



Aplicación de modelos mixtos en el análisis de recuentos de células CD4 para pacientes con VIH.

Application of Mixed Models in Analyzing CD4 Cell Counts for HIV Patients.

Recibido: 29/11/2024 - Aceptado: 09/04/2025 - Publicado: 15/09/2025

Autores

Heber Esteban Bermúdez González*

Estadístico. BSc.
Centro de Servicios de Salud,SENA.
Medellín - Colombia
hebermudez@sena.edu.co

Mónica Liliana Arteaga Sierra

Matemáticas aplicadas. MSc.
Centro de Servicios de Salud,SENA.
Medellín - Colombia
marteaga@sena.edu.co

*Autor de correspondencia



RESUMEN

Este estudio evalúa la eficacia de modelos mixtos para analizar el recuento de la glucoproteína CD4 en pacientes con VIH y supresión inmune avanzada, utilizando datos del estudio 193A del Grupo de Ensayo Clínico del SIDA (ACTG) en Estados Unidos. El objetivo fue identificar los factores que influyen en la variabilidad de los recuentos de CD4 y evaluar la eficacia de diferentes tratamientos antirretrovirales. Se aplicaron modelos mixtos que consideraron efectos fijos y aleatorios para variables como edad, sexo, tipo de tratamiento y tiempo de seguimiento. Se compararon varios modelos, destacando el modelo con interacción tratamiento-sexo y términos cuadráticos para el tiempo, que incorporó intercepto y pendientes aleatorias. Este modelo mostró un mejor ajuste según los criterios AIC (Criterio de Información de Akaike) y BIC (Criterio de Información Bayesiano), capturando de manera efectiva la tendencia decreciente en los recuentos de CD4. El análisis demuestra que los modelos mixtos son herramientas poderosas para entender la complejidad de los datos longitudinales en contextos clínicos, proporcionando una base sólida para optimizar la toma de decisiones terapéuticas en pacientes con SIDA.

Palabras clave: Análisis de supervivencia, Estadística longitudinal, Inferencia estadística, Modelos jerárquicos, Variabilidad intra-paciente, VIH/SIDA.

ABSTRACT

This study evaluates the effectiveness of mixed models in analyzing the CD4 glycoprotein count in HIV patients with advanced immune suppression, using data from the AIDS Clinical Trials Group (ACTG) study 193A in the United States. The objective was to identify factors influencing CD4 count variability and to assess the effectiveness of different antiretroviral treatments. Mixed models were applied, considering fixed and random effects for variables such as age, sex, type of treatment, and follow-up time. Several models were compared, highlighting the model with treatment-sex interaction and quadratic terms for time, which incorporated random intercepts and slopes. This model showed the best fit according to AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion) criteria, effectively capturing the declining trend in CD4 counts. The analysis demonstrates that mixed models are powerful tools for understanding the complexity of longitudinal data in clinical contexts, providing a solid foundation for optimizing therapeutic decision-making in AIDS patients.

Keywords: Survival analysis, Longitudinal statistics, Statistical inference, Hierarchical models, Intra-patient variability, HIV/AIDS.



INTRODUCCIÓN

En el contexto de la inmunología, la glucoproteína CD4 es un componente crítico del sistema inmunológico humano. Esta proteína de superficie se encuentra en diversas células inmunitarias, incluyendo linfocitos T auxiliares, monocitos, macrófagos y células dendríticas. La función principal de los linfocitos T CD4+ es coordinar las respuestas inmunitarias adaptativas a través de la secreción de citocinas y el contacto célula a célula. Su función se realiza principalmente mediante la activación de otras células del sistema inmunológico, como los linfocitos T citotóxicos CD8+, que son responsables de eliminar células infectadas y patógenos. La disminución en el número de células CD4, como se observa en infecciones por VIH no tratadas, compromete gravemente la capacidad del organismo para combatir infecciones, aumentando la susceptibilidad a enfermedades que, en condiciones normales, serían controladas eficazmente.

El recuento de CD4 se utiliza comúnmente como un marcador para evaluar el estado del sistema inmunológico en pacientes infectados con el VIH. Los modelos mixtos o modelos jerárquicos se han aplicado para analizar los recuentos de CD4 debido a su capacidad para manejar datos con efectos fijos y aleatorios. Estos modelos son particularmente útiles en estudios longitudinales donde se realizan mediciones repetidas en los mismos individuos, permitiendo capturar la variabilidad intra e interindividual.

A pesar de la importancia del recuento de CD4+, existe una comprensión limitada de los factores que influyen en su variabilidad. La investigación ha demostrado que factores como el tipo de tratamiento antirretroviral, el tiempo de seguimiento y características demográficas como la edad y el sexo pueden afectar los recuentos de CD4+. Sin embargo, el grueso de los estudios desarrollados a la fecha ha utilizado métodos estadísticos que no capturan adecuadamente la complejidad de los datos longitudinales, lo que limita la capacidad para identificar patrones y relaciones significativas.

En este estudio, se presenta una investigación de los recuentos de CD4 en pacientes con supresión inmune avanzada utilizando modelos mixtos o modelos jerárquicos. Los datos provienen del estudio 193A del Grupo de Ensayo Clínico del SIDA (ACTG) realizado en Estados Unidos, en el cual los pacientes fueron asignados aleatoriamente a combinaciones duales o triples de inhibidores del VIH-1. Este análisis es crucial para mejorar la comprensión de los factores que influyen en el recuento de CD4 y la eficacia de los tratamientos contra el VIH-1 en pacientes con SIDA, así como proporcionar una metodología estadística robusta para analizar y hacer inferencias sobre esta enfermedad, lo cual es fundamental para avanzar en la investigación de esta.

El objetivo de este estudio es aplicar modelos mixtos para analizar cómo factores como el tipo de tratamiento, la semana de seguimiento, el sexo y la edad del paciente influyen en el recuento de CD4. Además, se exploran posibles interacciones entre estos factores mediante la construcción de varios modelos con diferentes niveles de complejidad. La selección del mejor modelo se basa en varios criterios, incluyendo el AIC y BIC, la correlación entre los valores reales y ajustados, y el MSE.

MARCO TEÓRICO

La importancia del recuento de células CD4+ radica en su capacidad para reflejar el estado inmunológico de los pacientes. Este recuento es un marcador crucial en el manejo de la infección por VIH, ya que permite evaluar tanto la progresión de la enfermedad como la efectividad del tratamiento antirretroviral (ART). Estudios han demostrado que el recuento de CD4+ es un parámetro fundamental para iniciar la terapia antirretroviral (1,2).

