Aplicación del modelo mixto semiparamétrico para dosificar ciclosporina a partir de variables clínicas y genéticas

Trabajo de titulación presentado por:
Ruth Elizabeth Flores Rodríguez

para optar al grado de:
Licenciado Estadístico

y al título profesional de:
Ingeniero Estadístico

Directora:
Dra. María Pilar Sánchez Olavarría

Co-director:
Dr. Germán Ibacache Pulgar

Co-directora:
Dr. Jana Stojanova

Valparaíso, Chile, 2017

Dedico este trabajo a mi esposo Marco y a mis hijos Francisco, Belén y Alejandro con mucho amor.

Índice general

Lista de Figuras **V**

Lista de Tablas **VII**

Lista de Abreviaturas **IX**

Agradecimientos **X**

Resumen **XII**

1. Introducción **1**

1.1. Objetivos 2

1.1.1. Objetivo general 2

1.1.2. Objetivos específicos 3

1.2. Motivación 3

1.2.1. Motivación del punto de vista farmacéutico 3

1.2.2. Motivación del punto de vista estadístico 6

2. Modelos Estadísticos **7**

2.1. Modelo Mixto (MM) 7

2.1.1. Especificación del modelo 9

2.1.2. Estimación del modelo 16

2.1.2.1. Máxima verosimilitud (MV) 16
2.1.2.2. Máxima verosimilitud restringida (MVR) . . 17
2.1.3. Criterio de información 18
2.1.4. Diagnóstico del modelo 19
2.1.4.1. Diagnóstico residual 19
2.1.4.2. Residuo condicional 19
2.1.4.3. Diagnóstico para efectos aleatorios 20
2.1.4.4. Coeficientes de correlación intraclase (CCI) . 20
2.2. Modelo Mixto Semiparamétrico (MMS) 21
2.2.1. Especificación del modelo 21
2.2.2. Estimación de los Parámetros 22
2.2.3. Estimación del modelo usando un método mixto 24

3. Metodología **26**
3.1. Identificación de la base datos 26
3.2. Exploración de los datos 27
3.3. Análisis estadístico 28

4. Resultados **32**
4.1. Base de datos 32
4.2. Análisis exploratorio 33
4.2.1. Análisis de las variables clínicas 34
4.2.1.1. Curvas de los datos de pacientes 35
4.2.2. Análisis de las variables genéticas 37
4.3. Análisis de los Modelos 40
4.3.1. Aplicación Modelo Mixto (MM) 41
4.3.1.1. Modelo vacío 41
4.3.1.2. Análisis univariante del Modelo Mixto (MM) 42

4.3.1.3. Análisis del Modelo Mixto (Modelo 1) 43
4.3.1.4. Interacciones 44
4.3.1.5. Análisis del Modelo Mixto (Modelo Final) . . 48
4.3.1.6. Diagnóstico Modelo Mixto (MM) 50
4.3.2. Aplicación del Modelo Mixto Semipamétrico (MMS) . 52
4.3.2.1. Diagnóstico Modelo Mixto Semiparamétrico (MMS) 53
4.3.2.2. Diagnóstico parte paramétrica. 53
4.3.2.3. Diagnóstico parte semiparamétrica 57
4.3.3. Comparación de criterio de información (AIC) 59
4.3.3.1. Curvas ajustadas con spline 60

5. Conclusiones **63**

Bibliografía **65**

Índice de Figuras

1.1. Descripción de un polimorfismo 5

4.1. Gráficos de spaghetti de la concentración de ciclosporina en la sangre de los pacientes trasplantados en el Hospital San Juan de Dios de Santiago, 2015 35
4.2. Gráfico de dispersión individual de la concentración de ciclosporina en la sangre para ($C0$) de los pacientes trasplantados en el Hospital San Juan de Dios de Santiago, 2015 36
4.3. Gráfico de dispersión individual de la concentración de ciclosporina en la sangre para ($C2$) de los pacientes trasplantados en el Hospital San Juan de Dios de Santiago, 2015 37
4.4. Gráfico de interacción entre los polimorfismos en C0 46
4.5. Gráfico de interacción entre los polimorfismos en C2 48
4.6. Gráfico de residuos v/s valores ajustados del Modelo final de C0 50
4.7. Gráfico Q-Q plot de los residuos estandarizados del Modelo final de C0 50
4.8. Gráfico Q-Q normal de los residuos aleatorios del Modelo final usando el Modelo mixto 51
4.9. Gráficos de residuos v/s valores ajustados del Modelo mixto semiparamétrico 54

