Tagarro García y Pablo Rojo Conejo

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE MEDICINA

# Caracterización de pacientes perinatalmente infectados por VIH-1 con un inicio precoz del tratamiento antirretroviral

TESIS DOCTORAL
POR COMPENDIO DE PUBLICACIONES

SARA DOMÍNGUEZ RODRÍGUEZ

Madrid, 2021



# Caracterización de pacientes perinatalmente infectados por VIH-1 con un inicio precoz del tratamiento antirretroviral

Esta memoria ha sido presentar para optar al grado de Doctor en Medicina por la Universidad Complutense de Madrid por la licenciada:

Sara Domínguez Rodríguez

## DIRECTORES DE TESIS

Pablo Rojo Conejo
Doctor en Medicina. Profesor Asociado del Departamento de Pediatría, Facultad de Medicina, Universidad Complutense de Madrid. Médico Adjunto, Servicio de Pediatría. Hospital Universitario 12 Octubre, Madrid.

VºBº del Director de Tesis:

Alfredo Tagarro García
Doctor en Medicina. Profesor Titular del Departamento de Pediatría, Facultad de Medicina, Universidad Europea de Madrid. Médico Adjunto, Servicio de Pediatría. Hospital Universitario Infanta Sofía, Madrid.

VºBº del Director de Tesis:

*A mi madre, mi padre y mi hermano.*

# Agradecimientos

Esta tesis no hubiese sido posible sin la ayuda y apoyo de los que trabajan y comparten o han compartido la vida conmigo.

Quería empezar a agradecer a todas las personas que se sorprendieron e intentaron reconducir mi decisión de estudiar biología por no ser de las carreras con más salidas profesionales que existían en el momento. Todos y cada uno de ellos, sin muchos saberlo, me hicieron tener más ganas de comenzar el reto que supondría dedicarme a la investigación.

Gracias Noemí por creer en mí desde muy pequeñas y celebrar tanto mis aciertos como reírnos juntas de mis fallos, aún desde la distancia. A María, Sara, Carlos y Jon, por ser mi apoyo en la universidad y después de ella. Por esas eternas horas en la biblioteca y por el después haciendo cantar a las ranas.

Gracias, a la Dra. Paloma Talayero, de la que he heredado su amor incondicional por la investigación intrahospitalaria. A la Dra. África Holguín, que me dio mi primera oportunidad en investigación y me dio a conocer este virus del VIH del que luego no pude escapar. A todo CoRISpe por acogerme siempre con una sonrisa.

A Patricia Rojas, gracias por ser el optimismo personificado, por darme paz en tiempos de guerra y margaritas cuando más se necesitaban. Porque me encanta esa forma tuya de ver la vida y cómo vas a por ella, estoy segura de que Martín y Violeta heredarán todo eso de ti. El resto espero que sea de mi querido Guillermo.

A Patricia Álvarez, por ser muchas veces mi guía. Gracias por tu empatía, tu paciencia y todos tus consejos. La P de PCR para mí siempre será de Patricia y no de la Polimerasa.

Gracias a las chicas del CNIO del grupo impronunciable (Rebeca, Lola, Lidia, Esther, Evangelina, Paulina, Dori, Noelia, Mirari y Marta). Sin vosotras no sabría ni la mitad de las cosas que se ahora de bioestadística, de bioinformática, y sobre todo de lo que significa ser un equipo de investigación y de amigas.

Gracias a todo el grupo de enfermedades infecciosas pediátricas del Hospital 12 de Octubre (Dani, Cristina, Elisa, Lourdes, David, Estrella…). Por contar conmigo para cada una de vuestras locas hipótesis y por tratarme tan bien siempre. Al grupo de EMPIRICAL (Cinta, Luis, Lilit, Lola, Álvaro, Miquel, Alfredo y Pablo), sin vosotros nada hubiese sido posible.

A Carlos por ser el único psicólogo cuerdo que conoceré y por valorar tanto mi trabajo. Por tu disponibilidad siempre a todo y por dar humor a este doctorado. A Itziar, por estar presente siempre incluso sin ser del mismo Hospital y por esa sonrisa de oreja a oreja que lo cura todo.

Gracias a mis directores de tesis Alfredo y Pablo. Alfredo, sin tu ayuda no tendría nada que escribir ni defender. Gracias por poner siempre una nota humor en el piano y por amenizar viajes de avión a congresos infinitos. De ti he aprendido a tener paciencia y a apuntar muy alto con los sueños. A ti Pablo, gracias por enseñarme que se puede ser el mejor de los líderes de grupo y hacerlo con un cariño y respeto enorme. Por siempre confiar en mi criterio y por acompañarme en todo este recorrido. Muchas gracias a esa marmita de optimismo en la que te caíste de pequeño.

Gracias al grupo de todas las caras, a Diyu, Ana, Navarro, Santi, Marcos, Barañano, Bistra, Ángel y Franco. Por escucharme, apoyarme y tomarse todas las cervezas que hiciesen falta cuando lo necesitaba y cuando no también. Habéis sido y sois mi colchón en la vida, incluyendo este doctorado. Gracias a vosotros tengo que callarme cada vez que alguien dice eso de "los amigos y los dedos de una mano". Este doctorado también es gracias a vosotros y os lo dedico.

También gracias a Carolina, quién ha sido mi brazo derecho, izquierdo, cabeza y corazón durante este tiempo poniendo cordura a todo lo que hacía y ser casi mi compañera de piso.

A Javi por ser, más que un cuñado, un segundo hermano.

A mis abuelos, a Juan, Carmen, Nati y Carlos. Porque siempre habéis estado orgullosos, incluso cuando no se entendía muy bien a lo que me dedicaba y por vuestro apoyo incondicional. Sin vosotros no hubiese podido lograr nada de lo que tengo.

A mis padres. Gracias por nunca perder la oportunidad de demostrarme lo orgullosos que estáis de mí. Gracias por quererme como soy, por intentar restar todos mis defectos y potenciar todas mis virtudes. Gracias por ayudarme en cada una de mis decisiones y por ser mis mejores amigos siempre. A mi hermano Ángel, por sus mil preguntas sobre este doctorado y por estar siempre cuando te he necesitado sin ni siquiera llegar a pedírtelo.

A mis pequeños Gordo y Elvis. Porque no sabemos nada de la lealtad, del amor y del compañerismo verdadero si no lo hemos experimentado con un animal.

Y gracias, Cristóbal. Por ser el mejor compañero de vida. Por aceptar que soy distinta a ti y reírte todavía cuando no consigues entenderme. Gracias por hacerlo todo fácil, por no permitir que diese importancia excesiva a lo que no lo tuviera y no dejarme caer nunca en las quejas. Por apoyarme y darme luz cuando más lo necesito.  Esta tesis es nuestra.

Por último, este trabajo va dedicado a todas las personas que conviven o convivieron con el VIH y con el estigma y desolación que hayan podido sentir. Todavía queda mucho por descifrar de este virus y espero que mi trabajo haya contribuido en algo.

*"The show must go on"*-Freddie Mercury 

# Abreviaturas

3TC: Lamivudina

ANN: *Artificial Neural Network* o Redes neuronales artificiales

ARV: Fármacos antirretrovirales

AZT: Zidovudina

CD4: células linfocitos T CD4

HR: Hazard Ratio

IC: Intervalo de confianza

ITIAN: inhibidores de la transcriptasa inversa análogos de nucleósido

ITINN: inhibidores de la transcriptasa inversa no análogos de nucleósido

KNN: *K-nearest neighbor*

TAR: Tratamiento antirretroviral

Log: logaritmo en base decimal

ML: *Machine learning* o aprendizaje automático

NB: Naïve Bayes

NVP: Nevirapina

PBMC: Célula mononuclear de sangre periférica

RF: Random Forest

RIC: Rango intercuartílico

RL: Regresión logística

SIDA: Síndrome de inmunodeficiencia adquirida

SVM: *Support Vector Machine*

VIH-1: Virus de la inmunodeficiencia humana de tipo I

# Publicaciones y Congresos

La presente tesis doctoral, de acuerdo con el informe correspondiente autorizado por los directores de tesis, y en cumplimiento con la normativa aprobada por el Órgano Responsable del Programa de Doctorado, se presenta como un compendio de cuatro publicaciones. Tres artículos ya han sido publicados, el cuarto ha sido enviado para su publicación en el momento de redacción de esta memoria. Las referencias completas de los artículos que constituyen el cuerpo de la tesis son las siguientes:

1. **Reduced Time to Suppression Among Neonates with HIV Initiating Antiretroviral Therapy Within 7 Days After Birth**. Domínguez-Rodríguez S, Tagarro A, Palma P, Foster C, Puthanakit T, Jupimai T, Cotugno N, Ananworanich J, Zangari P, Nastouli E, Muñoz-Fernández MÁ, Navarro ML, Giaquinto C, Rossi P, Kuhn L, Rojo P. 2019 J Acquir Immune Defic Syndr. doi: 10.1097/QAI.0000000000002188.
   (Factor de Impacto 3.86, 1er cuartil, Enfermedades Infecciosas 22/89)
2. **The CARMA Study. Early infant ART: timing impacts on total HIV-1 DNA quantitation 12 years later.** Foster C, Domínguez-Rodríguez S, Tagarro A, Gkouleli T, Heaney J, Watters S, Bamford A, Fidler K, Navarro M, De Rossi A, Palma P, Nastouli E, Rossi P, Giaquinto C, Rojo P on behalf of the EPIICAL. 2020 JPIDS doi:10.1093/jpids/piaa071
   (Factor de Impacto 2.12, segundo cuartil, Pediatría 40/128)
3. **Faster initial viral decay in female children living with HIV**. Domínguez-Rodríguez S, Serna-Pascual M, Foster C, Palma P, Nastouli E, De Rossi A, Seoane J, Rossi P, Giaquinto C, Tagarro A, Rojo P on behalf of the EPIICAL Project. JPIDS 2020 doi: 10.1093/jpids/piaa174
   (Factor de Impacto 2.12, segundo cuartil, Pediatría 40/128)
4. **Machine learning outperformed logistic regression classification even with limit sample size: A model to predict pediatric HIV mortality and clinical progression to AIDS**. Domínguez-Rodríguez S, Serna-Pascual M, Oletto A, Barnabas S, Zuidewind P, Dobbels E, Danaviah S, Behuhuma O, Lain MG, Vaz P, Fernandez S, Nhampossa T, Lopez-Varela E, Otwombe K, Liberty A, Violari A, Maiga AI, Rossi P, Giaquinto C, Kuhn L, Rojo P, Tagarro A. Submitted to Pediatric Infectious Disease Journal 2021
   (Factor de Impacto 2.126, segundo cuartil en Pediatria)

Los resultados expuestos han sido presentados en los siguientes congresos científicos internacionales:

1. Tagarro A, Dominguez Rodriguez S, Puthanakit T, Palma P, Foster C, Jupimai T, Cotugno N, Ananworanich J, Jimenez de Ory S, Zangari P, Navarro ML, Rossi P, Nastuoli E, Giaquinto C, Rojo P. "Neonatal ART< 7 days vs 7-28 days reduced time to suppression". Comunicación oral en 2019 Conference on Retroviruses and Opportunistic Infections (CROI) (ID 2436), Seattle, EEUU.
2. The CARMA Study: Children on early suppressive ART: total HIV-1 DNA quantitation 12 years post ART initiation. C Foster, S Dominguez, A Tagarro, E Nastouli, R Gkouleli, J Heaney, P Palma, P Rossi, C Giaquinto, P Rojo on behalf of the EPIICAL consortium. Póster en 10th IAS Conference on HIV Science (IAS 2019), Mexico City, Mexico
3. Tagarro A, Domínguez-Rodríguez S, Violari A, Cotton M, Mhampossa T, Klein N, Ramsagar N, Janse Van Rensburg A, Behuhuma O, Vaz P, Oletto A, Palma P, Rossi P, Rojo P, on behalf of the EPIICAL Consortium. Poor outcome in early treated HIV-perinatally infected infants in Africa. Póster en 2020 Conference on Retroviruses and Opportunistic Infections (CROI) (Poster 00806), Boston, EEUU.

4. Domínguez-Rodríguez S, Tagarro A, Serna Pascual M, Otwombe K, Violari A, Fernández S, Nhampossa T, Lain M, Vaz P, Behuhuma O, Danaviah S, Dobbels E, Barnabas S, Cotugno N, Zangari P, Palma P, Oletto A, Nardone A, Nastouli E, Spyer M, Kuhn L, Rossi P, Giaquinto C, Rojo P on behalf of EPIICAL consortium. A Machine Learning Approach for Predicting Probability of Death or Disease Progression in an Early-Treated Pediatric African cohort. Póster en 2020 International AIDS Conference 2020, San Francisco, EEUU.

5. Domínguez-Rodríguez S, Serna-Pascual M, Foster C, Palma P, Nastouli E, De Rossi A, Seoane J, Rossi P, Giaquinto C, Tagarro A, Rojo P on behalf of the EPIICAL Project. Faster initial viral decay in female children living with HIV. Abstract Nº 1617 en 2021 Conference on Retroviruses and Opportunistic Infections (CROI), Chicago, EEUU. Aceptado como presentación Science Spotlight

6. Domínguez-Rodríguez S, Barnabas S, Behuhuma O, Lain MG, Nhampossa T, Otwombe K, Maiga AI, Rossi P, Giaquinto C, Kuhn L, Serna-Pascual M, Rojo P, Tagarro A on behalf of EPIICAL Consortium. MAGPIE: A Machine-learning app to predict pediatric HIV mortality and progression. Pendiente de aceptación en 2021 International AIDS Conference 2021, Virtual.

Y en los siguientes congresos nacionales:

1. Domínguez-Rodríguez S, Jiménez de Ory S, Rojo P, Collado P, Fortuny C, Mur A, Moreno D, Roa M.Á, Angulo R, Mellado P, Penin M, Gavilán M, Méndez P, Soler-Palacín, Navarro M, Ramos JT, Tagarro A. "El inicio de TAR antes de la primera semana de vida disminuye el tiempo hasta la supresión viral". Comunicación oral en X Congreso Nacional de GeSIDA 2018, Madrid



# Índice


Resumen ..... 13

Summary ..... 19

Introducción ..... 25

Hipótesis ..... 35

Objetivos ..... 39

Materiales y Métodos ..... 43

Población de estudio ..... 45

- Población de estudio para el capítulo I ..... 45
- Población de estudio para el capítulo II ..... 45
- Población de estudio para el capítulo III ..... 46
- Población de estudio para el capítulo IV ..... 46

Métodos de laboratorio ..... 47

- Métodos de laboratorio para el capítulo I ..... 47
- Métodos de laboratorio para el capítulo II y III ..... 47

Análisis estadístico ..... 48

- Análisis estadísticos empleados en el capítulo I ..... 48
- Análisis estadísticos empleados en el capítulo II ..... 51
- Análisis estadísticos empleados en el capítulo III ..... 52
- Análisis estadísticos empleados en el capítulo IV ..... 53

Capítulo I: ..... 57

Justificación y Objetivos ..... 59

Resultados ..... 59

Capítulo II ..... 73

Justificación y Objetivos ..... 75

Resultados ..... 75

Capítulo III: ..... 81

Justificación y Objetivos ..... 82

Resultados ..... 82

Capítulo IV: ..... 93
Justificación y Objetivos ..... 95
Resultados ..... 96
Discusión ..... 123
Conclusiones ..... 133
Bibliografía ..... 136

# Resumen

## Introducción

Desde el inicio de la epidemia de la infección por el VIH-1, la transmisión vertical del virus de una madre infectada a su hijo se ha considerado un problema de salud pública de primera magnitud. A pesar del avance de las políticas y medidas de prevención de la transmisión vertical madre-hijo, se estima que anualmente 150,000 niños aún se infectan con VIH-1. El tratamiento antirretroviral precoz ha conseguido ser eficaz en la reducción de mortalidad y morbilidad, en restaurar la función del sistema inmunológico y conseguir una supresión viral duradera. Sin embargo, aún se desconoce cuánto de precoz ha de ser el inicio del TAR y los factores asociados a su eficacia en el descenso de carga viral en niños. Asimismo, el TAR por sí solo no es capaz de eliminar la infección debido a la incapacidad de acceso al reservorio viral, lugares donde permanece el virus, produciéndose una infección latente. Conocer los factores asociados a un reservorio bajo podría arrojar luz sobre la remisión del VIH-1 en pacientes pediátricos.

## Objetivos

El objetivo general de esta memoria es estudiar las características clínicas y virológicas de los pacientes perinatalmente infectados por VIH-1 con un inicio precoz del tratamiento antirretroviral.

De forma específica, esta memoria tiene como objetivos: 1) determinar el beneficio a largo plazo de iniciar TAR en la primera semana de vida en comparación con iniciar entre la segunda y cuarta semana en neonatos que iniciaron el TAR en el primer mes de vida, 2) describir la asociación entre el sexo y el descenso de la carga viral y si este descenso es un factor predictor del reservorio a largo plazo, 3) identificar los factores predictores de un reservorio viral bajo, y por último, 4) desarrollar una herramienta que permita predecir la mortalidad y progresión al SIDA de los pacientes perinatalmente infectados por VIH-1.

## Materiales y métodos

Las poblaciones de estudio sobre las que se basan los capítulos que componen esta memoria se engloban dentro del proyecto EPIICAL (*Early-treated Perinatally HIV-infected Individuals: Improving Children's Actual Life with Novel Immunotherapeutic Strategies*) cuya finalidad principal es establecer una colaboración científica para liderar tratamientos efectivos para

conseguir la remisión del VIH-1 en niños. Este proyecto ha estudiado cohortes pediátricas tanto en Europa como en África para diseñar y dirigir estudios en niños con VIH-1. Las cohortes estudiadas son las siguientes: para el objetivo 1, Cohorte internacional (Italia, UK, Tailandia, España-Cohorte Nacional Pediátrica Española de la red CoRISpe) de 44 neonatos tratados en el primer mes de vida (capítulo I); para los objetivos 2, 3 y 4: *Child and Adolescent Reservoir Measurements on Early Suppressive ART* (CARMA), consistente en pacientes pediátricos europeos con infección vertical por VIH-1 tratados precozmente y estudiados de forma transversal (capítulos II y III); y para el objetivo 4: *Early Antiretroviral Treatment in HIV Children* (EARTH), consistente en lactantes africanos con infección vertical tratados muy precozmente y seguidos prospectivamente desde su nacimiento (capítulo IV).

### Análisis estadísticos

En el capítulo I, para describir el efecto del tiempo de iniciación del TAR en el tiempo a la supresión, se empleó un modelo multivariable flexible semiparamétrico de supervivencia con censura por intervalos.

En el capítulo II se realizó un modelo generalizado mixto con distribución de Poisson incluyendo género como variable de interacción para determinar el efecto del género en el descenso de la carga viral en función del tiempo al TAR.

En el capítulo III se determinó la asociación entre factores predictivos como el tiempo al inicio del TAR, carga viral pre-TAR, porcentaje de CD4 pre-TAR usando un modelo generalizado aditivo mixto.

En el capítulo IV se compararon siete algoritmos de aprendizaje automático (regresión logística (RL), random forest (RF), Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Naïve Bayes (NB), Artificial Neural Network (ANN) y Elastic Net) para la predicción de mortalidad y progresión a SIDA.

## Resultados

De los 44 neonatos tratados antes de un mes, 19 (43.2%) iniciaron el TAR en la primera semana de vida y 25 (56.6%) entre el día 8-28. En aquellos niños tratados en la primera semana fue cuatro veces más probable la supresión que en aquellos tratados entre el día 8-28 (Hazard

Ratio (HR) 4,01 (Intervalo de confianza (IC) 95% 1,7-9,5). Por cada semana de retraso en el inicio del TAR, la probabilidad de supresión se redujo un 35% (HR 0,65 (IC 95% 0,46-0,92)). La asociación entre la edad al inicio del TAR y el tiempo a la supresión fue lineal. Sin embargo, la asociación entre la carga viral pre-ART y el tiempo a la supresión se definió por una primera fase con una asociación mínima (cargas virales <5 logaritmo en base decimal (logs) copias/mL) seguida de una fase exponencial por encima de 5 logs copias/mL.

En el segundo artículo se modelizaron las curvas longitudinales de descenso de carga viral en función del tiempo desde el inicio del TAR y el sexo. Se pudo observar que existieron diferencias en la asociación entre las cargas virales y tiempo en TAR con respecto al género. Las niñas presentaron descensos de carga viral en la fase I más rápido que los niños (coeficiente de interacción -0,01±0,001, p-valor=0,040) independientemente de la edad al inicio del TAR, porcentajes de CD4 pre-TAR o cargas virales pre-TAR. Además, se observó que este descenso de carga viral era capaz de predecir el reservorio viral. Pendientes más acusadas estaban asociadas a reservorios más bajos medidos 12 años después.

En el tercer artículo se observó que, de los 40 niños y adolescentes estudiados, 11 tuvieron un reservorio viral bajo definido como menos de 10 copias de DNA de VIH-1 por millón de células mononucleares de sangre periférica (PBMCs). Los factores asociados a presentar un reservorio bajo fueron una menor edad al inicio del TAR (coeficiente 0,10±0,001, p-valor<0,001), niveles bajos de cargas virales pre-TAR (coeficiente 0,35±0,01, p-valor<0,001), menor porcentaje de CD4 al inicio del TAR (coeficiente 0,03±0,01, p-valor=0,049), y puntuaciones de western blot menores. Mientras que la asociación entre la edad al inicio del TAR y el reservorio era lineal (a mayor edad al inicio del TAR, mayor reservorio), la asociación entre la carga viral y reservorio fue notable solo por encima de 6 logs, así como por encima de 40% de CD4.

En el cuarto capítulo y último trabajo que compone la memoria se comparan siete algoritmos para predecir la mortalidad y progresión a SIDA en niños precozmente tratados. El algoritmo con mejor rendimiento fue *Random Forest* con una precisión en la clasificación del 82,8%, sensibilidad 78% y área bajo la curva de 0,73. Este algoritmo tuvo mejor capacidad de clasificación de los pacientes que morían o progresaban a SIDA que otros algoritmos como la regresión logística comúnmente utilizada para analizar datos clínicos. Este modelo fue incluido en una aplicación web a la que se puede acceder mediante el siguiente enlace: https://rserver.h12o.es/pediatria/MAGPIE/. Username: user, Password:0000.

## Conclusiones

El inicio precoz del TAR en la primera semana de vida condiciona una supresión viral y control de la viremia cuatro veces más rápida que aquellos niños que iniciaron entre los días 8-28 de vida.

El sexo podría ser un factor determinante de la eficacia del tratamiento y control de la viremia en niños precozmente tratados. Las niñas presentaron descensos de cargas virales más rápidos que los niños independientemente de la edad al inicio del TAR, porcentajes de CD4 pre-TAR o cargas virales pre-TAR. Además, un descenso de carga viral acusado es predictor de un menor tamaño de reservorio a largo plazo.

Un reservorio bajo de ADN total de VIH-1 se asoció con una edad temprana al inicio del TAR y cargas virales bajas pre-TAR. El impacto de un inicio precoz del tratamiento en el tamaño del reservorio persiste después de una década de terapia supresora.

Los algoritmos de aprendizaje automático son capaces de predecir de forma más eficiente la mortalidad y progresión a SIDA de los niños perinatalmente infectados y precozmente tratados que los modelos tradicionales. En concreto, el algoritmo que mejor rendimiento tuvo fue *Random Forest*.

# Summary

## Background

Since the beginning of the epidemic, HIV mother-to-child transmission has been recognised as a major public health issue. Despite the great efforts done to reduce HIV infection among children, in 2015, more than 150,000 children became infected with HIV. Early combination antiretroviral therapy (ART) initiation in vertically infected infants with HIV has had a major impact on clinical outcomes reducing morbidity and mortality. Moreover, early initiation of ART leads to a significantly smaller reservoir size and a better immune reconstitution. However, definitions of early vary a great deal and literature reporting on ART initiation in the first year of life rarely reports on outcomes when treatment is initiated in the neonatal period. Also, there is limited information about factors associated with the rapid decline of the viral load in infants.

The ART alone is unable to confer a cure, a consequence of an inaccessible latent pool of HIV-1-infected reservoir cells. Identifying factors and biomarkers, associated with limited viral reservoir size will potentially aid identification of participants for future proof-of-concept studies aimed at inducing ART-free viral remission.

## Objectives

The main aim of this thesis is to study the clinic and virologic characteristics of the perinatally early treated HIV infected children. More specifically, this thesis aims to: i) investigate the association between the timing of ART initiation and time-to-suppression among infants who tested HIV-positive and initiated ART within the first 28 days of life; ii) describe the association between sex and viral decay and iii) if the viral decay is a proxy of long-term HIV reservoir; iv) identify the clinical, virologic, and immunologic factors associated with a small reservoir size; and iv) to develop a tool that allows to predict mortality and/or progression to AIDS in the perinatally HIV infected children

## Materials and Methods

### Study population

The study population of this thesis are framed in the EPIICAL Consortium (*Early-treated Perinatally HIV-infected Individuals: Improving Children's Actual Life with Novel*

*Immunotherapeutic Strategies*). The overall goal of this project is to establish new, early scientific efforts leading to treatments aimed at HIV remission in children. This project has two well-established cohorts: CARMA (a cross-sectional multicentric study cohort of children who commenced ART <2 years and remained suppressed within 12 months and remained suppressed (viral load, (VL) <50 copies/ml) for >5 years), to perform chapter II and II, and EARTH (a prospective cohort study of infants born with HIV and treated before 90 days of life in South Africa and Mozambique), to perform chapter IV. In addition, we collected a retrospective subset of 44 neonates treated within 28 days of life from four different cohorts (Rome Cohort (Italy), Imperial College Hospital (London, United Kingdom), Chulalongkorn University (Thailand), and Spanish National HIV Paediatric Network-CoRISpe (Spain)) to perform chapter I.

Statistical Analysis

In chapter I, we described the association between age at ART and time to suppression with a multivariable flexible semiparametric survival model with interval censoring.

In chapter II we assessed the effect of sex in the viral decay due to ART by using a generalized Poisson mixed regression model including sex as an interaction term.

In chapter III we performed a generalized additive model to investigate the type of association between predictive factors as age at ART, pre-ART viral load and percentage of CD4 and the reservoir size.

Finally, we compared seven machine learning algorithm including the conventional logistic regression, random forest (RF), Support Vector Machine (SVM), K-Nearest Neighbour (KNN), Naïve Bayes (NB), Artificial Neural Network (ANN), and Elastic Net to predict a combined outcome of 1-year mortality and clinical progression to AIDS.

## Results

Nineteen of the forty infants included in chapter I (43.2%) initiated ART within 7 days of life and 25 (56.8%) from 8 to 28 days. Infants treated within 7 days were 4-fold more likely to suppress earlier than those treated after 7 days [Hazard ratio (HR) 4.01 (1.7–9.5)]. For each week the ART initiation was delayed, the probability of suppression decreased by 35% (HR 0.65 [0.46–0.92]). Age at ART start was linearly associated with time-to-suppression. However,

a linear association with normally distributed residuals was not found between baseline viral load and time-to-suppression, with no association found when baseline viral loads were ≤5 log(10) copies/mL, but with exponential increase in time-to-suppression with >log5 copies/mL at baseline. Starting ART within 7 days of life led to 4-fold faster time to viral suppression, in comparison to initiation from 8 to 28 days.

In chapter II, we observed statistically significant differences in viral decay between female and male infants. Females presented faster phase-I viral decay (interaction coefficient, -0.01±0.001, p-value=0.040) regardless of their age at ART initiation, baseline %CD4, or baseline RNA levels. Also, for each month elapsed under ART, females had faster viral decay than male. In addition, viral decay was a predictor of long-term HIV DNA reservoir measured at a median of 12 years after ART initiation.

In chapter III, forty perinatally HIV infected infants were described, of those 27 (67.5%) were female and 21 (52.5%) Black/Black African. The reservoir size was measured at a median of 12 (IQR, 7.3– 15.4) years after ART initiation. Eleven had total HIV-1 DNA <10 copies/106 PBMCs. HIV-1 DNA levels were positively associated with age and VL at ART initiation, baseline CD4%, and Western blot antibody score. Age at ART initiation presented a linear association (coefficient = 0.10 ± 0.001, P ≤ .001), the effect of VL (coefficient = 0.35 ± 0.1, P ≤ .001) noticeable >6 logs. The effect of CD4% (coefficient = 0.03 ± 0.01, P = .049) was not maintained >40%.

Finally, in chapter IV, we compared seven state-of-the-art machine learning algorithms to predict mortality and/or progression to AIDS in early treated patients. Random Forest ranked as the best algorithm in terms of accuracy (82·8%), sensitivity (78%), and AUC (0·73). Regarding specificity and sensitivity, RF showed better performance than the other algorithms in the external validation and the highest AUC. LR showed lower performance compared with RF, SVM, or KNN. The RF model was included in the shiny app, which allowed to output the classification of risk and the probability of the result. (https://rserver.h12o.es/pediatria/MAGPIE/). Username: user, Password:0000.

## Conclusions

Early initiation of ART, even in the first week, is associated with a better viremia control and this may be associated with a smaller reservoir size.

The sex could be a determinant factor of the ART efficacy and viremia control in early treated infants. Females presented faster viral decay in phase I than the males adjusted by the age at ART initiation, percentage of CD4, and pre-ART viral loads. Besides, the more pronounced is the slope of the viral decay, the smaller the DNA reservoir was estimated according to this model. A reduce reservoir size was associated also with younger age and lower %CD4 at ART initiation. The impact of early-infant treatment on reservoir size persists after a decade of suppressive therapy.

The outcome of early-treated infants with HIV can be predicted using an ML model, which integrates standard clinical, virological, and immunological data. The online tool can be useful for health workers to identify children with high risk and implement appropriate countermeasures.

# Introducción

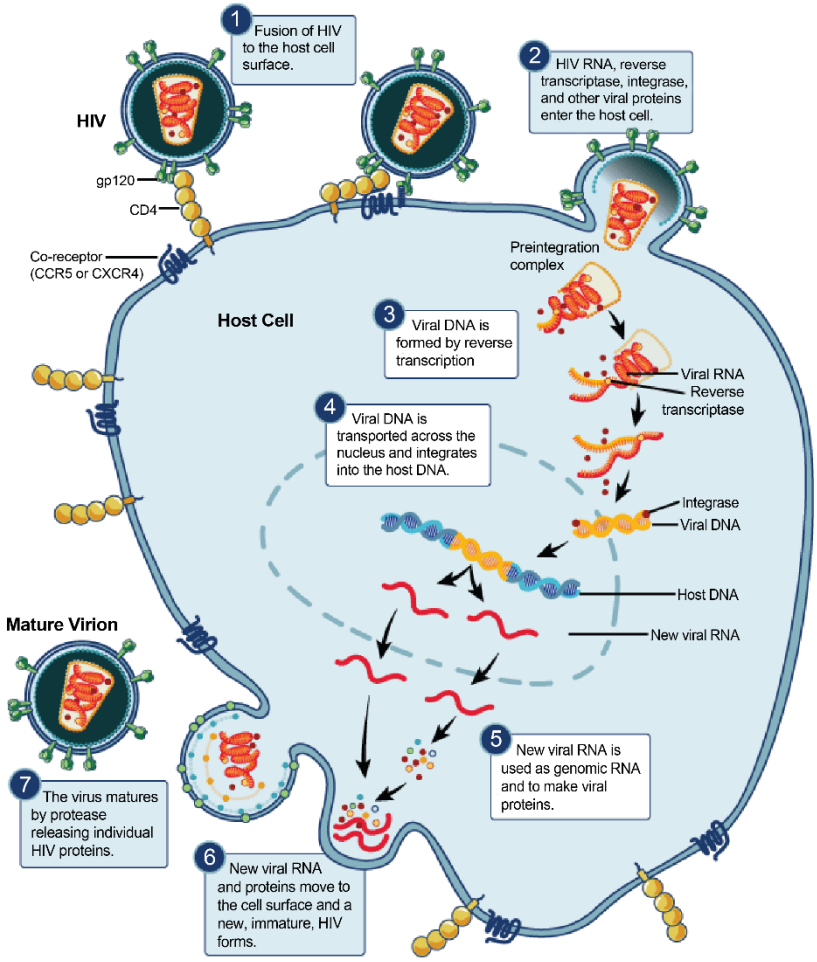
El virus de la inmunodeficiencia humana (VIH-1) es el agente infeccioso causante del síndrome de inmunodeficiencia adquirida (SIDA). Una de las principales características del VIH-1 es su gran heterogeneidad genética, que surge como resultado de la combinación de una alta tasa de replicación y sus elevadas tasas de mutación y recombinación (1). Asimismo, estos viriones pueden presentar un amplio rango de fenotipos respecto a virulencia, tropismo celular, velocidad de replicación, composición antigénica y susceptibilidad a los fármacos antirretrovirales (ARV), características que dan al virus la capacidad de escapar al sistema inmune y al tratamiento (2). Esta enorme diversidad se observa tanto a nivel de individuo como a nivel poblacional (3).

El ciclo infectivo y replicativo del VIH-1 se resume en las siguientes etapas (Figura 1). Primero se produce la interacción entre el virus con la célula diana (linfocitos T CD4 principalmente) por la unión de la glicoproteína viral con el receptor celular CD4 y de otras proteínas de membrana celulares que actúan como correceptores del virus. Después, se produce la fusión de las membranas del virus y de la célula y se libera al citoplasma la cápside viral. Una vez dentro de la célula, se inicia la transcripción inversa del ARN genómico viral y la formación de ADN por la acción de la transcriptasa inversa. Este ADN se transporta al núcleo celular y se integra en el genoma por la acción de la integrasa viral. Por último, se transcriben los genes virales y se produce la traducción de proteínas precursoras virales. Una vez que estas proteínas se ensamblan formando un virión, éste sale de la célula por gemación arrastrando parte de la membrana de la célula huésped produciendo su depleción. Por último, los viriones maduran gracias a la proteasa viral que procesa las proteínas precursoras para formar un virión infectivo (4).

Dentro de un individuo, variantes circulantes del virus podrán integrarse en el genoma celular en forma de *provirus*, donde permanecerán latentes constituyendo un reservorio viral. Estos provirus latentes pueden revertir su estado a productivos después de una estimulación. Además del reservorio latente, existe un reservorio activo que se compone de una población viral muy limitada productora de ARN viral, proteínas y posiblemente viriones (5). La localización del reservorio está relacionada con el sistema nervioso central, sistema reproductor y órganos linfoides, siendo las células T CD4+, aquellas por las que el virus tiene un mayor tropismo (6). Esta extraordinaria variabilidad genética del VIH-1, junto con la capacidad de integración en el genoma celular, tiene implicaciones clínicas y epidemiológicas

importantes y representa un enorme desafío para el diseño de fármacos ARV, vacunas y terapias efectivas.

**Figura 1.** Ciclo replicativo del VIH




Adaptado de National Institutes of Allergy and Infectious Diseases (NAID), EEUU. HIV Replication Cycle http://phil.cdc.gov/phil/details.asp ID#: 18162).

A pesar del éxito de las políticas y medidas de prevención de la transmisión vertical madre-hijo de la infección por el VIH-1 en los países desarrollados, en 2019 el 20% de las madres embarazadas infectadas por VIH-1 no tuvieron acceso al tratamiento antirretroviral (TAR) y aproximadamente 150,000 niños se infectaron por VIH-1 en el mundo (7). El curso natural de la infección por VIH-1 en edad pediátrica tiene particularidades propias. La mayoría de los

niños se infectan por transmisión vertical, en una etapa de pleno desarrollo y diferenciación celular. Esta transmisión vertical puede ocurrir en diferentes puntos en el tiempo: intraútero, intraparto o a través de la lactancia. La transmisión del virus es un mecanismo multifactorial. La infección intraútero es mediada por microtransfusiones sanguíneas de la madre al feto a través de la placenta y favorecida por otras infecciones del líquido amniótico. Estas microtransfusiones se vuelven más frecuentes por las contracciones uterinas en el momento del parto, lo que aumenta las probabilidades de transmisión del virus. Además, a través del canal del parto el feto tiene contacto mucocutáneo con la sangre materna, el líquido amniótico y las secreciones cervicovaginales (8). Por último, los lactantes se pueden infectar durante todo el período de la lactancia a través de la leche materna. En ausencia de tratamiento, el curso de la enfermedad en los niños infectados intraútero es diferente presentando una supervivencia significativamente menor que aquellos infectados después del nacimiento (9).

El VIH-1 provoca un grave trastorno de la inmunidad, que se manifiesta clínicamente por infecciones graves y por una mayor predisposición para el desarrollo de algunas enfermedades (10). Comparado con adultos, la infección aguda por VIH-1 en niños se caracteriza por la presencia de altas cargas virales que conllevan a un alto riesgo de progresión a SIDA y/o la muerte (11). Además, los niños tienen un sistema inmune diferente al de los adultos, con un timo más activo y una gran capacidad para la regeneración inmune (12). Ello se traduce en que los niños infectados con VIH-1 tratados en los primeros meses de vida presentan un tamaño de reservorio muy limitado y una alta diversidad viral comparado con los adultos precozmente tratados (13). Además, un control virológico en niños actúa como prevención de la activación del reservorio y el mantenimiento de un reservorio bajo en el tiempo (14).

El TAR ha conseguido ser eficaz en la reducción de mortalidad, morbilidad, restaurar la función del sistema inmunológico y conseguir una supresión viral duradera (15). En ausencia del TAR, la progresión de la enfermedad es más rápida en niños infectados por transmisión vertical que en adultos, adolescentes o niños infectados por otras vías (16). El ensayo clínico aleatorizado CHER (17), realizado en población pediátrica infectada por VIH-1, demostró el beneficio del inicio inmediato del TAR después del diagnóstico comparado con el tratamiento deferido acorde a las guías de la OMS del 2006 basadas en criterios clínicos e inmunológicos

(18). En estas guías se recomendaba iniciar el TAR en aquellos niños un porcentaje de células CD4 menor al 20%, o en el caso de mayores de doce años, menor al 25% o un recuento total de menos de 1000 células/$mm^3$ y con un estadio severo C definido por el Centro de Control y Prevención de Enfermedades (CDC). El estudio CHER reportó una reducción del 76% en la mortalidad y del 75% en la progresión a SIDA en los pacientes con tratamiento inmediato. Estos resultados conllevaron a la revisión de las guías de la OMS en el 2008, recomendando el inicio del TAR en todos los niños infectados por VIH-1 menores de un año independientemente de cual fuera su estado inmunológico o clínico. Posteriormente dichas recomendaciones se generalizaron a toda la población pediátrica y adulta.