Un aumento adecuado en el recuento de células CD4+ está significativamente asociado con una mayor supervivencia en pacientes con VIH (3,4). El recuento de células CD4+ al inicio del tratamiento ART puede servir como un indicador del momento de detección del VIH. Un recuento bajo de CD4+ (≤ 350 células/ μ l) generalmente señala una detección



tardía, lo cual se asocia con un mayor riesgo de progresión rápida a SIDA y una mayor mortalidad relacionada. Por lo tanto, un inicio temprano del tratamiento ART, reflejado en recuentos de CD4+ más altos, es esencial para mejorar los resultados clínicos y reducir la mortalidad (2,4,5).

En cuanto al análisis del recuento de células CD4+ en pacientes con VIH, la eficacia de los modelos mixtos ha sido explorada en la literatura. Estos modelos son especialmente útiles para analizar datos longitudinales en contextos clínicos, ya que pueden manejar la correlación entre observaciones repetidas en el tiempo para un mismo individuo e incorporar efectos tanto fijos como aleatorios (6). Algunos estudios han demostrado que estos modelos capturan de manera efectiva la variabilidad en los recuentos de CD4+, evaluando no solo los efectos del tiempo, sino también las interacciones entre diferentes covariables. Por ejemplo, un estudio que analizó datos de pacientes con VIH en China utilizando modelos mixtos encontró diferencias significativas en los recuentos de CD4+ según el género y la edad, capturando variaciones en la progresión de la enfermedad y la respuesta inmunológica individualizada al ART (7).

Un estudio similar realizado en 21 países de África subsahariana analizó las diferencias de género en la supresión de la carga viral y las tasas de mortalidad en personas bajo tratamiento antirretroviral (ART), utilizando modelos mixtos para evaluar tanto la carga viral como los recuentos de CD4+ (8). Estos modelos también permitieron predecir el tiempo necesario para la supresión viral y los cambios en los niveles de CD4+ en bebés con VIH (9). Asimismo, en España, un estudio con pacientes de la red CoRIS empleó modelos mixtos para comparar los cambios en la relación CD4/CD8 a lo largo del tiempo, evaluando el impacto de diferentes terapias antirretrovirales de primera línea en la recuperación a largo plazo de dicha relación en personas con VIH (5).

En otro estudio realizado en San Petersburgo, Rusia, que investigó el impacto del consumo excesivo de alcohol en mujeres con VIH, los modelos mixtos de efectos múltiples mostraron que las mujeres que

bebían alcohol en exceso tenían en promedio 41 células CD4/mm³ menos que aquellas que no consumían alcohol excesivamente. Además, reportaban más síntomas del VIH y menor adherencia al tratamiento antirretroviral (10). Un estudio similar sobre los determinantes de la falta de adherencia al ART en mujeres con VIH/SIDA identificó factores como el nivel educativo, la edad, los ingresos, y aspectos personales como el apoyo sociofamiliar y el estigma percibido, entre otros (1,2).

El uso de criterios de información como el AIC (Criterio de Información de Akaike) y el BIC (Criterio de Información Bayesiano) ha sido fundamental para comparar la eficacia de diferentes modelos mixtos. Estos criterios permiten seleccionar el modelo que mejor se ajusta a los datos, penalizando la complejidad del modelo para evitar el sobreajuste (1). Aunque los modelos mixtos ofrecen una herramienta poderosa para el análisis de datos complejos, presentan ciertos desafíos, como la selección adecuada del modelo y la interpretación de interacciones complejas entre variables. (11,12).

METODOLOGÍA

A continuación, describe detalladamente la metodología empleada en el estudio. Se abordan aspectos clave como la descripción del estudio clínico, el análisis estadístico realizado, y el modelamiento utilizado para interpretar los datos.

Descripción del estudio clínico

El conjunto de datos usado corresponde al estudio 193A del Grupo de Ensayo Clínico del SIDA realizado en Estados Unidos. Participaron pacientes con SIDA que presentaban supresión inmune avanzada, definida como un recuento de la glucoproteína CD4 igual o inferior a 50 células/mm³. Los pacientes fueron asignados aleatoriamente a una de cuatro combinaciones de tratamientos diarios que incluían 600 mg de zidovudina. Las mediciones del recuento de CD4 se realizaron al inicio del estudio y cada 8 semanas durante el seguimiento.



Variables del estudio

- **CD4:** Recuento transformado como $\log(\text{CD4} + 1)$, que es la variable de respuesta.
- **ID:** Identificación del paciente.
- **Tratamiento:** Variable categórica codificada como:
 - 1 = Zidovudina alternando mensualmente con 400 mg de didanosina.
 - 2 = Zidovudina más 2.25 mg de zalcitabina.
 - 3 = Zidovudina más 400 mg de didanosina.
 - 4 = Zidovudina más 400 mg de didanosina y 400 mg de nevirapina.
- **Semana:** Tiempo desde el inicio del tratamiento.
- **Edad:** Edad en años del paciente.
- **Sexo:** Sexo del paciente, codificado como M (Masculino) y F (Femenino).

Análisis descriptivo de datos

Se realizó un análisis descriptivo inicial para explorar las características de la población estudiada y la distribución de las variables. Esto incluyó la generación de estadísticas descriptivas y gráficos para las variables. La distribución del recuento de CD4 por tratamiento y sexo se evaluó utilizando histogramas y boxplots. A continuación, se pueden observar una parte del conjunto de datos obtenidos del estudio clínico.

Tabla 1. Tabla datos.

ID	Tratamiento	Edad	Sexo	Semana	CD4
1	2	36.4271	M	0.0000	3.135494
1	2	36.4271	M	7.5714	3.044522
1	2	36.4271	M	15.5714	2.772589
1	2	36.4271	M	23.5714	2.833213
1	2	36.4271	M	32.5714	3.218876
1	2	36.4271	M	40.0000	3.044522
2	4	47.8467	M	0.0000	3.068053
2	4	47.8467	M	8.0000	3.891820
2	4	47.8467	M	16.0000	3.970292
2	4	47.8467	M	23.0000	3.610918

Fuente: Elaboración propia.

El conjunto de datos consta de 5036 observaciones y 6 columnas, que corresponden al ID del paciente, tratamiento, edad, sexo, semana de tratamiento y la variable de respuesta, que es el conteo transformado de la glucoproteína CD4. El conjunto de datos incluye un total de 1309 pacientes. Se determinó que el número máximo de seguimientos para un paciente en el estudio fue de 9, mientras que el número mínimo fue de 1. Es importante destacar que aquellos pacientes con solo una observación no proporcionaron información sobre los cambios longitudinales a lo largo de las semanas de estudio. Sin embargo, estos datos fueron útiles para estimar las variaciones y efectos aleatorios entre pacientes.