4.10. Gráficos Q-Q plot de los residuos del Modelo mixto semiparamétrico 55
4.11. Gráficos de caja de los datos atípicos del Modelo mixto semiparamétrico 55
4.12. Gráficos Q-Q plot de los residuos sin datos atípicos del Modelo mixto semiparamétrico 56
4.13. Gráficos de residuos ajustados del Modelo mixto semiparamétrico 57
4.14. Gráficos Q-Q plot de los splines del Modelo mixto semiparamétrico 57
4.15. Gráficos de caja de los datos atípicos de los splines del Modelo mixto semiparamétrico 58
4.16. Gráficos Q-Q plot de los splines sin datos atípicos del Modelo mixto semiparamétrico 59
4.17. Gráfico de dispersión individual con el ajuste Spline del Modelo mixto semiparamétrico, aplicado a la concentración de ciclosporina en la sangre ($C0$) de los pacientes trasplantados en el Hospital San Juan de Dios de Santiago, 2015 61
4.18. Gráfico de dispersión individual con el ajuste Spline del Modelo mixto semiparamétrico, aplicado a la concentración de ciclosporina en la sangre ($C2$) de los pacientes trasplantados en el Hospital San Juan de Dios de Santiago, 2015 62

Índice de Tablas

2.1. Ejemplo de estructura multinivel en dos niveles 9

3.1. Variables de la base de datos. 27

4.1. Estructura de la base de datos. 33
4.2. Distribución de las variables clínicas (N=65). 34
4.3. Distribución según condición de género de la concentración de ciclosporina en la sangre 34
4.4. Distribución de frecuencia de los polimorfismos 38
4.5. Análisis descriptivo de los polimorfismos 39
4.6. Modelo vacío del Modelo mixto de la concentración de ciclosporina en la sangre para C0 y C2 aplicado en los pacientes trasplantados en el Hospital San Juan de Dios, 2015 41
4.7. Análisis univariante del Modelo mixto de la concentración de ciclosporina en la sangre para C0 y C2 en los pacientes trasplantados en el Hospital San Juan de Dios, 2015 42
4.8. Modelo mixto con las variables explicatorias seleccionadas del análisis univariante, aplicado en C0 y C2 en los pacientes trasplantados (Modelo 1). 43
4.9. Interacción de los polimorfismos en C0 45
4.10. Interacción entre los polimorfismos en C2 47

4.11. Modelo mixto con interacción, aplicado en C0 y C2 en los pacientes trasplantados en el Hospital San Juan de Dios, 2015 (Modelo final) 49
4.12. Modelo mixto semiparamétrico con splines, aplicado en C0 y C2 en los pacientes trasplantados en el Hospital San Juan de Dios, 2015 52
4.13. Comparación de Modelos 60

Lista de Abreviaturas

MM	Modelo mixto.
MMS	Modelo mixto semiparamétrico.
SNP	Polimorfismo nucleótido único.
MV	Máxima verosimilitud.
MVR	Máxima verosimilitud restringida.
MVP	Máxima verosimilitud penalizada.
AIC	Criterio de información Akaike.
CCI	Coeficiente de correlación intraclase.
$D.E$	Desviación estándar o Error estándar.

Agradecimientos

En este trabajo de titulación quiero agradecer primeramente a Dios por darme la oportunidad de estudiar y lograr terminar esta carrera, solo mi Dios conoce todo el esfuerzo para llegar hasta aquí, para él sea este triunfo, porque sin él no somos nada, todo viene de parte de Dios.

A mi esposo Marco por su paciencia, comprensión y por todo el apoyo que cada día me entrego incondicional durante todos los años de estudio, también a mis hijos por cada una de sus palabras que confortaban mi alma, haciéndome sentir que valía la pena todo este sacrificio.

A mis padres Elizabeth y Arturo por todo su apoyo y preocupación durante todo este proceso, también a mis hermanos Pablo y David por estar siempre conmigo y a todos mis familiares que siempre estuvieron pendiente de mí, especialmente a mi tía Cecilia.

A mis amigas y compañeras de universidad, con las cuales compartimos largas jornadas de estudio y trabajo, días de alegrías y anécdotas, días de entretenimiento y distracción que quedaran en nuestras vidas. Estephany Holmstrons y Joselin Díaz.

A mi querido amigo David Muñoz que también fue mi compañero de universidad, quien alegró esos momentos de universidad con todas sus locuras, quien siempre estuvo entregándome ánimo para lograr terminar esta carrera, sobre todo en los momentos más difíciles de mi vida.

A la estimada profesora María Pilar Sánchez quien fue la que propuso este proyecto de investigación, gracias por todo su apoyo y paciencia durante todo este proceso de titulación, también agradecer al profesor Germán Ibacache quien me ayudo con sus consejos y conocimiento para avanzar en este trabajo y a la Dra. Jana Stojanova quien proporciono la base datos para realizar este proyecto. Además agradecer a cada uno de los profesores del Instituto de Estadística por enseñarme e instruirme a lo largo de todos estos años.

¡Muchas gracias a todos!

Resumen

El presente estudio tiene como objetivo aplicar el modelo mixto semiparamétrico en datos longitudinales a una muestra de pacientes a los que se les administró dosis de ciclosporina y a los cuales se les consigno las variables clínicas y genéticas (polimorfismos) asociados a este fármaco.