El inicio precoz del TAR está asociado a un rápido control viral y posterior control de la viremia (15,19), a la presencia de reservorios virales menores y a una reconstitución inmune mejor (19–23). Actualmente existen muy pocos estudios en niños centrados en comparar resultados virológicos comparando tratar antes y después (17,24–26). Sin embargo, la definición de tratamiento precoz es muy variable. La literatura al respecto fluctúa desde los primeros tres meses de vida a seis meses de vida, siendo muy escasa cuando se trata del periodo neonatal (primeros 28 días de vida) (15).

Para poder llevar a cabo un tratamiento precoz es necesario tener también en consideración un diagnóstico temprano. Debido a los considerables avances en las técnicas diagnósticas *in situ* o *point-of-care* en las áreas de maternidad y su actual disponibilidad en países de alta prevalencia, es necesario determinar con cuánta rapidez debería ser prescrito el TAR en el periodo neonatal y si iniciar en los primeros días de vida supone un beneficio a largo plazo para estos pacientes. Mientras que en Europa y Estados Unidos se diagnostica el VIH-1 en los primeros días de vida, en África Subsahariana está recomendado el diagnóstico entre la 4-6 semana durante la primera visita posnatal. A pesar de esa recomendación, se estima que se realiza el test del VIH-1 a menos del 20% de los niños expuestos (27,28).

Existen modelos que explican la dinámica de la infección del VIH-1 a partir de cambios temporales en la carga viral desde el inicio del TAR. En particular, se ha descrito como los diferentes subconjuntos de células infectadas, como las células T CD4, descienden a diferente velocidad y en momentos distintos asociados al descenso de la carga viral tras el inicio del TAR. Este descenso es representado por una curva bifásica compuesta por una fase rápida, que dura semanas, seguida de una fase más lenta que puede durar meses (Figura 2) (29–31).

Estas dos fases en el descenso de la carga viral están asociadas a dos principales fuentes celulares de viriones: células de memoria de vida corta y células *naïve* de vida larga.

En la primera fase (primeras semanas después del inicio del TAR), se produce un descenso de la carga viral rápido y exponencial. Al inicio de la infección existe una tasa de replicación del VIH-1 muy alta llegando, en ocasiones, hasta cargas virales del orden de $10^{6-7}$ copias de ARN por mililitro. En esta fase inicial rápida, el VIH-1 preferiblemente infecta a células T CD4 activadas o de memoria (29,32), donde se replica y provoca su depleción por efecto citopático directo del virus (debido al ensamblaje y salida) o efecto indirecto que implica la inducción a muerte celular. Además, existe un bloqueo en la regeneración linfocitaria (33). Ni los virus circulantes ni las células infectadas productoras de la mayoría de los virus en plasma sobreviven por mucho tiempo. Asimismo, el TAR previene nuevas infecciones a células susceptibles (34). Por tanto, este descenso tan acusado después del TAR en la carga viral en plasma es debido a dos procesos independientes: la eliminación de los viriones circulantes en plasma y la depleción de las células infectadas (35).

En la segunda fase que dura varios meses, este descenso de la carga viral va a un ritmo más lento y se asocia a la depleción de células T CD4 *naïve* infectadas cuya vida media es más prolongada. Finalmente, quedarán solo células "durmientes" con una infección latente y la carga viral permanecerá suprimida por el TAR (Figura 2). De esta manera, se producen dos fases diferentes de descenso de la carga viral. La velocidad de descenso de carga viral, es decir, el perfil de estas curvas, es usado como indicador de la eficacia del tratamiento (36).

Figura 2. Curva de descenso de carga viral desde el inicio del TAR

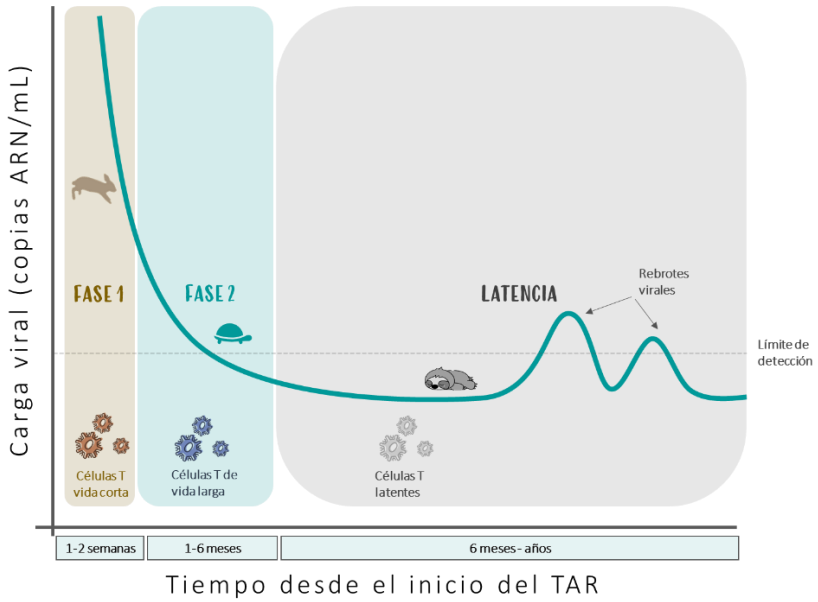



*Figura adaptada de Luzuriaga K et al. 1999* (37)

Sin embargo, existe variabilidad en las respuestas virológicas al tratamiento en los diferentes pacientes. Esta heterogeneidad puede ser debida a distintos factores genéticos, inmunológicos, microbioma, componentes hormonales, sociales y de comportamiento (38). En concreto, diversos estudios han demostrado el papel que juega el género en la prevención del VIH-1, transmisión, vulnerabilidad y respuesta (39). Estas diferencias han sido muy poco estudiadas en población pediátrica, cuyo análisis podría arrojar luz a la posible asociación del sexo y el descenso de la carga viral desde las primeras etapas de desarrollo. Algunos autores han descrito cargas virales más bajas (40), linfocitos CD4 más elevados (41) y mayor número de controladores de élite en niñas (42). A pesar de ello, hasta la fecha no existe ningún estudio que analice el efecto del género en las fases de descenso de carga viral desde el inicio de tratamiento.

A pesar de que el TAR ha sido una revolución en el aumento de la supervivencia, por sí mismo, por el momento, es incapaz de curar debido a la incapacidad de acceso al reservorio. Esto conlleva una infección latente y la necesidad de un tratamiento crónico. Sin embargo, el diagnóstico precoz en niños y un inicio temprano del TAR después de la infección minimiza el tamaño del reservorio viral en células de memoria T CD4+ limitando la evolución virológica de la infección por el VIH-1. Son varios los casos de pacientes precozmente tratados y bien controlados inmuno-virológicamente, en los que, tras la interrupción del TAR, se produce una

respuesta virológica o rebote viral en menos de 2-4 semanas (43). Sin embargo, el caso del bebé Mississippi, paciente en el que accidentalmente se interrumpió el TAR a los 18 meses de edad y estuvo 27 meses con carga viral indetectable y bajo reservorio viral, proporcionó un nuevo paradigma (44). Este caso generó una prueba de concepto, sugiriendo que el tamaño de reservorio en una infección perinatal podría permitir una remisión parcial o total del TAR. En el 2013 otro caso similar fue reportado en el que se describía como una adolescente de origen francés infectada al nacimiento había permanecido aparentemente con buena salud durante 12 años desde su última toma de antirretrovirales (45). El último caso, reportado en el 2019, fue un niño de la cohorte CHER (46). Estos tres casos se suman a la cohorte VISCONTI de 20 pacientes, tratados precozmente y con discontinuación del TAR sin efectos clínicos aparentes (47).

Diversos estudios han confirmado la hipótesis de que la reducción del tiempo hasta el inicio del tratamiento está asociada a un menor tamaño de reservorio (19,21,48,49). Conocer los factores virológicos, inmunológicos y clínicos asociados a un menor reservorio viral podría permitir identificar pacientes para futuros estudios de estrategias terapéuticas. Los pacientes pediátricos infectados por VIH-1 con un excelente y prolongado control de la infección podrían ser los mejores candidatos para probar nuevas estrategias de tratamiento sin TAR y enfocados a una "cura funcional" (supresión de la replicación viral en ausencia de TAR) o "*esterilización*" (cuando se eliminan los provirus con replicación competente). Por ejemplo, anticuerpos neutralizantes y vacunas adyuvantes entre otras (14).

Hasta que esta cura y remisión del VIH-1 no sea posible, los niños infectados por VIH-1 tendrán que recibir a lo largo de su vida medicación crónica y en ocasiones precisarán de cambios en el tratamiento debido a resistencia virológica o falta de adherencia, dando lugar a eventos de fracaso terapéutico virológico, empeoramiento clínico o incluso el fallecimiento. Además, los problemas sociales, incluyendo adicciones, desempleo o desestructuración de las familias y la falta de recursos de países de altas prevalencias presentan un desafío en el manejo clínico de estos pacientes. En países como Sudáfrica o Mozambique existe una mortalidad infantil significativamente alta en los primeros 3-6 meses desde el inicio al TAR (50).

Existen numerosos estudios que utilizan modelos para predecir el riesgo de mortalidad o de progresión clínica a SIDA. La edad avanzada al diagnóstico, un inicio tardío del TAR, alta carga viral, bajo peso para la edad, bajos linfocitos CD4 y género femenino son los factores que se

han visto asociados tanto a mortalidad como a progresión. Sin embargo, muchos de estos modelos se basan en modelos clásicos multivariables de regresiones logísticas. Estos modelos tienen complejas asunciones como la linealidad entre las interacciones de los predictores y la variable respuesta, y el número limitado de predictores según el tamaño muestral.

Los algoritmos basados en aprendizaje automático o *machine learning* (ML) ofrecen una amplia gama de ventajas. Entre ellas, la habilidad de manejar un gran número de predictores y la combinación de predictores explorando diferentes tipos de asociaciones. A pesar de sus ventajas, la aplicación de ML en el campo del VIH-1 es limitada. Las herramientas que permitan predecir pacientes con un riesgo alto de muerte o progresión clínica pueden alertar a profesionales clínicos menos especializados en países de recursos limitados para mejorar el manejo clínico de estos pacientes.

# Hipótesis

1. Los pacientes tratados durante la primera semana de vida serán capaces de controlar el virus mejor y tendrán una probabilidad mayor de suprimir antes que aquellos que inician el TAR entre el día 8 y 28 de vida.

2. El control de la viremia en las primeras fases de la curva de decremento de carga viral es diferente según el género del paciente también en población pediátrica. Además, un mejor control de la viremia post-TAR está asociado a un menor reservorio viral a largo plazo.

3. La edad al TAR, carga viral al diagnóstico, porcentaje de CD4 y seroprevalencia son factores asociados al tamaño de reservorio. Esta asociación no es lineal, es decir, la tasa del aumento o descenso puede cambiar a medida que la variable cambia, siguiendo un patrón.

4. La implementación de un clasificador basado en aprendizaje automático mejoraría el rendimiento conseguido por métodos tradicionales como la regresión logística. El diseño de una herramienta que incluya este clasificador puede ayudar en el manejo clínico de pacientes en países con recursos limitados.

# Objetivos

El objetivo general de esta memoria es estudiar las características clínicas y virológicas de los pacientes perinatalmente infectados por VIH-1 con un inicio precoz del tratamiento antirretroviral.

1. Determinar el beneficio a largo plazo de iniciar muy precozmente el tratamiento en la primera semana de vida y la asociación entre el tiempo al inicio del TAR y el tiempo hasta la supresión viral (cargas virales menores de 50 copias/mL) en los niños perinatalmente infectados tratados en los primeros 28 meses de vida. (Capítulo I)

2. Determinar si el sexo es un factor influyente en el descenso de carga viral desde el inicio del TAR en niños perinatalmente infectados. Identificar si la velocidad de descenso de la carga viral es un factor predictor del reservorio a largo plazo en los niños perinatalmente infectados con cargas virales suprimidas por el TAR. (Capítulo II)

3. Identificar los factores que puedan tener una influencia en el establecimiento de un reservorio bajo en los niños perinatalmente infectados con cargas virales suprimidas por el TAR. (Capítulo III)

4. Desarrollar una herramienta capaz de predecir mortalidad o progresión al SIDA en niños perinatalmente infectados por VIH-1 para mejorar el manejo clínico de estos pacientes en países de recursos limitados. (Capítulo IV)

# Materiales y Métodos

## Población de estudio

Las poblaciones de estudio sobre las que se basan los capítulos que componen esta memoria se engloban dentro del proyecto EPIICAL (*Early-treated Perinatally HIV-infected Individuals: Improving Children's Actual Life with Novel Immunotherapeutic Strategies*). La finalidad principal de este proyecto es establecer una colaboración científica para liderar tratamientos efectivos para conseguir la remisión del VIH-1 en niños. Este proyecto ha establecido cohortes pediátricas tanto en Europa como en África para diseñar y dirigir estudios en niños con VIH-1. Entre ellas, la cohorte europea multicéntrica transversal CARMA con pacientes pediátricos infectados con VIH-1 precozmente tratados y la cohorte africana multicéntrica prospectiva EARTH con pacientes pediátricos infectados VIH-1 precozmente tratados.

### Población de estudio para el capítulo I

Para realizar este estudio se recogieron datos de forma retrospectiva de 44 pacientes pediátricos de cuatro cohortes (Cohorte de Roma (Italia), Cohorte del Hospital Imperial College de Londres (Reino Unido), Cohorte de la Universidad de Chulalongkorn (Tailandia) y Cohorte Nacional de pacientes pediátricos con infección VIH en España (CoRISpe). Se incluyeron en el estudio aquellos pacientes tratados en los primeros 28 días de vida con al menos dos medidas de carga viral disponible y sin interrupciones de TAR durante al menos dos años de seguimiento. En aquellos niños que recibieron triple terapia incluyendo NVP como profilaxis y continuaron con TAR en los 15 días siguientes, se consideró la fecha de inicio de tratamiento el inicio de la profilaxis (51).

### Población de estudio para el capítulo II

La población de estudio analizada en este capítulo proviene de la cohorte CARMA. CARMA es una cohorte multicéntrica transversal con pacientes pediátricos VIH-1 de 7 centros pediátricos europeos (Inglaterra (3), España (2) e Italia (2)) para la investigación clínica del VIH-1. En esta cohorte están incluidos 40 niños infectados con VIH-1 en seguimiento que comenzaron el TAR en los primeros 24 meses de vida y llegaron a suprimir virológicamente (<400copias/mL) antes del año después de iniciar el TAR y durante al menos 5 años. Se excluyeron aquellos niños coinfectados con hepatitis y/o tuberculosis, con presencia de malignidades o enfermedades concomitantes con

inmunosupresores. Para cada sujeto se recogió información demográfica, datos al diagnóstico, datos del TAR, datos inmunológicos y datos virológicos.

Para este estudio se seleccionaron de esta cohorte aquellos pacientes que tuvieran un descenso consistente (sin fallos virológicos) después de la supresión y tuvieran al menos cinco medidas de carga viral por debajo del límite de detección de las técnicas de medidas de carga viral para ser más específicos en la estimación de los parámetros. Un total de 25 pacientes fueron incluidos.

### Población de estudio para el capítulo III

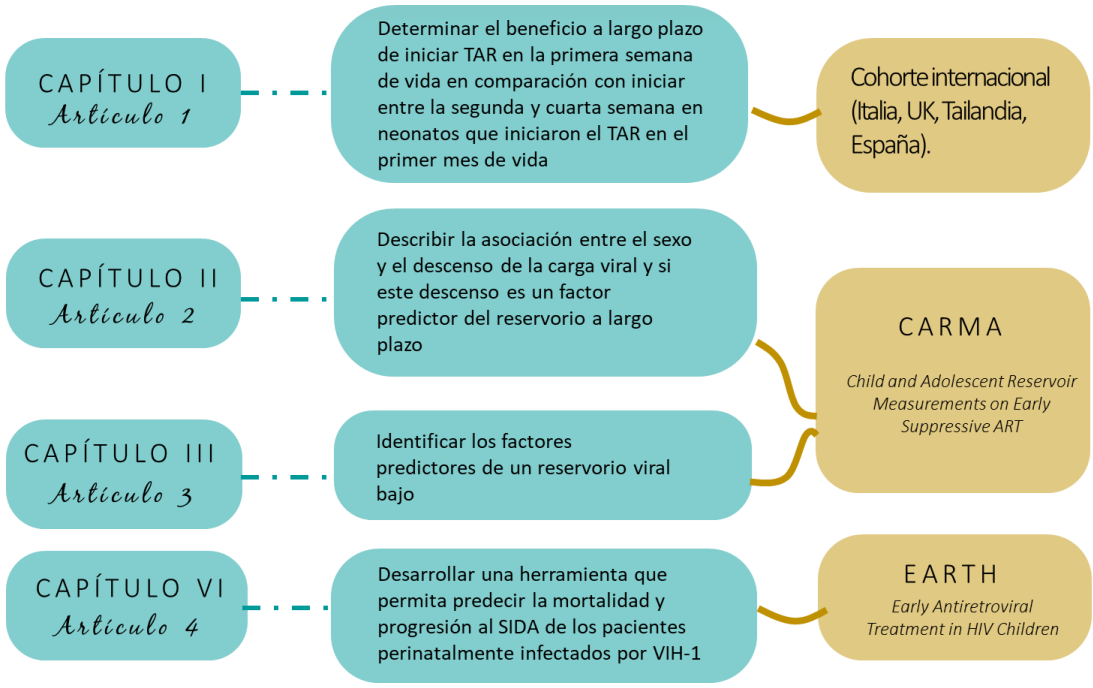
La población de estudio analizada en este capítulo proviene de la cohorte CARMA previamente descrita en el apartado anterior.

El total de los pacientes de esta cohorte (n=40) fueron incluidos en el estudio.

### Población de estudio para el capítulo IV

La población de este estudio proviene de la cohorte EARTH (Early treated Perinatally HIV Infected individuals: Improving Children's Actual Life). EARTH es una cohorte multicéntrica prospectiva que recluta lactantes con infección vertical por VIH-1 de tres centros de Sudáfrica (Africa Health Research Institute (AHRI), Durban; Family Centre For Research With Ubuntu (FAMCRU), Stellenbosch; Perinatal HIV Research Unit (PHRU), Soweto), dos centros de Mozambique (Centro de Investigación en Salud de Manhiça (CISM) – ISGlobal, Manhiça; Fundação Ariel Glaser contra o SIDA Pediatrico, Mozambique) y un centro de Mali (Bamako Hospital, Mali). En esta cohorte, actualmente en reclutamiento activo, están incluidos más de 200 pacientes perinatalmente infectados y tratados en los primeros 3 meses de vida o en los primeros 3 meses después del diagnóstico. Para cada sujeto se recogió información demográfica, datos de la madre, datos al diagnóstico, datos del TAR, datos inmunológicos y datos virológicos. Se incluyeron en este estudio solo aquellos pacientes con datos disponibles de carga viral y porcentaje de CD4 al inicio del estudio.

Figura 3. Resumen de capítulos, objetivos y poblaciones de estudio




## Métodos de laboratorio

### Métodos de laboratorio para el capítulo I

Debido al carácter retrospectivo del estudio, las técnicas de cuantificación de carga viral variaron en el tiempo. Para aquellos datos de cuantificación de carga viral con fecha previa al año 2003, se utilizó *Amplicor HIV-1 monitor; Roche Diagnostic System, Branchburg, NJ* en todos los laboratorios, cuyo límite de detección es 400 copias/mL. A partir del 2003, los laboratorios usaron técnicas ultrasensibles con un límite de detección de entre 20 y 50 copias/mL. Los recuentos de linfocitos T CD4 y CD8 se realizaron mediante citometría de flujo.

### Métodos de laboratorio para el capítulo II y III

ADN total de VIH-1: El reservorio total de ADN del VIH-1 se midió en PBMCs mediante PCR cuantitativa en tiempo real. Para la extracción de ácidos nucleicos se utilizó la plataforma QUIAsymphony usando el DSP Virus/Pathogen Mini Kit de Qiagen acorde al fabricante. Para la cuantificación se generaron curvas estándar con diluciones seriadas de 10. El límite de cuantificación es 10 copias/millon PBMCs y el límite de detección 1 copia/ millón de PBMC. Todas las pruebas fueron realizadas por triplicado.

Para determinar el estadio serológico se utilizó una prueba de antígeno-anticuerpo de cuarta generación (*Abbott ARCHITECT Combo Assay*). Para detectar respuestas específicas de anticuerpo contra diez tipos diferentes de antígeno (gp160, gp120, p66, p55, p51, gp41, p39, p31, p24 y p17), se realizó un *Western blot* (*kit 2.2 Medical Systems, Génova, Italia)* siguiendo las recomendaciones del fabricante. La lectura de las bandas fue analizada usando ImageJ software y la intensidad de cada banda fue calculada para cada antígeno. Aquellas bandas con una intensidad ≥50% a la intensidad del control se le asignó una puntuación de 1, y 0.5 a aquellas inferiores al 50%. La puntuación total de Western blot se calculó sumando el número de respuestas positivas o intermedias (de 0 a 10).

## Análisis estadístico

Todos las figuras y análisis fueron realizados mediante el software R. Las pruebas de hipótesis fueron contrastadas utilizando el nivel de significación al 5%, determinando como significativo aquellos p-valores menores a 0,05.

### Análisis estadísticos empleados en el capítulo I

#### Análisis descriptivo

Para describir las características sociodemográficas, clínicas, virológicas e inmunológicas, se utilizaron frecuencias para las variables categóricas y medianas junto con el rango intercuartílico (RIC Q1-Q3) para las variables continuas. Estas características se compararon en los pacientes tratados antes de la primera semana de vida y en los 8-28 días de vida. Para realizar las comparaciones se utilizaron pruebas chi-cuadrado para frecuencias y U-Mann Whitney para comparar variables numéricas.

#### Variables principales

Las variables principales analizadas en este estudio fueron: el tiempo hasta supresión (dos o más medidas consecutivas de carga viral con menos de 50 copias/mL), tiempo hasta el primer fallo virológico (más de dos medidas consecutivas de más de 400 copias/mL precedidas de supresión), tiempo hasta la primera respuesta subóptima

(más de dos medidas consecutivas de entre 50-400 precedidas y seguidas de supresión), tiempo hasta el primer blip (una única medida de carga viral de entre 50-400 copias/mL precedida y seguida de supresión) y tiempo hasta el primer repunte/spike (una única medida de más de 400 copias precedida y seguida de supresión).

Tiempo hasta evento virológico – Modelos de supervivencia

Para analizar el efecto del tiempo de iniciación al TAR y la probabilidad de cada uno de los eventos virológicos en el tiempo se realizó un análisis de supervivencia flexible semiparamétrico con uso de *splines* censurado por intervalos.

Los modelos de supervivencia describen la probabilidad del tiempo que tarda en producirse un evento aleatorio. Estos modelos se interpretan a partir de la probabilidad de que en un momento en el tiempo el evento de interés ocurra dado que no ha ocurrido aún. Esta probabilidad recibe el nombre de tasa de riesgo o Hazard Ratio.

Los modelos de supervivencia paramétricos son aquellos que hacen predicciones sobre las características de la población a la que la muestra pertenece. Es decir, sobre el riesgo de que el evento ocurra cuando todas las variables independientes sean cero. Los modelos semiparamétricos, por el contrario, no hacen esta asunción sobre la función basal y es estimada a partir de los datos. El modelo más conocido es el modelo de Cox. El modelo de Cox es un modelo semiparamétrico con riesgos proporcionales. Esto significa que asume que los riesgos entre dos individuos cualesquiera de la muestra son constantes a lo largo de su periodo de seguimiento/supervivencia. En otras palabras, se asume que la curva de riesgo sigue el mismo perfil a lo largo del tiempo. Sin embargo, esta asunción podría no ser cierta cuando entendemos que el riesgo del evento puede variar con el tiempo en el conjunto de la muestra. Para ello se pueden utilizar modelos flexibles implementando *splines*. Estos splines o cortes a trozos consisten en dividir el rango del predictor en un número de subintervalos (k). Para cada una de estas divisiones se aplica una función de suavizado o smoothing que hace que los extremos se aproximen. Cuando en un análisis de supervivencia no se conocen los tiempos exactos en los que ha

ocurrido el evento, pero se sabe que se ha producido dentro de un intervalo determinado se aplica la censura por intervalos.

Asociación entre la edad al inicio del TAR y el tiempo hasta la supresión

Para analizar el tipo de asociación entre la edad al inicio del TAR y el tiempo hasta la supresión se realizó un modelo generalizado aditivo mixto. Para el componente fijo se seleccionaron la carga viral basal, % CD4 basales y la edad al inicio del TAR con *splines*. Para el componente aleatorio se incluyó la combinación de TAR al inicio.

Un modelo lineal generalizado es aquel en el que la variable respuesta depende de funciones desconocidas de algunas variables predictoras (funciones de suavizado). El interés de estos modelos se basa en inferir sobre estas funciones de suavizado para conocer el tipo de asociación entre los predictores y las respuestas. Un tipo de función de suavizado son los *splines* que son curvas definidas a trozos mediante polinomios, es decir, se divide el rango de la variable predictora en trozos y sobre cada uno de ellos se ajusta un modelo polinómico. El conjunto de todos estos trozos o *splines* formará una curva. Estos modelos son especialmente interesantes para casos en los que los datos no están relacionados de una manera lineal simple, sino que la relación entre la variable respuesta y predictores se describe de una forma desconocida.

En general los modelos se basan en el análisis de relaciones lineales entre la variable respuesta y ciertas variables predictoras o explicativas. Una de las asunciones de este tipo de modelos es que el error no explicado derivado de esta predicción tiene la propiedad de seguir una distribución homogénea. Es decir, se comete similar error en la predicción en todas las observaciones que componen el modelo. Sin embargo, los modelos mixtos permiten la variabilidad de los errores del modelo. Este tipo de modelos son muy útiles cuando los sujetos parten de una variabilidad aleatoria en diferentes niveles, por ejemplo, regímenes TAR de inicio diferentes.

## Análisis estadísticos empleados en el capítulo II

### Descenso de carga viral

Para analizar la tendencia del descenso de la carga viral en las diferentes fases posteriores al inicio del TAR se realizó un modelo de regresión segmentada. La regresión segmentada o regresión por pedazos es un método de regresión en el que la variable predictora o explicativa es partida en intervalos ajustando en cada intervalo una línea o curva a los datos. Este tipo de regresión es útil cuando la asociación entre la variable independiente y dependiente es diferente en varios segmentos, por ejemplo, cuando queremos determinar un cambio en la tendencia. Aquellos cambios en la tendencia fueron considerados como los puntos de cambio entre las dos fases.

### Asociación entre el descenso de carga viral y las poblaciones celulares

Para modelar el descenso de la carga viral y estimar las diferentes poblaciones celulares, se utilizó el paquete de R ushr (52). Mediante esta herramienta se procesan los datos y mediante un algoritmo se estima la duración de la vida de las poblaciones celulares.

### Asociación entre el descenso de carga viral y el sexo

Para analizar el efecto del sexo en el descenso de la carga viral debido al TAR, se utilizó un modelo de regresión mixto basado en la distribución de Poisson (distribución para variables respuestas basadas en contajes).

En este modelo se incluyó el porcentaje basal de CD4 como variable de confusión y el sexo como término de interacción. Además, se incluyó como intercepta aleatoria, es decir, elemento variable del modelo, el propio identificador del sujeto.

La interacción en estadística ocurre cuando la asociación entre la variable respuesta y una variable explicativa se ve modificada por la presencia de una tercera. Es decir, decimos que existe interacción entre dos factores cuando el efecto de uno de ellos sobre la variable respuesta no es el mismo en todos los niveles del otro factor. Esto significa que la variable explicativa en cuestión y la variable de interacción están correlacionadas entre sí y comparten varianza.

## Análisis estadísticos empleados en el capítulo III

### Variables principales

Las variables principales analizadas en este estudio fueron: tamaño del reservorio (ADN total/mL medido en millón de PBMC), reservorio bajo (≤10 copias/millón de PBMC), tiempo hasta el primer blip (una única medida de carga viral de entre 50-400 copias/mL precedida y seguida de supresión) y tiempo hasta el primer repunte/spike (una única medida de más de 400 copias precedida y seguida de supresión).

### Análisis descriptivo

Para describir las características sociodemográficas, clínicas, virológicas e inmunológicas, se utilizaron frecuencias para las variables categóricas y medianas junto con el RIC para las variables continuas. Estas características se compararon en los pacientes con reservorio ≤10 copias/millón de PBMC *versus* aquellos con >10 copias/millón de PBMC. Para realizar las comparaciones se utilizaron pruebas chi-cuadrado para frecuencias y U-Mann Whitney para comparar variables numéricas.

### Asociación entre factores clínicos, virológicos e inmunológicos y el tamaño del reservorio

Para describir la asociación entre los diferentes factores estudiados y el tamaño de reservorio del VIH-1 se realizó un modelo multivariable aditivo mixto. Para seleccionar las variables se empleó un método por pasos hacia atrás basado en el criterio de información de Akaike. Mediante este método, se seleccionaron en el modelo final aquellas variables cuya eliminación supusiese una pérdida de información en el modelo predictivo.

## Análisis estadísticos empleados en el capítulo IV

### Variables principales

La variable respuesta para este estudio fue mortalidad o progresión clínica a SIDA al año de ser incluidos en el estudio. La progresión clínica a SIDA fue definida como presentar un estadio OMS de SIDA III o IV después de la visita de inicio del estudio. Las variables incluidas como predictoras fueron iguales en el conjunto de algoritmos comparados y fueron: edad al reclutamiento (días), sexo (femenino/masculino), z-score de peso para la edad calculada usando la referencia de crecimiento según la OMS (53), pretérmino (si/no), edad al diagnóstico de VIH (días), edad al TAR (días), régimen de TAR de inicio, carga viral pre-TAR, porcentaje de CD4 pre-TAR, eventos severos en la madre (incluyendo cambio en el trabajo, separación o divorcio, nueva pareja, pérdida del hogar o mudanza o alguna muerte en la familia), eventos de salud en la madre, adherencia al TAR de la madre, porcentaje de CD4 y carga viral materna.

### Generación de los diferentes conjuntos de datos analizados

Todos los participantes incluidos en el estudio (n=100) fueron aleatoriamente divididos en un conjunto de entrenamiento usado para generar los modelos (~70% del total, n=71) y un conjunto de validación para determinar el rendimiento de los modelos (el ~30% restante, n=29). Esta partición fue equilibrada por la variable respuesta (mortalidad y/o progresión).

Modelos

Para cada uno de los seis modelos de machine learning se generó un ajuste automatizado de los parámetros para seleccionar aquellos con mejor rendimiento mediante una validación cruzada de diez iteraciones.

- Regresión Logística: Este algoritmo es considerado una extensión de la regresión lineal, usado para modelar una variable dicotómica (ausencia o presencia de un evento). Con la regresión logística se predice la probabilidad de ocurrencia de un determinado evento.
- Random Forest: El algoritmo *Random Forest* está construido por muchos árboles de decisión de la misma manera que un bosque está formado por números árboles. Para clasificar una nueva muestra, el vector inicial compuesto por las diferentes variables explicativas pasa por todos los árboles de decisión que componen el algoritmo. Cada uno de estos árboles da una clasificación de la respuesta, es decir, un voto. Una vez que todos los votos han sido generados, el algoritmo selecciona la clase más popular de esta mayoría de votos para decidir la clasificación final.
- Support Vector Machine: Esta técnica se basa en la construcción de un hiperplano. Este hiperplano es una delimitación que separa el conjunto de observaciones en función de la variable respuesta en un espacio dimensional con n dimensiones. La distancia entre el hiperplano y la instancia más cercana será la distancia marginal de la clase. El hiperplano óptimo es aquel que maximiza la distancia entre las clases y es capaz de diferenciar mejor al conjunto. Cuando los datos no pueden separarse de forma lineal, el hiperplano se construye mediante funciones no lineales denominadas *kernels*. Unas de las más utilizadas son las funciones *kernel* radiales.
- Naïve Bayes: Este algoritmo se basa en el teorema de Bayes. Este clasificador calcula la probabilidad de un evento basado en los valores observados de las variables predictoras y conocimiento previo de la frecuencia de que estos valores se den cuando el evento ocurre. El algoritmo asume que cada una de las variables predictoras contribuyen independientemente a esa probabilidad. La clasificación final se obtiene combinando la información previa y las

observaciones de los predictores comparando la frecuencia de que ocurra el evento.

- K-nearest neighbor: Este algoritmo es una variante del algoritmo Naïve Bayes. La diferencia es que K-nearest neighbor no se basa en probabilidades previas, sino que asume que eventos similares ocurren según proximidad. La implementación de este algoritmo es precedida por una técnica de conglomerados para determinar la distancia de similitud y luego relacionar la proximidad de un nuevo evento a un evento conocido.
- Artificial neural network: Las redes neuronales artificiales están, como su nombre indica, inspiradas por la actividad neuronal del cerebro. Estos algoritmos se representas por nodos (análogos a neuronas) agrupados en matrices denominadas capas. Diferentes capas se interconectan entre ellas para generar conocimiento como las capas o cortezas cerebrales. Los nodos y extremos (conexiones) tienen pesos ajustados según la frecuencia de estas, permitiendo al modelo reconocer patrones en conjuntos de datos.
- Elastic net: Es un modelo de regresión penalizada optimizada. Estos algoritmos están basados en la regresión lineal simple. Sin embargo, penalizan a los predictores menos informativos mediante la reducción de sus coeficientes o efectos generando un modelo generalizado. Estos modelos son particularmente interesantes cuando hay sobredimensión de predictores.

### Validación de los modelos

Para comparar el rendimiento de los modelos se generó una matriz de confusión para cada uno de ellos determinando su precisión, sensibilidad y especificidad. Se calculó el área bajo la curva (AUC) de cada uno de os modelos tanto aparente (determinada usando el conjunto de entrenamiento) como actual (determinada usando el conjunto de validación).

### Aplicación web

El algoritmo con mejor rendimiento se incluyó en una aplicación web utilizando Shiny R. El código de esta aplicación está disponible en doi: 10.5281/zenodo.8475.

# Capítulo I:

Los neonatos que inician el tratamiento antirretroviral precozmente en los primeros 7 días de vida tienen tiempos de supresión reducidos

## Justificación y Objetivos

El inicio del TAR de forma precoz en los niños verticalmente infectados con VIH-1 tiene un impacto clínico notable. El estudio CHER demostró el beneficio de tratar a los niños lo antes posible desde el momento de diagnóstico con respecto a aplazar el inicio hasta cumplir criterios clínicos (17,24). Asimismo, un tratamiento precoz está significativamente asociado con un tamaño de reservorio viral menor y una mejor reconstitución inmune (19–22).

Existen diversos estudios que muestran un control de la viremia más rápido o tiempo a la supresión menor en aquellos pacientes que inician antes el TAR (15,54,55). Sin embargo, existe controversia en la definición de "precoz", que varía desde los tres a seis meses de vida. Además, la literatura referente al periodo neonatal (7-28 días de vida) es muy escasa. En una revisión reciente sobre población pediátrica precozmente tratada (43), se plantearon diferentes preguntas sobre las que todavía existe poca evidencia. Este capítulo aborda tres de las seis preguntas que plantea esta revisión: ¿cuánto de rápida es la supresión viral en neonatos que inician el TAR muy precozmente?; ¿ es diferente la tasa de supresión en aquellos niños que inician antes el TAR con respecto a los que lo hacen más tarde;? y el control virológico a largo plazo, ¿es diferente en aquellos niños que inician precozmente el TAR.?

Considerando que actualmente el diagnóstico *point-of-care* en las áreas de maternidad es cada vez más frecuente en entornos de alta prevalencia de VIH-1, conocer los beneficios asociados a tratar en los primeros días de vida podría ser de gran relevancia clínica.

Los objetivos de este estudio son:

- Comparar el tiempo hasta la supresión viral en niños perinatalmente infectados con VIH-1 que iniciaron TAR en los primeros 7 días con aquellos que iniciaron entre el día 8-28 de vida.
- Describir la asociación entre el tiempo de inicio del TAR y el tiempo a la supresión en niños tratados entre los días 7 y 28 de vida perinatalmente infectados con VIH-1.

## Resultados

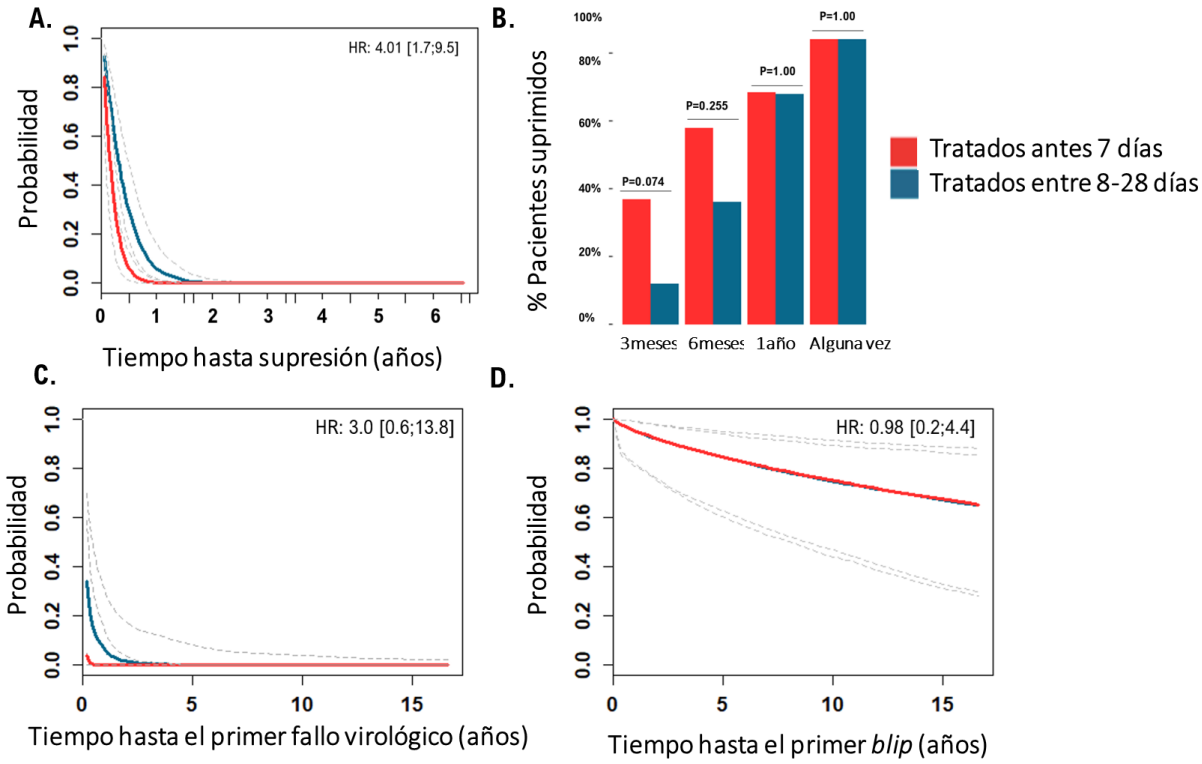
- De los 44 neonatos incluidos en este estudio, 25/44 (57%) fueron niñas y 13/44 (35%) fueron pretérmino. La mediana de peso al nacimiento fue de 2,4 kg (RIC 2,1-2,9). Estos pacientes presentaban una mediana de tiempo en seguimiento de 7,4 años (RIC 2,0-

1,1) con una mediana de tiempo entre visitas con cargas virales disponibles de 3,7 meses (RIC 2,9;5,6).