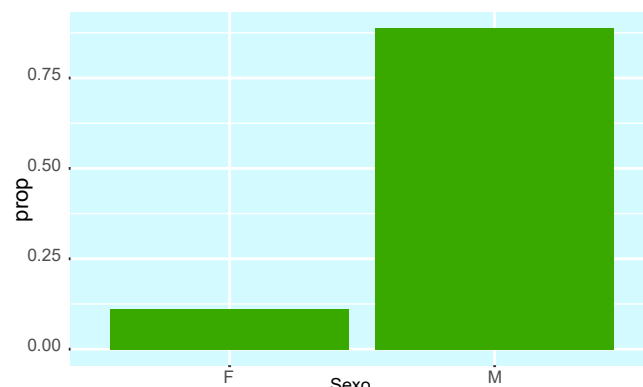
La cantidad de pacientes que recibieron cada uno de los tratamientos se detalla a continuación:

- 325 pacientes recibieron el tratamiento 1
- 324 pacientes recibieron el tratamiento 2
- 330 pacientes recibieron el tratamiento 3
- 330 pacientes recibieron el tratamiento 4

Se observó que la proporción de pacientes que recibieron cada tratamiento fue relativamente equilibrada, lo cual resultó útil para mantener un balance adecuado de las clases de tratamiento, aspecto importante en el estudio.

Se realizó un análisis de la distribución de los pacientes por sexo para evaluar el balance de género en el estudio. La Figura 1 muestra la proporción de pacientes de sexo masculino y femenino en la muestra.

Figura 1. Proporción de pacientes por sexo.



Fuente: Elaboración propia.

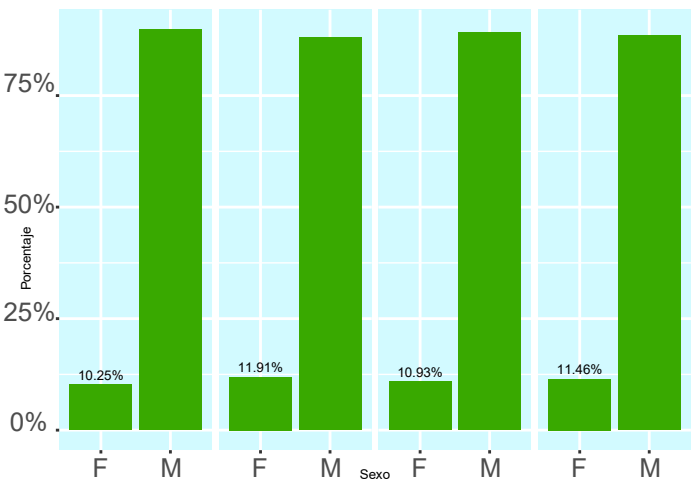
Como se puede observar en la Figura 1, existe un desequilibrio significativo en la distribución del sexo. La mayoría de los pacientes son de sexo masculino, representando aproximadamente el 87.62% de la



muestra, mientras que los pacientes de sexo femenino constituyen solo el 12.38%. Este desequilibrio de sexo es importante de considerar, ya que influye en los resultados y la interpretación del estudio, especialmente si existen diferencias en la respuesta al tratamiento entre hombres y mujeres.

En la Figura 2, se presenta la distribución de sexo por tratamiento, donde se puede apreciar este desbalance por tratamiento.

Figura 2. Distribución del sexo por tratamiento.

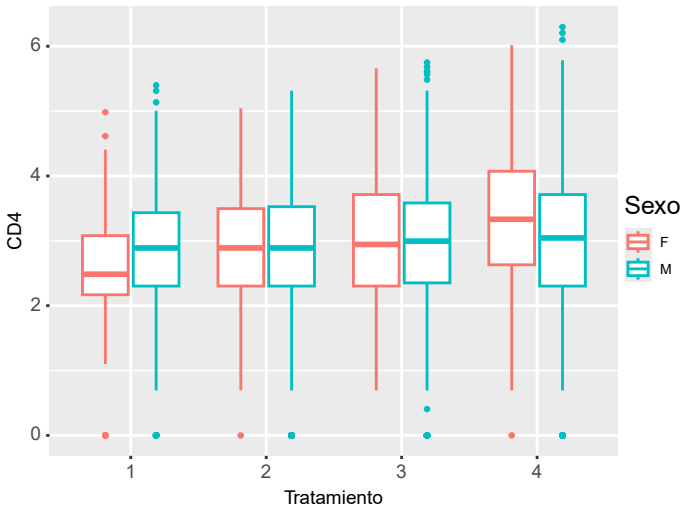


Fuente: Elaboración propia.

Este desbalance del sexo se consideró al interpretar los resultados del modelo, ya que la baja proporción de datos de sexo femenino puede influir en la estimación de los efectos del tratamiento. La baja cantidad de datos correspondientes a mujeres implica que el modelo podría tener un mejor desempeño en el sexo masculino. No obstante, se optó por no aplicar metodologías de remuestreo debido a que la cantidad de datos femeninos disponibles fue considerada suficiente para realizar inferencias razonables y evaluar el modelo adecuadamente.

En la Figura 3, se presenta el conteo de CD4 por tratamiento y género. Este gráfico permite visualizar posibles diferencias en la respuesta al tratamiento entre hombres y mujeres.

Figura 3. Conteo de CD4 por tratamiento.

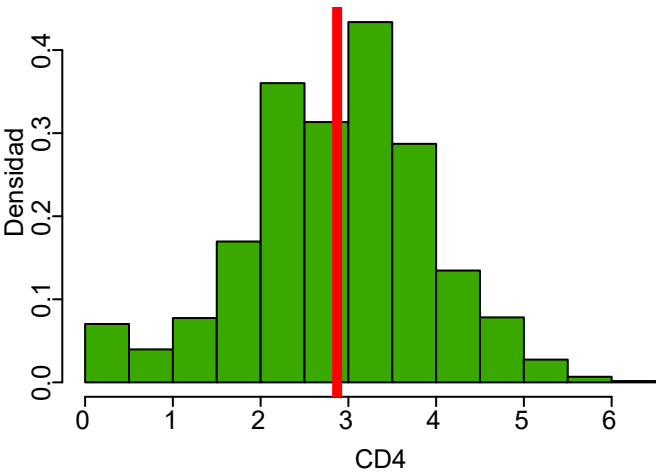


Fuente: Elaboración propia.

Se observó que los tratamientos 1 y 4 presentaron comportamientos levemente diferentes cuando se subdividieron por género. Aunque estas diferencias no fueron claramente evidentes, resultó importante incluir la interacción entre tratamiento y género en el modelo mixto para los efectos fijos, y verificar si esta interacción era significativa y mejoraba el ajuste del modelo.