- La mediana de edad al inicio del TAR fue de 15,5 días (RIC 0-24,2). Un total de 19/44 (43%) iniciaron el TAR en los primeros siete días y 25/44 (57%) entre el día 8 y el 28. Treinta y dos niños recibieron profilaxis, todos ellos continuaron con triple terapia de TAR. La mayoría de los neonatos (30/44, 70%) recibieron como primer TAR un régimen de dos inhibidores de la transcriptasa inversa análogos de nucleósido (ITIAN) junto a nevirapina (NVP) y el 30% restante recibieron dos ITIAN y un inhibidor de la proteasa. El grupo tratado en la primera semana de vida fue más frecuentemente tratado con el régimen 2 ITIAN+NVP que el grupo del tratamiento diferido (94% vs. 52%, p-valor=0,008). La combinación zidovudina (AZT), lamivudina (3TC) y NVP fue la más frecuente (27/44, 63%).
- Al inicio del TAR, la mediana de linfocitos CD4 fue 2.766 células/$mm^3$ (RIC 2.126-3.368) en recuento total y 46% células/$mm^3$ (RIC 42,2-52,5) en porcentaje. La mediana del recuento de linfocitos T CD8 fue de 1.102 células/$mm^3$ (RIC 837; 1491) con una ratio CD4/CD8 de 2.5 (RIC 1,6- 3,1). No se observaron diferencias entre los grupos de estudio. Sin embargo, el grupo que inició en los primeros 7 días presentaba mayor carga viral al momento del TAR que aquel que inició posteriormente (4,4 logaritmos vs. 3,3 logaritmos, p-valor=0,020).
- La mayoría de los niños llegaron a suprimir durante el tiempo de seguimiento (84%) en una mediana de tiempo de 5,6 meses después de recibir el TAR. No se observaron diferencias en la proporción de niños en supresión entre ambos grupos. Del total de niños que suprimieron la carga viral, 10/37 (22%) lo hicieron en los primeros tres meses de vida y 20/37 (45%) en los primeros seis meses de vida.
- Un tercio de los pacientes (12/37) que suprimieron, sufrieron un fallo virológico durante el tiempo de seguimiento, 1/37 con una respuesta subóptima (al menos dos medidas de carga viral de entre 50-400 copias/mL precedidas y seguidas de supresión), 19% con algún *blip* (al menos una única medida de 50-400 copias/mL precedidas y seguidas de supresión*)* y 22% con un *spike*.
- La probabilidad acumulada de alcanzar la supresión virológica en seis y doce meses después del inicio del TAR fue estimada en 49% (RIC, 35- 62,2) y 77% (RIC, 60,7- 88) respectivamente. Aquellos niños tratados en los primeros siete días tuvieron cuatro

veces más probabilidad de suprimir antes que aquellos tratados entre el día 8 y 28 (HR: 4,01 (IC 95%, 1,7;9,5)) según el modelo semiparamétrico flexible con *splines*.

Figura 4. Tiempo a evento virológico acorde al grupo de inicio del TAR




Panel A) Probabilidad de supresión en función del tiempo hasta supresión en años en niños tratados antes de 7 días (rojo) comparado con aquellos tratados entre los 8-28 días (azul).

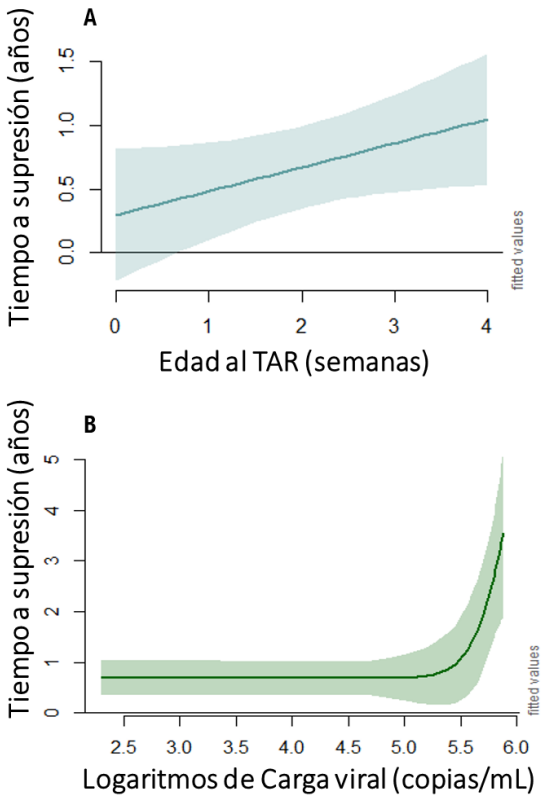
Panel B) Porcentaje de niños en supresión acorde al tiempo de seguimiento en los dos grupos de inicio al TAR. Los p-valores fueron calculados por la prueba exacta de Fisher.

Panel C) Probabilidad de fallo virológico acorde al tiempo hasta el primer fallo virológico (años) en niños tratados antes de 7 días (rojo) comparado con aquellos tratados entre los 8-28 días (azul).

Panel D) Probabilidad de *blip* acorde al tiempo hasta el primer *blip* (años) en niños tratados antes de 7 días (rojo) comparado con aquellos tratados entre los 8-28 días (azul). El suavizado o *smooth* de los modelos aplicados en los paneles A, C y D se realizó mediante un modelo predictivo flexible semiparamétrico usando *splines*. El Hazard Ratio fue ajustado por la carga viral al inicio del TAR y el régimen de TAR al inicio.

- Por cada semana de retraso en el inicio del TAR la probabilidad de supresión disminuyó un 35% (HR 0,65 (IC 95%, 0,46, 0,92)) ajustando por la carga viral y % CD4 pre-TAR.
- No se observaron diferencias significativas entre los grupos respecto a la probabilidad de fallo virológico (HR 0,87 (IC 95%, 0,57,1,32) o presentar algún *blip* (HR 0,98 IC 95%, (0,2, 4,4)).
- El tipo de asociación entre el tiempo hasta supresión y la edad al inicio del TAR fue descrita mediante un modelo generalizado aditivo (ver Anexo). Tanto la edad al inicio del TAR como la carga viral pre-TAR como variable de confusión fueron incluidas en el modelo como efectos fijos y el régimen de TAR como intercepta aleatoria. La carga viral fue introducida en el modelo mediante *splines* con 5 nodos. Se observó una relación lineal entre la edad al inicio del TAR y el tiempo a la supresión viral (coeficiente 0,2 ± 0,1, p-valor=0,048).
- Sin embargo, se observó una relación no-normal entre la carga viral pre-TAR y el tiempo a la supresión. En aquellos niños con cargas virales por debajo de 5 logaritmos, no se observa una asociación con el tiempo a supresión mientras que en aquellos con cargas virales mayores a 5 logaritmos aparece una asociación con pendiente exponencial.

Figura 5. Predicción de la asociación entre el tiempo hasta supresión y edad al TAR y carga viral pre-TAR en niños que suprimen durante el tiempo de seguimiento (n=37)




Panel A) Predicción de la asociación entre la edad al TAR (semanas) y el tiempo a supresión (años). Panel B) Predicción de la asociación entre la carga viral pre-TAR (copias/mL) y el tiempo a supresión (años). Se utilizó un modelo generalizado mixto con edad al inicio del TAR y carga viral como efectos fijos y régimen de TAR al inicio como intercepta aleatoria.

CLINICAL SCIENCE

# Reduced Time to Suppression Among Neonates With HIV Initiating Antiretroviral Therapy Within 7 Days After Birth

*Sara Domínguez-Rodríguez, MSc,[a,b,c] Alfredo Tagarro, PhD,[a,b,c] Paolo Palma, PhD,[d] Caroline Foster, PhD,[e] Thanyawee Puthanakit, PhD,[f,g] Thidarat Jupimai, PhD,[f,g] Nicola Cotugno, PhD,[d] Jintanat Ananworanich, PhD,[h,i,j] Paola Zangari, PhD,[d] Eleni Nastouli, PhD,[k] María Ángeles Muñoz-Fernández, PhD,[l] María Luisa Navarro, PhD,[b,m] Carlo Giaquinto, PhD,[n] Paolo Rossi, PhD,[d] Louise Kuhn, PhD,[o] and Pablo Rojo, PhD,[a,b] on behalf of the EPIICAL Consortium*

**Abstract:** There are limited data on infants with HIV starting antiretroviral therapy (ART) in the neonatal period. We investigated the association between the timing of ART initiation and time-to-suppression among infants who tested HIV-positive and initiated ART within the first 28 days of life. The effect was estimated using cumulative probability flexible parametric spline models and a multivariable generalized additive mixed model was performed to test nonlinear associations. Forty-four neonates were included. Nineteen (43.2%) initiated ART within 7 days of life and 25 (56.8%) from 8 to 28 days. Infants treated within 7 days were 4-fold more likely to suppress earlier than those treated after 7 days [Hazard ratio (HR) 4.01 (1.7–9.5)]. For each week the ART initiation was delayed, the probability of suppression decreased by 35% (HR 0.65 [0.46–0.92]). Age at ART start was linearly associated with time-to-suppression. However, a linear association with normally distributed residuals was not found between baseline viral load and time-to-suppression, with no association found when baseline viral loads were ≤5 log(10) copies/mL, but with exponential increase in time-to-suppression with > log5 copies/mL at baseline. Starting ART within 7 days of life led to 4-fold faster time to viral suppression, in comparison to initiation from 8 to 28 days.

**Key Words:** HIV, infants, viral load, viral suppression, neonates, newborns, early treated

(*J Acquir Immune Defic Syndr* 2019;82:483–490)

Received for publication July 5, 2019; accepted September 2, 2019.

From the [a]Pediatric Infectious Diseases Unit, Fundación para la Investigación Biomédica del Hospital 12 de Octubre, Universidad Complutense, Madrid, Spain; [b]Spanish National HIV-1 Paediatric Network (CoRISPe); [c]Universidad Europea de Madrid, Madrid, Spain; [d]Research Unit in Congenital and Perinatal Infections, Division of Immunology and Infectious Diseases, Academic Department of Pediatrics (DPUO), Bambino Gesù Children's Hospital, Rome, Italy; [e]Imperial College Healthcare NHS Trust, London, United Kingdom; [f]Department of Pediatrics, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand; [g]Center of Excellence for Pediatric Infectious Diseases and Vaccines, Faculty of Medicine, Chulalongkorn University, Bangkok, Thailand; [h]SEARCH, The Thai Red Cross AIDS Research Centre, Bangkok, Thailand; [i]Henry M. Jackson Foundation for the Advancement of Military Medicine, Bethesda, MD; [j]Department of Global Health, University of Amsterdam, Amsterdam, the Netherlands; [k]Department of Population, Policy and Practice, UCL GOS Institute of Child Health, London, United Kingdom; [l]Immuno-Molecular Laboratory, Immunology Department, Hospital General Universitario Gregorio Marañón (HGM), Instituto de Investigación Sanitaria Gregorio Marañón, CoRISpe, Spanish HIV HGM BioBank, Madrid, Spain; [m]Hospital General Universitario Gregorio Marañón, Instituto de Investigación Sanitaria Gregorio Marañón, (IISGM), CoRISpe, Spain. Universidad Complutense, Madrid, Spain; [n]PENTA Onlus, Padova, Italy; and [o]Gertrude H. Sergievsky Center, College of Physicians and Surgeons, Columbia University Irving Medical Center, New York, NY.

This work is part of the EPIICAL project (http://www.epiical.org/), supported by PENTA-ID foundation (http://penta-id.org/), funded through an independent grant by ViiV Healthcare UK. The work in Thailand was partially supported by the US NIAID (R01AI114236), the Thailand MOPH-US CDC Collaboration and the Active Case Management network. J.A. has received honoraria for participating in advisory meetings for ViiV Healthcare, Gilead, Merck, Roche, and AbbVie.

Presented in CROI 2019; March 4–7, 2019; Seattle, WA.

The authors have no conflicts of interest to disclose.

A.T., S.D.-R., and P.R. conceptualized design the study. S.D.-R., A.T., and L.K. participated in drafting the manuscript and the interpretation of the results. S.D.-R. performed statistical analysis. All co-authors participated in discussions about the interpretation of findings, were involved in the preparation and critical review of the final manuscript.

S.D.-R. and A.T. have contributed equally.

The content of this manuscript is solely the responsibility of the authors and does not necessarily represent the official views of any of the institutions mentioned above, the U.S. Department of the Army or the U.S. Department of Defence. The investigators have adhered to the policies for protection of human subjects as prescribed in AR-70.

Correspondence to: Alfredo Tagarro, PhD, Pediatric Infectious Diseases Unit, Fundación para la Investigación Biomédica del Hospital 12 de Octubre Av. de Córdoba, s/n, 28041 Madrid, Spain (e-mail: alfredotagarro@gmail.com).

Copyright © 2019 Wolters Kluwer Health, Inc. All rights reserved.

## INTRODUCTION

Early combination antiretroviral therapy (ART) initiation in vertically infected infants with HIV has had a major impact on clinical outcomes. The landmark Infants with HIV Early anti-Retroviral (CHER) study clearly demonstrated the benefits of starting ART as soon as possible after infant diagnosis compared to deferring ART until clinical criteria are met.[1,2] Moreover, early initiation of ART leads to a significantly smaller reservoir size and a better immune reconstitution.[3–9] Early ART may affect HIV pathogenesis in favorable ways to help achieve future HIV remission in some individuals.[10,11]

Most studies show faster time to viral suppression and longer sustained viral control when compared with treatment that started later.[5,12–15] However, definitions of early vary a great deal, with early ART frequently including infants started in the first 3–6 months of life. Literature reporting on ART initiation in the first year of life rarely reports on



Copyright © 2019 Wolters Kluwer Health, Inc. Unauthorized reproduction of this article is prohibited.

outcomes when treatment is initiated in the neonatal period (the first 28 days of life).[12]

A recent review described individual viral trajectories in a handful of infants starting ART in the neonatal period.[12] However, numbers of cases were too few to distinguish outcomes by the timing of ART initiation within this very early period and might have been biased toward the over-representation of infants who actually survived to be reported retrospectively.

Currently, point-of-care diagnosis in maternity wards is increasingly available in high prevalence settings.[16,17] With increasing access to very early ART initiation, there is a crucial need to understand just how rapidly treatment should be started within the neonatal period, and whether very early initiation within 7 days of life has any long-term benefit. Here, we investigated the association between the timing of ART initiation and time to viral suppression among HIV-perinatally infected neonates treated in the first 28 days of life.

## METHODS

### Study Population

This is a retrospective analysis of longitudinal data pooled from 4 cohorts: Rome Cohort (Italy),[18] Imperial College Hospital (London, United Kingdom),[19] Chulalongkorn University (Thailand),[20] and Spanish National HIV Paediatric Network-CoRISpe (Spain).[21] Each cohort has established and maintained standardized databases for pediatric HIV for many years. The analysis was approved by the Ethical Committee of Hospital 12 de Octubre (17/398).

Perinatally HIV-infected infants aged ≤28 days at ART initiation, with at least 2 viral load (VL) measurements and without interruptions in ART within 2 years of follow-up were included. Those infants who received triple-drug postnatal prophylaxis including nevirapine and subsequently transitioned to treatment doses of ART within 15 days were considered for the purpose of analysis as starting ART from date of prophylaxis initiation.[22] Intrapartum prophylaxis was defined as intravenous intrapartum AZT administered to the mother during delivery.

### Endpoints and Secondary Variables

The primary endpoints were mortality and time to VL suppression up to 15 years of follow-up. Other virological control markers were also analyzed: virological failure, suboptimal response, blips, and spikes. Definitions were as follows: suppression was ≥2 consecutive measurements ≤50 copies/mL; virological failure was ≥2 consecutive measurements >400 copies/mL preceded by suppression; suboptimal viral response was ≥2 consecutive measurements 50–400 copies/mL preceded and followed by suppression; blips were a single measurement 50–400 copies/mL preceded and followed by suppression, and spikes were single measurements >400 copies/mL preceded and followed by suppression.

Baseline (defined as the closest measurement to start of ART) sociodemographic, clinical, virological, and immunological features, were included in the analysis. Baseline VL and CD4 measurements were restricted to those on the day of ART start or before. VL measurements, CD4%, CD8%, CD4:CD8 ratio, and lymphocyte counts were routinely collected at each follow-up visit.

### Laboratory Methods

Before 2003, standard HIV RNA testing (Amplicor HIV-1 monitor; Roche Diagnostic Systems, Branchburg, NJ) was used in all laboratories, with a routine lower limit of detection (LOD) of <400 copies/mL and ultrasensitive LOD of <50 copies/mL. Since 2003, participating laboratories used ultrasensitive measurements with a LOD of 20–50 copies/mL. CD4 and CD8 analyses were performed by flow cytometry at each site.

### Statistical Analysis

Baseline characteristics, including ART initiation date, were described with frequencies for categorical variables and means (SD) or medians [interquartile range (IQR)] for continuous variables, both in the total population and stratified by timing of ART initiation in ≤7 days and >7 days after birth. $\chi^2$ or Fisher tests were applied to assess differences across groups for categorical variables and Student *t* test or Mann–Whitney *U* tests for continuous variables.

For time-to-event-analysis, the main effect of ART initiation on time to each virologic event (suppression, virological failure, blip, and suboptimal response) was estimated by cumulative probability flexible parametric spline models with an interval-censored method. This approach, implemented in survival and flexsurv R package,[23] allows to model nonproportional hazards, a feature that is particularly useful in cases where the given outcome variable may have different effect during the time to follow-up. Moreover, this model also allows a built-in-choice to include splines so that it is possible to handle nonlinear assumptions. Hazard ratios (HRs) and 95% confidence intervals were estimated either for treatment groups or age at ART initiation as a continuous variable. A univariable and multivariable model were designed. The multivariable model was adjusted by baseline VL and ART regimen at treatment initiation. Variable selection for the multivariable model was performed with a stepwise Akaike Information Criterion variable selection. The site/cohort was not selected as confounder according to this criterion.

To test the type of association between age at ART initiation and time-to-suppression, only infants who ever suppressed in the follow-up were selected. To analyze possible nonlinear associations, a multivariable generalized additive mixed model implemented in *gamm4* package[24] was performed. Fixed component of the model was built including baseline VL, baseline CD4, and age at ART initiation with smoothing splines. To model individual treatment differences that provide variability, ART drug combination at initiation was included as a random intercept according to prior clinical knowledge. Backward stepwise elimination was applied to reach the final multivariable model. In those variables with significant smoothing terms, knots were selected according to a smoothing selection optimization implemented in mgcv R



Copyright © 2019 Wolters Kluwer Health, Inc. All rights reserved.

Copyright © 2019 Wolters Kluwer Health, Inc. Unauthorized reproduction of this article is prohibited.

**TABLE 1.** Baseline Sociodemographic, Viral, and Immunologic Characteristics of 44 Children Starting ART Within 7 Days vs. Those Starting ART From 8 to 28 of Life

| | Overall N = 44 | ART within 7 days N = 19 | ART after 7 days N = 25 | *P** | Odd Ratio (95% CI) | *P*† |
|---|---|---|---|---|---|---|
| Gender: | | | | 0.295 | | |
| Male | 19 (43.2%) | 6 (31.6%) | 13 (52.0%) | | 0.44 (0.12 to 1.51) | 0.195 |
| Region | | | | **0.036** | | |
| Spain | 23 (52.3%) | 11 (57.9%) | 12 (48.0%) | | | |
| Italy | 3 (6.82%) | 2 (10.5%) | 1 (4.00%) | | | |
| United Kingdom | 5 (11.4%) | 4 (21.1%) | 1 (4.00%) | | | |
| Thailand | 13 (29.5%) | 2 (10.5%) | 11 (44.0%) | | | |
| Birth weight | | | | | | |
| Median (IQR) | 2440 (2130; 2922) | 2330 (1300; 2720) | 2515 (2258; 2945) | 0.157 | 1.00 (1.00; 1.00) | **0.046** |
| Preterm birth | | | | 1.000 | | |
| No | 24 (64.9%) | 11 (64.7%) | 13 (65.0%) | | Ref. | |
| Yes | 13 (35.1%) | 6 (35.3%) | 7 (35.0%) | | 1.01 (0.25 to 4.08) | 0.984 |
| Intrapartum prophylaxis | | | | 1.000 | | |
| No | 11 (32.4%) | 6 (33.3%) | 5 (31.2%) | | Ref. | |
| Yes | 23 (67.6%) | 12 (66.7%) | 11 (68.8%) | | 0.91 (0.20 to 4.03) | 0.906 |
| Prophylaxis type | | | | 0.104 | | |
| No | 7 (15.9%) | 4 (21.1%) | 3 (12.0%) | | | |
| Monotherapy | 5 (11.4%) | 0 (0.00%) | 5 (20.0%) | | | |
| Triple therapy | 32 (72.7%) | 15 (78.9%) | 17 (68.0%) | | | |
| Age at first ART *(d)* | | | | **<0.001** | | |
| Median (IQR) | 15.5 (0.00; 24.2) | 0.00 (0.00; 2.50) | 23.0 (19.0; 27.0) | | | |
| ART regimen at initiation | | | **0.008** | | | |
| 2 NRTI+1 NVP | 30 (69.8%) | 17 (94.4%) | 13 (52.0%) | | Ref. | |
| 2 NRTI+1PI | 13 (30.2%) | 1 (5.56%) | 12 (48.0%) | | **0.08 (0.00 to 0.46)** | **0.003** |
| ART at initiation | | | | **0.023** | | |
| 3TC+ABV+NVP | 3 (6.98%) | 1 (5.56%) | 2 (8.00%) | | | |
| 3TC+AZT+NVP | 27 (62.8%) | **16 (88.9%)** | 11 (44.0%) | | | |
| 3TC+AZT+LPVr | 7 (16.3%) | 0 (0.00%) | **7 (28.0%)** | | | |
| Other | 7 (15.9%) | 2 (10.5%) | 5 (20.0%) | | | |
| Baseline CD4 count | | | | 0.986 | | |
| Median (IQR) | 2766 (2126; 3368) | 2810 (2185; 3245) | 2540 (2106; 4179) | | 0.98 (0.71; 1.36) | 0.933 |
| Baseline CD8 count | | | | 0.589 | | |
| Median (IQR) | 1102 (837; 1491) | 1050 (854; 1265) | 1131 (844; 1792) | | 1.00 (1.00; 1.00) | 0.501 |
| Baseline CD4/CD8 | | | | 0.423 | | |
| Median (IQR) | 2.5 (1.6; 3.1) | 2.6 (1.7 to 3.0) | 2.3 (1.5; 3.1) | | 1.17 (0.84; 1.62) | 0.323 |
| Baseline viral load | | | | **0.020** | | |
| Median (IQR) | 17,000 (962; 174,882) | 26,938 (14,785; 259,276) | 1970 (782; 26,840) | | **9.2 (1.5 to 56.7)** | **0.018** |
| log(10) Median (IQR) | 4.2 (2.9; 5.2) | 4.4 (4.2; 5.4) | 3.3 (2.9; 4.4) | | | |

Statistical significant values are provided in bold.
**P*-value for the distribution test (either continuous *U*-Mann Whitney or categorical $\chi^2$).
†*P*-value for the odds ratio (Wald-test).

package,[25] computing an estimate of residual variance based on differing residuals that are near neighbors according to the (numeric) covariates of the smooth. Plots were built with the ggplot2 R package[26] and itsadug R package.[27] R software was used for all analysis.[28]

## RESULTS

### Participants Characteristics

A total of 44 neonates born between 2000 and 2017 were included in this study. Of these newborns, 25/44 (57%) were female and 13/44 (35%) were preterm defined as <37 gestational weeks. Median birth weight was 2440 grams (IQR 2130; 2922 g). The median duration of follow-up was 7.4 (IQR 2.0; 11.1) years and the median time between VL checks was 3.7 (2.9–5.6) months. No infant died during follow-up. Median age at ART initiation was 15.5 (IQR 0.0; 24.2) days. Nineteen (43%) infants initiated ART at age ≤ 7 days and 25 (60%) initiated ART from 8 to 28 days. Thirty-two infants received prophylaxis, all of them started triple ART therapy after prophylaxis. Most neonates (30/44, 70%) received their first ART as 2 nucleoside reverse transcriptase

Copyright © 2019 Wolters Kluwer Health, Inc. All rights reserved.



Copyright © 2019 Wolters Kluwer Health, Inc. Unauthorized reproduction of this article is prohibited.

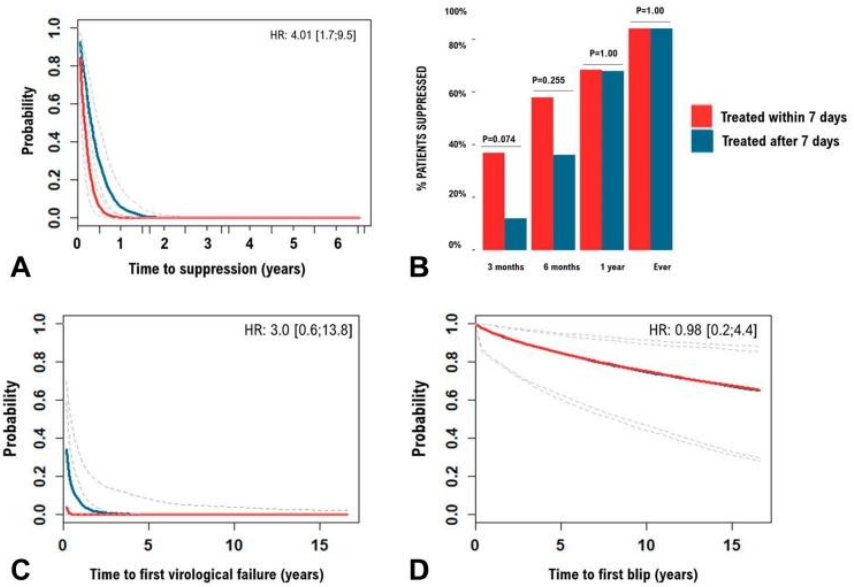



**FIGURE 1.** Time to virological event analysis according to ART timing groups. A, Probability of suppression according to the time-to-suppression in years in infants treated within less than 7 days (red) and those treated from 8 to 28 days (blue). Smoothed lines were predicted according to a semiparametric flexible spline model. Hazard Ratio was adjusted by baseline VL and ART regimen at initiation. B, Percentage of suppressed infants according to the time of follow-up in both ART timing groups. *P*-values were calculated by Fisher exact tests. C, Probability of virological failure according to time-to-first-virological failure, in years, in infants treated within 7 days (red) and those treated from 8 to 28 days (blue). Smoothed lines were predicted according to a semiparametric flexible spline model adjusted by baseline VL and ART regimen at initiation. D, Probability of blip according to time-to-first-blip in years in infants treated within less than 7 days (red) and those treated at 8–28 days (blue). Smoothed lines were predicted according to a semiparametric flexible spline model adjusted by baseline VL and ART regimen at initiation.

inhibitors (NRTIs) plus NVP and 30% received NRTIs and protease inhibitors. The group treated within 7 days of life was treated initially more often with NRTIs and NNRTIs than the group treated after 7 days (94% vs. 52%, $P = 0.008$). The combination AZT + 3 TC and NVP was the most frequent regimen (27/44; 63%) (Table 1). Infants required ART switches a median of 2.5 times with no differences between groups of age at ART initiation ($P = 0.219$).

At ART initiation, the median CD4 count was 2766 (IQR 2126; 3368) cells/mm$^3$, CD4 percentage was 46% (IQR 42.2; 52.5), median CD8 count was 1102 (IQR 837; 1491) cells/mm$^3$, and the median CD4/CD8 ratio was 2.5 (IQR 1.6; 3.1). No differences were found between groups (Table 1). Median log(10) baseline VL was 4.2 (IQR 2.9; 5.2). The group initiating ART within 7 days had a higher baseline VL than the group initiating ART after 7 days (4.4 vs. 3.3, $P = 0.020$).

Overall, 37/44 (84%) infants suppressed during the follow-up period at a median time of 5.6 months after receiving ART. No differences in the proportions were found between participants treated within 7 days and after 7 days (84% vs. 84%, $P = 1.00$). Among infants who suppressed, 10/37 (22%) suppressed in the first 3 months after ART initiation, 20/37 (45%) in the first 6 months, and 30/37 (68%) in the first year. Infants spent almost 68% (IQR 36.2; 91.5) of the follow-up time suppressed. One-third of infants (12/37) with prior viral suppression experienced subsequent virological failure during the follow-up, 1/37 (3%) patients had a suboptimal viral response, 7/37 (19%) presented any blip, and 8/37 (22%) had a spike. There were no significant differences between groups in virological failure, suboptimal response, blip, or spikes (Table 2).

## Time to Event Analysis

Using the flexible parametric spline model, the cumulative probability of achieving virological suppression by 6 and 12 months after ART initiation was estimated at 49% (35–62.2) and 77% (60.7–88) respectively. Infants treated within 7 days were 4-fold more likely to suppress earlier than those treated after 7 days [HR 4.01 (1.7–9.5)] (Fig. 1A). Most infants of both groups reached suppression after 1 year of follow-up (Fig. 1B).

When age at ART initiation was introduced in the model as a continuous variable, age at ART had a HR of 0.65 (0.46–0.92) after adjustment for pre-ART VL and ART regimen. In other words, for each week the ART initiation was delayed, the probability of suppression decreased by 35%.

No significant association was found between the probability of first virological failure and the age at ART initiation groups [HR: 3.0 (0.6–13.8)] (Fig. 1C), even when introducing age at ART initiation as a continuous variable [HR: 0.87 (0.57–1.32)]. Likewise, no significant association was found between the probability of first blip and the age at ART initiation groups [HR: 0.98 (0.2–4.4)] (Fig. 1D), or as



Copyright © 2019 Wolters Kluwer Health, Inc. All rights reserved.

Copyright © 2019 Wolters Kluwer Health, Inc. Unauthorized reproduction of this article is prohibited.

**TABLE 2.** Virologic Features of Participants During Follow-up

| | Overall N = 44 | ART within 7 days N = 19 | ART after 7 days N = 25 | P* | Odd Ratio (95% CI) |
|---|---|---|---|---|---|
| Follow-up time (yrs) | | | | 0.749 | |
| Median (min–max) | 7.37 (1.91 to 11.1) | 8.31 (2.61 to 10.8) | 6.55 (1.45 to 11.3) | | 1.01 (0.90 to 1.13) |
| Have ever been suppressed ≤50 cp/mL | | | | 1.00 | |
| No | 7 (15.9%) | 3 (15.8%) | 4 (16.0%) | | Ref. |
| Yes | 37 (84.1%) | 16 (84.2%) | 21 (84%) | | 1.01 (0.18 to 6.1) |
| Time to reach suppression, months | | | | 0.053 | |
| Median (min–max) | 5.6 (2.43 to 11.0) | 3.75 (1.6 to 6.4) | 7.6 (3.6 to 10.8) | | 0.91 (0.8 to 0.99) |
| % Time suppressed | | | | 0.506 | |
| Median (min–max) | 67.8 (36.2 to 91.5) | 82.7 (46.3 to 91.3) | 59.4 (31.6 to 91.9) | | 1.01 (0.99 to 1.02) |
| Any suboptimal viral response ≥2 consecutive VL≥ 50–400 | | | | 1.00 | |
| No | 36 (97.3%) | 16 (100%) | 20 (95.2%) | | |
| Yes | 1 (2.70%) | 0 (0.00%) | 1 (4.76%) | | |
| No. of suboptimal viral responses | | | | 0.553 | |
| Mean (SD) | 0.73 (1.49) | 0.50 (1.07) | 0.86 (1.70) | | 0.82 (0.41 to 1.66) |
| Blips Single VL≥ 50–400 | | | | 1.00 | |
| No | 30 (81.1%) | 13 (81.2%) | 17 (81.0%) | | Ref. |
| Yes | 7 (18.9%) | 3 (18.8%) | 4 (19.0%) | | 0.99 (0.16 to 5.57) |
| No. of blips/years of follow-up | | | | 0.650 | |
| Mean (SD) | 0.22 (0.2) | 0.18 (0.05) | 0.26 (0.3) | | |
| Spikes Single VL≥ 400 | | | | 0.254 | |
| No | 29 (78.4%) | 11 (68.8%) | 18 (85.7%) | | Ref. |
| Yes | 8 (21.6%) | 5 (31.2%) | 3 (14.3%) | | 2.61 (0.51 to 15.9) |
| No. of spikes/years of follow-up | | | | 0.476 | |
| Mean (SD) | 0.1 (0.02) | 0.1 (0.03) | 0.1 (0.02) | | |
| Virological failure | | | | 1.000 | |
| No | 23 (65.7%) | 11 (68.8%) | 12 (63.2%) | | Ref. |
| Yes | 12 (34.3%) | 5 (31.2%) | 7 (36.8%) | | 0.79 (0.18 to 3.31) |
| Time to first virological failure | | | | 0.323 | |
| Median (min–max) | 2.40 (1.01 to 9.61) | 2.81 (1.65 to 8.66) | 1.32 (0.62 to 9.07) | | 1.01 (0.89 to 1.16) |
| No. of virological failures/years of follow-up | | | | 0.905 | |
| Mean (SD) | 0.14 (0.1) | 0.14 (0.1) | 0.14 (0.1) | | 1.0 (1.0 to 1.0) |

*P-value for the distribution test (either continuous U-Mann Whitney or categorical $\chi^2$).

a continuous variable [HR: 1.03 (0.62–1.72)]. Because few numbers of events were observed for a suboptimal response (n = 1) or spike (n = 8) in this population, there was an insufficient sample size to explore an association.

### Association Between Time to Suppression and Age at ART Initiation in Infants Who Suppressed

The type of association between time to suppression and age at ART was tested in those infants who suppressed (n = 37) using a generalized additive mixed model. Age at ART and baseline VL as a confounder were included as fixed effects and ART regimen as random intercepts. VL was smoothed by splines with a significant smooth term ($P$ = 0.035) with 5 knots according to the smoothness selection.

Evaluating associations between age at ART initiation in weeks and time to viral suppression a normal linear relationship was noted [β = 0.2 (SD 0.1); $P$ = 0.048] (Fig. 2). Whereas a non-normally linear association was detected in the association between log VL at baseline and time to VL suppression. Specifically, for infants with a baseline VL below 5 log(10), no association was noted with time to viral suppression. However, when baseline VL was at or above 5 log(10), a strong effect with an exponential slope was observed.

## DISCUSSION

In newborns who started ART within the first 28 days of life, those who started ART within 7 days of life had 4 times higher chance of viral suppression than those who started at age from 8 to 28 days. The probability of achieving suppression decreased by 35% for each week ART initiation was delayed in the neonatal period, after adjustment for baseline VL and ART regimens at initiation. These significant results were mainly driven by the very first months of follow-up. Eventually, there was a similar proportion of suppressed infants after 1 year of follow-up regardless of when ART was

Copyright © 2019 Wolters Kluwer Health, Inc. All rights reserved.



Copyright © 2019 Wolters Kluwer Health, Inc. Unauthorized reproduction of this article is prohibited.

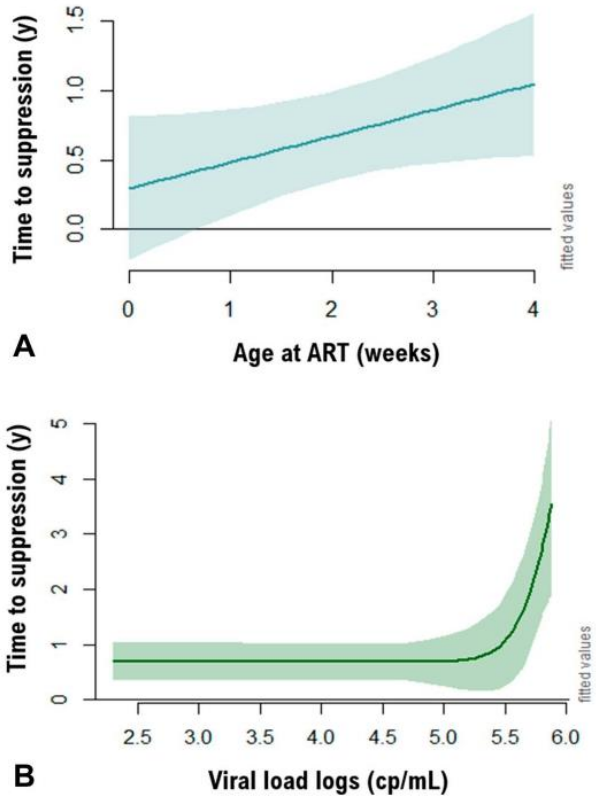



**FIGURE 2.** Prediction of the association between time to suppression and age at ART initiation or baseline viral load in infants who suppressed during the follow-up (n = 37). Association prediction between age at ART initiation (A) and baseline viral load (B) and time to suppression in a total of 37 infants. Generalized additive mixed model with age at ART initiation and baseline viral load as fixed effects and ART regimen at initiation as random effects.

started. This implies that there are other factors driving time to virologic suppression beyond the age at ART initiation.

Other virological endpoints examined, including ever achieving viral suppression and markers of lack of sustained viral control after suppression did not differ between the 2 ages at ART start groups. Current clinical practice generally favors rapid initiation of ART, sometimes based on a presumptive diagnosis, and our results support this approach. We hypothesize that a more rapid decline in VL may have a beneficial effect on the size of the viral reservoir. Age at which suppression is attained, which is linked but independent of age at ART initiation, has been shown to be an important predictor of the size of the viral reservoir.[29] The other virological endpoints we examined may be less sensitive markers of these important early viral dynamics.

The virological endpoints we observed in this multi-site cohort of perinatally-infected neonates in Europe and Thailand are different than has been reported in other cohorts of very early-treated infants. In a Canadian cohort, 4 of 12 neonates treated within 72 hours of birth achieved sustained virological control.[30] In a South African cohort, only half of the 22 HIV-infected neonates treated remained in care and only 2 of these had a VL <50 copies when last seen.[31] No deaths were observed in our cohort which is encouraging. However, we selected only infants with at least 2 VL measurements which are likely to be biased toward those surviving and continuing to engage in care for longer.

The association between age at ART initiation, baseline CD4, and baseline VL with time-to-suppression has been previously described.[32–38] However, most of the studies have assumed that these dependent variables followed a normal linear distribution by applying conventional linear or logistic regressions to test this relationship, which could provide some loss of interpretability. The use of smoothing functions fitting relaxed assumptions allowed us to identify 2 different associations in this population. First, we identified a positive linear association between time to suppression and age at ART initiation. Second, the analysis showed a novel finding in the relationship between time to suppression and baseline VL. The dynamics of this relationship suggest that below 5 log(10), baseline VL has no influence in time to suppression. In these infants, time to treatment is the only factor we found influencing time to suppression. However, when baseline VL is over 5 log(10), an exponential linear relationship is observed, suggesting a strong influence in time to suppression. In these infants, not only the age at ART initiation is making a difference in time to suppression, but also the VL. In the study conducted by Frigati et al,[31] baseline VL of the 22 participants was 4.444 copies/mL. In our study, VL was higher, especially because of the 19 patients who initiated earlier. The range was also wider, as well as in a South African study,[39] reaching >6 logs. The reason for these differences is unclear and may be related with the sample size. In any case, very high VL seems to have an influence in time to suppression and should be considered in models. One may wonder whether these newborns with high VL and longer time to suppression are those infected in utero or may have a more immature immune function. Further studies are warranted to evaluate this possibility. Our study has several limitations. Data are retrieved from an observational cohort and the age at ART initiation is dependent on a range of clinical and social factors beyond our control. These include, for example, when the HIV-infected mother came to clinical attention, what infant diagnosis programs and treatment recommendations were in place at the time, how the mother engaged with the health service and adhered to medical advice, etc. Because of the nature of the databases, we also have some missing data and very limited information on these factors including markers of the mother's health and the timing of infant diagnosis. The assessment of suppression is based on having 2 VL measurements available raising concern as to whether this frequency of measurement can truly reflect the extent of virologic suppression. We cannot rule out that the faster time to suppression observed in the earlier treated infants is because of confounding factors and not the timing of ART. However, on the parameters on which we have data, we did not observe significant differences between neonates starting ART within the first 7 days of life



Copyright © 2019 Wolters Kluwer Health, Inc. All rights reserved.

Copyright © 2019 Wolters Kluwer Health, Inc. Unauthorized reproduction of this article is prohibited.

and those starting later, except for baseline ART VL. Interestingly, baseline ART VL was higher in those starting ART earlier than in those starting slightly later, which may reflect the dynamics of primary infection.[40] This potential bias was addressed adjusting by baseline VL and ART regimens at initiation.

In conclusion, our data support the potential benefits of starting ART within 7 days of birth even in comparison to starting from 8 to 28 days after birth. We observed a faster time to viral suppression, which may result in a favorable impact on the viral reservoir. Our results lend support to the clinical guidance to implement universal birth testing to diagnose intrauterine-infected neonates as early as possible and to clinical guidance that encourages rapid initiation of ART either presumptively or at diagnosis.

## ACKNOWLEDGMENTS

*The authors thank all the infants, families, and all who care for the infants. Other collaborators include attending physicians in charge of participants included. Spain: Pilar Collado (Hospital Clínico Universitario Lozano Blesa, Zaragoza, Spain), Claudia Fortuny (Hospital Sant Joan de Dèu, Barcelona, Spain), Antonio Mur (Hospital del Mar, Barcelona, Spain), David Moreno (Hospital Carlos Haya, Málaga, Spain), Miguel Ángel Roa (Hospital de Móstoles, Madrid, Spain), Raquel Angulo (Hospital de El Ejido, Almería, Spain), María José Mellado (Hospital Universitario La Paz, Madrid, Spain), María Penín (Hospital Príncipe de Asturias, Madrid, Spain), César Gavilán (Hospital San Juan de Alicante, Alicante, Spain), María Méndez (Hospital Germans Trias i Pujol, Barcelona, Spain), and Pere Soler-Palacín (Hospital Vall d'Hebron, Barcelona, Spain). Data Manager: Jiménez De Ory, Santiago. Data collection supported by the Spanish National HIV Paediatric Network (CoRISpeS).*

*Thailand: Thitiporn Borkird, Rachanee Saksawad (Hat Yai Hospital, Songkhla, Thailand), Pope Kosalaraksa, Chanasda Kakkaew (Srinagarind Hospital, Khon Kaen University, Khon Kaen, Thailand), Suparat Kanjanavanit, Siripim Kamphaengkham (Nakornping Hospital, Chiang Mai, Thailand), Mark de Souza, Panadda Sawangsinth (SEARCH, Thai Red Cross AIDS Research Centre, Bangkok).*

*United Kingdom: Clinicians: Hermione Lyall, Gareth Tudor-Williams, Caroline Foster and Sam Walters (Imperial College Healthcare NHS Trust, United Kingdom) with data collection supported by the national Collaborative HIV Paediatric Study (CHIPS) (www.chipscohort.ac.uk).*

*Italy: Ilaria Pepponi (Laboratory, Hospedale Bambino Gèsu), Jennifer Faudella (Administrative work, Hospedale Bambino Gèsu). Attending physicians in charge of infants included from the Rome sites are: Stefania Bernardi, Hyppolite Tchidjou, Emma C Manno, Michela Di Pastena.*

## REFERENCES

1. Violari A, Cotton MF, Gibb DM, et al. Early antiretroviral therapy and mortality among HIV-infected infants. *N Engl J Med.* 2008;359:2233–2244.
2. Cotton MF, Violari A, Otwombe K, et al. Early time-limited antiretroviral therapy versus deferred therapy in South African infants infected with HIV: results from the children with HIV early antiretroviral (CHER) randomised trial. *Lancet.* 2013;382:1555–1563.
3. Persaud D, Palumbo PE, Ziemniak C, et al. Dynamics of the resting CD4(+) T-cell latent HIV reservoir in infants initiating HAART less than 6 months of age. *AIDS.* 2012;26:1483–1490.
4. Luzuriaga K, Tabak B, Garber M, et al. HIV type 1 (HIV-1) proviral reservoirs decay continuously under sustained virologic control in HIV-1–Infected children who received early treatment. *J Infect Dis.* 2014;210:1529–1538.
5. van Zyl GU, Bedison MA, van Rensburg AJ, et al. Early antiretroviral therapy in South African children reduces HIV-1-Infected cells and cell-associated HIV-1 RNA in blood mononuclear cells. *J Infect Dis.* 2015;212:39–43.
6. Tagarro A, Chan M, Zangari P, et al. Early and highly suppressive antiretroviral therapy are main factors associated with low viral reservoir in european perinatally HIV-infected children. *J Acquir Immune Defic Syndr.* 2018;79:269–276.
7. Kuhn L, Paximadis M, Da Costa Dias B, et al. Age at antiretroviral therapy initiation and cell-associated HIV-1 DNA levels in HIV-1-infected children. *PLoS One.* 2018;13:e0195514.
8. Resino S, Resino R, Bellon JM, et al. Clinical outcomes improve with highly active antiretroviral therapy in vertically HIV type-1–infected children. *Clin Infect Dis.* 2006;43:243–252.
9. Klein N, Palma P, Luzuriaga K, et al. Early antiretroviral therapy in children perinatally infected with HIV: a unique opportunity to implement immunotherapeutic approaches to prolong viral remission. *Lancet Infect Dis.* 2015;15:1108–1114.
10. Frange P, Faye A, Avettand-Fenoël V, et al. HIV-1 virological remission lasting more than 12 years after interruption of early antiretroviral therapy in a perinatally infected teenager enrolled in the French ANRS EPF-CO10 paediatric cohort: a case report. *Lancet HIV.* 2016;3:e49–e54.
11. Violari A, Cotton MF, Kuhn L, et al. A child with perinatal HIV infection and long-term sustained virological control following antiretroviral treatment cessation. *Nat Commun.* 2019;10:412.
12. Shiau S, Abrams EJ, Arpadi SM, et al. Early antiretroviral therapy in HIV-infected infants: can it lead to HIV remission? *Lancet HIV.* 2018;5:e250–e258.
13. Chiappini E, Galli L, Tovo PA, et al. Five-year follow-up of children with perinatal HIV-1 infection receiving early highly active antiretroviral therapy. *BMC Infect Dis.* 2009;9:140.
14. Chiappini E, Galli L, Tovo PA, et al. Virologic, immunologic, and clinical benefits from early combined antiretroviral therapy in infants with perinatal HIV-1 infection. *AIDS.* 2006;20:207–215.
15. Shiau S, Strehlau R, Technau KG, et al. Early age at start of antiretroviral therapy associated with better virologic control after initial suppression in HIV-infected infants. *AIDS.* 2016;31:1.
16. Sandbulte MR, Gautney BJ, Maloba M, et al. Infant HIV testing at birth using point-of-care and conventional HIV DNA PCR: an implementation feasibility pilot study in Kenya. *Pilot Feasibility Stud.* 2019;5:18.
17. Dunning L, Kroon M, Fourie L, et al. Impact of birth HIV-PCR testing on the uptake of follow-up early infant diagnosis services in Cape Town, South Africa. *Pediatr Infect Dis J.* 2017;36:1159–1164.
18. Rocca S, Zangari P, Cotugno N, et al. Human immunodeficiency virus (HIV)-Antibody repertoire estimates reservoir size and time of antiretroviral therapy initiation in virally suppressed perinatally HIV-infected children. *J Pediatr Infect Dis Soc.* 2018. doi:10.1093/jpids/piy080.
19. Collins IJ, Foster C, Tostevin A, et al. Clinical status of adolescents with perinatal HIV at transfer to adult care in the UK/Ireland. *Clin Infect Dis.* 2017;64:1105–1112.
20. Phongsamart W, Hansudewechakul R, Bunupuradah T, et al. Long-term outcomes of HIV-infected children in Thailand: the Thailand pediatric HIV observational database. *Int J Infect Dis.* 2014;22:19–24.
21. Jiménez de Ory S, González-Tomé MI, Fortuny C, et al. New diagnoses of human immunodeficiency virus infection in the Spanish pediatric HIV Cohort (CoRISpe) from 2004 to 2013. *Medicine (Baltimore).* 2017;96:e7858.
22. Mirochnick M, Nielsen-Saines K, Pilotto JH, et al. Nevirapine concentrations in newborns receiving an extended prophylactic regimen. *J Acquir Immune Defic Syndr.* 2008;47:334–337.

*Copyright © 2019 Wolters Kluwer Health, Inc. All rights reserved.*



Copyright © 2019 Wolters Kluwer Health, Inc. Unauthorized reproduction of this article is prohibited.

23. Jackson C. flexsurv : a platform for parametric survival modeling in *R*. *J Stat Softw*. 2016;70:1–33.
24. Simon Wood FS. *Package Gamm4*. 2017. Available at: https://cran.r-project.org/web/packages/gamm4/index.html. Accessed January 2, 2019.
25. Wood SN, Pya N, Safken B. Smoothing parameter and model selection for general smooth models (with discussion). *J Am Stat Assoc*. 2016;111:1548–1575.
26. Wickham H. *Ggplot2 : Elegrant Graphics for Data Analysis*. New York, NY: Springer-Verlag.
27. van Rij J, Wieling M, Baayen RH, et al. *{itsadug} : Interpreting Time Series and Autocorrelated Data Using GAMMs*; 2017.
28. R Development Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: Organization: R Foundation for Satistical Computing; 2008.
29. Uprety P, Chadwick EG, Rainwater-Lovett K, et al. Cell-associated HIV-1 DNA and RNA decay dynamics during early combination antiretroviral therapy in HIV-1-Infected infants. *Clin Infect Dis*. 2015;61:1862–1870.
30. Bitnun A, Samson L, Chun TW, et al. Early initiation of combination antiretroviral therapy in HIV-1-Infected newborns can achieve sustained virologic suppression with low frequency of CD4+ T cells carrying HIV in peripheral blood. *Clin Infect Dis*. 2014;59:1012–1019.
31. Frigati L, Cotton M, Rabie H. Antiretroviral therapy for the management of HIV in children. *South Afr Med J* 2014;104:898.
32. Hoenigl M, Chaillon A, Moore DJ, et al. Rapid HIV viral load suppression in those initiating antiretroviral therapy at first visit after HIV diagnosis. *Sci Rep*. 2016;6:32947.
33. Ford N, Migone C, Calmy A, et al. Benefits and risks of rapid initiation of antiretroviral therapy. *AIDS*. 2018;32:17–23.
34. Phillips AN, Staszewski S, Weber R, et al. HIV viral load response to antiretroviral therapy according to the baseline CD4 cell count and viral load. *JAMA*. 2001;286:2560.
35. Jiamsakul A, Kariminia A, Althoff KN, et al. HIV viral load suppression in adults and children receiving antiretroviral therapy-results from the IeDEA collaboration. *J Acquir Immune Defic Syndr*. 2017;76:319–329.
36. Luzuriaga K, McManus M, Mofenson L, et al. A trial of three antiretroviral regimens in HIV-1–Infected children. *N Engl J Med*. 2004;350:2471–2480.
37. Duong T, Judd A, Collins IJ, et al. Long-term virological outcome in children on antiretroviral therapy in the UK and Ireland. *AIDS*. 2014;28:2395–2405.
38. Chan MK, Goodall R, Judd A, et al. Predictors of faster virological suppression in early treated infants with perinatal HIV from Europe and Thailand. *AIDS*. 2019:1.
39. Technau KG, Strehlau R, Patel F, et al. 12-month outcomes of HIV-infected infants identified at birth at one maternity site in Johannesburg, South Africa: an observational cohort study. *Lancet HIV*. 2018;5:e706–e714.
40. Shearer WT, Quinn TC, LaRussa P, et al. Viral load and disease progression in infants infected with human immunodeficiency virus type 1. Women and Infants Transmission Study Group. *N Engl J Med*. 1997;336:1337–1342.




*Copyright © 2019 Wolters Kluwer Health, Inc. All rights reserved.*

Copyright © 2019 Wolters Kluwer Health, Inc. Unauthorized reproduction of this article is prohibited.

# Capítulo II

El descenso de carga viral de VIH-1 es más rápido en niñas que en niños con infección perinatal que inician el tratamiento

## Justificación y Objetivos

El sexo es un factor que ha sido asociado con el curso de diferentes enfermedades infecciosas (56). Existe evidencia de que, en humanos, las mujeres presentan una inmunidad innata más fuerte y una respuesta celular inmune más elevada ante la estimulación antigénica que los hombres (57).

Las diferencias debidas al sexo en adultos infectados por el VIH-1 ha sido ampliamente estudiada. Se han reportado diferencias en perfiles genéticos, fenotipos celulares inmunes, capacidad para el mantenimiento de la latencia y la composición diferente del microbioma (38). Comparado con los hombres, las mujeres tienen menores cargas virales en plasma (40) y mayor recuento total basal de CD4 (41). Además, algunos autores han observado también menores niveles de carga viral, mayores porcentajes de CD4 y menores tiempos hasta supresión en niñas (42). Asimismo, en población pediátrica, existe mayor número de controladores de élite niñas que niños (58).

El TAR lleva a un descenso de la carga viral modelado en dos fases diferentes; una fase inicial rápida exponencial (fase I) y una fase más lenta (fase II). Considerando que la medida de las tasas de descenso de carga viral se considera uno de los indicadores de eficacia del TAR, podría ser importante conocer el efecto del sexo en las diferentes fases del descenso de la carga viral en niños tratados precozmente.

Los objetivos de este estudio son:

- Comparar el descenso de carga viral en niños y niñas en la fase I
- Describir la asociación entre el descenso de carga viral y el reservorio a largo plazo

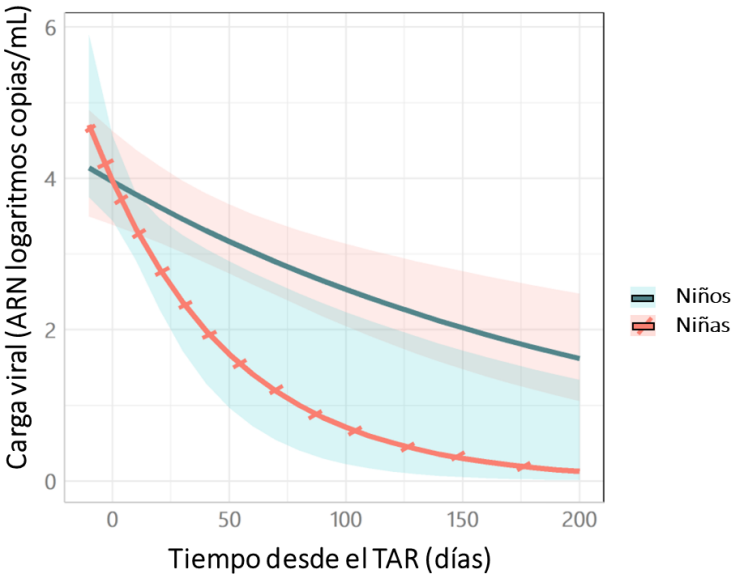
## Resultados

- Del total de 25 pacientes incluidos en este estudio, 17/25 (68%) fueron niñas y 8/25 (32%) niños.
- La mediana de tiempo al diagnóstico de VIH-1 fue de 4,75 meses (RIC, 2,2-5,9) y de tiempo al inicio del TAR 4,7 meses (RIC, 3,07-7,19). No se observaron diferencias entre sexos. Un total de 10/25 (40%) fueron tratados con dos ITIAN y un inhibidor de la

proteasa y 15/25 (60%) con dos ITIAN y un ITINN. Todos los pacientes fueron tratados antes de la era de inhibidores de la integrasa.

- Ambos sexos presentaron similares cargas virales al diagnóstico con una mediana global de 5,7 logaritmos copias/mL (RIC, 5,17; 6,0) y tiempos a supresión de 3,67 meses (2,26-5,51).
- Las niñas presentaron fases 1 de descenso de carga viral más cortas que los niños con puntos de corte menores (3 meses (1,66-4,59) *vs*, 6,79 meses (5,14-9,94), p-valor=0,023). Sin embargo, el descenso de carga viral hasta supresión (<50 copias seguidas) fue similar (-0,67 [-0,77- -0.49] vs. -0.82 [-2.24- -0.43], p=0,244).
- Las pendientes de descenso de carga viral calculadas por regresión segmentada (ver Anexo) fueron más negativas en las niñas en la fase-1 (-17,8 [-28- -10,76] vs, -6,87 [-10,06- -5,99], p-valor=0,016), Sin embargo, no se encontraron diferencias significativas en la fase II.
- Utilizando un modelo mixto se observó cómo, al inicio del TAR, ambos sexos presentaban similares cargas virales, pero por cada mes en tratamiento, las niñas presentaron descensos más acusados que los niños (coeficiente de interacción -0,01 ±0,001, p-valor=0,040).

Figura 6. Asociación entre la interacción con sexo y el descenso de carga viral




Modelo generalizado mixtos con interacción del sexo ajustado por el porcentaje de CD4 pre-TAR.

- El descenso de carga viral en la fase I fue un predictor significativo del tamaño de reservorio medido a los 12,2 años (RIC, 8,0-15,6). Por cada punto que la pendiente (tiempo/logaritmo carga viral) aumentaba, se observó un 5% de incremento en el tamaño de reservorio medido como ADN total/millón células PBMC. Este modelo estaba ajustado por la edad al inicio del TAR, sexo y porcentaje de CD4.

*Journal of the Pediatric Infectious Diseases Society*

BRIEF REPORT

# Faster Initial Viral Decay in Female Children Living With HIV

**Sara Domínguez-Rodríguez,[1] Miquel Serna-Pascual,[1] Caroline Foster,[2] Paolo Palma,[3] Eleni Nastouli,[4] Anita De Rossi,[5] Javier Seoane,[6] Paolo Rossi,[3] Carlo Giaquinto,[5] Alfredo Tagarro,[1] and Pablo Rojo[1]; on behalf of the EPIICAL Project**

[1]Pediatric Infectious Diseases Unit, Fundación para la Investigación Biomédica del Hospital 12 de Octubre, Madrid, Spain, [2]Imperial College Healthcare NHS Trust, London, UK, [3]Division of Immune and Infectious Diseases, IRCCS Ospedale Pediatrico Bambino Gesu, Rome, Italy, [4]Department of Population, Policy and Practice, UCL Great Ormond Street Institute of Child Health, London, UK, [5]Department of Surgery, Oncology and Gastroenterology, Section of Oncology and Immunology, University of Padova, Padova, Italy, and [6]Terrestrial Ecology Group, Department of Ecology, Universidad Autónoma de Madrid, Madrid, Spain

Limited data exist regarding sex bias and viral decay in children with HIV. We investigated the sex differences in viral decay in 25 perinatally HIV-infected children. Females presented faster phase I viral decay regardless of their age at antiretroviral therapy (ART) initiation, baseline CD4 percentages, or baseline RNA levels. Also, for each month elapsed under ART, females had faster viral decay than males.

**Key words.** HIV-1; pediatrics; reservoir; sex; viral decay.

Sex is an important associated factor in several infectious diseases, starting from the beginning of life [1]. There is evidence showing that, in humans, females present stronger humoral and cellular immune responses to infection or antigenic stimulation than males [2].

In adult HIV infection, sex differences are well documented. Significant differences are reported in genetic profiles, immune cell phenotypes, latency maintenance capabilities, and microbiome composition [3]. Compared with men, women living with HIV presented lower plasma viral loads (VLs) [4] and higher CD4 counts [5]. Likewise, some authors described significantly lower HIV RNA levels, higher CD4 percentages, and lower time to suppression in female children compared with males [6]. Furthermore, elite controllers in childhood are more frequently females [7].

Treatment with combination antiretroviral therapy (ART) drives to a decay of plasma HIV RNA that can be described in 2 phases. The majority of the virus in plasma is cleared during an initial exponential rapid decline followed by a slower second phase. The observed pattern is explained by 2 different cellular HIV sources: short-lived CD4 T cells and other long-lived cells with HIV-1 provirus, respectively. Since viral decay rates are used as a measure of therapy efficacy, it may be relevant to determine factors that affecting in the different viral decay phases.

In this pilot study, we investigated whether viral decay in phase I is different between females and males in children, and if viral decay is a predictive factor of DNA reservoir size.

## METHODS

CARMA (Child and Adolescent Reservoir Measurements on early suppressive ART) is a multicentre study performed within the EPIICAL (Early treated Perinatally HIV-infected individuals: Improving Children's Actual Life) Consortium. CARMA gathered data from perinatally HIV-infected children from 7 European centers. In this study, children treated with ART within the first 24 months of life who suppressed within the first year of ART initiation and had remained suppressed for at least 5 years were included [8]. Viral suppression was considered as reaching VL <400 cp/mL in 2 consecutive measurements. We selected patients who presented a consistent decreasing sequence toward suppression and had at least 5 measurements above the detection threshold to improve parameter identifiability.

To analyze the viral decay trends in phase I post ART initiation, a flexible piecewise regression model was applied. Time breakpoints were considered to be located where the viral decay phase changes took place. Slopes for each phase were calculated using this model. The ushr R package was used to model the decline and estimate the lifespans of short- and long-lived cells [9]. An alternative Poisson mixed model was used to assess the effect of sex on phase I viral decay, which included baseline CD4 percentages and the interaction between sex and time since ART initiation as fixed effects and subjects as a random intercept. To determine the strength of the association between viral decay in phase I and DNA reservoir, a Poisson regression model was applied adjusting by age at ART, sex, and CD4 percentages. Variable selection for the models was done by a backward stepwise procedure using the Akaike index criterion. Variables such as ART regimen at initiation, ethnicity, or baseline age were not included in the best-fitted model.

The continuous variables were described as medians and interquartile ranges and absolute numbers and relative frequencies when categorical. Comparisons between sexes were assessed by the Wilcoxon test and Fisher exact test, respectively.

Received 14 October 2020; editorial decision 22 December 2020; accepted 28 December 2020; Published online January 25, 2021.

Corresponding Author: Sara Domínguez Rodríguez. E-mail: sara.dominguez.r@gmail.com.

**Journal of the Pediatric Infectious Diseases Society 2021;XX(XX):0–0**

© The Author(s) 2021. Published by Oxford University Press on behalf of The Journal of the Pediatric Infectious Diseases Society. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com.

DOI: 10.1093/jpids/piaa174



Downloaded from https://academic.oup.com/jpids/advance-article/doi/10.1093/jpids/piaa174/6118585 by sara.dominguez.r@gmail.com on 26 January 2021

## RESULTS

A total of 25 ART-treated HIV-infected children were included. Most of the patients were female (n = 17, 68%). The median time for HIV diagnosis was 4.75 months (2.2; 5.9), and the median age for ART initiation was 4.7 months (3.07; 7.19) with no differences found between sexes ($P$ = .954). All the patients were treated before the integrase inhibitors (INI) era. Likewise, both sexes presented similar VL at diagnosis (baseline) with an overall median of 5.7 logs (5.17; 6.0) and 3.67 months to suppression (2.26; 5.51), $P$ = .883. Only 1 of the 25 female patients has had a virological failure (2 consecutive VL measurements higher than 400 copies/mL after suppression).

Females presented a shorter phase I viral decay duration with earlier breakpoints than males (3 months [1.66; 4.59] vs 6.79 months [5.14; 9.94], $P$ = .023). The viral decay slopes calculated with the piecewise regression were also significantly more negative in females than in males (−17.8 [−28; −10.76] vs −6.87 [−10.06; −5.99], $P$ = .016). However, no differences were found in the viral decay slope for phase II (−0.03 [−0.03; −0.01] vs −0.03 [−0.05; −0.01], $P$ = .258) or viral decay to suppression (−0.67 [−0.77; −0.49] vs −0.82 [−2.24; −0.43], $P$ = .244). According to the model proposed by Morris et al [9], no differences were found in cell subpopulations. The model found 50% biphasic decays in males compared with 35.5% in females ($P$ = .667). No differences were found between males and females in short lifespan (0.30 [0.20; 0.38] years vs 0.18 [0.14; 0.37] years, $P$ = .915). Although not significant, females presented suggestively higher long lifespan cells than males (1.78 [1.37; 14.5] years vs 4.29 [2.64; 5.66] years, $P$ = .394).

At ART initiation, both sexes presented similar VLs, but for each month elapsed, females had higher viral decay rates than males (interaction coefficient = −0.01 ± 0.001, $P$ = .040) (Figure 1).

The viral decay in phase I was a significant predictor for DNA reservoir size measured at a median age of 12.2 years [8.0; 15.6]. For each point in the slope (time/$\log_{10}$ VL) increased, a 5% increase in DNA reservoir size was predicted (exponential coefficient: 1.05 [1.04–1.06], $P = 2.5 \times 10^{-16}$) adjusted for age at ART, sex, and CD4 percentages. In other words, the less pronounced is the slope (more positive), the higher the DNA reservoir was estimated according to this model.

## DISCUSSION

In this pilot analysis of children initiating ART and suppressing within the first year of therapy, we observed that, regardless of age at ART initiation, baseline CD4 percentages, or baseline RNA levels, females may present faster phase I viral decay. Furthermore, the viral decay slope predicted the reservoir size more than a decade later.

Female and male children presented similar baseline VL at ART initiation. Moreover, both sexes presented similar times to suppression, consistent with what was already published in a pediatric cohort from the same consortium where viral decay to suppression was modeled [10]. However, the phase I viral decay was significantly faster in females than in males. These results are in concordance with the previous evidence suggesting that female children may control viremia better than males [7]. These data support that there are also gender differences in viral decay rates during childhood. Hormones are unlikely to be the

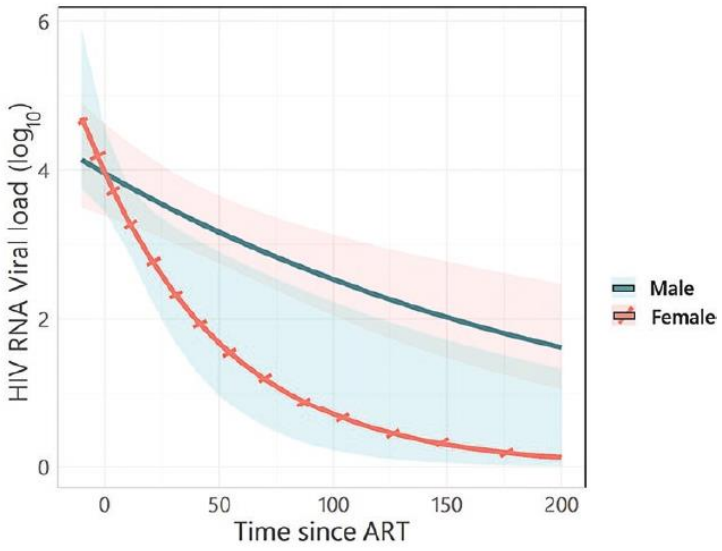



**Figure 1.** Mixed model effects for sex interaction association with viral decay adjusted by CD4 percentages.





Downloaded from https://academic.oup.com/jpids/advance-article/doi/10.1093/jpids/piaa174/6118585 by sara.dominguez.r@gmail.com on 26 January 2021

major drivers for such differences, whereas sex-biased gene expression mechanisms [6], ART adherence, or different CD4 population between sexes can play a critical role.

Although the time to suppression was similar between sex groups analyzed, the key finding of the present study is that females have more pronounced phase I viral decay. We also considered the effect of viral decay on DNA reservoir size and found that viral decay in phase I can be a significant predictor together with age at ART, sex, and CD4 percentages.

As a limitation of the study, it is possible that our observations could be biased for the unbalanced numbers between sexes and the limited sample size. We were not able to control for adherence or specific ART drugs although we included regimens as a possible confounder. Moreover, these results could be different from modern INI-based therapy. Despite these limitations, our results may help to understand the role of sex in response to ART in this susceptible pediatric population.

In conclusion, the viral decay in phase I was faster in female perinatally HIV-infected children, and this slope can be a predictor of long-term reservoir size.

## Notes

***Acknowledgments***. We thank the EPIICAL Consortium team and the families and patients to participate in the study. We also thank the AIDA group (Applied Intelligence and Data Analysis) from Escuela Técnica Superior de Ingeniería de Sistemas Informáticos de la Universidad Politécnica de Madrid and, in particular, to Angel Panizo for the assistance with the computational resources.

***Financial support***. The EPIICAL project is funded through an independent grant by ViiV Healthcare, UK. The funders had no role in study design, data collection, analysis, and interpretation, or in manuscript preparation.

***Potential conflicts of interest***. All authors: No reported conflicts.

All authors have submitted the ICMJE Form for Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

## References

1. Muenchhoff M, Goulder PJR. Sex differences in pediatric infectious diseases. J Infect Dis **2014**; 209:S120–6.
2. Fish EN. The X-files in immunity: sex-based differences predispose immune responses. Nat Rev Immunol **2008**; 8:737–44.
3. Scully EP. Sex differences in HIV infection. Curr HIV/AIDS Rep **2018**; 15:136–46.
4. Sterling TR, Vlahov D, Astemborski J, Hoover DR, Margolick JB, Quinn TC. Initial plasma HIV-1 RNA levels and progression to AIDS in women and men. N Engl J Med **2001**; 344:720–5.
5. Maskew M, Brennan AT, Westreich D, et al. Gender differences in mortality and CD4 count response among virally suppressed HIV-positive patients. J Womens Health (Larchmt) **2013**; 22:113–20.
6. Ruel TD, Zanoni BC, Ssewanyana I, et al. Sex differences in HIV RNA level and CD4 cell percentage during childhood. Clin Infect Dis **2011**; 53:592–9.
7. Vieira VA, Zuidewind P, Muenchhoff M, et al. Strong sex bias in elite control of paediatric HIV infection. AIDS **2019**; 33:67–75.
8. Foster C, Domínguez-Rodríguez S, Tagarro A, et al. The CARMA Study: early infant antiretroviral therapy—timing impacts on total HIV-1 DNA quantitation 12 years later. J Pediatric Infect Dis Soc **2020**: piaa071. doi:10.1093/jpids/piaa071. Epub ahead of print. PMID: 32678875.
9. Morris SE, Dziobek-Garrett L, Yates AJ. ushr: understanding suppression of HIV in R. BMC Bioinformatics **2020**; 21:1–8.
10. Schröter J, Anelone AJN, Yates AJ, de Boer RJ. Time to viral suppression in perinatally HIV-infected infants depends on the viral load and CD4 T-cell percentage at the start of treatment. J Acquir Immune Defic Syndr **2020**; 83:522–9.

Downloaded from https://academic.oup.com/jpids/advance-article/doi/10.1093/jpids/piaa174/6118585 by sara.dominguez.r@gmail.com on 26 January 2021

# Capítulo III:

El tratamiento antirretroviral precoz en niños tiene un impacto en la cuantificación del ADN total del reservorio de VIH-1 12 años después

## Justificación y Objetivos

El TAR ha sido una revolución en el incremento de la supervivencia, sin embargo, por sí mismo no es capaz de curar la infección. La incapacidad para acceder al conjunto viral latente en las células de reservorio y la amplia variabilidad del virus hace que el TAR se convierta en un tratamiento crónico.

Tanto en niños como en adultos se sabe que el inicio precoz del TAR minimiza el tamaño del reservorio en las células CD4 de memoria y limita la evolución molecular del virus dentro de los mismos (59,60). Los niños, por sus características inmunológicas, poseen una gran capacidad de regeneración inmune y un estado de activación inmune menor (12,13).

Algunos niños con inicios precoces de TAR se mantienen seronegativos para VIH-1 después de perder los anticuerpos maternos y no presentan respuestas celulares específicas para VIH-1, lo que les hace potencialmente más susceptible a un menor control de la viremia virológico si el TAR se interrumpe (61).

Identificar factores y biomarcadores asociados con un tamaño de reservorio bajo puede arrojar luz en la identificación de participantes para futuros estudios cuyo objetivo sea lograr la remisión completa sin TAR (14).

El objetivo de este estudio es:

- Identificar factores asociados a presentar un menor reservorio viral (ADN total de VIH-1) a largo plazo en niños y adolescentes perinatalmente infectados y virológicamente suprimidos.

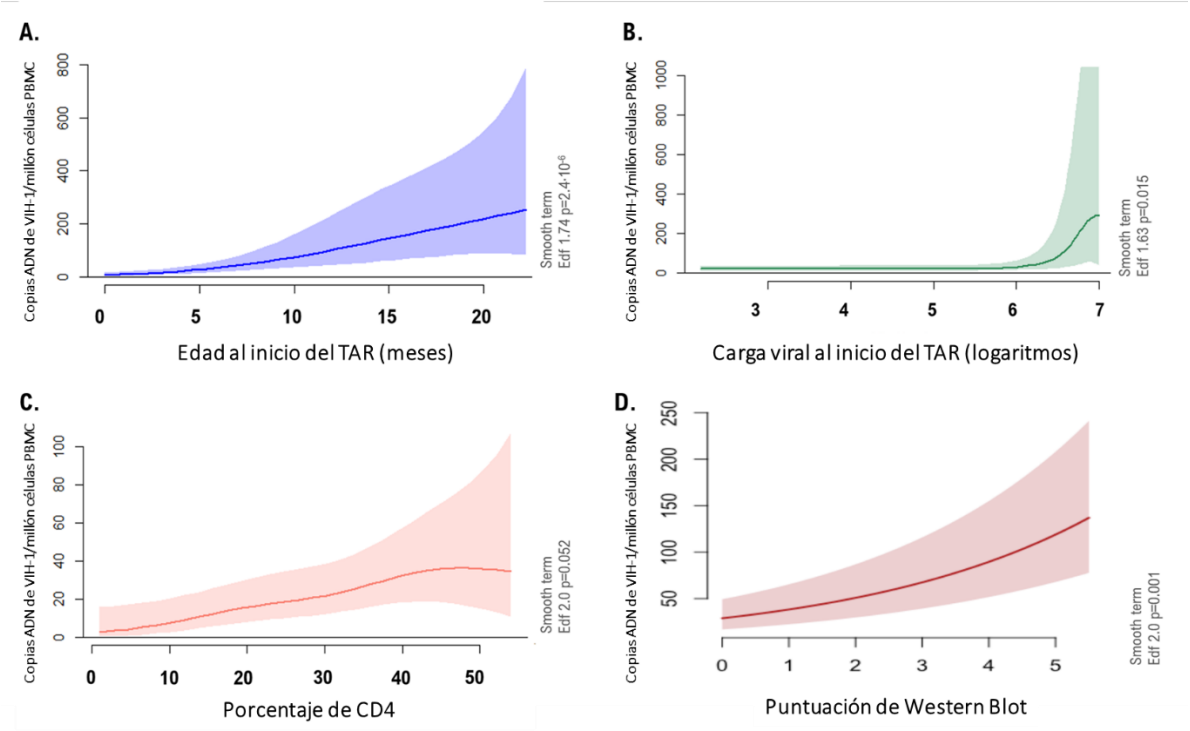
## Resultados

- De los 40 niños incluidos en este estudio, 27 (67,5%) fueron niñas, y 21 (52,5%) de raza negra. Comenzaron el TAR a una mediana de 4,08 meses (RIC, 0,25- 6,23). Previo al TAR, los niños presentaban una mediana de recuento total de CD4 de 1.310 células/µL (RIC, 478- 2.376), una mediana de porcentaje total de CD4 de 30,5% (RIC, 12,2- 42,5) y una mediana de carga viral de 5,28 logaritmos copias/mL (RIC, 4,07- 5,70). La mediana de tiempo hasta la supresión fue de 4,69 meses (RIC, 2,52- 6,26).
- Al momento del análisis, los niños llevaban una mediana de 12 años (RIC, 7,3; 15,4) desde la iniciación del TAR. En ese momento, los niños presentaron una mediana de

recuento total de CD4 de 872 células/µL (RIC, 646- 1.053) y una mediana de porcentaje total de CD4 de 41,0% (RIC, 33,8- 46,2).