Para una mejor comprensión de la variable respuesta, se presenta a continuación un histograma de su distribución. La Figura 4 muestra la distribución del conteo de CD4 transformado.

Figura 4. Histograma CD4.



Fuente: Elaboración propia.

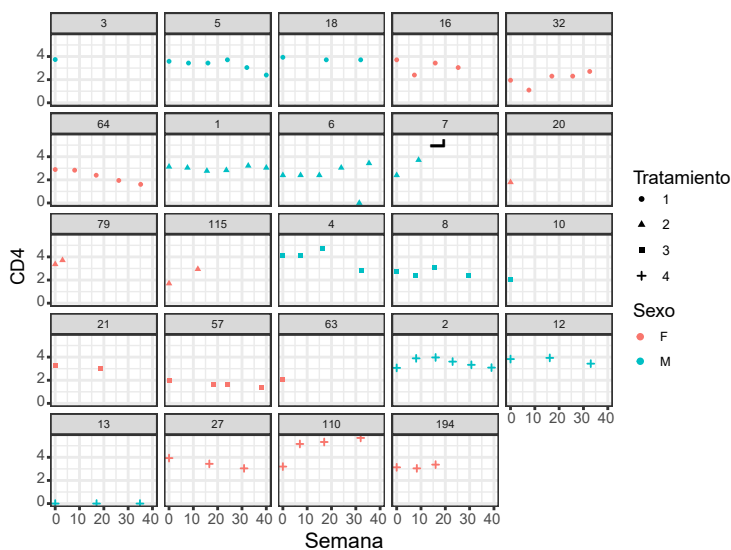


Se observó que esta variable se comportó aproximadamente como una distribución normal, con una media cercana al valor de 3. Sin embargo, también fue evidente que la distribución presenta una cola izquierda más pesada, lo que podría indicar una asimetría en los datos. Esta asimetría debe ser considerada al realizar análisis estadísticos y al interpretar los resultados.

A continuación, se seleccionaron aleatoriamente algunos pacientes para observar cómo se comporta el fenómeno a nivel individual, con el objetivo de identificar posibles variaciones o variables importantes a considerar en los efectos aleatorios intrapacientes.

En la Figura 5, se pueden observar las mediciones de CD4 a lo largo del tiempo para los pacientes seleccionados, diferenciados por sexo y tratamiento.

Figura 5. Mediciones de CD4 por paciente, diferenciadas por sexo y tratamiento.



Fuente: Elaboración propia.

Se observó que en el tratamiento 1, las mediciones de CD4 comenzaron ligeramente más bajas para las mujeres que para los hombres, lo cual se nota en las parcelas 32 y 64, donde las mediciones iniciales están un poco por debajo de 4, mientras que en las parcelas 3, 5 y 18, correspondientes a hombres, las mediciones iniciales están más cerca de 4. En general, es difícil apreciar a simple vista el efecto del

tratamiento debido a la variabilidad entre pacientes. Sin embargo, se destacó que las mediciones para las mujeres fueron un poco más bajas en la mayoría de los tratamientos. En particular, en las parcelas 32, 64, 115, 57, 63 y 194, que corresponden a mujeres, se observaron mediciones de CD4 inferiores a 4, mientras que para los hombres estas mediciones fueron más cercanas a 4. En consecuencia, se consideró el sexo como una variable importante para los efectos aleatorios.

Cabe destacar que, en la mayoría de los pacientes, a lo largo de las semanas, los recuentos de CD4 disminuyeron, lo cual es preocupante para la salud del paciente. Sin embargo, es difícil evitar esta situación debido a su condición de salud. En la parcela 110, se observó que la paciente con tratamiento 4 mostró una leve mejoría en el recuento de CD4, al igual que en la parcela 194. No obstante, no se pudo afirmar si esta mejoría fue debida al efecto del tratamiento 4, ya que en la parcela 27, las observaciones fueron decrecientes. Por tanto, fue necesario incluir pendientes aleatorias en el modelamiento para capturar estos casos.

También se tuvo en cuenta los bajos conteos de algunos pacientes, como el paciente 13, un hombre con tratamiento 4, quien presentó niveles de CD4 inferiores a los demás pacientes en esta muestra. Por lo tanto, fue importante considerar un intercepto aleatorio para los pacientes en el modelamiento, ya que, para algunos de ellos, sus conteos comenzaron en niveles muy distintos.

Algunas observaciones mostraron un patrón que podría ser descrito por una función cuadrática en relación con alguna de las variables continuas. Este comportamiento se observó en las parcelas 5, 64, 2 y 110. Por lo tanto, fue relevante considerar la inclusión de un término cuadrático en el modelo y comparar su desempeño con otros modelos.

Modelos Mixtos Considerados

Se desarrollaron y compararon varios modelos mixtos para analizar la relación entre el recuento de CD4 y los factores mencionados:



Modelo 1: Modelo Lineal con Intercepto Aleatorio

A continuación, se presenta el modelo matemático que se ajustó. En primer lugar, se consideró la combinación lineal de todas las covariables involucradas y únicamente el intercepto aleatorio.

Ecuación 1

$$CD4_{ij} \sim N(\mu_{ij}, \sigma_{CD4}^2)$$

$$\mu_{ij} = \beta_0 + \beta_1 Tratamiento_{2i} + \beta_2 Tratamiento_{3i} + \beta_3 Tratamiento_{4i} + \beta_4 Edad_{ij} + \beta_5 Sexo_{Mi} + \beta_6 Semana_{ij} + b_{0i}$$

$$b_{0i} \sim N(0, \sigma_{b0}^2)$$

CD4 corresponde al conteo transformado de la glucoproteína CD4 y donde el vector de parámetros para este modelo es:

Ecuación 2.

$$\Theta = (\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6, \sigma_{CD4}, \sigma_{b0})^T$$

Modelo 2: Modelo Lineal con Intercepto y Pendiente Aleatorio.

A continuación, se presenta el segundo modelo matemático que se ajustó. En este modelo, se consideraron tanto el intercepto como la pendiente aleatoria para capturar la variabilidad en el tiempo.

Ecuación 3.

$$CD4_{ij} \sim N(\mu_{ij}, \sigma_{CD4}^2)$$

$$\mu_{ij} = \beta_0 + \beta_1 Tratamiento_{2i} + \beta_2 Tratamiento_{3i} + \beta_3 Tratamiento_{4i} + \beta_4 Edad_{ij} + \beta_5 Sexo_{Mi} + \beta_6 Semana_{ij} + b_{0i} + b_{1i} Semana_{ij}$$

$$\begin{pmatrix} b_{0i} \\ b_{1i} \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{b0}^2 & \sigma_{b0b1} \\ \sigma_{b0b1} & \sigma_{b1}^2 \end{pmatrix} \right)$$

CD4 corresponde al conteo transformado de la glucoproteína CD4 y donde el vector de parámetros para este modelo es:

Ecuación 4.

$$\Theta = (\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6, \sigma_{CD4}, \sigma_{b0}, \sigma_{b1}, \sigma_{b0b1})^T$$

Modelo 3: Modelo con Interacción Tratamiento-Sexo, Intercepto y Pendiente Aleatorios.