- Un total de 11 (27,5%) niños presentaron menos de 10 copias ADN VIH-1/millón de PBMCs. De estos 11 niños, 5 estuvieron por debajo del límite de detección (<1 copia ADN VIH-1/millón PBMC). Los 35 niños con reservorios por encima de 10 copias ADN VIH;1/millón de PBMCs presentaron una mediana de 50,9 copias ADN VIH-1/millón de PBMCs (RIC, 25,3- 117,3).
- Un total de 39/40 pacientes presentaban datos sobre análisis serológicos. De ellos, diez pacientes fueron seronegativos por el test VIH anticuerpo-antígeno de cuarta generación en el momento del análisis. La mitad (5/10) de los niños con menos de 10 copias ADN VIH-1/millón de PBMCs y 5/29 (17,2 %) con un reservorio mayor de 10 copias ADN VIH-1/millón de PBMCs fueron seronegativos (p-valor=0,087).
- Los niños con reservorio bajo de VIH-1 fueron aquellos diagnosticados antes de VIH (3,6 meses (RIC, 0,34- 4,26) *vs.* 4,9 meses (RIC, 3,1- 6,5)) y con un inicio del TAR más temprano (3,6 meses (RIC, 0,1; 4,5) *vs.* 4,6 meses (RIC, 2,66; 7,6)). No se observaron diferencias significativas en la asociación entre el tamaño del reservorio y estadio clínico de SIDA o el tiempo a la supresión.
- No se encontraron diferencias significativas entre la edad al inicio del TAR y cargas virales al inicio del TAR y el estado serológico al análisis 12 años después.
- Como factores asociados al reservorio únicamente fueron seleccionados la edad al inicio del TAR, carga viral al inicio del TAR y porcentaje de CD4 al inicio del seguimiento. La edad al inicio del TAR presentó una relación positiva lineal con el total de reservorio (coeficiente 0,1 ±0,001, p-valor<0,001), es decir, a mayor edad al inicio del TAR mayor reservorio. El efecto de la carga viral al inicio del TAR fue solo notable para cargas virales por encima de 6 logaritmos (coeficiente 0,35 ±0,1, p-valor<0,001). El efecto del porcentaje de CD4 al inicio del seguimiento fue solo notable hasta el 40% (coeficiente 0,03±0,01, p-valor=0,049). A mayor porcentaje de CD4, mayor reservorio entre 0-40% mientras que a partir del 40% se observa poco efecto en el tamaño del reservorio. La puntuación de *western blot* fue significativamente asociada con el total de reservorio. A mayor puntuación de western blot mayor reservorio viral (coeficiente 0,15±0,01, p-valor<0,001).

Figura 7. Modelo generalizado aditivo mixto multivariable para describir la asociación entre la edad al TAR (Panel A), carga viral al inicio del TAR (Panel B), porcentaje de CD4 al inicio del TAR (Panel C), Puntuación de Western blot (Panel D) y el tamaño de reservorio (copias de ADN de VIH-1/millón células PBMC).




Las diferentes combinaciones de regímenes de TAR fueron incluidas como interceptas aleatorias. Las áreas sombreadas indican el intervalo de confianza al 95%. Los términos de alisado o *smooth* están especificados en cada panel describiendo los grados de libertad efectivos (edf) para cada una de las variables incluidas en el modelo.

*Journal of the Pediatric Infectious Diseases Society*

ORIGINAL ARTICLE

# The CARMA Study: Early Infant Antiretroviral Therapy—Timing Impacts on Total HIV-1 DNA Quantitation 12 Years Later

**Caroline Foster,[1] Sara Domínguez-Rodríguez,[2] Alfredo Tagarro,[2] Triantafylia Gkouleli,[3] Judith Heaney,[3,4] Sarah Watters,[3] Alasdair Bamford,[5,6] Katy Fidler,[7] Marisa Navarro,[8] Anita De Rossi,[9] Paolo Palma,[10] Eleni Nastouli,[3,6] Paolo Rossi,[10] Carlo Giaquinto,[9] and Pablo Rojo[2]; for the Early Treated Perinatally HIV Infected Individuals: Improving Children's Actual Life (EPIICAL) Consortium**

[1]Imperial College Healthcare NHS Trust, London, United Kingdom; [2]Pediatric Infectious Diseases Unit, Fundación para la Investigación Biomédica del Hospital 12 de Octubre, Universidad Complutense, Madrid, Spain; [3]Department of Population, Policy and Practice, University College London Great Ormond Street Institute of Child Health, London, United Kingdom; [4]Advanced Pathogen Diagnostics Unit, University College London Hospitals NHS Foundation Trust, London, United Kingdom; [5]Great Ormond Street Hospital for Children NHS Trust, London, United Kingdom; [6]University College London Great Ormond Street Institute of Child Health, London, United Kingdom; [7]Brighton and Sussex Medical School, Falmer, United Kingdom; [8]Pediatric Infectious Disease Unit, Institute of Investigation, Hospital General Universitario Gregorio Marañón, Universidad Complutense, Madrid, Spain; [9]Department of Surgery, Oncology and Gastroenterology, Section of Oncology and Immunology, University of Padova, Padova, Italy; and [10]Division of Immune and Infectious Diseases, Istituto di Ricovero e Cura a Carattere Scientifico, Ospedale Pediatrico Bambino Gesu, Rome, Italy

***Background.*** Strategies aimed at antiretroviral therapy (ART)–free remission will target individuals with a limited viral reservoir. We investigated factors associated with low reservoir measured as total human immunodeficiency virus type 1 (HIV-1) DNA in peripheral blood mononuclear cells (PBMCs) in perinatal infection (PaHIV).

***Methods.*** Children from 7 European centers in the Early Treated Perinatally HIV Infected Individuals: Improving Children's Actual Life (EPIICAL) consortium who commenced ART aged <2 years, and remained suppressed (viral load [VL] <50 copies/mL) for >5 years were included. Total HIV-1 DNA was measured by quantitative polymerase chain reaction per million PBMCs. Factors associated with total HIV-1 DNA were analyzed using generalized additive models. Age, VL at ART initiation, and baseline CD4% effects were tested including smoothing splines to test nonlinear association.

***Results.*** Forty PaHIV, 27 (67.5%) female 21 (52.5%) Black/Black African, had total HIV-1 DNA measured; median 12 (IQR, 7.3–15.4) years after ART initiation. Eleven had total HIV-1 DNA <10 copies/$10^6$ PBMCs. HIV-1 DNA levels were positively associated with age and VL at ART initiation, baseline CD4%, and Western blot antibody score. Age at ART initiation presented a linear association (coefficient = 0.10 ± 0.001, $P \leq .001$), the effect of VL (coefficient = 0.35 ± 0.1, $P \leq .001$) noticeable >6 logs. The effect of CD4% (coefficient = 0.03 ± 0.01, $P = .049$) was not maintained >40%.

***Conclusions.*** In this PaHIV cohort, reduced total HIV-1 DNA levels were associated with younger age and lower VL at ART initiation. The impact of early-infant treatment on reservoir size persists after a decade of suppressive therapy.

**Key words.** adolescents; children; early treated; HIV-1; HIV reservoir; viral suppression.

Despite the very significant gains made in the prevention of perinatal transmission of human immunodeficiency virus type 1 (HIV-1), globally >150 000 infected infants are born annually. Antiretroviral therapy (ART) has revolutionized survival; however, alone it is unable to confer a cure, a consequence of an inaccessible latent pool of HIV-1–infected reservoir cells. In both adults and infants, initiation of ART soon after infection minimizes the reservoir size in resting memory $CD4^+$ T cells and limits the evolution of HIV-1 within reservoirs [1, 2]. In infants, early ART, typically defined as within 8–12 weeks of birth, is possible but challenging; those infected in utero can be identified within a few days of life, and those with intrapartum HIV-1 transmission by 2–6 weeks of age [3]. Rare case reports in perinatally infected individuals demonstrate that long-term viral control can be influenced by very early ART, although alone this is almost always insufficient for sustained viral control when subsequently stopped [4, 5]. Children differ immunologically from adults, with a more active thymus and greater capacity for immune regeneration, a reduced central memory T-cell compartment, and low immune activation state, a condition not conducive for HIV-1 reservoir seeding [6, 7]. Children starting ART within 3 months of birth remain frequently HIV-1 seronegative after losing maternal antibodies and lack HIV-specific cellular immune responses, making them potentially more susceptible to uncontrolled viral replication when ART is interrupted [8, 9].

Received 12 February 2020; editorial decision 28 May 2020; accepted 3 June 2020

Correspondence: Caroline Foster, MA, MBBS, MRCPCH, Department of Paediatric Infectious Diseases, Imperial College Healthcare NHS Trust, London W2 1NY, UK (caroline.foster5@nhs.net).

**Journal of the Pediatric Infectious Diseases Society 2020;XX(XX):1–7**

© The Author(s) 2020. Published by Oxford University Press on behalf of The Journal of the Pediatric Infectious Diseases Society. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs licence (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial reproduction and distribution of the work, in any medium, provided the original work is not altered or transformed in any way, and that the work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

DOI: 10.1093/jpids/piaa071

Downloaded from https://academic.oup.com/jpids/article-abstract/doi/10.1093/jpids/piaa071/5873261 by guest on 31 July 2020

Identifying factors and biomarkers, such as seronegativity, associated with limited viral reservoir size will potentially aid identification of participants for future proof-of-concept studies aimed at inducing ART-free viral remission [10].

The CARMA study (Child and Adolescent Reservoir Measurements on Early Suppressive ART) is part of the existing EPIICAL consortium (Early Treated Perinatally HIV Infected Individuals: Improving Children's Actual Life; www.epiical.org), an international collaboration whose aim is to select promising HIV-1 therapeutic strategies and to identify candidates among perinatally HIV-1–infected children and adolescents and is described in detail previously [10].

The primary endpoint analyses of the CARMA cohort are presented, identifying factors that influence the establishment of a low total HIV DNA level in perinatally HIV-infected children and adolescents on suppressive ART.

## MATERIALS AND METHODS

### Design

This was a multicenter, cross-sectional study enrolling subjects from 7 European pediatric HIV clinical research centers contributing to EPIICAL, in England (3), Spain (2), and Italy (2).

### Participants

Forty perinatally HIV-1–infected children who commenced ART within the first 24 months of life were included. Viral suppression (viral load [VL] <400 copies/mL) had to be achieved within 12 months of ART initiation, maintained for a minimum of 5 years with a VL of <50 copies/mL confirmed at enrollment. Children with current viral hepatitis or tuberculosis coinfection, malignancy, pregnancy, or concomitant immunosuppressive therapy were excluded. For each participant, data collection included demographics; age at HIV diagnosis, at ART start, and at time of analysis; and ART drug regimens ever received. Immunology (CD4 total, CD8 total, CD4% and CD8%, and CD4:CD8 ratio) and VL were recorded from diagnosis, ART initiation, throughout treatment, and at time of reservoir analysis. Viral load "blips" (defined as a rise in plasma VL from 50 to 399 copies/mL returning to <50 copies/mL on repeat sampling) and single annual viral "spikes" (defined as VL 400–999 copies/mL returning to <50 copies/mL on repeat sampling) were permitted.

### Laboratory Analyses

#### *Total HIV-1 DNA*

Total HIV-1 DNA was measured in purified peripheral blood mononuclear cells (PBMCs) by real-time quantitative polymerase chain reaction (qPCR). Total nucleic acids were extracted from PBMCs on the QIAsymphony platform using the DSP Virus/Pathogen Mini Kit (Qiagen) according to the manufacturer's protocol. Primers were designed for HIV-1 long terminal repeat and the genomic reference gene pyruvate dehydrogenase as previously described [11]. Quantification of cell-associated HIV-1 DNA by qPCR was determined using 20 µL of extract and a standard curve with known copy numbers in 10-fold dilutions. The limit of quantitation is 10 copies/$10^6$ PBMCs and the limit of detection is 1 copy/$10^6$ PBMCs. Assays were performed in triplicate.

#### *Fourth-Generation Antibody/Antigen*

Serological status was assessed by plasma venous sampling using the fourth-generation Abbott ARCHITECT HIV Ag/Ab Combo assay.

#### *Western Blot: HIV-Specific Antibody Characterization*

Plasma samples were tested as previously described [12]. Human immunodeficiency virus Western blot (WB) kit 2.2 (Medical Systems, Genova, Italy) was used, following the manufacturer's instructions, to detect specific antibody (Ab) responses, in 20 µL of plasma, against 10 different HIV-1 viral proteins (gp160, gp120, p66, p55, p51, gp41, p39, p31, p24, and p17) as previously described [13]. Nitrocellulose strips containing HIV proteins incorporated through electrophoretic blotting were tested with plasma samples and with controls. The presence or absence of Abs to HIV was determined by comparing each nitrocellulose strip to the assay control strips tested with the nonreactive, strong reactive, and weak reactive controls. All of the WB strips were analyzed using the ImageJ software (ImageJ 1.52a; National Institutes of Health). The band intensity was calculated for each viral antigen in each sample, and compared with the intensity calculated in the strong reactive control. A score of 1 was assigned when the intensity of the band was ≥50% of those calculated in the strong positive control for the same antigen, whereas 0.5 was assigned when the band intensity ranged from 10% to 49% of those found in the positive control. Finally, negative results (WB score = 0) were assigned for band intensity ranging from 0% to 9% of those found in the positive control. A WB score was assigned to each patient by adding up the number of positive and weak responses (from 0 to 10).

### Statistical Methods

Data were displayed as a whole and summarized by HIV-1 DNA reservoir (groups ≤10 copies vs >10 copies/$10^6$ PBMCs). HIV-1 DNA reservoir was dichotomized: low (≤10) vs high (>10) as per previous studies and reflecting the limit of assay quantitation. In multivariable models, HIV-1 DNA was used as a continuous variable. In case of continuous variables, median and interquartile range (IQR) were assessed for nonparametric variables and mean and standard deviation for parametric variables. In summary tables of categorical variables, counts and percentages were reported and comparisons were assessed using Fisher exact test. For normally distributed continuous variables, Student *t* test was performed, and Mann-Whitney *U* test when nonparametric.





Downloaded from https://academic.oup.com/jpids/article-abstract/doi/10.1093/jpids/piaa071/5873261 by guest on 31 July 2020

To describe the association between the HIV-1 reservoir and the different effect of each clinical, virological, and immunological feature, a generalized additive mixed model was applied. Backward stepwise elimination was applied to reach the final multivariable model, and Akaike information criteria (AIC) was used to identify the best-fitting model.

To analyze possible nonlinear associations, each variable was included in the model with smoothing splines. Smoothing parameters defining degrees of freedom (*df*) were selected according to nested AIC comparisons. ART combination at initiation was included as random intercept to correct for random variance across regimens. R software version 3.5.2 software was used for all analyses (R Development Core Team, 2008).

#### Ethical Considerations

The CARMA study was approved by ethics committees within each country and informed assent/consent obtained from adolescents and/or parent(s)/legal guardian in accordance with country-specific law. Each participant received a unique study number, under which data were pseudo-anonymized.

## RESULTS

Forty perinatally HIV-1–infected children, 27 (67.5%) female, 21 (52.5%) Black/Black African, and 13 (32.5%) white, commenced ART at a median age of 4.08 (IQR, 0.25–6.23) months, CD4 count 1310 (IQR, 478–2376) cells/μL, CD4% 30.5 (IQR, 19.2–42.5), and $\log_{10}$ VL 5.28 (IQR, 4.07–5.70) copies/mL (Table 1). Median time to viral suppression was 4.69 (IQR, 2.52–6.26) months. In analysis at a median of 12 (IQR, 7.3–15.4) years after ART initiation, median CD4 count was 872 (IQR, 646–1053) cells/μL, CD4% 41.0 (IQR, 33.8–46.2), and total HIV-1 DNA <10 copies/$10^6$ PBMCs in 11 (27.5%) children. Of these 11 children, 5 were below the lower limit of detection (<1 copy/$10^6$ PBMCs), the remaining 6 children having low levels of detectable total HIV-1 DNA (1–9 copies/$10^6$ PBMCs). In the 35 children with a quantifiable total HIV-1 DNA level (≥1 copy/$10^6$ PBMCs), the median was 50.9 (IQR, 25.3–117.3) copies/$10^6$ PBMCs. Serological analyses were available for 39/40 children. Ten children were seronegative on fourth-generation HIV Ab/antigen (Ag) testing at time of analysis. Five of 10 (50%) children with total HIV-1 DNA <10 copies/$10^6$ PBMCs and 5 of 29 (17.2%) children with a reservoir >10 copies/$10^6$ PBMCs were seronegative ($P = .087$) (Table 1). Only 1 child with a reservoir below the limit of detection (<1 copy/$10^6$ PBMCs) had positive HIV serology. There was a suggestive association of children with a reservoir of <10 copies/$10^6$ PBMCs being Ab/Ag nonreactive compared with those with higher reservoir size ($P = .087$; Table 1).

Viral load at diagnosis was significantly lower in those with total HIV-1 DNA <10 copies/$10^6$ PBMCs ($\log_{10}$ VL, 5.70 [IQR, 5.25–6.00] copies/mL vs 4.66 [IQR, 4.18–5.33] copies/mL; $P = .030$; Table 1). Children with a reservoir of <10 copies/$10^6$ PBMCs were likely to be younger at HIV diagnosis (3.61 [median (IQR), 0.34–4.26] months vs 4.92 [IQR, 3.08–6.49] months) and at ART start (3.6 [median (IQR), 0.10–4.49] months vs 4.6 [IQR, 2.66–7.64] months), although this did not reach statistical significance. No differences were found in patients with low vs high reservoir related to clinical stage (AIDS vs non-AIDS) or time to viral suppression. The median age at ART initiation was 4.6 (IQR, 3.11–7.04) months in those with positive/equivocal serology and 3.1 (IQR, 0.21–3.80) months in those who were seronegative (Table 1). There was no association of age at ART with $\log_{10}$ VL at ART start and serological status at analysis 12 years later ($\log_{10}$ VL, 5.3 [IQR, 4.04–5.8] copies/mL in those with positive/equivocal serology and 5.02 [IQR, 4.22–5.4] copies/mL in those who were seronegative).

The number of patients ever experiencing a blip and/or spike and time from viral suppression to first blip and/or spike were comparable between groups (Table 1). Immunological parameters (total/% CD4, CD8, and CD4:CD8) did not differ between groups at diagnosis, ART initiation, viral suppression, or 12 years later at time of reservoir analysis.

Only age and VL at ART initiation and baseline CD4% were selected to be included in the multivariable model to predict reservoir size. Total HIV-1 DNA levels were positively associated with age and VL at ART initiation and baseline CD4% (Figure 1). While age at ART initiation presented a linear association (coefficient = 0.10 ± 0.001, $P \leq .001$, $df = 2$), the effect of VL (coefficient = 0.35 ± 0.10, $P \leq .001$, $df = 3$) only was noticeable for individuals with VLs at ART initiation above 6 logs. The effect of baseline CD4% (coefficient = 0.03 ± 0.01, $P = .049$) was not maintained above 40%. Western blot score was significantly positively associated with HIV DNA reservoir adjusted for age and considering drug combinations as random intercepts ($P < .001$; Figure 1).

## DISCUSSION

Of 40 HIV-1 perinatally infected children from 7 different sites across 3 European countries initiating ART in infancy (median age, 4 months), with viral suppression maintained for more than a decade, just over one-quarter had very low/undetectable total HIV-1 DNA levels (<10 copies/$10^6$ PBMCs). Patients with a lower reservoir size commenced ART earlier after birth and at a lower baseline VL. The association of earlier ART initiation following primary infection and a reduced reservoir size is well-established in adults and is increasingly recognized for children, further supported by this study. The impact of pretreatment VL on future reservoir size is less established. Baseline VL is associated with time to viral suppression, and subsequent risk of viral rebound in adult populations and with persistent residual low-level viremia on ultrasensitive assays





Downloaded from https://academic.oup.com/jpids/article-abstract/doi/10.1093/jpids/piaa071/5873261 by guest on 31 July 2020

**Table 1. Immunological and Virological Parameters Compared by Reservoir Size**

| Parameter | All (N = 40) | >10 Copies/$10^6$ PBMCs (n = 29) | <10 Copies/$10^6$ PBMCs (n = 11) | *P* Value |
|---|---|---|---|---|
| Region, No. (%) | | | | .062 |
| Italy | 16 (40.0) | 9 (31.0) | 7 (63.6) | |
| Spain | 5 (12.5) | 3 (10.3) | 2 (18.2) | |
| United Kingdom | 19 (47.5) | 17 (58.6) | 2 (18.2) | |
| Sex, female, No. (%) | 27 (67.5) | 17 (58.6) | 10 (90.9) | .068 |
| Race/ethnicity, No. (%) | | | | .663 |
| White | 13 (32.5) | 8 (27.6) | 5 (45.5) | |
| Black/Black African | 21 (52.5) | 16 (55.15) | 5 (45.5) | |
| Other | 6 (15.0) | 5 (17.25) | 1 (9.09) | |
| Age at HIV diagnosis, mo | 4.17 (2.19–6.32) | 4.92 (3.08–6.49) | 3.61 (0.34–4.26) | .084 |
| AIDS diagnosis, No. (%) | | | | .486 |
| No | 25 (62.5) | 17 (58.6) | 8 (72.7) | |
| Yes | 15 (37.5) | 12 (41.4) | 3 (27.3) | |
| Age at ART start, mo | 4.08 (0.25–6.23) | 4.59 (2.66–7.64) | 3.61 (0.10–4.49) | .071 |
| Follow-up, y | 13.0 (8.19–16.1) | 13.6 (8.29–16.4) | 10.7 (8.99–15.0) | .892 |
| Virology | | | | |
| VL at diagnosis, log copies/mL | 5.60 (4.98–5.93) | 5.70 (5.25–6.00) | 4.66 (4.18–5.33) | .030 |
| VL at ART start, log copies/mL | 5.28 (4.07–5.70) | 5.48 (4.91–5.80) | 4.64 (4.00–5.18) | .119 |
| Time to viral suppression, mo | 4.69 (2.52–6.26) | 4.59 (3.21–6.46) | 4.79 (1.93–5.11) | .844 |
| VL blip. No. (%) | 11 (27.5) | 10 (34.5) | 1 (9.09) | .233 |
| Time to blip, y | 3.7 (2.35–7.26) | 3.31 (2.23–7.30) | 6.48 (6.48–6.48) | .723 |
| Viral spike, No. (%) | 5 (12.5) | 3 (10.3) | 2 (18.2) | .603 |
| Time to spike, y | 1.53 (0.86–3.53) | 3.52 (2.53–7.97) | 0.61 (0.48–0.73) | .200 |
| Immunology | | | | |
| At diagnosis | | | | |
| CD4 count | 1515 (637–2235) | 1510 (643–2120) | 1621 (827–2955) | .721 |
| CD4 % | 31.0 (18.0–38.0) | 34.0 (18.0–38.0) | 30.5 (25.2–39.0) | 1.000 |
| CD8 count | 1406 (720–2170) | 985 (639–1931) | 2170 (2115–3226) | .077 |
| CD8 % | 32.0 (25.0–40.0) | 29.0 (22.5–40.0) | 35.0 (33.0–47.0) | .157 |
| At ART start | | | | |
| CD4 count | 1310 (478–2376) | 948 (478–2284) | 1506 (988–2965) | .590 |
| CD4 % | 30.5 (19.2–42.5) | 27.0 (19.0–41.0) | 33.0 (30.0–45.0) | .482 |
| CD8 count | 1068 (795–2129) | 960 (695–1688) | 2115 (1380–2170) | .166 |
| CD8 % | 31.0 (26.5–40.0) | 30.0 (25.0–40.0) | 34.0 (30.8–36.5) | .500 |
| CD4/CD8 | 0.84 (0.60–1.57) | 0.82 (0.59–1.98) | 0.89 (0.64–0.94) | .749 |
| At viral suppression | | | | |
| CD4 count | 1710 (924–2263) | 1680 (891–2120) | 2028 (1535–2735) | .223 |
| CD4 % | 31.0 (25.4–44.0) | 30.0 (23.5–42.5) | 39.5 (30.0–46.2) | .161 |
| CD8 count | 1064 (826–1732) | 1064 (795–1716) | 1297 (878–1715) | .780 |
| CD8 % | 27.5 (18.5–32.2) | 28.0 (19.0–32.0) | 24.0 (18.0–32.0) | .784 |
| At reservoir analysis | | | | |
| CD4 count | 872 (646–1053) | 877 (719–1043) | 630 (452–1080) | .440 |
| CD4 % | 41.0 (33.8–46.2) | 41.0 (36.0–46.0) | 39.0 (32.0–46.5) | .606 |
| CD8 count | 584 (402–708) | 636 (434–712) | 426 (292–642) | .142 |
| CD8 % | 26.0 (23.0–30.0) | 26.0 (23.0–30.0) | 25.0 (23.0–30.0) | .761 |
| CD4/CD8 | 1.54 (1.28–1.88) | 1.54 (1.31–1.88) | 1.54 (1.28–2.04) | .774 |
| Serology at reservoir analysis (n = 39) | | n = 29 | n = 10 | |
| Anti-CMV IgG, No. (%) | | | | 1.000 |
| Negative | 10 (25.6) | 8 (27.6) | 2 (20.0) | |
| Positive | 29 (74.4) | 21 (72.4) | 8 (80.0) | |
| HIV Ag/Ab at reservoir analysis, No. (%) | | | | .087 |
| Equivocal/positive | 29 (74.4) | 24 (82.8) | 5 (50.0) | |
| Nonreactive | 10 (25.6) | 5 (17.2) | 5 (50.0) | |
| Ag/Ab | 4.47 (0.74–26.8) | 6.02 (2.05–27.8) | 1.20 (0.10–3.21) | .087 |
| Western blot score | 1.0 (0.5–2.0) | 1.5 (0.5–2.0) | 0.75 (0.0–1.75) | .269 |

Data are presented as median (IQR) unless otherwise indicated.
Abbreviations: Ag/Ab, antigen/antibody; ART, antiretroviral therapy; CMV, cytomegalovirus; HIV, human immunodeficiency virus; IgG, immunoglobulin G; PBMCs, peripheral blood mononuclear cells; VL, viral load.



Downloaded from https://academic.oup.com/jpids/article-abstract/doi/10.1093/jpids/piaa071/5873261 by guest on 31 July 2020

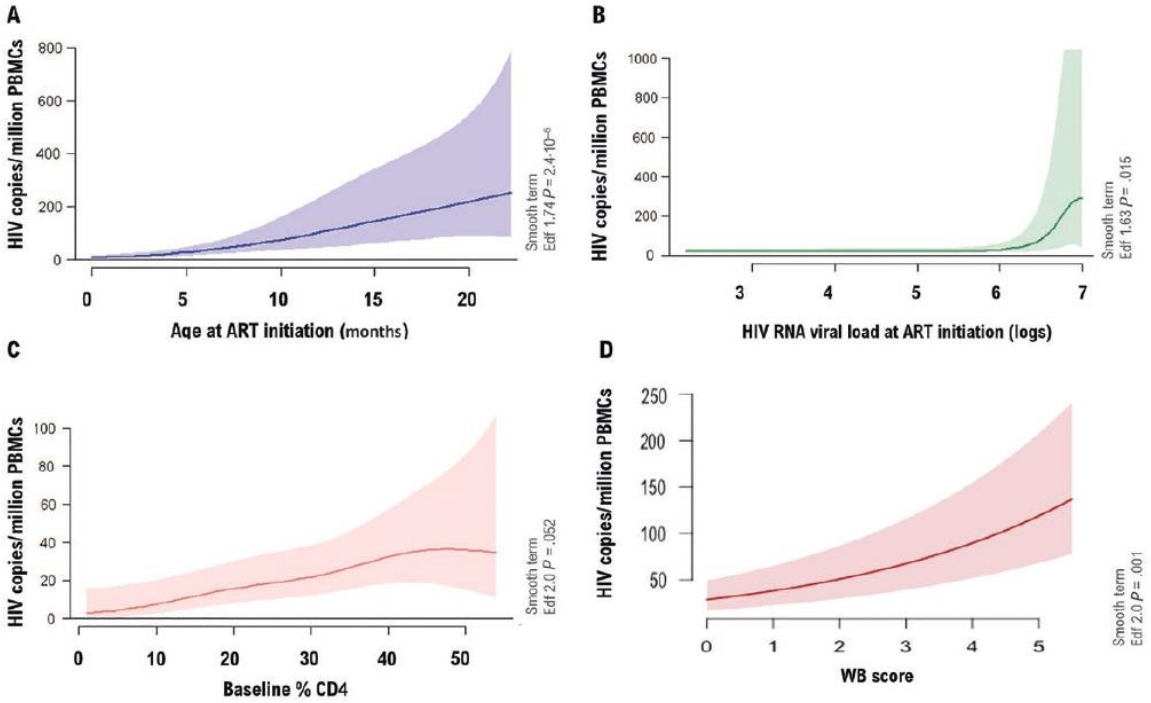



**Figure 1.** Generalized additive mixed multivariable model plots showing the association between the age at antiretroviral therapy (ART) (*A*), human immunodeficiency virus (HIV) RNA viral load (VL) at ART initiation (*B*), baseline % CD4 (*C*), and Western blot (WB) score (*D*) and the HIV reservoir size (HIV copies per million peripheral blood mononuclear cells [PBMCs]). ART combinations at initiation were included as random intercepts. The shaded areas indicate the 95% confidence intervals. Smooth terms are specified showing the effective degrees of freedom (Edf) for each of the variables included in the model. While age at ART initiation and WB score presented a linear association, the effect of VL was only noticeable for individuals with VLs at ART initiation above 6 logs. The effect of baseline CD4% was not maintained above 40%.

[14, 15]. Perinatally HIV-1–infected infants have higher plasma VLs compared with older children and, despite higher CD4%, a longer time to viral suppression [16]. In a perinatal cohort from Europe and Thailand of >400 infants initiating therapy before 6 months of age, younger age, higher CD4%, and lower VL were independent predictors of viral suppression on multivariate analysis [17]. Starting ART even earlier, within the first 7 days of life, is associated with more rapid viral suppression [18]. The rate of viral suppression following ART initiation and length of time on suppressive ART has been shown to be associated with a reduced reservoir size, with a recent perinatal European cohort demonstrating a 10% reduction in total HIV-1 DNA for every year suppressed [11]. Reservoir size in children has been shown to continue to decline for more than a decade, but in adults appears to plateau after around 4 years of suppressive therapy [13, 19].

One-quarter of participants in our cohort are HIV-1 seronegative, 12 years after initiating suppressive ART, a rate comparable to other early-treated perinatal cohorts [9]. Previous studies have reported a correlation between earlier age at ART initiation and subsequent seronegativity (seropositivity odds ratio, 13.7 [95% confidence interval, 3.1–60.2]; $P = .001$) ART initiated 12–24 weeks vs <12 weeks) and with lower pretreatment VL [8, 9], although the latter association was not supported by our data, possibly due to an inadequate sample size and treatment initiation at a later stage (median age, 4 months) than in other cohorts [20, 21]. However higher Western blot score in multivariant analysis was strongly related to reservoir size, as reported in other cohorts [13, 21–23]. The close correlation found between the HIV-DNA and the magnitude of HIV-specific Ab response is consistent with the hypothesis that persistence of specific Abs is dependent on the HIV viral antigen exposure. Infancy represents a period of relative immunological immaturity characterized by the presence of maternal Abs in the infant's blood that may hide the antigens. Suppression of viral replication through early ART in this period of life gives these children distinct HIV-specific humoral responses, which persists over time even after long-term viral suppression [24]. Quantitative HIV-1–specific antibodies have been shown to correlate and predict plasma HIV-1 RNA and cell-associated DNA levels [13, 23] and may potentially be an adjunctive tool to screen children, particularly in low-income settings, using small blood volumes for future





Downloaded from https://academic.oup.com/jpids/article-abstract/doi/10.1093/jpids/piaa071/5873261 by guest on 31 July 2020

intervention studies aimed at ART-free remission [23]. It remains unclear as to whether seronegativity will confer advantage in future treatment interruption studies in children, or potentially increase susceptibility to uncontrolled viral replication.

Our study has several limitations, being a cross-sectional analysis of reservoir dynamics in children and adolescents commencing ART a decade ago with limited numbers reducing analytical power. However, this reflects the clinical reality; very small numbers of children who start ART in infancy achieve and consistently maintain viral control for many years. For this reason, the study entry criteria were broad: ART initiation before the age of 2 years with viral suppression for at least 5 years. Yet of >400 children aged 7–18 years in pediatric care in 7 European centers, <1 in 10 fulfilled the criteria. The multicenter, multicountry enrollment, the length of sustained viral suppression, and pretreatment VL data frequently absent in perinatal cohorts adds weight to our data. Although quantification of total HIV-1 DNA overestimates the size of the reservoir, lack of universal standards make between-study comparisons complex, particularly as many data sets are single-cohort studies. Subsequent immunological and virological secondary endpoint analyses in the CARMA cohort are ongoing; cell-associated HIV-1 RNA and DNA, near full-length single-genome sequencing to infer on intact provirus, and the association of natural killer cell function with HIV-1 DNA and T-cell markers are to be presented at the Conference on Retroviruses and Opportunistic Infections 2020 (www.croiconference.org/abstracts/search-abstracts/). Following this initial cross-sectional analysis, the CARMA participants are being offered enrollment in a longitudinal follow-on study, looking at changing reservoir dynamics as they enter early adult life. Additional cohorts of early-treated children on sustained suppressive ART from South Africa and Thailand are being invited to join the CARMA cohort to explore potential differences in high-prevalence settings.

Despite variations between studies, time to ART start, pretreatment VL, and length of viral suppression remain consistent markers associated with the likelihood of a smaller reservoir size. Children and adolescents who started ART in the first months of life, with a lower VL and who have remained suppressed, may be preferred candidates for future intervention studies aiming for ART-free remission.

## CONCLUSIONS

One-quarter of early-treated, perinatally infected children have low reservoir 12 years after enrollment. Initiation of ART at a younger age and lower plasma VL were associated with a lower HIV-1 viral reservoir in PBMCs a decade later in the context of sustained virological suppression. While the clinical benefits of early-infant ART are profound, whether a smaller latent reservoir can confer future ART-free advantage remains an elusive question for a population that currently faces a lifetime on ART.

## Supplementary Data

Supplementary materials are available at *Journal of the Pediatric Infectious Diseases Society* online.

## Notes

***Acknowledgments.*** The writing group would like to thank all of the young people who contributed their time and blood samples to the Child and Adolescent Reservoir Measurements on Early Suppressive ART (CARMA) study. The writing group would also like to thank Viviana Gianuzzi, Annalisa Landi, Francesca Rocchi, Maria Grazia Lain, Alessandra Nardone, and Silvia Faggion for their support in the trial regulatory aspects.

***Financial support.*** CARMA was supported by the EPIICAL (Early-Treated Perinatally HIV-Infected Individuals: Improving Children's Actual Life With Novel Immunotherapeutic Strategies) project, funded through an independent grant by ViiV Healthcare United Kingdom. This work is part of the EPIICAL project (http://www.epiical.org/), supported by the PENTA-ID foundation (http://penta-id.org/).

Recruiting site Partners are listed in Supplementary Appendix 1, Virology and Immunology Partners are listed in Supplementary Appendix 2 and the members of the EPIICAL Consortium are listed in Supplementary Appendix 3.

***Potential conflicts of interest.*** All authors: No reported conflicts of interest. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

## References

1. Klein N, Palma P, Luzuriaga K, et al. Early antiretroviral therapy in children perinatally infected with HIV: a unique opportunity to implement immunotherapeutic approaches to prolong viral remission. Lancet Infect Dis **2015**; 15:1108–14.
2. Foster C, Pace M, Kaye S, et al; CHERUB Investigators. Early antiretroviral therapy reduces HIV DNA following perinatal HIV infection. AIDS **2017**; 31:1847–51.
3. Tagarro A, Chan M, Zangari P, et al; EPIICAL Consortium. Early and highly suppressive ART are the main factors associated with low viral reservoir in European perinatally HIV infected children. JAIDS **2018**; 79:269–76.
4. Persaud D, Gay H, Ziemniak C, et al. Absence of detectable HIV-1 viremia after treatment cessation in an infant. N Engl J Med **2013**; 369:1828–35.
5. Pasvol TJ, Foster C, Fidler S. Novel therapies/hopes for HIV cure in perinatally acquired HIV-positive adolescents. Curr Opin HIV AIDS **2018**; 13:281–7.
6. Muenchhoff M, Prendergast AJ, Goulder PJ. Immunity to HIV in early life. Front Immunol **2014**; 5:391.
7. Ananworanich J, Puthanakit T, Suntarattiwong P, et al; HIV-NAT 194 Study Group. Reduced markers of HIV persistence and restricted HIV-specific immune responses after early antiretroviral therapy in children. AIDS **2014**; 28:1015–20.
8. Payne H, Mkhize N, Otwombe K, et al. Reactivity of routine HIV antibody tests in children who initiated antiretroviral therapy in early infancy as part of the Children with HIV Early Antiretroviral Therapy (CHER) trial: a retrospective analysis. Lancet Infect Dis **2015**; 15:803–9.
9. Fidler K, Foster C, Lim E, et al. Reactivity of routine HIV antibody tests in children with perinatally-acquired HIV-1 in England: cross sectional analysis. Pediatr Infect Dis J **2019**; 38:146–8.
10. Palma P, Foster C, Rojo P, et al. The EPIICAL project: an emerging global collaboration to investigate immunotherapeutic strategies in HIV-infected children. J Virus Erad **2015**; 1:134–9.
11. Tagarro A, Chan M, Zangari P, et al. Early and highly suppressive antiretroviral therapy are main factors associated with low viral reservoir in European perinatally HIV-infected children. J Acquir Immune Defic Syndr **2018**; 79:269–76.
12. Rocca S, Zangari P, Cotugno N, et al. Human immunodeficiency virus (HIV) antibody repertoire estimates reservoir size and time of antiretroviral therapy initiation in virally suppressed perinatally HIV-infected children. J Pediatric Infect Dis Soc **2019**; 8:433–8.
13. Luzuriaga K, Tabak B, Garber M, et al. HIV type 1 (HIV-1) proviral reservoirs decay continuously under sustained virologic control in HIV-1-infected children who received early treatment. J Infect Dis **2014**; 210:1529–38.
14. O'Connor J, Smith C, Lampe FC, et al; UK Collaborative HIV Cohort (CHIC) Study. Durability of viral suppression with first-line antiretroviral therapy in patients with HIV in the UK: an observational cohort study. Lancet HIV **2017**; 4:e295–302.



Downloaded from https://academic.oup.com/jpids/article-abstract/doi/10.1093/jpids/piaa071/5873261 by guest on 31 July 2020

15. McKinnon E, Castley A, Payne L, et al. Determinants of residual viraemia during combination HIV treatment: impacts of baseline HIV RNA levels and treatment choice. HIV Med **2016**; 17:495–504.
16. Ásbjörnsdóttir KH, Hughes JP, Wamalwa D, et al. Differences in virologic and immunologic response to antiretroviral therapy among HIV-1-infected infants and children. AIDS **2016**; 30:2835–43.
17. European Pregnancy and Paediatric HIV Cohort Collaboration (EPPICC) and Early-Treated Perinatally HIV-Infected Individuals: Improving Children's Actual Life With Novel Immunotherapeutic Strategies (EPIICAL) Study Groups. Predictors of faster virological suppression in early treated infants with perinatal HIV from Europe and Thailand. AIDS **2019**; 33:1155–65.
18. Tagarro A, Dominguez-Rodriguez S, Palma P, et al; EPIICAL Consortium. Neonatal ART <7 days versus 7–28 days reduced time to viral suppression. In: Conference on Retroviruses and Opportunistic Infections, Seattle, WA, 4–7 March 2019. Abstract 44.
19. Uprety P, Patel K, Karalius B, et al; Pediatric HIV/AIDS Cohort Study (PHACS). Human immunodeficiency virus type 1 DNA decay dynamics with early, long-term virologic control of perinatal infection. Clin Infect Dis **2017**; 64:1471–8.
20. Kuhn L, Schramm DB, Shiau S, et al. Young age at start of antiretroviral therapy and negative HIV antibody results in HIV-infected children when suppressed. AIDS **2015**; 29:1053–60.
21. Martínez-Bonet M, Puertas MC, Fortuny C, et al. Establishment and replenishment of the viral reservoir in perinatally HIV-1-infected children initiating very early antiretroviral therapy. Clin Infect Dis **2015**; 61:1169–78.
22. Persaud D, Patel K, Karalius B, et al; Pediatric HIV/AIDS Cohort Study. Influence of age at virologic control on peripheral blood human immunodeficiency virus reservoir size and serostatus in perinatally infected adolescents. JAMA Pediatr **2014**; 168:1138–46.
23. McManus M, Henderson J, Gautam A, et al; PACTG 356 Investigators. Quantitative human immunodeficiency virus (HIV)-1 antibodies correlate with plasma HIV-1 RNA and cell-associated DNA levels in children on antiretroviral therapy. Clin Infect Dis **2019**; 68:1725–32.
24. Cotugno N, Morrocchi E, Rinaldi S, et al; EPIICAL Consortium. Early antiretroviral therapy-treated perinatally HIV-infected seronegative children demonstrate distinct long-term persistence of HIV-specific T-cell and B-cell memory. AIDS **2020**; 34:669–80.

Downloaded from https://academic.oup.com/jpids/article-abstract/doi/10.1093/jpids/piaa071/5873261 by guest on 31 July 2020

# Capítulo IV:

## Modelo de predicción usando aprendizaje automatizado para mortalidad y progresión clínica a SIDA en niños perinatalmente infectados con VIH-1

## Justificación y Objetivos

A pesar del éxito que supuso el TAR y las medidas de prevención de la transmisión vertical madre-hijo, todavía existe una tasa de infecciones por VIH en niños y una alta tasa de mortalidad y morbilidad en los primeros 3-6 meses desde el inicio del TAR. En África Subsahariana las cifras de mortalidad alcanzan un 21% en Sudáfrica y 26% en Mozambique (62).

Diversos factores han sido asociados a un incremento en la mortalidad o progresión a SIDA en niños con VIH-1 en TAR. Aquellos más reportados son: una avanzada enfermedad al diagnóstico (63), mayor edad al inicio del TAR (64), cargas virales altas al inicio del TAR (65), menor puntuación z en peso para la edad (64), menor porcentaje de CD4 (65) y sexo femenino (64). Frecuentemente estos factores de riesgo fueron reportados mediante aproximaciones tradicionales y más comúnmente utilizadas como la regresión logística. Esto es mayoritariamente por ser un modelo que no demanda conocimientos computacionales o avanzados conocimientos estadísticos.

La inteligencia artificial, en concreto el aprendizaje automatizado, está extensamente usado en varios campos de la medicina por su gran poder de predicción y habilidad para manejar complejos conjuntos de datos (66). Estos sistemas están basados en generar modelos de predicción mediante el estudio o aprendizaje de un conjunto de datos donde la respuesta es conocida y seguidamente su posterior validación en otro conjunto de datos donde la respuesta es desconocida. A diferencia de la regresión logística, estos modelos permiten incluir un amplio número de predictores, permiten la correlación entre predictores y además detectan relaciones no lineales complejas en los datos (67).

A pesar de ello, la aplicación de estos modelos en el campo del VIH-1 es limitada o no es accesible para los clínicos.

Los objetivos de este estudio son:

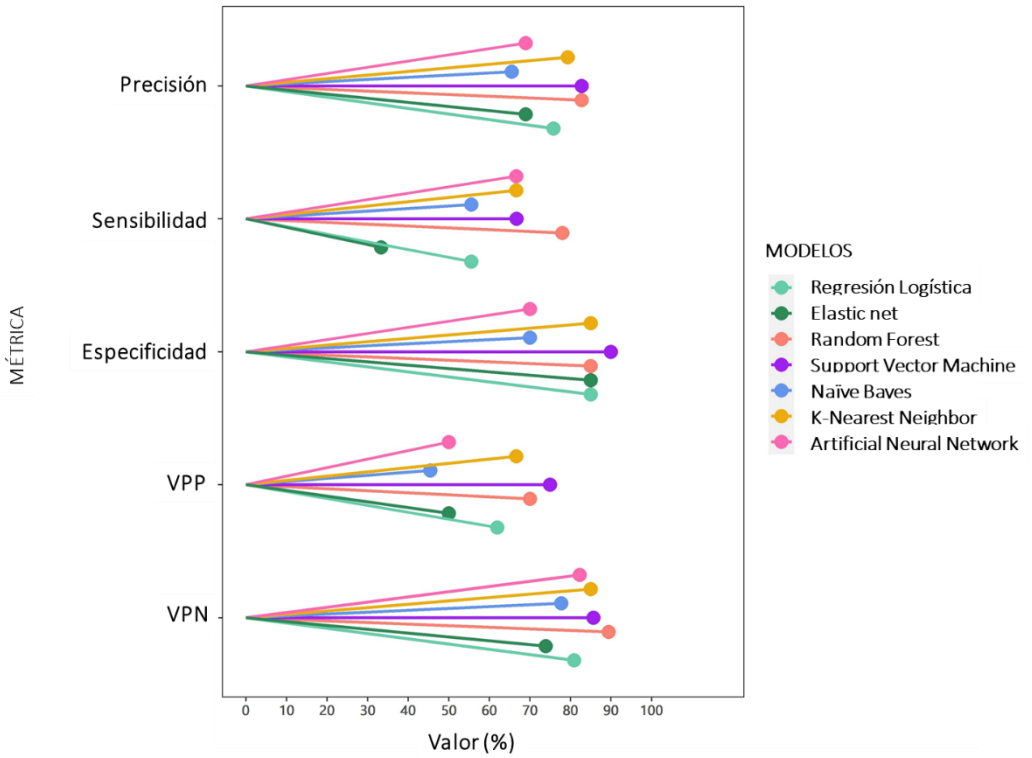
- Comparar siete algoritmos de predicción para seleccionar el de mayor rendimiento para predecir mortalidad y/o progresión a SIDA en niños africanos con infección vertical por VIH-1 que comenzaron TAR muy precozmente utilizando los factores más ampliamente asociados a mortalidad/progresión en la literatura.

- Implementar el algoritmo de mayor rendimiento en una aplicación web gratuita para que pueda utilizarse por los clínicos

## Resultados

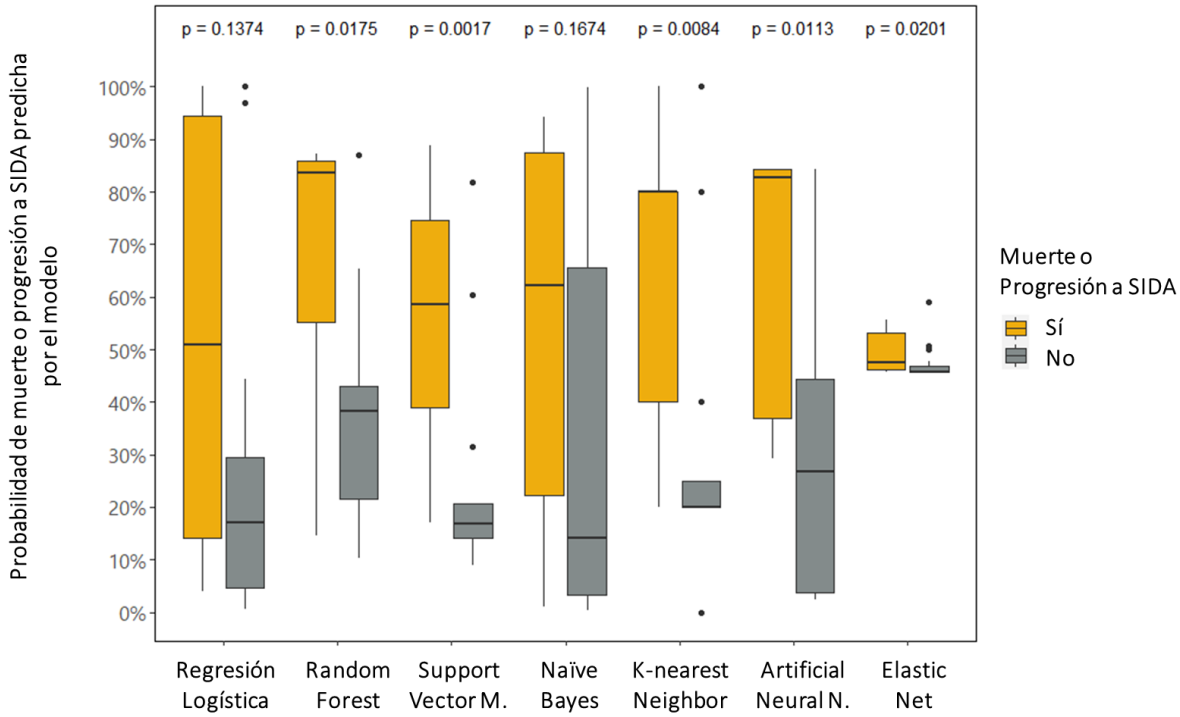
- Del total de los 100 niños incluidos en el estudio, 33 (33%) murieron o progresaron a SIDA en el primer año de seguimiento. Concretamente, 22% murieron, 14% progresaron y 3% progresaron y luego murieron.
- Los pacientes que murieron o progresaron a SIDA tuvieron menor puntuación z de peso para la edad (-1,35 (RIC, -2,10 - -0,36) *vs.* -1,96 (RIC, -2,86- -1,17), p-valor=0,050), mayores cargas virales al inicio del seguimiento (5,13 logs (RIC, 3,82-6,1) *vs.* 6,35 (RIC, 5,59-6,85), p-valor<0,001), fueron diagnosticados de VIH-1 a mayor edad (1,0 días de vida (RIC, 0-32) *vs.* 33 (RIC, 29-45), p-valor=0,002) y empezaron el TAR más tarde (4 meses (RIC, 0-36) *vs.* 33 (RIC, 30-48), p-valor=0,006). Sin embargo, no hubo diferencias en la distribución de sexos, CD4 o problemas sociales o de salud de las madres.
- El algoritmo RF junto con SVM presentaron una precisión mayor en la clasificación de los pacientes que murieron o progresaron a SIDA que el resto de los algoritmos (RF: 82,8%, SVM: 82,8%, KNN:79,3%, RL: 75·9%, Elastic net: 69%, ANN: 69%, NB:65,5%).
- El algoritmo RF fue el que presentó mayor sensibilidad (RF:78%, SVM: 66,7%, KNN: 66,7%, ANN: 66,7%, NB: 55,6%, RL:55,6% y Elastic net: 33,3%) y alta especificidad (SVM:90%, RF:85%, Elastic net: 85%, RL: 85%, ANN: 70% y NB: 70%).
- El algoritmo RF junto con SVM presentaron un poder predictivo positivo alto (75% y 70%) así como un poder predictivo negativo alto (89,5% y 85,7%).

Figura 6. Rendimiento de los modelos utilizando los diferentes algoritmos




- El algoritmo RF presentó el área bajo la curva mayor (73,2% (67,2-79,1)).
- En el conjunto de validación se obtuvieron las probabilidades de muerte/progresión arrojadas por el modelo y se compararon entre aquellos pacientes que realmente murieron/progresaron durante el año de seguimiento para describir la capacidad de clasificación. El algoritmo más comúnmente utilizado en la investigación biomédica, la regresión logística, no logró clasificar significativamente los pacientes (no hubo diferencias en las probabilidades reales de muerte/progresión en los pacientes de los 2 grupos creados por el modelo de RL) (p-valor=0.137). Sin embargo, el modelo RF sí que pudo clasificar significativamente a los pacientes (p-valor=0.017).

Figura 7. Probabilidad de muerte o progresión a SIDA predicha por los diferentes modelos en el conjunto de datos de validación N=29.




La probabilidad de muerte/progresión arrojada por los diferentes modelos en el conjunto de datos de validación fue comparada entre los pacientes que murieron o progresaron (amarillo) *versus* los que no lo hicieron (gris) mediante una prueba U-Mann Whitney.

El modelo de predicción basado en RF fue implementado en la aplicación web MAGPIE a la cual se puede acceder desde el link https://rserver.h12o.es/pediatria/MAGPIE/ (username: user, password:0000).

The Pediatric Infectious Disease Journal

# Machine learning outperformed logistic regression classification even with limit sample size: A model to predict pediatric HIV mortality and clinical progression to AIDS

--Manuscript Draft--

| | |
|---|---|
| **Manuscript Number:** | PIDJ-221-172 |
| **Article Type:** | HIV Reports |
| **Corresponding Author:** | Sara Dominguez<br>Hospital Universitario Ramon y Cajal<br>Madrid, SPAIN |
| **First Author:** | Sara Dominguez |
| **Order of Authors:** | Sara Dominguez |
| | Miquel Serna-Pascual |
| | Andrea Oletto |
| | Shaun Barnabas |
| | Peter Zuidewind |
| | Els Dobbels |
| | Siva Danaviah |
| | Osee Behuhuma |
| | Maria Grazia Lain |
| | Paula Vaz |
| | Sheila Fernández-Luis |
| | Tacilta Nhampossa |
| | Elisa Lopez-Varela |
| | Kennedy Otwombe |
| | Afaaf Liberty |
| | Avy Violari |
| | Almoustapha Issiaka Maiga |
| | Paolo Rossi |
| | Carlo Giaquinto |
| | Louise Kuhn |
| | Pablo Rojo |
| | Alfredo Tagarro |
| **Manuscript Region of Origin:** | SPAIN |
| **Abstract:** | Background<br><br>Logistic regression (LR) is the most common prediction model in medicine. In recent years, supervised machine learning (ML) methods have gained popularity. However, there are many concerns about ML utility for small sample sizes. In this study, we aim to compare the performance of 7 algorithms in the prediction of 1-year mortality and clinical progression to AIDS in a small cohort of infants living with HIV from South Africa and Mozambique.<br><br>Methods<br><br>The data set (n=100) was randomly split into 70% training and 30% validation set. |

Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation

| | |
|---|---|
| | Seven algorithms (LR, Random Forest (RF), Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Naïve Bayes (NB), Artificial Neural Network (ANN), and Elastic Net) were compared. The variables included as predictors were the same across the models including sociodemographic, virologic, immunologic, and maternal status features. For each of the models, a parameter tuning was performed to select the best-performing hyperparameters using 5 times repeated 10-fold cross-validation. A confusion-matrix was built to assess their accuracy, sensitivity, and specificity.<br><br>Results<br><br>RF ranked as the best algorithm in terms of accuracy (82,8%), sensitivity (78%), and AUC (0,73). Regarding specificity and sensitivity, RF showed better performance than the other algorithms in the external validation and the highest AUC. LR showed lower performance compared with RF, SVM, or KNN.<br><br>Conclusions<br><br>The outcome of children living with perinatally acquired HIV can be predicted with considerable accuracy using ML algorithms. Better models would benefit less specialized staff in limited resources countries to improve prompt referral in case of high-risk of clinical progression. |
| **Funding Information:** | PENTA Foundation — Not applicable |

Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation

# ABSTRACT

## Background

Logistic regression (LR) is the most common prediction model in medicine. In recent years, supervised machine learning (ML) methods have gained popularity. However, there are many concerns about ML utility for small sample sizes. In this study, we aim to compare the performance of 7 algorithms in the prediction of 1-year mortality and clinical progression to AIDS in a small cohort of infants living with HIV from South Africa and Mozambique.

## Methods

The data set (n=100) was randomly split into 70% training and 30% validation set. Seven algorithms (LR, Random Forest (RF), Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Naïve Bayes (NB), Artificial Neural Network (ANN), and Elastic Net) were compared. The variables included as predictors were the same across the models including sociodemographic, virologic, immunologic, and maternal status features. For each of the models, a parameter tuning was performed to select the best-performing hyperparameters using 5 times repeated 10-fold cross-validation. A confusion-matrix was built to assess their accuracy, sensitivity, and specificity.

## Results

RF ranked as the best algorithm in terms of accuracy (82,8%), sensitivity (78%), and AUC (0,73). Regarding specificity and sensitivity, RF showed better performance than the other algorithms in the external validation and the highest AUC. LR showed lower performance compared with RF, SVM, or KNN.

## Conclusions

The outcome of children living with perinatally acquired HIV can be predicted with considerable accuracy using ML algorithms. Better models would benefit less specialized staff in limited resources countries to improve prompt referral in case of high-risk of clinical progression.

## INTRODUCTION

There is still a high rate of mortality in children living with HIV in the first months after antiretroviral (ART) initiation in Sub-Saharan Africa. These rates reported being as high as 21% in South Africa and 26% in Mozambique [1].

Several risk factors are associated with increased childhood mortality and clinical progression to AIDS. Advanced disease at HIV diagnosis [2], older age at ART initiation [3], higher baseline HIV viral load [4], low weight-for-age Z-score,[3], low % CD4 [4], and female sex [3] are the factors reported most. However, the risk models that yielded these risk factors were built using logistic regression (LR) models, considered the standard and most used approach for binary classification in the context of low-dimensional data.

Modern ML algorithms use statistical, probabilistic and optimization methods to detect patterns or rules from large and complex datasets [5]. Specifically, supervised ML systems create prediction models by studying a dataset where the outcome is known and subsequently validate this model in a dataset where the outcome is unknown. Unlike classical multivariable LR, ML allows the inclusion of a large number of predictors, that may exceed the number of observations, and allows the combination of these predictors in flexible linear and nonlinear ways [6]. By relaxing model assumptions, ML may improve risk classification.

Despite its advantages, the application of ML other than LR in the HIV field has been limited [7–11]. Several studies have claimed the better performance of machine learning versus logistic regression in other medical fields [12–14]. Despite these advantages, many researchers question the true benefit of these methodologies, as they rightly point out that large amounts of data are usually required to yield accurate results with modern ML algorithms.[15]. However, the performance was only measured in terms of area under the curve (AUC), a metric that can be misleading to test the performance of a model [16].

Interpretability of ML results has also been a source of concern for medical researchers, as in many cases, complex data analyses need to be performed requiring high-level programming skills and requiring complex interpretations. However, the implementation of automated ML in easy packages have emerged in recent years helping non-experts to use ML off-the-shelf [17].

To our knowledge, no study has addressed whether novel ML algorithms outperform traditional models in limited sample sizes in a validation set assessing not only AUC but other performing metrics including the discrimination ability of several algorithms in terms of probability class given by the model.

In this study, we did not aim to explain or describe the course of infection in these children but selecting the best predicting algorithm.

## METHODS

### Study population

This analysis was performed within an international prospective multicenter Early Antiretroviral Treatment in HIV Children (EARTH) Cohort, embedded in the EPIICAL Consortium (Early treated Perinatally HIV Infected individuals: Improving Children's Actual Life Project). Patients were recruited from South Africa (Africa Health Research Institute (AHRI), Durban; Family Centre For Research With Ubuntu (FAMCRU), Stellenbosch; Perinatal HIV Research Unit (PHRU), Soweto) and Mozambique (Centro de Investigación en Salud de Manhiça (CISM) – ISGlobal, Manhiça; Fundação Ariel Glaser contra o SIDA Pediatrico, Mozambique). Inclusion criteria included infants living with perinatally acquired HIV who begin ART before 3 months of age, or breastfed infants diagnosed with HIV under three months of age that started ART within three months of diagnosis were enrolled in the study.

In this analysis, we included a total of 100 children with a minimum follow-up time of 12 months and had data on viral load (VL) or percentage of CD4 T-cells (%CD4) available at enrollment. Infants and mothers' demographic, clinical, virological, and immunological variables at enrollment were included as possible predictors in the dataset. The combined primary outcome was mortality or clinical progression to AIDS 12 months after enrolment. Clinical progression to AIDS was defined as having a clinical WHO AIDS stage III or IV events after enrolment visit.

The variables included as predictors were the same for all the models: age at recruitment, sex, weight-for-age (WAZ) z-score at enrolment calculated using the WHO Growth Reference [18],

preterm birth, age at HIV diagnosis, age at ART initiation, initiation ART regimen, pre-ART VL, baseline %CD4, mother severe life events (including change in employment, separation or relationship break-up, new partner, loss of home or move, or death in the family) or health issues, mother's adherence to ART at enrolment, and mother's last CD4 count and VL measurements.

The EARTH study was approved by ethics committees within each country and informed consent obtained from the parent(s)/legal guardian following country-specific law. Each participant received a unique study number, under which data were pseudo-anonymized.

## Models

Logistic Regression (LR): This algorithm is considered an extension of linear regression, used to model a dichotomous variable that usually represents the presence or absence of an event. In other words, LR can predict the likelihood of occurrence of a certain outcome [19].

Random Forest (RF): The RF algorithm is built by many decision trees, as a forest is formed by a collection of many trees. To classify a new sample, the input vector of that sample goes through all the decision trees of the forest. Each decision tree gives a classification outcome, i.e., a vote. After a large number of trees is generated, the forest chooses the majority vote for the most popular class and chooses the classification outcome [20].

Support Vector Machine (SVM): The SVM technique constructs an optimal hyperplane. In other words, a delimitation to separate the training set according to the outcome into a n-dimensional space, where n is the number of features. The distance between this hyperplane and the nearest instance is the marginal distance for a class. The optimal hyperplane is the one that maximizes this distance for both classes. When the data are not linearly separable, the hyperplane is constructed by a non-linear function (kernel) [21]. For the purposes of this research, the radial basis kernel was used.

Naïve Bayes (NB): The NB algorithm is based on the probabilistic Bayes' theorem. This classifier calculates the probability of an event based on observed values of relevant variables and prior knowledge on the frequency at which these values occur when the event takes place. The

algorithm assumes that each of the variables contributes independently to that probability. The final classification is produced by combining the prior knowledge and observations of predictor variables and comparing the likelihood of an event occurring.

K-nearest neighbour (KNN): The KNN algorithm can be considered a variant of NB classifier. Unlike NB, the KNN algorithm does not require prior probabilities. This algorithm assumes that similar events exist in close proximity. In other words, the values of their predictor variables are similar. The implementation of this algorithm is preceded by a clustering technique, used to determine this similarity and therefore the proximity of a new event to a known event.

Artificial Neural Networks (ANN): The ANNs are machine learning algorithms inspired by the neural activity of the brain. [22] ANNs are represented by nodes (analogous of neurons) grouped in matrices called layers, which interconnect to other layers to generate knowledge, just as neuronal layers in the brain. Nodes and edges (connections) have weights that are adjusted according to the frequency in which the connections are made, allowing the model (analogous of the brain itself) to recognize patterns in future datasets characteristic of a given class.

Elastic Net (EN): The EN is the most recent and optimized penalized regression algorithm used in the machine learning field. These algorithms are based on simple linear regression. However, they 'penalize' the least informative variables by reducing or even eliminating their coefficients, providing a simple and more generalizable model, which helps to reduce the dimensions of the model and its tendency to overfit training data [23].

### Training and validation of the models

All the included participants were randomly divided into a training dataset used to generate the models (70% of the original dataset), and a validation dataset used to assess the performance of the models (remaining 30%). Partitions were balanced by the outcome class, which yielded a training set with an n=71, and a validation set with an n=29. Both the training and validation data set was compared to ensure that the validation set was representative of the whole population. Chi-squared tests were used to compare categorical variables and Mann-Whitney U test to

compare continuous variables. Absolute numbers and frequencies were assessed for the categorical variables and medians and interquartile ranges [IQR] for the continuous.

For each of the novel 6 ML models, a parameter tuning was performed in the training set with 10 tune grids to select the best-performing hyperparameters of each model using 5 times repeated 10-fold cross-validation. To deal with class imbalance, a down-sampling was performed, in other words, randomly sub-setting all the classes in the training set so that their class frequencies match the least prevalent class. A p-value <0·05 was taken as statistically significant.

A confusion matrix was built for each of the models to assess their accuracy, sensitivity, and specificity. The area under the curve (AUC) of the receiver operating characteristic curves (ROCs) was also determined for each model [24]. Both apparent AUC (determined using the training set) and actual AUC (determined in the validation set) were calculated. We calculated optimism as the difference between the apparent AUC and validated AUC.

All missing values were imputed independently for the training and the testing dataset, Imputation was done by using a non-parametric approach based on Random Forest (RF) algorithm. The Out-of-bag (OOB) imputation error estimate was assessed as the normalized root mean squared error (NRMSE) and the proportion of falsely classified (PFC) outcomes.

All the analyses were performed using the R language. The caret R package [25] was used to implement the splitting of the data set, model parameter tuning, model training, and confusion matrix validation.

## RESULTS

A total of 100 children living with perinatally acquired HIV who received early treatment were included in the study. From those, 33/100 (33%) died or clinically progressed to AIDS within 12-months of follow-up. Specifically, 22% died, 14% progressed, and 3% progressed and then died. The description of the infants' features according to the primary outcome is described in Supplementary Table 1.

The distribution of each randomly divided data sets (training and testing) was compared in Table 1. Training and testing data sets did not differ significantly across the variables included in the models. In the imputation of the training data set, the NRMSE was 0·22 for continuous and the PFC 0·14 for categorical, and for the testing data set NRMSE was 0·01 and the PFC 0·0001. We did not find differences between the imputed dataset and the complete case data set (Supplementary Table 2). The tuning parameters for training each of the algorithms tested are described in Table 2.

RF and SVM presented higher accuracy than the rest of the algorithms (RF: 82·8%, SVM: 82·8%, KNN:79·3%, LR: 75·9%, GLMnet: 69%, ANN: 69%, NB:65·5%). RF was the most sensitive algorithm (RF:78%, SVM: 66·7%, KNN:66·7%, ANN: 66·7%, NB: 55·6%, LR:55·6%, and GLMnet: 33.3%). Most algorithms presented similar specificity, except ANN and NB (SVM:90%, RF:85%, GLMnet:85%, LR: 85%, ANN: 70%, and NB: 70%). SVM and RF presented the highest PPV in the validation (SVM: 75%, RF:70%). Likewise, RF (89·5%) and SVM (85·7%) presented the highest NPV in the validation. (Figure 1). Complete case analysis was performed, and RF was also the best performing model.

The RF was the algorithm with the largest AUC of the ROC curve (73.2% (67·2-79·1)) followed by the KNN algorithm (69·6% (36·2-76·1)). (Figure 2). The AUC optimism was higher in RF (0.2), LR (0.19), and NB (0.18) than in the rest of the algorithms like KNN (0.05), ANN (0.03), and SVM (0.02).

The probability of having the primary combined outcome calculated by each of the algorithms is described in Figure 3. LR and NB algorithms failed to assess significantly different probabilities for each of the outcome results. RF, SVM, KNN, ANN, and GLMnet were able to assess significantly different outcome probabilities.

## DISCUSSION

The results obtained in this study indicate that the outcome of children living with perinatally acquired HIV who received early treatment can be predicted with considerable accuracy (>80%) by using novel ML algorithms that are able to integrate clinical, virological, and immunological data. The RF algorithm outperformed in sensitivity all other tested ML methods, including LR, commonly used in the medical field, even trained in a small sample size (n=71).

Due to its flexibility, ML is broadly considered to be a better analytical tool than traditional statistical modelling, such as logistic regression. One of the main reasons why this may be is due to ML can overcome issues of multiple and correlated predictors, non-linear relationships, and interactions between predictors and endpoints that can lead to misassumptions and hyperparametric scenarios in traditional regression models [26]. However, LR has been the standard of care for binary classification. This is mainly because LR does not demand computational or statistical expertise and also because novel ML are believed to be data-hungry [15].

The limited sample size hurdle was very present in our study since acquiring hundreds or thousands of patients' samples in some fields such as pediatric HIV is rarely feasible, particularly in the case were a follow-up is needed in countries with limited resources. We have shown that ML learning approaches including RF, SVM, ANN, or KNN can offer a performance advantage over traditional LR methods even in small sample sizes in terms of accuracy, sensitivity, and specificity. Of special interest, our results have shown that the RF model yielded significantly different outcome probabilities for each class unlike the other models applied. These means that the outcome probabilities given by this model could help the physicians to identify those children who will death or progress better than with the standard LR. These findings agree with other studies where RF performed better than LR [27,28].

The AUC optimism given by the models were higher in all the models due to the limited sample size. Also, the RF, LR and NB models presented the highest differences between AUC validated and apparent AUC. However, in our study, LR did not present the lowest optimism as reported in other studies [15]. The data hungriness or performance should not be only evaluated with the AUC metric [16], which would be not the best measurement for all prediction goals including, for example,

where maximizing sensitivity could save patients' lives. We believe that with small data any algorithm will be penalized in robustness and the performance metric should be evaluated depending on the outcome of interest. In the clinical field, there are complex associations and highly correlated data, issues that RF copes better than traditional methods.

Interpretability of ML results has also been a source of concern for medical researchers, as in many cases, the complex data analyses performed do not yield clear and simple relationships between variables and outcomes. Therefore, we believe that traditional LR may be the most suitable model for studies that aim to not just predict or classify outcomes, but, rather, to measure associations between specific factors and an event.

In limited resources countries where HIV patients' care is decentralized and no specialized staff is attending children, accurate models are needed to alert the clinician if the patient is at risk of clinical progression. In this case, clinicians may consider referring the patients to a higher level of care or a different follow-up schedule.

In terms of limitations, the main drawback of this research is the limited number of observations, which likely diminishes the prediction capacity of all the trained models, designed to work with a much larger number of events. The design where the training and validation set are separated could exacerbate the limitation of sample size. This could be solved with methods like nested cross-validations in which the validation set also remained completely blind for the model. However, we decided to use a unique built-in R package as caret to reduce programming complexity and these novel methods are not implemented automatically. Further, we employed the machine learning classifiers without extensive optimizations, leading to the possibility that our results could be improved by performing these adjustments. Because of the same reason, we decided to follow the simplest way to implement modern ML.

The present study has several strengths. This study has brought ML close to the physicians by designing a digital comprehensive friendly app with direct applicability to healthcare. The collaboration of multidisciplinary staff such as physicians, bioinformaticians and biostatistician is of great value in identifying new and more efficient statistical methods to better interpret clinical data and aid clinical practice/patient's management.

In conclusion, we have shown how infant mortality and progression to AIDS in early-treated HIV-infected infants can be predicted more accurately with modern ML algorithms than with traditional LR. Specifically, the RF was the most precise algorithm in this study.

## Acknowledgments

We thank the EPIICAL Consortium team, all the staff involved in EARTH cohort, and thank to the families and participants who participate in the study. We also want to thank Angel Panizo for the artificial intelligent support.

## Declaration of interest

All the authors declare they have no conflict of interest for this study

## References


1. Ahmed I, Lemma S. Mortality among pediatric patients on HIV treatment in sub-Saharan African countries: a systematic review and meta-analysis. *BMC Public Health*. 2019;19(1):149.

2. Kabue MM, Buck WC, Wanless SR, et al. Mortality and clinical outcomes in HIV-infected children on antiretroviral therapy in Malawi, Lesotho, and Swaziland. *Pediatrics*. 2012;130(3).

3. Zanoni BC, Phungula T, Zanoni HM, et al. Risk factors associated with increased mortality among HIV infected children initiating antiretroviral therapy (ART) in South Africa. *PLoS One*. 2011;6(7).

4. Anigilaje EA, Aderibigbe SA. Mortality in a Cohort of HIV-Infected Children: A 12-Month Outcome of Antiretroviral Therapy in Makurdi, Nigeria. *Adv Med*. 2018;2018:1–11.

5. Obermeyer Z, Emanuel EJ. Predicting the future-big data, machine learning, and clinical medicine. *N Engl J Med*. 2016;375(13):1216–1219.

6. Mullainathan S, Spiess J. Machine learning: An applied econometric approach. In: *Journal of Economic Perspectives*. Vol 31. American Economic Association; 2017:87–106.

7. Balzer LB, Havlir D V, Kamya MR, et al. Machine Learning to Identify Persons at High-Risk of Human Immunodeficiency Virus Acquisition in Rural Kenya and Uganda. *Clin Infect Dis*. November 2019.

8. Zheng W, Balzer L, Vander Laan M, et al. Constrained binary classification using ensemble learning: an application to cost-efficient targeted PrEP strategies. In: *Statistics in Medicine*. Vol 37. John Wiley and Sons Ltd; 2018:261–279.

9. Krakower DS, Gruber S, Hsu K, et al. Development and validation of an automated HIV prediction algorithm to identify candidates for pre-exposure prophylaxis: a modelling study. *Lancet HIV*. 2019;6(10):e696–e704.

10. Zorn KM, Lane TR, Russo DP, et al. Multiple Machine Learning Comparisons of HIV Cell-

based and Reverse Transcriptase Data Sets. *Mol Pharm*. 2019;16(4):1620–1632.

11. Nan Y, Gao Y. A machine learning method to monitor China's AIDS epidemics with data from Baidu trends. *PLoS One*. 2018;13(7).

12. Couronné R, Probst P, Boulesteix AL. Random forest versus logistic regression: A large-scale benchmark experiment. *BMC Bioinformatics*. 2018;19(1):270.

13. Beunza J-J, Puertas E, García-Ovejero E, et al. Comparison of machine learning algorithms for clinical event prediction (risk of coronary heart disease). *J Biomed Inform*. 2019;97:103257.

14. van der Heide EMM, Veerkamp RF, van Pelt ML, et al. Comparing regression, naive Bayes, and random forest methods in the prediction of individual survival to second lactation in Holstein cattle. *J Dairy Sci*. 2019;102(10):9409–9421.

15. Van Der Ploeg T, Austin PC, Steyerberg EW. Modern modelling techniques are data hungry: A simulation study for predicting dichotomous endpoints. *BMC Med Res Methodol*. 2014;14(1):137.

16. Lobo JM, Jiménez-valverde A, Real R. AUC: A misleading measure of the performance of predictive distribution models. *Glob Ecol Biogeogr*. 2008;17(2):145–151.

17. Waring J, Lindvall C, Umeton R. Automated machine learning: Review of the state-of-the-art and opportunities for healthcare. *Artif Intell Med*. 2020;104:101822.

18. Bloem M. The 2006 WHO child growth standards. *Br Med J*. 2007;334(7596):705–706.

19. Hosmer DW, Lemeshow S, Sturdivant RX. *Applied Logistic Regression. 3rd Ed. Hoboken*. Wiley; 2013.

20. Breiman L. Random forests. *Mach Learn*. 2001;45(1):5–32.

21. Joachims, Thorsten. Making large-scale SVM learning practical. *Tech Reports*. 1998.

22. McCulloch WS, Pitts WH. A logical calculus of the ideas immanent in nervous activity. In: *Systems Research for Behavioral Science: A Sourcebook*. Taylor and Francis; 2017:93–96.

23. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw*. 2010;33(1):1–22.

24. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*. 2011;12(1):77.

25. Kuhn M. Building predictive models in R using the caret package. *J Stat Softw*. 2008;28(5):1–26.

26. Goldstein BA, Navar AM, Carter RE. Moving beyond regression techniques in cardiovascular risk prediction: applying machine learning to address analytic challenges. *Eur Heart J*. 2017;38(23).

27. Couronné R, Probst P, Boulesteix AL. Random forest versus logistic regression: A large-scale benchmark experiment. *BMC Bioinformatics*. 2018;19(1):270.

28. van der Heide EMM, Veerkamp RF, van Pelt ML, et al. Comparing regression, naive Bayes, and random forest methods in the prediction of individual survival to second lactation in Holstein cattle. *J Dairy Sci*. 2019;102(10):9409–9421.