A continuación, se presenta el tercer modelo matemático que se ajustó. En este modelo, se consideraron tanto el intercepto como la pendiente aleatoria, además de la interacción entre el tratamiento y el sexo.

Ecuación 5.

$$CD4_{ij} \sim N(\mu_{ij}, \sigma_{CD4}^2)$$

$$\mu_{ij} = \beta_0 + \beta_1 Tratamiento_{2i} + \beta_2 Tratamiento_{3i} + \beta_3 Tratamiento_{4i} + \beta_4 Edad_{ij} + \beta_5 Sexo_{Mi} + \beta_6 Semana_{ij} + \beta_7 (Tratamiento_{2i} \cdot Sexo_{Mi}) + \beta_8 (Tratamiento_{3i} \cdot Sexo_{Mi}) + \beta_9 (Tratamiento_{4i} \cdot Sexo_{Mi}) + b_{0i} + b_{1i} Semana_{ij}$$

$$\begin{pmatrix} b_{0i} \\ b_{1i} \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{b0}^2 & \sigma_{b0b1} \\ \sigma_{b0b1} & \sigma_{b1}^2 \end{pmatrix} \right)$$

CD4 corresponde al conteo transformado de la glucoproteína CD4 y donde el vector de parámetros para este modelo es:

Ecuación 6.

$$\Theta = (\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6, \beta_7, \beta_8, \beta_9, \sigma_{CD4}, \sigma_{b0}, \sigma_{b1}, \sigma_{b0b1})^T$$

Modelo 4: Modelo con Interacción Tratamiento-Sexo, Variable Cuadrática de Semana y Efectos Aleatorios.

A continuación, se presenta el cuarto modelo matemático que se ajustó. En este modelo, se consideraron tanto el intercepto como la pendiente aleatoria, la interacción entre tratamiento y sexo, y una variable cuadrática de semana.

Ecuación 7.

$$CD4_{ij} \sim N(\mu_{ij}, \sigma_{CD4}^2)$$

$$\mu_{ij} = \beta_0 + \beta_1 Tratamiento_{2i} + \beta_2 Tratamiento_{3i} + \beta_3 Tratamiento_{4i} + \beta_4 Edad_{ij} + \beta_5 Sexo_{Mi} + \beta_6 Semana_{ij} + \beta_7 Semana_{ij}^2 + \beta_8 (Tratamiento_{2i} \cdot Sexo_{Mi}) + \beta_9 (Tratamiento_{3i} \cdot Sexo_{Mi}) + \beta_{10} (Tratamiento_{4i} \cdot Sexo_{Mi}) + b_{0i} + b_{1i} Semana_{ij} + b_{2i} Semana_{ij}^2$$

$$\begin{pmatrix} b_{0i} \\ b_{1i} \\ b_{2i} \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{b0}^2 & \sigma_{b0b1} & \sigma_{b0b2} \\ \sigma_{b0b1} & \sigma_{b1}^2 & \sigma_{b1b2} \\ \sigma_{b0b2} & \sigma_{b1b2} & \sigma_{b2}^2 \end{pmatrix} \right)$$



CD4"# corresponde al conteo transformado de la glucoproteína CD4 y donde el vector de parámetros para este modelo es:

Ecuación 8.

$$\Theta = (\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6, \beta_7, \beta_8, \beta_9, \beta_{10}, \sigma_{CD4}, \sigma_{b0}, \sigma_{b1}, \sigma_{b2}, \sigma_{b0b1}, \sigma_{b0b2}, \sigma_{b1b2})^T$$

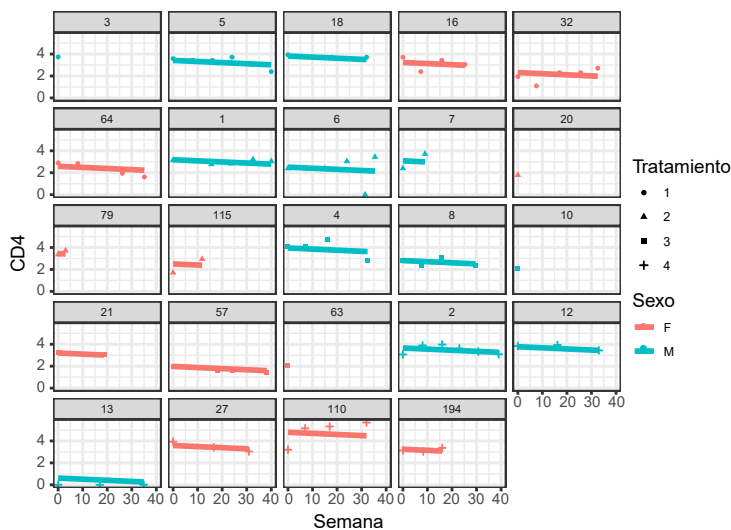
Entrenamiento de Modelos.

Se utilizó la función lme() del paquete nlme en R para ajustar los modelos a los datos. Los criterios de selección de modelos incluyeron el AIC, BIC, la correlación entre los valores reales y ajustados, y el error cuadrático medio (MSE).

Comparación de Modelos.

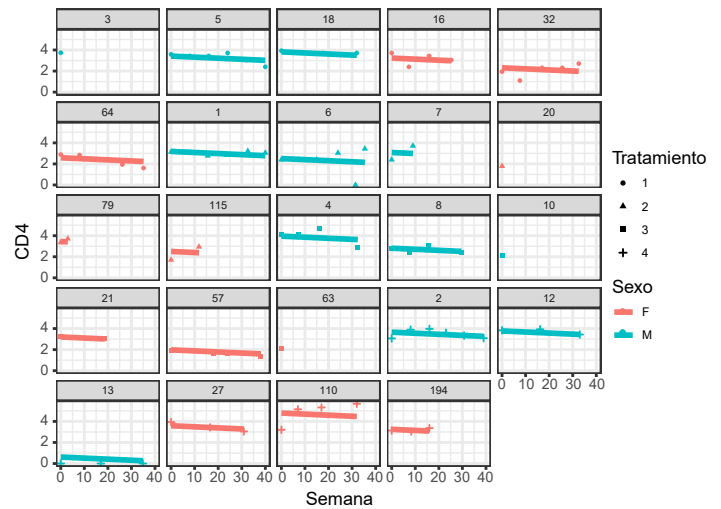
Los modelos se compararon utilizando pruebas de razón de verosimilitud y análisis de residuales. El modelo con el valor más bajo de AIC y BIC, la mayor correlación y el menor MSE se seleccionó como el mejor modelo ajustado.

Figura 6. Ajuste de modelo 1.



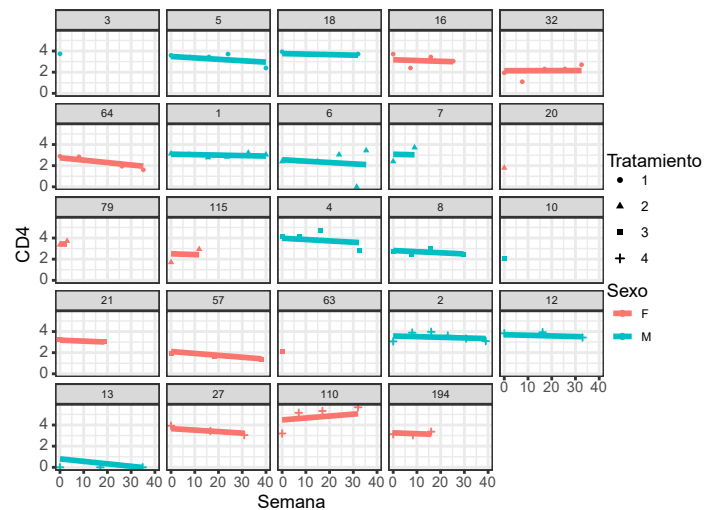
Fuente: Elaboración propia.

Figura 7. Ajuste de modelo 2.



Fuente: Elaboración propia.

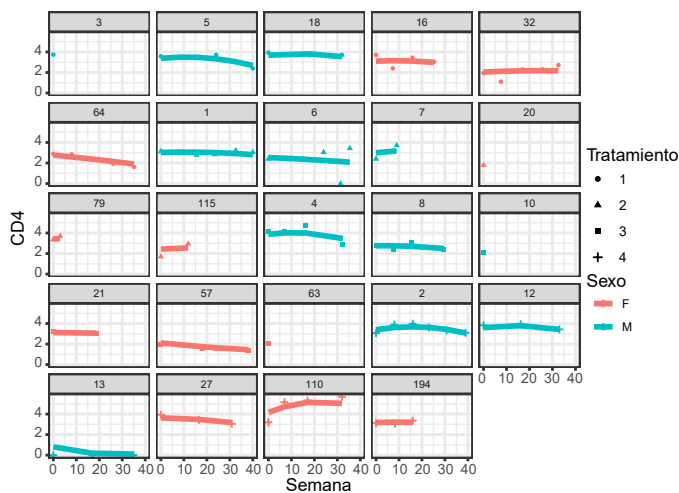
Figura 8. Ajuste de modelo 3.



Fuente: Elaboración propia.



Figura 9. Ajuste de modelo 4.



Fuente: Elaboración propia.

Tabla 2. Comparación de Métricas de Ajuste de Modelos.

Métricas	Modelo 1	Modelo 2	Modelo 3	Modelo 4
AIC	12281.35	12143.67	12147.18	12017.50
BIC	12340.05	12215.42	12238.50	12134.90
logLik	-6131.67	-6060.83	-6059.59	-5990.75

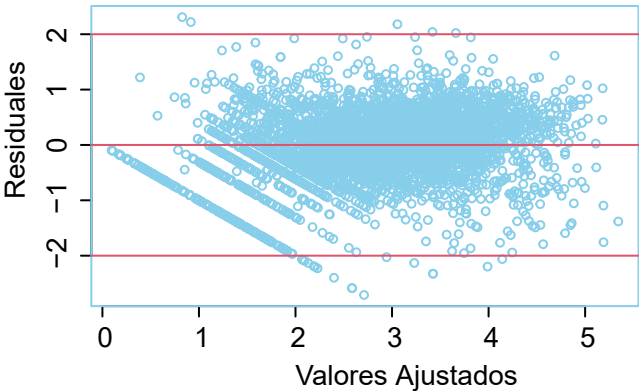
Fuente: Elaboración propia.

Análisis de Residuales y Normalidad de Efectos Aleatorios.

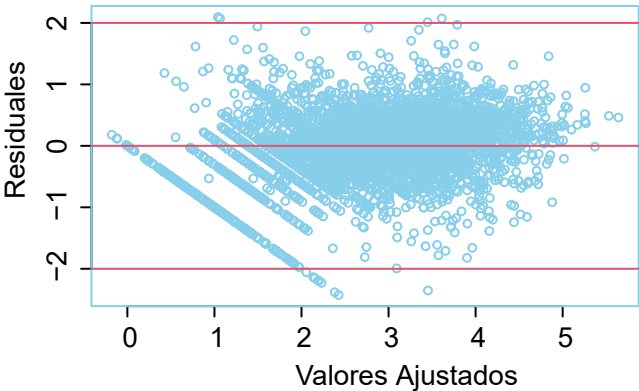
Se evaluaron los residuales para verificar el cumplimiento de los supuestos de homogeneidad de varianza y normalidad. Se generaron gráficos de valores ajustados vs valores reales y se calcularon las correlaciones y el MSE para cada modelo.

Figura 10. Distribución de los residuales de los modelos.