Figure 1. Algorithms performance in the validation set

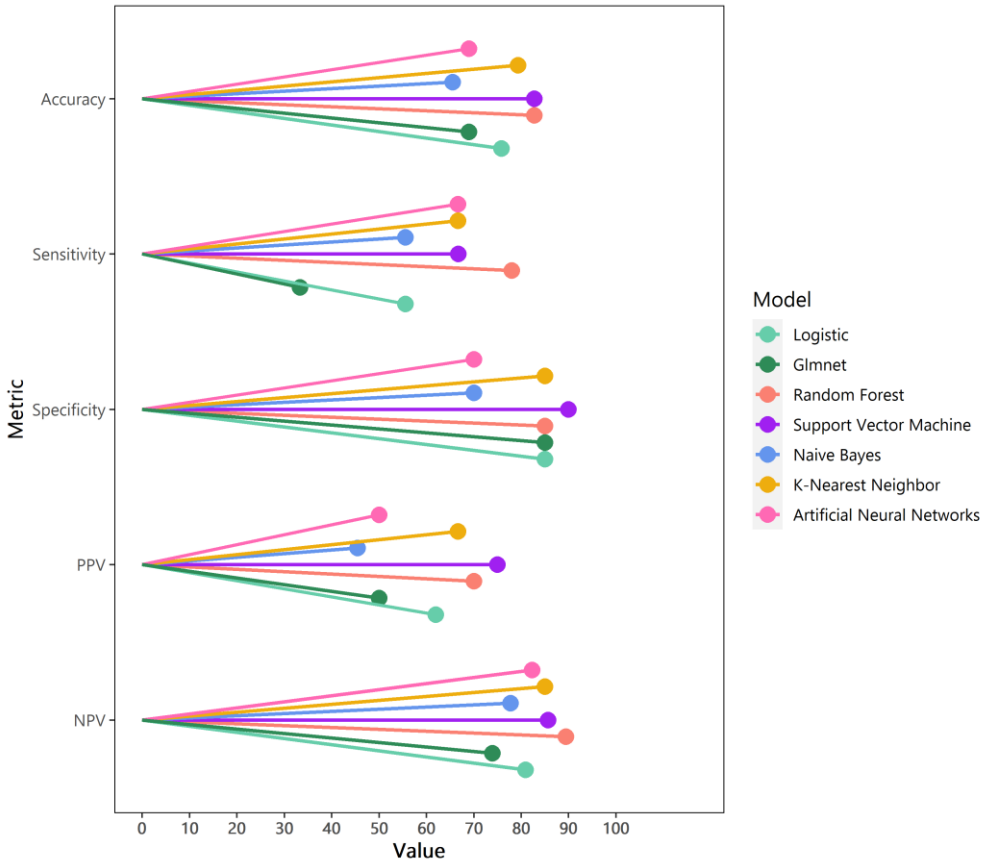


Figure 2. Algorithms receiving operating curve in the validation set

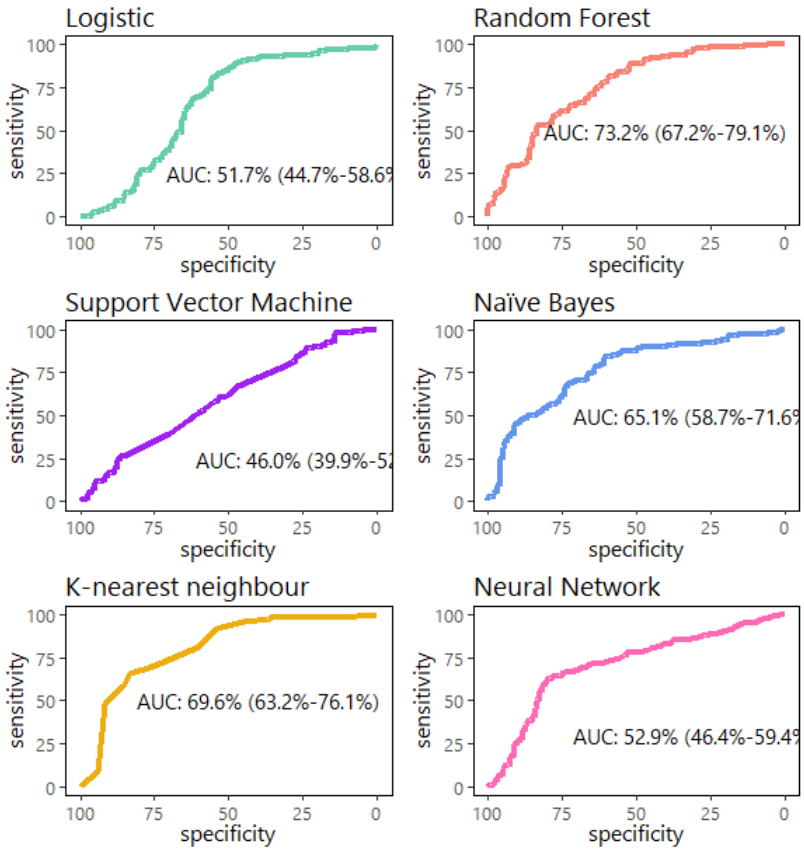


Figure 3. Probability of Death/Progression according to each algorithm in the validation set

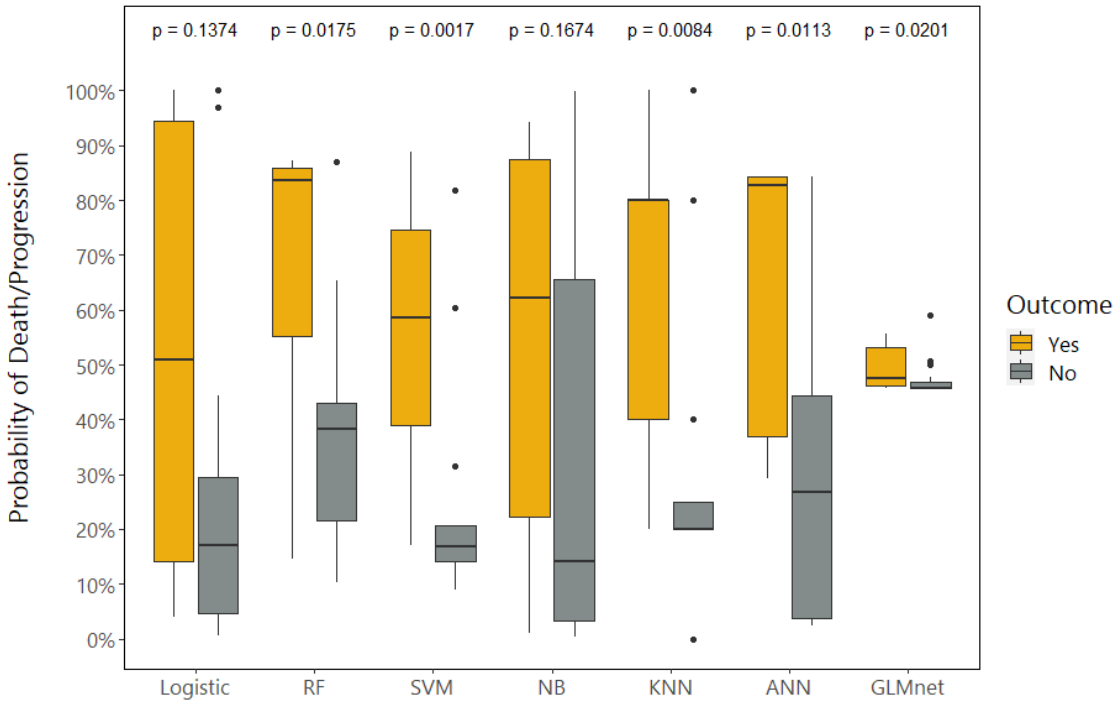


**Table 1**. Feature distribution according to the different data sets

| | **Training set** | **Testing set** | **p-value** |
|---|---|---|---|
| | ***N=71*** | ***N=29*** | |
| **Age at recruitment** | | | 0.316 |
| *Days (median, IQR)* | 36.0 [29.6;69] | 41.0 [30;89.1] | |
| **Gender** | | | 0.416 |
| Female | 35 (49.3%) | 11 (37.9%) | |
| Male | 36 (50.7%) | 18 (62.1%) | |
| **Weight-for-age at enrolment** | | | 0.805 |
| *z-score (median, IQR)* | -1.46 [-2.62;-0.87] | -1.18 [-2.98;-0.30] | |
| **Preterm birth** | | | 0.140 |
| No | 41 (57.7%) | 22 (75.9%) | |
| Yes | 30 (42.3%) | 7 (24.1%) | |
| **Age at HIV diagnosis** | | | 0.563 |
| *Days (median, IQR)* | 30.0 [0.00;35.5] | 31.0 [0.00;50.0] | |
| **Age at ART** | | | 0.195 |
| *Days (median, IQR)* | 32.0 [18.5;62.5] | 36.0 [23.0;82.0] | |
| **Initial ART regimen** | | | 0.558 |
| 3TC+ABC+LPVr | 33 (46.5%) | 14 (48.3%) | |
| 3TC+ABC+NVP | 0 (0.00%) | 1 (3.45%) | |
| 3TC+AZT+LPVr | 22 (31.0%) | 8 (27.6%) | |
| 3TC+AZT+NVP | 16 (22.5%) | 6 (20.7%) | |
| **Baseline viral load** | | | 0.350 |
| *Copies/mL (median, IQR)* | 609715 [36738;2570245] | 226844 [36295;1344319] | |
| **Baseline % CD4** | | | 0.587 |
| *Cell/mL (median, IQR)* | 36.9 [29.9;45.2] | 40.0 [28.0;47.0] | |
| **Maternal severe life events or health issues** | | | 1.000 |
| No | 34 (47.9%) | 14 (48.3%) | |
| Yes | 37 (52.1%) | 15 (51.7%) | |
| **Maternal adherence (self-reported at enrolment)** | | | 0.589 |
| Poor | 3 (4.23%) | 3 (10.3%) | |
| Intermediate low | 12 (16.9%) | 4 (13.8%) | |
| Intermediate high | 18 (25.4%) | 5 (17.2%) | |
| Good | 38 (53.5%) | 17 (58.6%) | |

ART: Antiretroviral; 3TC: Lamivudine; ABC: Abacavir; LPVr: Lopinavir boosted with ritonavir; NVP: Nevirapine; Maternal severe life events: change in employment, separation or relationship break-up, new partner, loss of home or move, or death in the family; Maternal adherence (Optimal: No ART dose missed; Intermediate low: 10-50% doses missed; Intermediate high: 50-90%; Good: >90%).

**Table 2**. Algorithm tuning parameters

| Algorithm | Tunning parameter |
|---|---|
| Logistic regression | - |
| Random forest | mtry=12 |
| Support Vector Machine | C=8; sigma=$4.69 \cdot 10^{-11}$ |
| Naïve Bayes | fL=0; adjust=1 |
| K-nearest neighbor | K=5 |
| Artificial Neural Network | Size=11; decay=0.1 |
| GLMNET | Alpha=0.8; lambda=0.21 |

Algorithm tuning parameters selected by repeated (5 times) 10-fold cross-validation in a grid. Mtry: Number of variables for splitting at each tree node in a random forest; C: regularization parameter that controls the trade off between the achieving a low training error and a low testing error; sigma: determines how fast the similarity metric goes to zero as they are further apart; fL: Laplace smoother; adjust: adjust the bandwidth of the kernel density; K=number of nearest neighbours; size: number of units in hidden layer; decay: regularization parameter to avoid over-fitting; alpha: regularization parameter; lambda: penalty on the coefficients

**Supplementary Table 1**. Study population characteristics according to the primary outcome (Death/Clinical progression to AIDS)

| | Total | Did not progress | Death/Progressed | p-value |
|---|---|---|---|---|
| | *N=100* | *N=67* | *N=33* | |
| **Site of recruitment** | | | | 0.002 |
| AHRI | 10 (10.0%) | 7 (10.4%) | 3 (9.1%) | |
| ARIEL | 12 (12.0%) | 6 (8.96%) | 6 (18.2%) | |
| CISM | 23 (23.0%) | 9 (13.4%) | 14 (42.4%) | |
| MALI | 1 (1.00%) | 0 (0.00%) | 1 (3%) | |
| PHRU | 28 (28.0%) | 23 (34.3%) | 5 (15.2%) | |
| FAMCRU | 26 (26.0%) | 22 (32.8%) | 4 (12.1%) | |
| **Age at recruitment** | | | | 0.268 |
| *Months (median, IQR)* | 1.18 [0.98;2.47] | 1.11 [0.77;2.48] | 1.38 [1.05;2.43] | |
| **Gender** | | | | 0.377 |
| Female | 45 (45%) | 33 (49.3%) | 12 (36.4%) | |
| Male | 55 (55%) | 34 (50.7%) | 21 (63.6%) | |
| **Weight at recruitment** | | | | |
| Weight-for-age (z.score) | -1.44 [-2.68;-0.53] | -1.35 [-2.10;-0.36] | -1.96 [-2.86;-1.17] | 0.050 |
| **Length at recruitment** | | | | 0.628 |
| *Cm (median, IQR)* | 51.8 [48.8;55.0] | 51.5 [48.0;54.0] | 52.0 [49.5;55.0] | |
| **Preterm birth** | | | | 0.391 |
| No | 62 (62%) | 18 (54.5%) | 44 (65.7%) | |
| Yes | 38 (38%) | 15 (45.4%) | 23 (34.3%) | |
| **Age at HIV diagnosis** | | | | 0.002 |
| *Days (median, IQR)* | 30.0 [0.00;35.4] | 1.0 [0.00;32.02] | 33.0 [30.0;48.0] | |
| **Age at ART initiation** | | | | 0.006 |
| *Months (median, IQR)* | 30.0 [0.00;40.6] | 4.0 [0.00;36.0] | 33.0 [29.0;45.1] | |
| **Initial ART regimen** | | | | 0.218 |
| 3TC+ABC+LPVr | 47 (47.0%) | 33 (49.3%) | 14 (42.4%) | |
| 3TC+ABC+NVP | 1 (1.00%) | 1 (1.49%) | 0 (0.00%) | |
| 3TC+AZT+LPVr | 30 (30.0%) | 16 (23.9%) | 14 (42.4%) | |
| 3TC+AZT+NVP | 22 (22.0%) | 17 (25.4%) | 5 (15.2%) | |
| **Pre-ART viral load** | | | | <0.001 |
| *Log10 Copies/mL (median, IQR)* | 5.56 [4.53;6.38] | 5.13 [3.82;6.1] | 6.35 [5.59;6.85] | |
| **Baseline % CD4** | | | | 0.279 |
| *% Cell/mm$^3$ (median, IQR)* | 37.2 [28.0;46.7] | 37.5 [28.2;48.3] | 37.0 [28.0;42.4] | |
| **Mother's severe life events or health issues** | | | | 0.362 |
| No | 47 (50.0%) | 31 (46.3%) | 16 (59.3%) | |
| Yes | 47 (50.0%) | 36 (53.7%) | 11 (40.7%) | |
| **Mother's adherence (self-reported at enrolment)** | | | | 0.960 |
| Poor | 5 (6.10%) | 3 (5.45%) | 2 (7.41%) | |
| Intermediate low | 11 (13.4%) | 7 (12.7%) | 4 (14.8%) | |
| Intermediate high | 18 (22.0%) | 12 (21.8%) | 6 (22.2%) | |
| Good | 48 (58.5%) | 33 (60.0%) | 15 (55.6%) | |
| **Mother's last CD4 count** | | | | 0.206 |
| *Cell/mL (median, IQR)* | 314 [199;488] | 344 [223;518] | 277 [173;443] | |
| **Mother's last viral load count** | | | | 0.467 |
| *Copies/mL (median, IQR)* | 4.5 [3.27;4.97] | 4.47 [3.22;4.93] | 4.53 [3.66; 4.99] | |
| **Time to follow-up** | | | | <0.001 |
| *Months (median, IQR)* | 12.1 [11.0;16.5] | 16.2 [11.5;16.8] | 5.51 [1.38;11.9] | |

ART: Antiretroviral; 3TC: Lamivudine; ABC: Abacavir; LPVr: Lopinavir boosted with ritonavir; NVP: Nevirapine; Maternal severe life events: change in employment, separation or relationship break-up, new partner, loss of home or move, or death in the family; Maternal adherence (Optimal: No ART dose missed; Intermediate low: 10-50% doses missed;

Intermediate high: 50-90%; Good: >90%). Africa Health Research Institute (AHRI), Durban (South Africa); Family Centre For Research With Ubuntu (FAMCRU), Stellenbosch (South Africa); Perinatal HIV Research Unit (PHRU), Soweto (South Africa); Centro de Investigación en Salud de Manhiça (CISM) – ISGlobal, Manhiça (Mozambique); Ariel Glaser Foundation (ARIEL) (Mozambique).

**Supplementary Table 2**. Feature distribution in the original and imputed data sets

| | Training data sets | | | Testing datasets | | |
|---|---|---|---|---|---|---|
| | **Original** | **Imputed** | **p-value** | **Original** | **Imputed** | **p-value** |
| | **N=71** | **N=71** | | **29** | **29** | |
| **Age at recruitment** | | | 1.00 | | | 1.00 |
| Days *(median, IQR)* | 36.0 [0.97;2.26] | 1.18 [0.97;2.26] | | 1.34 [1.02;2.92] | 1.34 [1.02;2.92] | |
| **Gender** | | | 1.00 | | | 1.00 |
| Female | 34 (48.6%) | 35 (49.3%) | | 11 (37.9%) | 11 (37.9%) | |
| Male | 36 (51.4%) | 36 (50.7%) | | 18 (62.1%) | 18 (62.1%) | |
| **Weight-for-age** | | | 1.00 | | | 1.00 |
| z-score *(median, IQR)* | -1.47 [-2.62;-0.86] | -1.46 [-2.62;-0.87] | | -1.18 [-2.98;-0.30] | -1.18 [-2.98;-0.30] | |
| **Preterm birth** | | | 1.00 | | | 1.00 |
| No | 40 (57.1%) | 41 (57.7%) | | 21 (75.0%) | 22 (75.9%) | |
| Yes | 30 (42.9%) | 30 (42.3%) | | 7 (25.0%) | 7 (24.1%) | |
| **Age at HIV diagnosis** | | | 1.00 | | | 1.00 |
| *Days (median, IQR)* | 30.0 [0.00;35.5] | 30.0 [0.00;35.5] | | 31.0 [0.00;50.0] | 31.0 [0.00;50.0] | |
| **Age at ART** | | | 0.928 | | | 1.00 |
| *Days (median, IQR)* | 32.5 [19.0;62.8] | 32.0 [18.5;62.5] | | 36.0 [23.0;82.0] | 36.0 [23.0;82.0] | |
| **Initiation ART regimen** | | | 1.00 | | | 1.00 |
| 3TC+ABC+LPVr | 33 (46.5%) | 33 (46.5%) | | 14 (48.3%) | 14 (48.3%) | |
| 3TC+ABC+NVP | 0 | 0 | | 1 (3.45%) | 1 (3.45%) | |
| 3TC+AZT+LPVr | 22 (31.0%) | 22 (31.0%) | | 8 (27.6%) | 8 (27.6%) | |
| 3TC+AZT+NVP | 16 (22.5%) | 16 (22.5%) | | 6 (20.7%) | 6 (20.7%) | |
| **Baseline viral load** | | | 0.920 | | | 1.00 |
| Copies/mL *(median, IQR)* | 391854 [30401;2454708] | 609715 [36738;2570245] | | 226844 [36295;1344319] | 226844 [36295;1344319] | |
| **Baseline % CD4** | | | 0.973 | | | 0.886 |
| Cell/mL *(median, IQR)* | 37.0 [28.9;45.4] | 36.9 [29.9;45.2] | | 38.6 [27.8;46.8] | 40.0 [28.0;47.0] | |
| **Maternal severe life events or health issues** | | | 0.939 | | | 1.00 |
| No | 33 (50.0%) | 34 (47.9%) | | 14 (50.0%) | 14 (48.3%) | |
| Yes | 33 (50.0%) | 37 (52.1%) | | 14 (50.0%) | 15 (51.7%) | |
| **Maternal adherence (self-reported at enrolment)** | | | 0.955 | | | 1.00 |
| Poor | 3 (4.48%) | 3 (4.23%) | | 3 (10.7%) | 3 (10.3%) | |
| Intermediate low | 9 (13.4%) | 12 (16.9%) | | 3 (10.7%) | 4 (13.8%) | |
| Intermediate high | 17 (25.4%) | 18 (25.4%) | | 5 (17.9%) | 5 (17.2%) | |
| Good | 38 (56.7%) | 38 (53.5%) | | 17 (60.7%) | 17 (58.6%) | |
| **Death/Clinical progression to AIDS** | | | 1.000 | | | 1.00 |
| No | 67 (67.0%) | 67 (67.0%) | | 9 (31.0%) | 9 (31.0%) | |
| Yes | 33 (33.0%) | 33 (33.0%) | | 20 (69.0%) | 20 (69.0%) | |

ART: Antiretroviral; 3TC: Lamivudine; ABC: Abacavir; LPVr: Lopinavir boosted with ritonavir; NVP: Nevirapine; Maternal severe life events: change in employment, separation or relationship break-up, new partner, loss of home or move, or death in the family; Maternal adherence (Optimal: No ART dose missed; Intermediate low: 10-50% doses missed; Intermediate high: 50-90%; Good: >90%

# Discusión

En esta memoria se describen algunas de las características de los pacientes perinatalmente infectados con VIH-1 y precozmente tratados. Es fundamental comprender los factores que determinan un control rápido de la viremia en niños, ya que esto podría tener implicaciones en la curación en un futuro próximo.

Existen estudios en población adulta que observan como iniciando el TAR en un tiempo cercano a la infección primaria, algunos individuos pueden conseguir periodos post-TAR en los que consiguen controlar la viremia cuando el TAR se discontinua (47,68,69). Casos como el del bebé Mississippi (44) la chica francesa (70) o el niño sudafricano (46) plantearon la posibilidad de observar resultados similares en otros niños precozmente tratados. Además, generaron gran interés en el tiempo óptimo de inicio del TAR. En el capítulo I se observó que, dentro de los niños recién nacidos que inician el TAR en el periodo neonatal, aquellos que iniciaron en la primera semana de vida tuvieron cuatro veces más probabilidad de llegar a una supresión viral que aquellos que iniciaron el TAR entre el octavo y el vigesimoctavo día. En otras palabras, aquellos neonatos que iniciaban el TAR en la primera semana de vida tenían un descenso de la carga viral más acusado que aquellos que iniciaban posteriormente, con una probabilidad mayor de llegar antes a controlar la viremia. Actualmente existen muy pocos estudios en niños centrados en comparar resultados virológicos comparando tratar antes y después (17,24–26), la mayoría de los cuales no fueron específicamente diseñados para responder a esta cuestión y no comprenden el periodo neonatal (15). Estos estudios describen las diferencias entre tratados precozmente y tarde con mucha variabilidad la definición de precoz, variando entre 3 y 6 meses (55,71,72). En nuestro estudio, pudimos observar que la probabilidad de alcanzar esta supresión disminuía un 35% por cada semana que se retrasaba el inicio del TAR, ajustando por carga viral y régimen de TAR al inicio del tratamiento. El inicio del TAR cercano al nacimiento no solo supone el proporcionar el tratamiento cerca del tiempo a la infección, sino también el hecho de tratar en un periodo de tiempo crítico a nivel de desarrollo, cuando el sistema inmune está madurando y es más quiescente (9,73). Por esta razón, en un periodo intenso de desarrollo, cada semana podría suponer un beneficio a largo plazo a la hora de controlar la viremia. Estos resultados apoyan la iniciativa de implementar pruebas universales al nacimiento para diagnosticar neonatos infectados tan temprano como sea posible ya que actualmente se estima que la cobertura de pruebas solo alcanza un 50% de la población expuesta (74).

Cuando observamos el modelo, pudimos observar que las principales diferencias entre ambos grupos de inicio al TAR fueron particularmente notables en los primeros meses de seguimiento, encontrándose similares proporciones absolutas de niños en supresión después de un año de seguimiento clínico y virológico. Este efecto probablemente pueda ser debido a que existen otros factores asociados a la supresión virológica aparte del tiempo al inicio del TAR. El hecho de que solo se observen diferencias en los primeros meses de seguimiento, podría también explicar por qué en otros estudios no se encuentran diferencias significativas esperables en resultados virológicos e inmunológicos entre tratados antes y después (75). Mientras que en estudios realizados en adultos se puede estimar el momento de la infección, en niños esto no es posible. En ocasiones se podría confundir la infección intrauterina con la intraparto o postparto. Iniciar el TAR en horas después del parto podría ser incluso tarde en relación con el tiempo de infección si la infección ha sido intra-útero (75).

Factores como la edad al TAR, recuentos basales de linfocitos CD4, y cargas virales al momento del inicio del TAR han sido asociados previamente al tiempo hasta la supresión (72,76–80). Sin embargo, estos estudios asumen una relación lineal de los predictores a la respuesta virológica, es decir, su asociación se podría representar por una línea recta asumiendo una tasa de asociación constante. El uso de modelos aditivos como el que se empleó en el capítulo I y III permiten describir otras clases de asociaciones en las que la relación no es constante. Nuestro estudio mostró como la asociación entre la carga viral y el tiempo a la supresión no era lineal. La dinámica de esta relación sugiere que por debajo de 5 logaritmos de carga viral al inicio no hay tanta relación con el tiempo a la supresión, mientras que a partir de cinco logaritmos se observa un incremento exponencial. Esto podría indicar que en los niños con menos de 5 logaritmos quizá el factor más limitante sea el tiempo hasta el TAR mientras que en aquellos niños con cargas virales más altas, la carga viral es un factor más relevante para tener en cuenta. Es interesante que este punto de corte de 5 logaritmos es similar al umbral de 30,000 copias/mL (4.5 logaritmos) reportado por otros estudios como asociado a un fallo virológico y predictor de progresión clínica a SIDA (81,82). Nuestro estudio apoya la idea de utilizar otros marcadores de progresión o de necesidad de cambio de régimen de TAR a parte de la carga viral (81), sobre todo cuando se observen cargas virales menores a 5 logaritmos.

Por tanto, nuestros resultados apoyan los beneficios de iniciar el TAR cuanto antes, a ser posible en la primera semana de vida para reducir el tiempo a la supresión. Esta reducción podría resultar en un impacto favorable a la hora de reducir el tamaño de reservorio viral (22), puesto que acorta el tiempo a la supresión y éste está relacionado con el tamaño a largo plazo del reservorio, como se demuestra en el capítulo III.

Este control de la viremia debido al inicio precoz del TAR está determinado por dos fases. Una primera fase exponencial y rápida y una segunda fase más lenta. En el segundo capítulo pudimos observar que, independientemente de la edad al inicio del TAR, el porcentaje pre-TAR de linfocitos CD4 y carga viral pre-TAR, las niñas presentaron un control más rápido de la viremia. Tanto los niños como las niñas presentaron similares cargas virales al inicio del TAR. A pesar de las diferencias encontradas entre niños y niñas en la fase I del descenso de carga viral, estas diferencias no se encontraron en los tiempos hasta supresión. Este hecho podría sugerir que, el tiempo a la supresión no tiene por qué coincidir con la primera fase del descenso de la carga viral. Probablemente porque la definición de supresión depende directamente del límite inferior de detección de las técnicas y del tiempo entre medidas de cargas virales, haciendo muy difícil determinar el momento exacto de supresión. Sin embargo, cuando se observa la primera fase de la curva de decremento de carga viral se puede observar cómo las niñas presentan pendientes más pronunciadas.

Nuestros resultados apoyan otros estudios en adultos que sugieren que las mujeres controlan mejor la viremia que los hombres (83,84). Además, estos resultados son consistentes con estudios en población pediátrica que observan mayor número de niñas en los grupos de progresores lentos, definidos como pacientes sin TAR que mantienen recuentos de CD4 por encima de 350 células/microlitro y no presentan estadios III o IV de la OMS (58). Estos datos sugieren que la causa que determina estas diferencias en sexo no es solo hormonal, como podría caber esperar si solo se observara en adultos, sino que quizá puedan estar involucrados la expresión de genes determinados por el sexo (42). Diversos estudios han asociado un mejor control de la viremia en mujeres y niñas con una respuesta aguda más intensa a interferón tipo I (85,86) y una menor tasa de replicación viral en mujeres comparado con hombres en las primeras fases de la infección por VIH-1 (87–89).

Sin embargo, pese a un mejor control inicial de la viremia en niñas y en mujeres, diversos estudios han asociado progresiones más rápidas y mayor mortalidad a este sexo (90–92).

Estas tasas de mortalidad y progresión mayores están relacionadas con una activación inmune mayor en mujeres después de controlar la viremia probablemente debidas a la expresión de genes estimulados por el interferón I como MX1 y ISG15 (86). Sin embargo, la mayoría de los estudios que relacionan mortalidad con el sexo femenino están realizados en población adulta con la limitación de que esta disparidad podría no reflejar una razón biológica sino un conjunto de factores sociodemográficos y psicosociales como una mayor exposición al virus debido a transfusiones por anemias, menos acceso al tratamiento o limitaciones étnicas en las mujeres con respecto a los hombres (93). En el capítulo IV no observamos diferencias significativas en mortalidad o progresión debidas al sexo habiendo incluso más porcentaje de niños que de niñas en el grupo de aquellos que progresaban o morían antes del primer año de seguimiento.

El VIH-1 establece reservorios virales dinámicos y diversos donde la replicación persiste renovando las poblaciones de virus activos en el organismo. La capacidad para caracterizar el reservorio viral es importante para evaluar y estudiar el estado virológico de los pacientes, así como determinar si ha habido una posible remisión. Además, limitar el reservorio podría ser un potencial método para erradicar células infectadas por el virus (94). Sin embargo, muchas de las técnicas disponibles en la rutina tienen limitaciones en pediatría como la necesidad de grandes cantidades de sangre, los costes elevados de la técnica, y laboratorios con personal cualificado (95). Por este motivo, es importante identificar factores y marcadores más simples, asociados con un reservorio bajo.

En el capítulo II se observó que la pendiente de la curva de decremento de carga viral fue un predictor del reservorio a largo plazo. A su vez, en el tercer capítulo pudimos observar como aquellos pacientes que iniciaron TAR antes, tuvieron un reservorio viral bajo a lo largo de su seguimiento. Esta asociación ha sido ampliamente estudiada en adultos y cada vez más reconocida en niños (19–21,48,49). Datos de estos cinco estudios que midieron el reservorio tanto en PBMC como en células T CD4 mostraron que el reservorio fue significativamente más bajo en aquellos niños que iniciaban el TAR antes (1.8-2.4 meses) comparado con aquellos que iniciaban después (6-155 meses). Sin embargo, el efecto de la carga viral previa al TAR todavía está poco establecido. Se sabe que la carga viral está asociada con el tiempo a la supresión viral y el subsecuente riesgo de futuros repuntes en población adulta (96,97). En otras palabras, un TAR precoz podría acortar el tiempo hasta supresión como se demuestra

en el capítulo I. Este descenso de la carga viral en la fase I, que es a su vez, marcador de una reducción de las células que contribuyen al descenso de la fase II, podría minimizar el tamaño del reservorio. En nuestro estudio que compone el capítulo III, pudimos observar cómo cargas virales superiores a 6 logaritmos pre-TAR estaban muy relacionadas con la presencia de un futuro reservorio alto. Este punto de corte de carga viral es similar a aquel reportado en el capítulo I (5 logaritmos) al describir la asociación entre la carga viral pre-TAR y el tiempo a la supresión.

Además de la carga viral pre-TAR, en el capítulo II y III se observó que el porcentaje de linfocitos CD4 pre-TAR son un predictor del reservorio viral a largo plazo. En el capítulo III se describió una asociación positiva, es decir a mayor porcentaje de CD4 pre-TAR, mayor reservorio a largo plazo. Esto contradice lo esperado, ya que una severidad mayor pre-TAR (menor porcentaje de CD4) predeciría un reservorio viral mayor (98,99). Este efecto podría explicarse por la ausencia de diana viral, es decir, presentar menor porcentaje de células T CD4 podría conllevar un menor número de células donde establecer el reservorio. Asimismo, los pacientes precozmente tratados presentan menor proliferación de las poblaciones de células T de memoria (100,101). Es importante tener en cuenta la peculiaridad de esta población precozmente tratada, con la mayoría de niños tratados en los primeros seis meses de vida, lo que podría indicar que la mayor parte de las células T CD4 fueran naïve, y por tanto menos susceptibles para la integración viral (102). Sería necesario estudiar esta asociación con medidas de reservorio hechas sobre ADN integrado en células T CD4 y no en PBMC totales para esclarecer este efecto.

Además, observamos que la seronegatividad estaba asociada a un inicio precoz del TAR. Se observó una asociación positiva entre el WB-score y el tamaño del reservorio apoyando lo ya reportado en otras cohortes (21,48,103). Esta estrecha relación entre el total de ADN del reservorio de VIH-1 y la magnitud de la respuesta de los anticuerpos específicos contra el VIH-1 es consistente con la hipótesis de que la persistencia de anticuerpos específicos es dependiente de la exposición con el antígeno y por tanto directamente relacionada con el tiempo de inicio del TAR. La primera infancia representa un periodo en el que el sistema inmune está en plena maduración y marcado por la presencia de anticuerpos maternos en sangre que podrían esconder los antígenos. Un tratamiento muy precoz lleva a una supresión de la replicación viral que impide al sistema inmune crear una respuesta humoral VIH-1

específica (104,105). Esta asociación sugiere que el estado serológico podría usarse como un marcador complementario para monitorizar pacientes, particularmente en países de recursos limitados (95).

A pesar de que los beneficios clínicos de tratar precozmente son claros, sigue siendo un interrogante si un tamaño de reservorio menor en estos niños podría conferir un futuro sin TAR y los niños podrían curarse de la infección.

En el otro extremo, todavía existen tasas altas de mortalidad en los primeros meses de inicio al TAR en África Subsahariana (62). En el cuarto de los capítulos que componen esta memoria se comparó el rendimiento de diferentes algoritmos, algunos basados en aprendizaje automático para predecir la mortalidad y progresión clínica en niños precozmente tratados. El algoritmo que tuvo mejor precisión en la predicción (>80%) fue *Random Forest*, alcanzando una sensibilidad mayor que otros algoritmos como la regresión logística comúnmente usada en el campo de la biomedicina.

La flexibilidad que ofrecen estos modelos hace que la inteligencia artificial y en concreto el aprendizaje automático sea una herramienta mejor que otros métodos tradicionales. Esta flexibilidad viene dada por la habilidad para trabajar con múltiples predictores y relaciones complejas entre las variable (106,107). A pesar de estas ventajas, existen muchos investigadores que cuestionan su utilidad por la gran cantidad de tamaño muestral que necesitan (108).

A diferencia de esto, nuestro estudio mostró que estas técnicas tienen mayor capacidad de predicción también en tamaños muestrales pequeños. Este resultado es particularmente importante en campos como es el estudio del VIH-1 pediátrico donde tener conjuntos de datos muy grandes no es viable.

Es importante remarcar que los algoritmos basados en aprendizaje automático como *Random Forest*, *Support Vector Machine*, Redes Neuronales, o *K-Nearest Neighbour* ofrecieron una ventaja considerable en la predicción comparado con métodos tradicionales como la regresión logística en términos de sensibilidad, especificidad y precisión. En concreto, *Random Forest* diferenciaba significativamente las probabilidades de mortalidad y progresión arrojadas por el modelo para los pacientes que habían muerto o progresado con respecto a los que no.

Debido a que estas técnicas no suelen ser de fácil interpretación y su desarrollo normalmente lleva complejos análisis de datos, nuestro estudio implementó el mejor modelo en una aplicación web amigable y sencilla llamada *MAGPIE* (**https://rserver.h12o.es/pediatria/MAGPIE/**; username: user, password: 0000). Esta aplicación web arroja la probabilidad de muerte o progresión a partir de un cuestionario simple en el que se introducen variables clínicas y virológicas. En países de recursos limitados donde el manejo de pacientes VIH-1 está descentralizado y hay pocos facultativos especialistas atendiendo a los niños, esta herramienta podría ser útil para alertar a los clínicos del riesgo de estos pacientes. Esta información podría servir para considerar referir a los pacientes a centros más especializados o aumentar el tiempo de seguimiento.

Nuestros estudios cuentan con varias limitaciones. Los primeros tres artículos están basados en datos obtenidos retrospectivamente de cohortes con diseño observacional. Esto indica que información como factores clínicos y sociales asociados de forma dependiente a la edad al inicio del TAR están fuera de nuestro control (ej.: programas de prevención, atención o tratamiento en el momento de la recogida de datos). Además, el escaso tamaño muestral dificulta el ajuste de los modelos por potenciales variables de confusión que quizá jueguen un papel importante en la predicción tanto del tiempo a la supresión como del tamaño del reservorio. Sin embargo, la naturaleza de esta enfermedad en esta población vulnerable hace que no se dispongan de tamaños muestrales muy elevados, siendo de vital importancia utilizar modelos flexibles de alto rendimiento como los empleados en estos artículos.

Otra de las posibles limitaciones de los dos primeros artículos es que los datos no fueron específicamente recogidos para estos objetivos o hipótesis, pudiéndose ver limitada la homogeneidad de los datos que provienen de diferentes cohortes, países o centros. Por ello, establecimos unos criterios de inclusión y exclusión claros a la hora de seleccionar pacientes para los estudios y contemplamos diferentes variables en el ajuste de los modelos de predicción para evaluar y remediar la heterogeneidad de la muestra.

Con respecto a los artículos II y III, es importante reseñar la limitación en la medida de reservorio. En ambos capítulos el tamaño de reservorio está medido como total de copias/mL de DNA viral por millón de células PBMC. Esta técnica, a pesar de que presenta una gran sensibilidad y es asequible económicamente, mide el DNA viral total, lo que incluye una gran cantidad de genoma defectivo. Es decir, genoma que no da lugar a viriones replicativos (109).

La cuantificación del ADN proviral tiene cuatro retos principales. Muchos de los provirus que se encuentran en los pacientes VIH en TAR presentan mutaciones y/o deleciones que les hacen defectivos o incapaces de replicarse. Además, no todo el genoma intacto, que, aunque no presenta mutaciones defectivas que impidan la replicación, es capaz de producir viriones puesto que necesita activaciones específicas. Por último, la frecuencia de células latentes es muy baja y la mayoría de los reservorios se encuentran en tejidos inaccesibles para la mayoría de las técnicas de muestreo (110). Futuros estudios con técnicas más específicas como la detección basada en PCR para medir ADN proviral intacto (IPDA) (111) o PCR cuantitativa quadruplex (Q4PCR) (112) podrán arrojar más luz sobre los marcadores y determinantes de un reservorio activo (110).

Con respecto al tercer y al último artículo, el principal inconveniente fue también el limitado número de observaciones, hecho que probablemente disminuye la capacidad de predicción de los diferentes algoritmos. En este trabajo no se puede saber cuánto efecto tiene el tamaño muestral y cómo de diferente es en cada uno de los algoritmos. Sin embargo, nuestro estudio es capaz de comparar estos sistemas diferentes con el mismo tamaño muestral y discriminar cuál tiene mayor poder de predicción en esta muestra lo que plantea que modelos basados en aprendizaje automático son mejores que los tradicionales incluso con poblaciones de tamaño reducido.

Nuestros estudios comparten una limitación común, y es que, dado que son estudios observacionales, las asociaciones no pueden ser asumidas como relaciones causales. Además, la infección por VIH-1 y en concreto la infección en población pediátrica está determinada por muchos factores que juegan un papel complejo difícil de modelizar. Por esta razón, los resultados deben ser interpretados con cautela a la espera de que estas conclusiones puedan ser reforzadas con estudios prospectivos y/o estudios aleatorizados.

Como fortalezas de nuestro estudio entendemos que este trabajo tiene carácter multidisciplinar puesto que engrana la experiencia clínica de los pediatras, virólogos e inmunólogos y la formación bioinformática y bioestadística de los metodólogos. Esta colaboración es de vital importancia para identificar y generar nuevas herramientas que permitan un mejor manejo clínico en la práctica del día a día.

# Conclusiones

1. El inicio durante la primera semana de vida en comparación con el inicio entre los días 8-28 de vida supone un incremento en la probabilidad de alcanzar la supresión virológica antes.
2. Las niñas presentan un descenso de carga viral debido al TAR más rápido que los niños en la primera fase independientemente de la edad al inicio del TAR, porcentaje de CD4 y cargas virales pre-TAR. Este descenso de carga viral en la primera fase, está asociado al tamaño de reservorio viral. Por cada punto que la pendiente en el descenso de carga viral (tiempo/logaritmo carga viral) aumentaba, se observó un 5% de incremento en el tamaño de reservorio medido como ADN total/millón células PBMC

3. Los factores predictores del tamaño del reservorio a largo plazo en nuestro conjunto de datos fue el tiempo al TAR, carga viral basal previa al TAR y la puntuación obtenida en el Western-blot y el porcentaje de CD4 basal. Nuestro modelo nos permitió observar dos tipos de asociaciones:
   - Una asociación lineal positiva entre la edad al TAR y el tiempo a supresión.
   - Una asociación no lineal entre la carga viral y el tiempo a supresión. La dinámica de esta asociación sugiere que, por debajo de cargas virales de 5 logaritmos, el tiempo al TAR es el principal factor asociado al tiempo a supresión. Sin embargo, para cargas virales superiores a 5 logaritmos, existe una relación exponencial entre la carga viral y el tiempo a la supresión.

4. La mortalidad y progresión clínica a SIDA puede predecirse mediante técnicas basadas en aprendizaje automático. Pese a un tamaño muestral pequeño, estas técnicas logran mayor poder de predicción que aquellas utilizadas tradicionalmente como la regresión logística. El modelo basado en el algoritmo Random Forest obtuvo una precisión >80%, una sensibilidad del 78% y una especificidad del 85% para predecir mortalidad y progresión a SIDA. Estos modelos de aprendizaje automático pueden incluirse en apps y utilizarse de forma sencilla para poder implementar las ventajas de esta predicción en la rutina clínica de países de recursos limitados.

# Bibliografía

1. Abram ME, Ferris AL, Shao W, Alvord WG, Hughes SH. Nature, position, and frequency of mutations made in a single cycle of HIV-1 replication. J Virol [Internet]. 2010 Oct 1 [cited 2019 Jan 8];84(19):9864–78. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20660205

2. Clapham PR, McKnight Á. HIV-1 receptors and cell tropism. Br Med Bull [Internet]. 2001 Sep 1 [cited 2019 Jan 8];58(1):43–59. Available from: https://academic.oup.com/bmb/article/58/1/43/337216

3. van Zyl G, Bale MJ, Kearney MF. HIV evolution and diversity in ART-treated patients. Retrovirology [Internet]. 2018 Dec 30 [cited 2019 Jan 8];15(1):14. Available from: https://retrovirology.biomedcentral.com/articles/10.1186/s12977-018-0395-4

4. Levy JA. Pathogenesis of human immunodeficiency virus infection. Microbiol Rev [Internet]. 1993 Mar;57(1):183–289. Available from: http://www.ncbi.nlm.nih.gov/pubmed/8464405

5. Abdel-Mohsen M, Richman D, Siliciano RF, Nussenzweig MC, Howell BJ, Martinez-Picado J, et al. Recommendations for measuring HIV reservoir size in cure-directed clinical trials. Nat Med [Internet]. 2020 Sep 1 [cited 2021 Jan 12];26(9):1339–50. Available from: https://www.nature.com/articles/s41591-020-1022-1

6. Brenchley JM, Hill BJ, Ambrozak DR, Price DA, Guenaga FJ, Casazza JP, et al. T-cell subsets that harbor human immunodeficiency virus (HIV) in vivo: implications for HIV pathogenesis. J Virol [Internet]. 2004 Feb [cited 2019 Jan 8];78(3):1160–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/14722271

7. UNAIDS. Progress towards theStart Free, Stay Free, AIDS Free targets [Internet]. 2020. Available from: https://www.unaids.org/sites/default/files/media_asset/start-free-stay-free-aids-free-2020-progress-report_en.pdf

8. Kourtis AP, Lee FK, Abrams EJ, Jamieson DJ, Bulterys M. Mother-to-child transmission of HIV-1: timing and implications for prevention. Lancet Infect Dis [Internet]. 2006 Nov;6(11):726–32. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1473309906706296

9. Tobin NH, Aldrovandi GM. Immunology Of Pediatric HIV Infection. Immunol Rev

[Internet]. 2013 Jul [cited 2021 Jan 12];254(1):143–69. Available from: /pmc/articles/PMC3737605/?report=abstract

10. Davies M-A, Gibb D, Turkova A. Survival of HIV-1 vertically infected children. Curr Opin HIV AIDS [Internet]. 2016 Sep [cited 2019 Jan 8];11(5):455–64. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27716730

11. Massanella M, Puthanakit T, Leyre L, Jupimai T, Sawangsinth P, de Souza M, et al. Continuous Prophylactic Antiretrovirals/Antiretroviral Therapy Since Birth Reduces Seeding and Persistence of the Viral Reservoir in Children Vertically Infected With Human Immunodeficiency Virus. Clin Infect Dis [Internet]. 2020 Jun 5 [cited 2021 Jan 12]; Available from: https://pubmed.ncbi.nlm.nih.gov/32504081/

12. Muenchhoff M, Prendergast AJ, Goulder PJR. Immunity to HIV in Early Life. Front Immunol [Internet]. 2014 Aug 12;5. Available from: http://journal.frontiersin.org/article/10.3389/fimmu.2014.00391/abstract

13. Ananworanich J, Puthanakit T, Suntarattiwong P, Chokephaibulkit K, Kerr SJ, Fromentin R, et al. Reduced markers of HIV persistence and restricted HIV-specific immune responses after early antiretroviral therapy in children. AIDS [Internet]. 2014 Apr;28(7):1015–20. Available from: http://journals.lww.com/00002030-201404240-00010

14. Palma P, Foster C, Rojo P, Zangari P, Yates A, Cotugno N, et al. The EPIICAL project: an emerging global collaboration to investigate immunotherapeutic strategies in HIV-infected children. J virus Erad [Internet]. 2015 [cited 2020 Jul 31];1(3):134–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26893908

15. Shiau S, Abrams EJ, Arpadi SM, Kuhn L. Early antiretroviral therapy in HIV-infected infants: can it lead to HIV remission? lancet HIV [Internet]. 2018 May 1 [cited 2018 Dec 28];5(5):e250–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/29739699

16. Blanche S, Tardieu M, Duliege A, Rouzioux C, Le Deist F, Fukunaga K, et al. Longitudinal study of 94 symptomatic infants with perinatally acquired human immunodeficiency virus infection. Evidence for a bimodal expression of clinical and biological symptoms. Am J Dis Child [Internet]. 1990 Nov [cited 2019 Jan

8];144(11):1210–5. Available from: http://www.ncbi.nlm.nih.gov/pubmed/1978551

17. Violari A, Cotton MF, Gibb DM, Babiker AG, Steyn J, Madhi SA, et al. Early Antiretroviral Therapy and Mortality among HIV-Infected Infants. N Engl J Med [Internet]. 2008 Nov 20 [cited 2018 Aug 2];359(21):2233–44. Available from: http://www.nejm.org/doi/abs/10.1056/NEJMoa0800971

18. World Health Organization. Antiretroviral therapy for HIV infection in infants and children: towards universal access — recommendations for a public health approach. Genova; 2006.

19. van Zyl GU, Bedison MA, van Rensburg AJ, Laughton B, Cotton MF, Mellors JW. Early Antiretroviral Therapy in South African Children Reduces HIV-1-Infected Cells and Cell-Associated HIV-1 RNA in Blood Mononuclear Cells. J Infect Dis [Internet]. 2015 Jul 1 [cited 2018 Aug 2];212(1):39–43. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25538273

20. Persaud D, Palumbo PE, Ziemniak C, Hughes MD, Alvero CG, Luzuriaga K, et al. Dynamics of the resting CD4(+) T-cell latent HIV reservoir in infants initiating HAART less than 6 months of age. AIDS [Internet]. 2012 Jul 31 [cited 2018 Aug 2];26(12):1483–90. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22555165

21. Luzuriaga K, Tabak B, Garber M, Chen YH, Ziemniak C, McManus MM, et al. HIV Type 1 (HIV-1) Proviral Reservoirs Decay Continuously Under Sustained Virologic Control in HIV-1–Infected Children Who Received Early Treatment. J Infect Dis [Internet]. 2014 Nov 15 [cited 2018 Aug 2];210(10):1529–38. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24850788

22. Tagarro A, Chan M, Zangari P, Ferns B, Foster C, De Rossi A, et al. Early and Highly Suppressive Antiretroviral Therapy Are Main Factors Associated With Low Viral Reservoir in European Perinatally HIV-Infected Children. JAIDS J Acquir Immune Defic Syndr [Internet]. 2018 Oct 26 [cited 2020 Feb 11];79(2):269–76. Available from: http://insights.ovid.com/crossref?an=00126334-201810010-00019

23. Kuhn L, Paximadis M, Da Costa Dias B, Loubser S, Strehlau R, Patel F, et al. Age at antiretroviral therapy initiation and cell-associated HIV-1 DNA levels in HIV-1-infected

children. Sluis-Cremer N, editor. PLoS One [Internet]. 2018 Apr 12 [cited 2019 Feb 1];13(4):e0195514. Available from: https://dx.plos.org/10.1371/journal.pone.0195514

24. Cotton MF, Violari A, Otwombe K, Panchia R, Dobbels E, Rabie H, et al. Early time-limited antiretroviral therapy versus deferred therapy in South African infants infected with HIV: results from the children with HIV early antiretroviral (CHER) randomised trial. Lancet [Internet]. 2013 Nov 9 [cited 2018 Aug 2];382(9904):1555–63. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24209829

25. Prendergast A, Mphatswe W, Tudor-Williams G, Rakgotho M, Pillay V, Thobakgale C, et al. Early virological suppression with three-class antiretroviral therapy in HIV-infected African infants. AIDS [Internet]. 2008 Jul 11 [cited 2021 Jan 20];22(11):1333–43. Available from: https://pubmed.ncbi.nlm.nih.gov/18580613/

26. Ásbjörnsdóttir KH, Hughes JP, Wamalwa D, Langat A, Slyker JA, Okinyi HM, et al. Differences in virologic and immunologic response to antiretroviral therapy among HIV-1-infected infants and children. AIDS [Internet]. 2016 Nov 28 [cited 2018 Aug 2];30(18):2835–43. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27603293

27. Kellerman S, Essajee S. HIV testing for children in resource-limited settings: What are we waiting for? PLoS Med [Internet]. 2010 Jul 20 [cited 2021 Jan 14];7(7):e1000285–e1000285. Available from: http://data.unaids.org/pub/Report/2008/20081201_3rd_stocktaking_summary_en.pdf

28. Ahmed S, Kim MH, Sugandhi N, Phelps BR, Sabelli R, Diallo MO, et al. Beyond early infant diagnosis: case finding strategies for identification of HIV-infected infants and children. [Internet]. Vol. 27 Suppl 2, AIDS (London, England). 2013 [cited 2021 Jan 14]. p. S235-45. Available from: https://europepmc.org/articles/PMC4122794

29. Ho DD, Neumann AU, Perelson AS, Chen W, Leonard JM, Markowitz M. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. Nature [Internet]. 1995 Jan 12 [cited 2021 Jan 14];373(6510):123–6. Available from: https://www.nature.com/articles/373123a0

30. Perelson AS, Essunger P, Cao Y, Vesanen M, Hurley A, Saksela K, et al. Decay characteristics of HIV-1-infected compartments during combination therapy. Nature

[Internet]. 1997 May 8 [cited 2021 Jan 14];387(6629):188–91. Available from: https://www.nature.com/articles/387188a0

31. Wu H, Ding AA. Population HIV-1 dynamics in vivo: Applicable models and inferential tools for virological data from AIDS clinical trials. Biometrics [Internet]. 1999 [cited 2021 Jan 14];55(2):410–8. Available from: https://pubmed.ncbi.nlm.nih.gov/11318194/

32. HELBERT MR, WALTER J, L'AGE J, BEVERLEY PCL. HIV infection of CD45RA + and CD45RO + CD4 + T cells. Clin Exp Immunol [Internet]. 1997 Feb;107(2):300–5. Available from: http://doi.wiley.com/10.1111/j.1365-2249.1997.280-ce1170.x

33. Dion ML, Poulin JF, Bordi R, Sylvestre M, Corsini R, Kettaf N, et al. HIV infection rapidly induces and maintains a substantial suppression of thymocyte proliferation. Immunity [Internet]. 2004 Dec [cited 2021 Jan 14];21(6):757–68. Available from: https://pubmed.ncbi.nlm.nih.gov/15589165/

34. Finzi D, Siliciano RF. Viral Dynamics in HIV-1 Infection. Cell [Internet]. 1998 May;93(5):665–71. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0092867400814270

35. Perelson AS, Neumann AU, Markowitz M, Leonard JM, Ho DD. HIV-1 Dynamics in Vivo: Virion Clearance Rate, Infected Cell Life-Span, and Viral Generation Time. Science (80- ) [Internet]. 1996 Mar 15;271(5255):1582–6. Available from: https://www.sciencemag.org/lookup/doi/10.1126/science.271.5255.1582

36. Haubrich RH, Riddler SA, Ribaudo H, DiRenzo G, Klingman KL, Garren KW, et al. Initial viral decay to assess the relative antiretroviral potency of protease inhibitor-sparing, nonnucleoside reverse transcriptase inhibitor-sparing, and nucleoside reverse transcriptase inhibitor-sparing regimens for first-line therapy of HIV infection. AIDS [Internet]. 2011 Nov 28 [cited 2020 Dec 10];25(18):2269–78. Available from: /pmc/articles/PMC3572727/?report=abstract

37. Luzuriaga K, Wu H, McManus M, Britto P, Borkowsky W, Burchett S, et al. Dynamics of Human Immunodeficiency Virus Type 1 Replication in Vertically Infected Infants. J Virol. 1999 Jan 1;73(1):362–7.

38. Scully EP. Sex Differences in HIV Infection. Curr HIV/AIDS Rep [Internet]. 2018 Apr

5;15(2):136–46. Available from: http://link.springer.com/10.1007/s11904-018-0383-2

39. Sherr L, Mueller J, Varrall R. Evidence-based gender findings for children affected by HIV and AIDS – a systematic overview. AIDS Care [Internet]. 2009 Aug 30;21(sup1):83–97. Available from: https://www.tandfonline.com/doi/full/10.1080/09540120902923105

40. Sterling TR, Vlahov D, Astemborski J, Hoover DR, Margolick JB, Quinn TC. Initial Plasma HIV-1 RNA Levels and Progression to AIDS in Women and Men. N Engl J Med [Internet]. 2001 Mar 8 [cited 2020 May 27];344(10):720–5. Available from: http://www.nejm.org/doi/abs/10.1056/NEJM200103083441003

41. Maskew M, Brennan AT, Westreich D, McNamara L, MacPhail AP, Fox MP. Gender differences in mortality and CD4 count response among virally suppressed HIV-positive patients. J Women's Heal. 2013 Feb 1;22(2):113–20.

42. Ruel TD, Zanoni BC, Ssewanyana I, Cao H, Havlir D V, Kamya M, et al. Sex differences in HIV RNA level and CD4 cell percentage during childhood. Clin Infect Dis [Internet]. 2011 Sep [cited 2020 May 11];53(6):592–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21840929

43. Shiau S, Strehlau R, Technau K-G, Patel F, Arpadi SM, Coovadia A, et al. Early age at start of antiretroviral therapy associated with better virologic control after initial suppression in HIV-infected infants. AIDS [Internet]. 2016 Nov 28 [cited 2018 Aug 2];31(3):1. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27828785

44. Persaud D, Gay H, Ziemniak C, Chen YH, Piatak M, Chun T-W, et al. Absence of Detectable HIV-1 Viremia after Treatment Cessation in an Infant. N Engl J Med [Internet]. 2013 Nov 7 [cited 2021 Jan 20];369(19):1828–35. Available from: https://www.nejm.org/doi/full/10.1056/nejmoa1302976

45. Frange P, Faye A, Avettand-Fenoël V, Bellaton E, Descamps D, Angin M, et al. HIV-1 virological remission lasting more than 12 years after interruption of early antiretroviral therapy in a perinatally infected teenager enrolled in the French ANRS EPF-CO10 paediatric cohort: a case report. Lancet HIV [Internet]. 2016 Jan [cited 2019 Apr 1];3(1):e49–54. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26762993

46. Violari A, Cotton MF, Kuhn L, Schramm DB, Paximadis M, Loubser S, et al. A child with

perinatal HIV infection and long-term sustained virological control following antiretroviral treatment cessation. Nat Commun [Internet]. 2019 Dec 1 [cited 2021 Feb 23];10(1):1–11. Available from: https://doi.org/10.1038/s41467-019-08311-0

47. Sáez-Cirión A, Bacchus C, Hocqueloux L, Avettand-Fenoel V, Girault I, Lecuroux C, et al. Post-Treatment HIV-1 Controllers with a Long-Term Virological Remission after the Interruption of Early Initiated Antiretroviral Therapy ANRS VISCONTI Study. Lifson J, editor. PLoS Pathog [Internet]. 2013 Mar 14 [cited 2021 Jan 14];9(3):e1003211. Available from: https://dx.plos.org/10.1371/journal.ppat.1003211

48. Martínez-Bonet M, Puertas MC, Fortuny C, Ouchi D, Mellado MJ, Rojo P, et al. Establishment and Replenishment of the Viral Reservoir in Perinatally HIV-1-infected Children Initiating Very Early Antiretroviral Therapy. Clin Infect Dis [Internet]. 2015 Oct 1 [cited 2019 Jan 8];61(7):1169–78. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26063721

49. McManus M, Mick E, Hudson R, Mofenson LM, Sullivan JL, Somasundaran M, et al. Early Combination Antiretroviral Therapy Limits Exposure to HIV-1 Replication and Cell-Associated HIV-1 DNA Levels in Infants. Sluis-Cremer N, editor. PLoS One [Internet]. 2016 Apr 22 [cited 2019 Jan 8];11(4):e0154391. Available from: https://dx.plos.org/10.1371/journal.pone.0154391

50. Ahmed I, Lemma S. Mortality among pediatric patients on HIV treatment in sub-Saharan African countries: A systematic review and meta-analysis [Internet]. Vol. 19, BMC Public Health. BioMed Central Ltd.; 2019 [cited 2020 Jul 31]. p. 149. Available from: https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-019-6482-1

51. Mirochnick M, Nielsen-Saines K, Pilotto JH, Pinto J, Jiménez E, Veloso VG, et al. Nevirapine concentrations in newborns receiving an extended prophylactic regimen. J Acquir Immune Defic Syndr [Internet]. 2008 Mar 1 [cited 2019 Jan 2];47(3):334–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18398973

52. Morris SE, Dziobek-Garrett L, Yates AJ. Ushr: Understanding suppression of HIV in R. BMC Bioinformatics. 2020 Feb 11;21(1):1–8.

53. Bloem M. The 2006 WHO child growth standards. Vol. 334, British Medical Journal.

BMJ Publishing Group; 2007. p. 705–6.

54. Chiappini E, Galli L, Tovo P-A, Gabiano C, Lisi C, Bernardi S, et al. Five-year follow-up of children with perinatal HIV-1 infection receiving early highly active antiretroviral therapy. BMC Infect Dis [Internet]. 2009 Dec 26 [cited 2018 Aug 2];9(1):140. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19709432

55. Chiappini E, Galli L, Tovo P-A, Gabiano C, Gattinara GC, Guarino A, et al. Virologic, immunologic, and clinical benefits from early combined antiretroviral therapy in infants with perinatal HIV-1 infection. AIDS [Internet]. 2006 Jan 9 [cited 2018 Aug 2];20(2):207–15. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16511413

56. Muenchhoff M, Goulder PJR. Sex Differences in Pediatric Infectious Diseases. J Infect Dis [Internet]. 2014 Jul 15;209(suppl 3):S120–6. Available from: https://academic.oup.com/jid/article-lookup/doi/10.1093/infdis/jiu232

57. Fish EN. The X-files in immunity: Sex-based differences predispose immune responses. Vol. 8, Nature Reviews Immunology. Nature Publishing Group; 2008. p. 737–44.

58. Vieira VA, Zuidewind P, Muenchhoff M, Roider J, Millar J, Clapson M, et al. Strong sex bias in elite control of paediatric HIV infection. AIDS. 2019 Jan 2;33(1):67–75.

59. Klein N, Palma P, Luzuriaga K, Pahwa S, Nastouli E, Gibb DM, et al. Early antiretroviral therapy in children perinatally infected with HIV: a unique opportunity to implement immunotherapeutic approaches to prolong viral remission. Lancet Infect Dis [Internet]. 2015 Sep [cited 2019 Apr 22];15(9):1108–14. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26187030

60. Foster C, Pace M, Kaye S, Hopkins E, Jones M, Robinson N, et al. Early antiretroviral therapy reduces HIV DNA following perinatal HIV infection. AIDS [Internet]. 2017 Aug 24 [cited 2020 Feb 11];31(13):1847–51. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28609403

61. Payne H, Mkhize N, Otwombe K, Lewis J, Panchia R, Callard R, et al. Reactivity of routine HIV antibody tests in children who initiated antiretroviral therapy in early infancy as part of the Children with HIV Early Antiretroviral Therapy (CHER) trial: a

retrospective analysis. Lancet Infect Dis [Internet]. 2015 Jul;15(7):803–9. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1473309915000870

62. Ahmed I, Lemma S. Mortality among pediatric patients on HIV treatment in sub-Saharan African countries: a systematic review and meta-analysis. BMC Public Health [Internet]. 2019 Dec 4 [cited 2020 Jan 24];19(1):149. Available from: https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-019-6482-1

63. Kabue MM, Buck WC, Wanless SR, Cox CM, McCollum ED, Caviness AC, et al. Mortality and clinical outcomes in HIV-infected children on antiretroviral therapy in Malawi, Lesotho, and Swaziland. Pediatrics. 2012 Sep;130(3).

64. Zanoni BC, Phungula T, Zanoni HM, France H, Feeney ME. Risk factors associated with increased mortality among HIV infected children initiating antiretroviral therapy (ART) in South Africa. PLoS One. 2011;6(7).

65. Anigilaje EA, Aderibigbe SA. Mortality in a Cohort of HIV-Infected Children: A 12-Month Outcome of Antiretroviral Therapy in Makurdi, Nigeria. Adv Med. 2018 Jun 19;2018:1–11.

66. Obermeyer Z, Emanuel EJ. Predicting the future-big data, machine learning, and clinical medicine. Vol. 375, New England Journal of Medicine. Massachussetts Medical Society; 2016. p. 1216–9.

67. Mullainathan S, Spiess J. Machine learning: An applied econometric approach. In: Journal of Economic Perspectives. American Economic Association; 2017. p. 87–106.

68. Namazi G, Fajnzylber JM, Aga E, Bosch RJ, Acosta EP, Sharaf R, et al. The Control of HIV After Antiretroviral Medication Pause (CHAMP) Study: Posttreatment Controllers Identified From 14 Clinical Studies. J Infect Dis [Internet]. 2018 Nov 5 [cited 2021 Jan 20];218(12):1954–63. Available from: https://academic.oup.com/jid/article/218/12/1954/5067282

69. Lodi S, Meyer L, Kelleher AD, Rosinska M, Ghosn J, Sannes M, et al. Immunovirologic control 24 months after interruption of antiretroviral therapy initiated close to HIV seroconversion. Arch Intern Med [Internet]. 2012 Sep 10 [cited 2021 Jan 20];172(16):1252–5. Available from:

https://jamanetwork.com/journals/jamainternalmedicine/fullarticle/1221717

70. Frange P, Faye A, Avettand-Fenoël V, Bellaton E, Descamps D, Angin M, et al. HIV-1 virological remission lasting more than 12 years after interruption of early antiretroviral therapy in a perinatally infected teenager enrolled in the French ANRS EPF-CO10 paediatric cohort: A case report. Lancet HIV [Internet]. 2016 Jan 1 [cited 2021 Jan 14];3(1):e49–54. Available from: https://pubmed.ncbi.nlm.nih.gov/26762993/

71. Goetghebuer T, Le Chenadec J, Haelterman E, Galli L, Dollfus C, Thorne C, et al. Short- and Long-term Immunological and Virological Outcome in HIV-Infected Infants According to the Age at Antiretroviral Treatment Initiation. Clin Infect Dis [Internet]. 2012 Mar 15 [cited 2018 Aug 2];54(6):878–81. Available from: https://academic.oup.com/cid/article/54/6/878/289914

72. Luzuriaga K, McManus M, Mofenson L, Britto P, Graham B, Sullivan JL. A Trial of Three Antiretroviral Regimens in HIV-1–Infected Children. N Engl J Med [Internet]. 2004 Jun 10 [cited 2019 Mar 22];350(24):2471–80. Available from: http://www.nejm.org/doi/abs/10.1056/NEJMoa032706

73. Goulder PJ, Lewin SR, Leitman EM. Paediatric HIV infection: The potential for cure [Internet]. Vol. 16, Nature Reviews Immunology. Nature Publishing Group; 2016 [cited 2021 Jan 20]. p. 259–71. Available from: https://www.nature.com/articles/nri.2016.19

74. Technau KG, Kuhn L, Coovadia A, Murnane PM, Sherman G. Xpert HIV-1 point-of-care test for neonatal diagnosis of HIV in the birth testing programme of a maternity hospital: a field evaluation study. Lancet HIV [Internet]. 2017 Oct 1 [cited 2021 Jan 20];4(10):e442–8. Available from: https://pubmed.ncbi.nlm.nih.gov/28711526/

75. Kuhn L, Strehlau R, Shiau S, Patel F, Shen Y, Technau KG, et al. Early antiretroviral treatment of infants to attain HIV remission. EClinicalMedicine [Internet]. 2020 Jan 1 [cited 2021 Jan 20];18. Available from: https://doi.org/10.1016/j.eclinm.2019.100241

76. Hoenigl M, Chaillon A, Moore DJ, Morris SR, Mehta SR, Gianella S, et al. Rapid HIV Viral Load Suppression in those Initiating Antiretroviral Therapy at First Visit after HIV Diagnosis. Sci Rep [Internet]. 2016 Dec 6 [cited 2019 Mar 21];6(1):32947. Available from: http://www.nature.com/articles/srep32947

77. Ford N, Migone C, Calmy A, Kerschberger B, Kanters S, Nsanzimana S, et al. Benefits and risks of rapid initiation of antiretroviral therapy. AIDS [Internet]. 2018 Jan [cited 2019 Mar 21];32(1):17–23. Available from: http://insights.ovid.com/crossref?an=00002030-201801020-00003

78. Phillips AN, Staszewski S, Weber R, Kirk O, Francioli P, Miller V, et al. HIV Viral Load Response to Antiretroviral Therapy According to the Baseline CD4 Cell Count and Viral Load. JAMA [Internet]. 2001 Nov 28 [cited 2019 Mar 21];286(20):2560. Available from: http://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.286.20.2560

79. Jiamsakul A, Kariminia A, Althoff KN, Cesar C, Cortes CP, Davies M-A, et al. HIV Viral Load Suppression in Adults and Children Receiving Antiretroviral Therapy-Results From the IeDEA Collaboration. J Acquir Immune Defic Syndr [Internet]. 2017 [cited 2019 Mar 22];76(3):319–29. Available from: http://www.ncbi.nlm.nih.gov/pubmed/28708808

80. Chan MK, Goodall R, Judd A, Klein N, Chiappini E, Klimkait T, et al. Predictors of faster virological suppression in early treated infants with perinatal HIV from Europe and Thailand. AIDS [Internet]. 2019 Feb 7 [cited 2019 Apr 1];1. Available from: http://www.ncbi.nlm.nih.gov/pubmed/30741823

81. Siberry GK, Harris DR, Oliveira RH, Krauss MR, Hofer CB, Tiraboschi AA, et al. Evaluation of Viral Load Thresholds for Predicting New World Health Organization Stage 3 and 4 Events in HIV-Infected Children Receiving Highly Active Antiretroviral Therapy. JAIDS J Acquir Immune Defic Syndr [Internet]. 2012 Jun;60(2):214–8. Available from: http://journals.lww.com/00126334-201206010-00019

82. PENPACT-1 (PENTA 9/PACTG 390) Study Team, Babiker A, Castro nee Green H, Compagnucci A, Fiscus S, Giaquinto C, et al. First-line antiretroviral therapy with a protease inhibitor versus non-nucleoside reverse transcriptase inhibitor and switch at higher versus low viral load in HIV-infected children: an open-label, randomised phase 2/3 trial. Lancet Infect Dis [Internet]. 2011 Apr [cited 2016 Dec 11];11(4):273–83. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1473309910703133

83. Napravnik S, Poole C, Thomas JC, Eron JJ. Gender Difference in HIV RNA Levels: A Meta-Analysis of Published Studies. JAIDS J Acquir Immune Defic Syndr [Internet].

2002 Sep;31(1):11–9. Available from: http://journals.lww.com/00126334-200209010-00002

84. Anastos K, Gange SJ, Lau B, Weiser B, Detels R, Giorgi J V., et al. Association of Race and Gender With HIV-1 RNA Levels and Immunologic Progression. JAIDS J Acquir Immune Defic Syndr [Internet]. 2000 Jul;24(3):218–26. Available from: http://journals.lww.com/00126334-200007010-00004

85. Meier A, Chang JJ, Chan ES, Pollard RB, Sidhu HK, Kulkarni S, et al. Sex differences in the Toll-like receptor–mediated response of plasmacytoid dendritic cells to HIV-1. Nat Med [Internet]. 2009 Aug 13;15(8):955–9. Available from: http://www.nature.com/articles/nm.2004

86. Chang JJ, Woods M, Lindsay RJ, Doyle EH, Griesbeck M, Chan ES, et al. Higher Expression of Several Interferon-Stimulated Genes in HIV-1-Infected Females After Adjusting for the Level of Viral Replication. J Infect Dis [Internet]. 2013 Sep 1;208(5):830–8. Available from: https://academic.oup.com/jid/article-lookup/doi/10.1093/infdis/jit262

87. Sterling TR, Lyles CM, Vlahov D, Astemborski J, Margolick JB, Quinn TC. Sex Differences in Longitudinal Human Immunodeficiency Virus Type 1 RNA Levels among Seroconverters. J Infect Dis [Internet]. 1999 Sep;180(3):666–72. Available from: https://academic.oup.com/jid/article-lookup/doi/10.1086/314967

88. Evans JS, Nims T, Cooley J, Bradley W, Jagodzinski L, Zhou S, et al. Serum Levels of Virus Burden in Early-Stage Human Immunodeficiency Virus Type 1 Disease in Women. J Infect Dis [Internet]. 1997 Apr;175(4):795–800. Available from: https://academic.oup.com/jid/article-lookup/doi/10.1086/513973

89. Meditz AL, MaWhinney S, Allshouse A, Feser W, Markowitz M, Little S, et al. Sex, Race, and Geographic Region Influence Clinical Outcomes Following Primary HIV-1 Infection. J Infect Dis [Internet]. 2011 Feb 15;203(4):442–51. Available from: https://academic.oup.com/jid/article-lookup/doi/10.1093/infdis/jiq085

90. Zanoni BC, Phungula T, Zanoni HM, France H, Feeney ME. Risk Factors Associated with Increased Mortality among HIV Infected Children Initiating Antiretroviral Therapy

(ART) in South Africa. Myer L, editor. PLoS One [Internet]. 2011 Jul 29;6(7):e22706. Available from: https://dx.plos.org/10.1371/journal.pone.0022706

91. Galvin SR, Cohen MS. The role of sexually transmitted diseases in HIV transmission. Nat Rev Microbiol [Internet]. 2004 Jan;2(1):33–42. Available from: http://www.nature.com/articles/nrmicro794

92. Farzadegan H, Hoover DR, Astemborski J, Lyles CM, Margolick JB, Markham RB, et al. Sex differences in HIV-1 viral load and progression to AIDS. Lancet [Internet]. 1998 Nov;352(9139):1510–4. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0140673698023721

93. Girum T, Wasie A, Lentiro K, Muktar E, Shumbej T, Difer M, et al. Gender disparity in epidemiological trend of HIV/AIDS infection and treatment in Ethiopia. Arch Public Heal [Internet]. 2018 Dec 17;76(1):51. Available from: https://archpublichealth.biomedcentral.com/articles/10.1186/s13690-018-0299-8

94. Richman DD, Margolis DM, Delaney M, Greene WC, Hazuda D, Pomerantz RJ. The Challenge of Finding a Cure for HIV Infection. Science (80- ) [Internet]. 2009 Mar 6;323(5919):1304–7. Available from: https://www.sciencemag.org/lookup/doi/10.1126/science.1165706

95. Rocca S, Zangari P, Cotugno N, De Rossi A, Ferns B, Petricone D, et al. Human Immunodeficiency Virus (HIV)-Antibody Repertoire Estimates Reservoir Size and Time of Antiretroviral Therapy Initiation in Virally Suppressed Perinatally HIV-Infected Children. J Pediatric Infect Dis Soc [Internet]. 2018 Aug 28 [cited 2019 Apr 22]; Available from: https://academic.oup.com/jpids/advance-article/doi/10.1093/jpids/piy080/5085568

96. O'Connor J, Smith C, Lampe FC, Johnson MA, Chadwick DR, Nelson M, et al. Durability of viral suppression with first-line antiretroviral therapy in patients with HIV in the UK: an observational cohort study. Lancet HIV [Internet]. 2017 Jul 1 [cited 2020 Aug 11];4(7):e295–302. Available from: https://pubmed.ncbi.nlm.nih.gov/28479492/

97. McKinnon E, Castley A, Payne L, Pummer S, Nolan D. Determinants of residual viraemia during combination HIV treatment: Impacts of baseline HIV RNA levels and

treatment choice. HIV Med [Internet]. 2016 Aug 1 [cited 2020 Aug 11];17(7):495–504. Available from: http://doi.wiley.com/10.1111/hiv.12323

98. Bachmann N, von Siebenthal C, Vongrad V, Turk T, Neumann K, Beerenwinkel N, et al. Determinants of HIV-1 reservoir size and long-term dynamics during suppressive ART. Nat Commun. 2019 Dec 1;10(1).

99. Kuhn L, Paximadis M, Da Costa Dias B, Loubser S, Strehlau R, Patel F, et al. Age at antiretroviral therapy initiation and cell-associated HIV-1 DNA levels in HIV-1-infected children. Sluis-Cremer N, editor. PLoS One [Internet]. 2018 Apr 12 [cited 2019 Mar 13];13(4):e0195514. Available from: http://www.ncbi.nlm.nih.gov/pubmed/29649264

100. Hong FF, Mellors JW. Changes in HIV reservoirs during long-term antiretroviral therapy [Internet]. Vol. 10, Current Opinion in HIV and AIDS. Lippincott Williams and Wilkins; 2015 [cited 2021 Feb 23]. p. 43–8. Available from: /pmc/articles/PMC4567251/

101. Ananworanich J, Schuetz A, Vandergeeten C, Sereti I, Souza Mark de, Rerknimitr R, et al. Impact of multi-targeted antiretroviral treatment on gut t cell depletion and hiv reservoir seeding during acute hiv infection. PLoS One [Internet]. 2012 Mar 30 [cited 2021 Feb 23];7(3). Available from: https://pubmed.ncbi.nlm.nih.gov/22479485/

102. Lee GQ, Lichterfeld M. Diversity of HIV-1 reservoirs in CD4+ T-cell subpopulations [Internet]. Vol. 11, Current Opinion in HIV and AIDS. Lippincott Williams and Wilkins; 2016 [cited 2021 Feb 23]. p. 383–7. Available from: /pmc/articles/PMC4915926/

103. Persaud D, Patel K, Karalius B, Rainwater-Lovett K, Ziemniak C, Ellis A, et al. Influence of age at virologic control on peripheral blood human immunodeficiency virus reservoir size and serostatus in perinatally infected adolescents. JAMA Pediatr [Internet]. 2014 Dec 1 [cited 2020 Aug 11];168(12):1138–46. Available from: https://pubmed.ncbi.nlm.nih.gov/25286283/

104. Zanchetta M, Anselmi A, Vendrame D, Rampon O, Giaquinto C, Mazza A, et al. Early therapy in HIV-1-infected children: effect on HIV-1 dynamics and HIV-1-specific immune response. Antivir Ther [Internet]. 2008;13(1):47–55. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18389898

105. Palma P, Romiti ML, Cancrini C, Pensieroso S, Montesano C, Bernardi S, et al. Delayed

early antiretroviral treatment is associated with an HIV-specific long-term cellular response in HIV-1 vertically infected infants. Vaccine [Internet]. 2008 Sep;26(40):5196–201. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0264410X08003447

106. Goldstein BA, Navar AM, Carter RE. Moving beyond regression techniques in cardiovascular risk prediction: applying machine learning to address analytic challenges. Eur Heart J. 2017;38(23).

107. Chen JH, Asch SM. Machine learning and prediction in medicine-beyond the peak of inflated expectations. Vol. 376, New England Journal of Medicine. Massachussetts Medical Society; 2017. p. 2507–9.

108. Van Der Ploeg T, Austin PC, Steyerberg EW. Modern modelling techniques are data hungry: A simulation study for predicting dichotomous endpoints. BMC Med Res Methodol [Internet]. 2014 Dec 22 [cited 2020 May 21];14(1):137. Available from: https://bmcmedresmethodol.biomedcentral.com/articles/10.1186/1471-2288-14-137

109. Henrich TJ, Deeks SG, Pillai SK. Measuring the Size of the Latent Human Immunodeficiency Virus Reservoir: The Present and Future of Evaluating Eradication Strategies. J Infect Dis [Internet]. 2017 Mar 15;215(suppl_3):S134–41. Available from: https://academic.oup.com/jid/article/215/suppl_3/S134/3779203

110. Abdel-Mohsen M, Richman D, Siliciano RF, Nussenzweig MC, Howell BJ, Martinez-Picado J, et al. Recommendations for measuring HIV reservoir size in cure-directed clinical trials. Nat Med [Internet]. 2020 Sep 7;26(9):1339–50. Available from: https://www.nature.com/articles/s41591-020-1022-1

111. Bruner KM, Wang Z, Simonetti FR, Bender AM, Kwon KJ, Sengupta S, et al. A quantitative approach for measuring the reservoir of latent HIV-1 proviruses. Nature [Internet]. 2019 Feb 7 [cited 2021 Jan 21];566(7742):120–5. Available from: https://www.nature.com/articles/s41586-019-0898-8

112. Gaebler C, Lorenzi JCC, Oliveira TY, Nogueira L, Ramos V, Lu CL, et al. Combination of quadruplex qPCR and next-generation sequencing for qualitative and quantitative analysis of the HIV-1 latent reservoir. J Exp Med [Internet]. 2019 Oct 1 [cited 2021 Jan 21];216(10):2253–64. Available from: https://pubmed.ncbi.nlm.nih.gov/31350309/