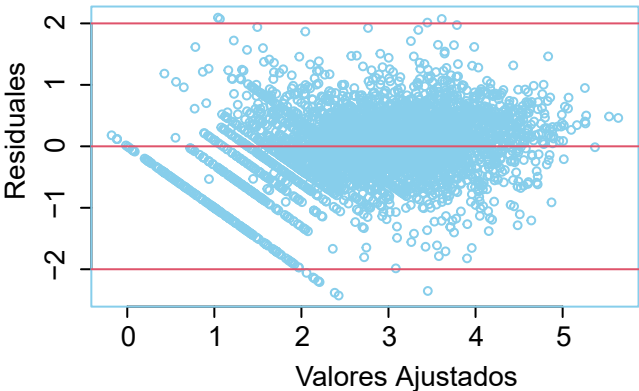
Residuales modelo1



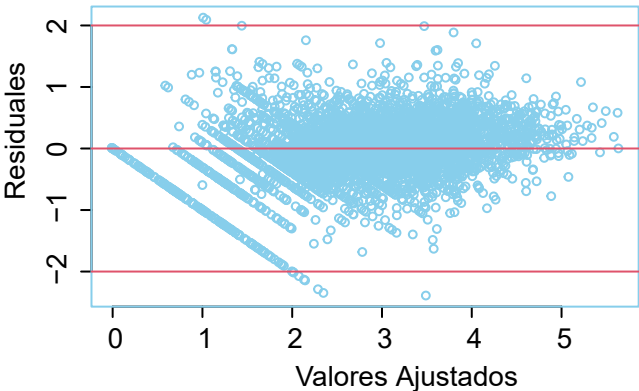
Residuales modelo2



Residuales modelo3



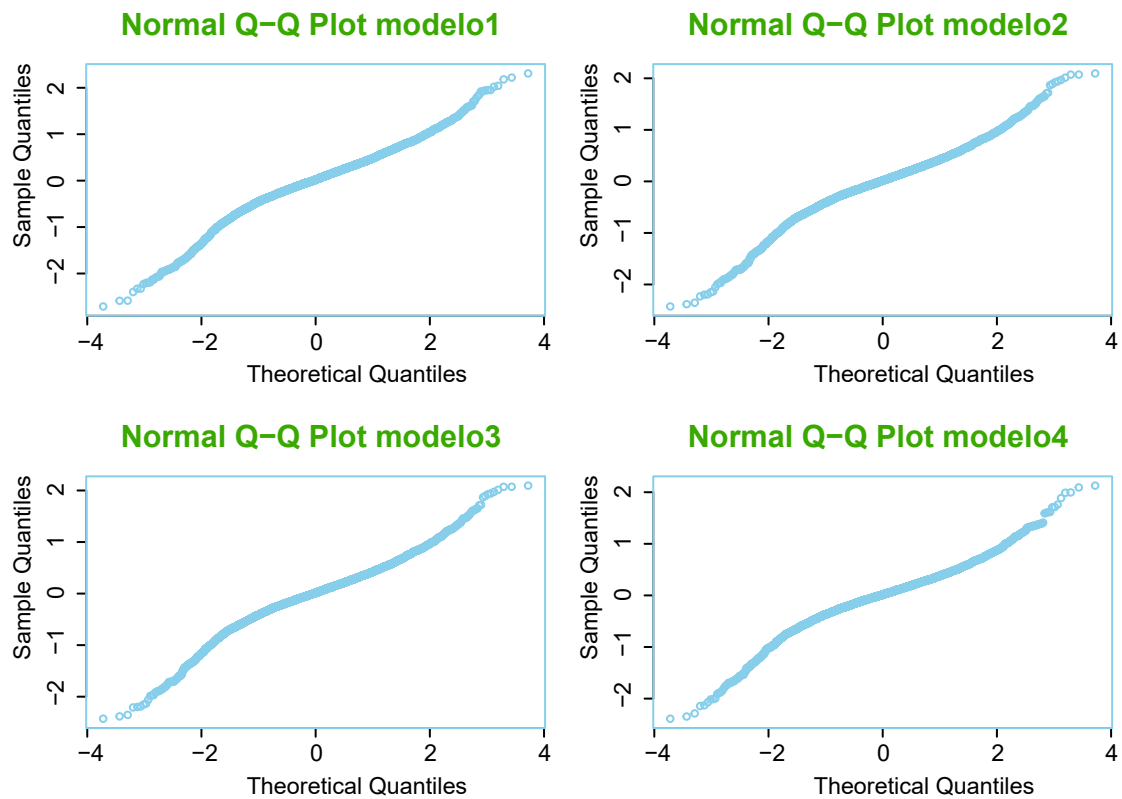
Residuales modelo4



Fuente: Elaboración propia.

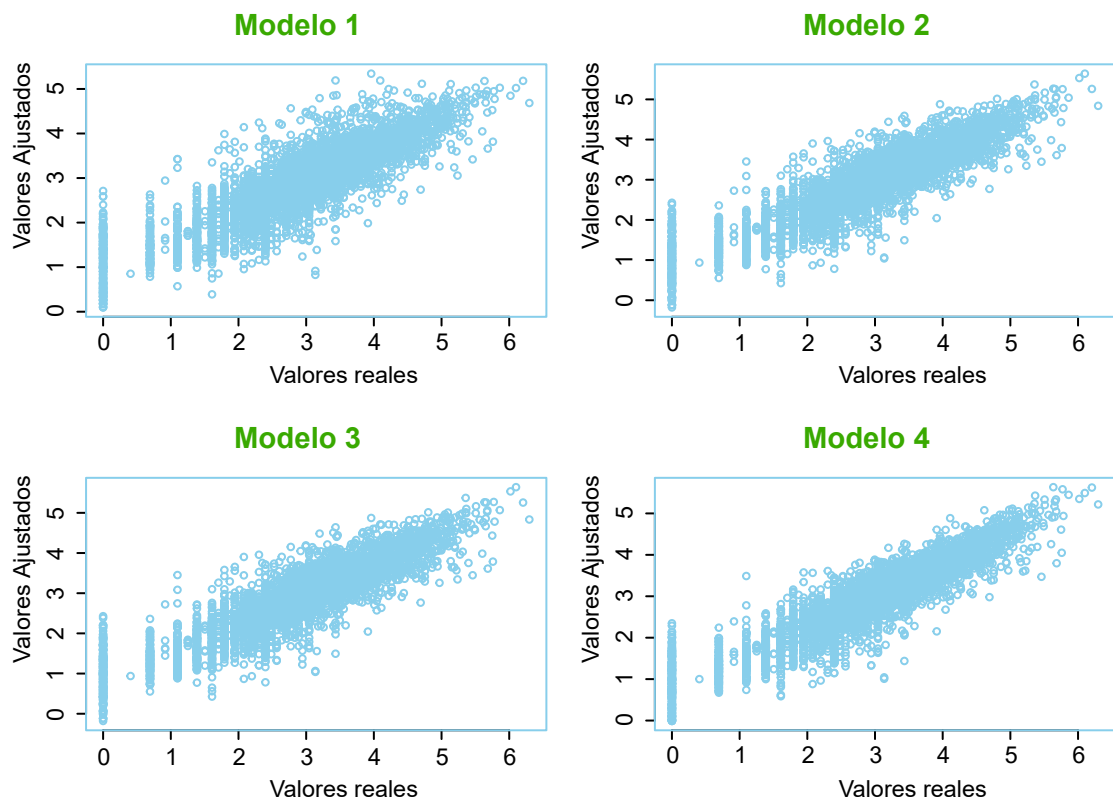


Figura 11. Gráfico QQ-plot, normalidad de los residuales.



Fuente: Elaboración propia.

Figura 12. De valores reales contra valores ajustados.



Fuente: Elaboración propia.



RESULTADOS Y DISCUSIÓN

Los resultados obtenidos a partir del ajuste de modelos mixtos para el recuento de la glucoproteína CD4 en pacientes con supresión inmune avanzada reflejan la capacidad de estos modelos para capturar tanto la variabilidad intraindividual como interindividual en los datos longitudinales. Se evaluaron cuatro modelos con diferentes niveles de complejidad, incluyendo interacciones entre variables y términos cuadráticos para capturar posibles efectos no lineales.

El Modelo 4, que incorpora la interacción entre tratamiento y sexo, además de un término cuadrático para la variable tiempo (Semana), demostró ser el más adecuado según los criterios AIC, BIC, error cuadrático medio (MSE), y la correlación entre valores predichos y observados.

Este modelo logró describir con precisión la tendencia decreciente en los recuentos de CD4 a lo largo del tiempo, ajustando también la heterogeneidad observada entre pacientes.

En particular, el coeficiente asociado con el tratamiento 4 ($\beta = 0.602$, $p < 0.01$) indica una mejora significativa en los recuentos de CD4 en comparación con otros tratamientos, sugiriendo que este tratamiento es más efectivo. Además, la interacción entre el tratamiento 4 y el sexo masculino ($\beta = -0.615$, $p < 0.01$) muestra que la eficacia del tratamiento 4 es significativamente menor en hombres en comparación con mujeres. El término cuadrático para el tiempo ($\beta = -0.000465$, $p < 0.001$) confirma la tendencia no lineal decreciente en los recuentos de CD4 a medida que avanza el tiempo.

El análisis de residuales y la verificación de la normalidad de los efectos aleatorios respaldan la idoneidad del modelo. Los residuales mostraron homogeneidad, y las pruebas de normalidad sugirieron la ausencia de desviaciones significativas en los efectos aleatorios.

A continuación, se presenta una tabla con los

coeficientes relevantes del Modelo 4:

Tabla 3. Estimaciones de Coeficientes, Errores Estándar y Valores p.

Coeficiente	Estimación (β)	Error Estándar	Valor p
Intercepto	2.339	0.184	< 0.001
Tratamiento 2	0.317	0.204	0.121
Tratamiento 3	0.236	0.202	0.243
Tratamiento 4	0.602	0.205	0.003
Edad	0.010	0.003	0.001
Sexo Masculino	0.220	0.160	0.169
Semana	0.005	0.003	0.036
(Semana ²)	-0.000465	0.000068	< 0.001
Tratamiento 2 Masculino	-0.338	0.217	0.119
Tratamiento 3 Masculino	-0.237	0.216	0.271
Tratamiento 4 Masculino	-0.615	0.218	0.005

Fuente: Elaboración propia.

Nota: La tabla presenta únicamente las estimaciones de los coeficientes fijos. Los coeficientes asociados a los efectos aleatorios no se muestran para mantener la claridad y concisión de la tabla.

CONCLUSIONES

Los modelos mixtos implementados han demostrado ser herramientas efectivas para el análisis de datos longitudinales en contextos clínicos, especialmente en pacientes con SIDA y supresión inmune avanzada. El modelo óptimo, que incluye interacciones y términos no lineales, proporciona una visión detallada de los factores que influyen en los recuentos de CD4, resaltando la importancia de considerar tanto las características demográficas como la evolución temporal de la enfermedad.

El tratamiento 4 mostró una eficacia significativa en la mejora de los recuentos de CD4, especialmente en mujeres, mientras que la interacción negativa con el sexo masculino sugiere una menor eficacia en hombres. La tendencia decreciente no lineal observada en los recuentos de CD4 a lo largo del tiempo tiene importantes implicaciones clínicas, sugiriendo la necesidad de ajustar y personalizar las intervenciones terapéuticas para optimizar la



respuesta al tratamiento. Estos hallazgos subrayan la utilidad de los modelos mixtos para mejorar la toma de decisiones clínicas basadas en datos y contribuyen a la optimización de estrategias terapéuticas en pacientes con VIH.

REFERENCIAS

1. Veil R, Poizot-Martin I, Reynes J, Goujard C, Seng R, Delobel P, et al. Virological and immunological impact of integrase inhibitor-based regimens initiated during primary HIV-1 infection. *AIDS*. 2020 Mar 15;34(4):493–500.
2. Peña SS, del Mar Pastor Bravo M, Tomás MÁC, Martínez PA, Guillén CP, Jiménez-Ruiz I. Factors affecting adherence to antiretroviral treatment in women with HIV: A sequential mixed model research design. *Enfermería Global*. 2021;20(2):18–34.
3. Marín-Palma D, Cardona-Arias JA, Hernández JC. Immunological factors related to HIV-1 in colombian patients. *Revista Ciencias de la Salud*. 2019;17(2):245–58.
4. Serrano-Villar S, De Lagarde M, Vázquez-Castellanos J, Vallejo A, Bernadino JI, Madrid N, et al. Effects of immunonutrition in advanced human immunodeficiency virus disease: A randomized placebo-controlled clinical trial (Promaltia study). *Clinical Infectious Diseases*. 2019 Jan 1;68(1):120–30.
5. Fernández-Felix M, Phd AM, Ciber Epidemiología) ;, Pública S, Msc FF, Serrano-Villar S, et al. Effects of first-line antiretroviral therapy on the CD4/CD8 ratio and CD8 cell counts in CoRIS: a prospective multicentre cohort study. 2020. Available from: www.thelancet.com/hivVol.
6. Casadiego Rincon EJ, Sosa Martinez JC. Generalized time-varying coefficient models for the analysis of longitudinal data in health: an application in HIV / AIDS and COVID-19. *SantoTomasUniversityJorunal*. 2021 May;
7. Tang L, Chen F, Ling Q, Li P, Ge L, Cai C, et al. HIV disease progression among heterosexually-infected individuals before the introduction of universal ART in China: A linear mixed-effects model. *Glob Health Med*. 2024;
8. Fernandez D, Ali H, Pals S, Alemnji G, Vasireddy V, Siberry GK, et al. Assessing sex differences in viral load suppression and reported deaths using routinely collected program data from PEPFAR-supported countries in sub-Saharan Africa. *BMC Public Health*. 2023 Dec 1;23(1).
9. 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. 2020. Available from: <http://penta-id.org/>
10. Capasso A, Brown JL, Safonova P, Belyakov N, Rassokhin V, DiClemente RJ. Heavy Alcohol Use is Associated with Lower CD4 Counts among Russian Women Living with HIV: A Multilevel Analysis. *AIDS Behav*. 2021 Nov 1;25(11):3734–42.
11. Demidenko E. *Mixed Models Theory and Applications with R Second Edition*. 2013.
12. Anderegg N, Panayidou K, Abo Y, Alejos B, Althoff KN, Anastos K, et al. Global Trends in CD4 Cell Count at the Start of Antiretroviral Therapy: Collaborative Study of Treatment Programs. *Clinical Infectious Diseases*. 2018 Mar 5;66(6):893–903.