La importancia de estudiar la ciclosporina con las distintas variables se debe a que este medicamento es uno de los principales fármacos utilizados en los procesos de trasplantes cuya función fundamental es prevenir el rechazo del órgano trasplantado y en donde la variabilidad genética cobra importancia dado que puede producir alteraciones en el metabolismo del fármaco. Esta variabilidad se puede investigar, incluyendo en la formulación del modelo la información de las variables genéticas asociadas al polimorfismo.

Para este estudio se usaran los datos de sesenta y cinco pacientes trasplantados en el Hospital San Juan de Dios ubicado en Santiago, compuesta de treinta y seis pacientes hombres (55,38 %) y veintinueve pacientes mujeres (44,62 %), con un rango de edad entre 25 a 74 años, a los cuales durante la hospitalización se les midió la concentración en la sangre de la dosis administrada de ciclosporina en distintos tiempos (días).

Capítulo 1

Introducción

Los modelos estadísticos son una representación de sucesos expresados en una ecuación matemática, por lo general representan una relación lineal entre una variable de respuesta y, con las variables explicativas x.

Muchos estudios de diferentes áreas tales como biológicas, biomédicas y clínicas utilizan los modelos estadísticos para realizar el análisis de los datos a través de modelos de regresión. Uno de los modelos más utilizados es el modelo mixto (MM), denominado modelo multinivel o jerárquico ("multilevel models" o "hierarchical models") (Seoane, 2014), este tiene una estructura particular de los datos en donde se incorporan distintos niveles (clases) que pueden estar ordenados o no jerárquicamente, también pueden tener datos de medidas repetidas a través del tiempo llamados datos longitudinales, siendo actualmente el MM, una herramienta potente para modelar la relación entre una variable respuesta y las variables explicativas en estudios longitudinales (Wu y Zhang, 2006). Otro de los modelos utilizados es el modelo mixto semiparamétrico (MMS), que es una extensión del MM, donde ambos modelos estiman los valores esperados de las observaciones con una parte de efectos fijos como las variancias y otra parte de efectos aleatorios como las covarian-

cias, permitiendo así el análisis de datos correlacionados, incompletos y con intervalos entre observaciones no constantes. El MMS permite modelar los datos de forma mas integra que el MM, ya que éste puede utilizar técnicas de suavizado con los splines para realizar el ajuste de los datos.

En este trabajo se examinaran los métodos existentes de los MM y MMS para combinar ambas técnicas para realizar el análisis de datos longitudinales, donde la implementación del MMS permitirá en este trabajo lograr estimar una dosis personalizada de ciclosporina en pacientes trasplantados a fin de minimizar los riesgos de toxicidad y su alta variabilidad, considerando la complejidad de los datos como las variables clínicas y genéticas, que en este caso en particular representa un modelo innovador para el cálculo de regímenes de dosificación y un ejemplo en la aplicación de este modelo.

1.1. Objetivos

Para el desarrollo de este trabajo se plantea un objetivo general y tres objetivos específicos, con el fin de trabajar con datos de ciclosporina en pacientes trasplantados.

1.1.1. Objetivo general

- Aplicar un modelo óptimo en datos longitudinales que sea capaz de dosificar ciclosporina en los pacientes trasplantados y que incluya las variables clínicas y genéticas.

1.1.2. Objetivos específicos

- Investigar y revisar el estado del arte del modelo mixto y el modelo mixto semiparamétrico.
- Analizar la base de datos de los pacientes trasplantados.
- Calcular y ajustar el modelo mixto y el modelo mixto semiparamétrico.

1.2. Motivación

A continuación en las subsecciones (1.2.1) y (1.2.2), se abordarán dos enfoques relevantes que motivan el desarrollo de este trabajo, uno de ellos es desde el punto de vista farmacéutico y el otro desde el punto de vista estadístico, ambos ámbitos permitirán la modelación de los datos bajo el MM y el MMS.

1.2.1. Motivación del punto de vista farmacéutico

La ciclosporina es una droga derivada del hongo Tolypocladium inflatum, descubierta en 1970 por el Dr. F. Hartmann Stähelin en Suiza, donde en 1983 fue aprobada para la utilización en el desarrollo de trasplantes de personas. Por lo tanto la incorporación de este fármaco permitió el incremento del número de trasplantes y el éxito de esta práctica (de Rada, 2008), provocando así un cambio importante en la farmacoterapia inmunosupresora, más del 80 % de los protocolos de administración a pacientes con trasplante renal, están basados en este fármaco (Oltra et al., 2002). También se indica en otras patologías asociadas a su condición, como por ejemplo, en la Artritis reumatoide, Esclerodermia, Vasculitis, Lupus eritematoso, entre otras.

Actualmente es uno de los fármacos inmunosupresores usado en el trasplante de órganos con el objetivo de reducir la actividad del sistema inmunitario del paciente y el riesgo de rechazo del órgano trasplantado.

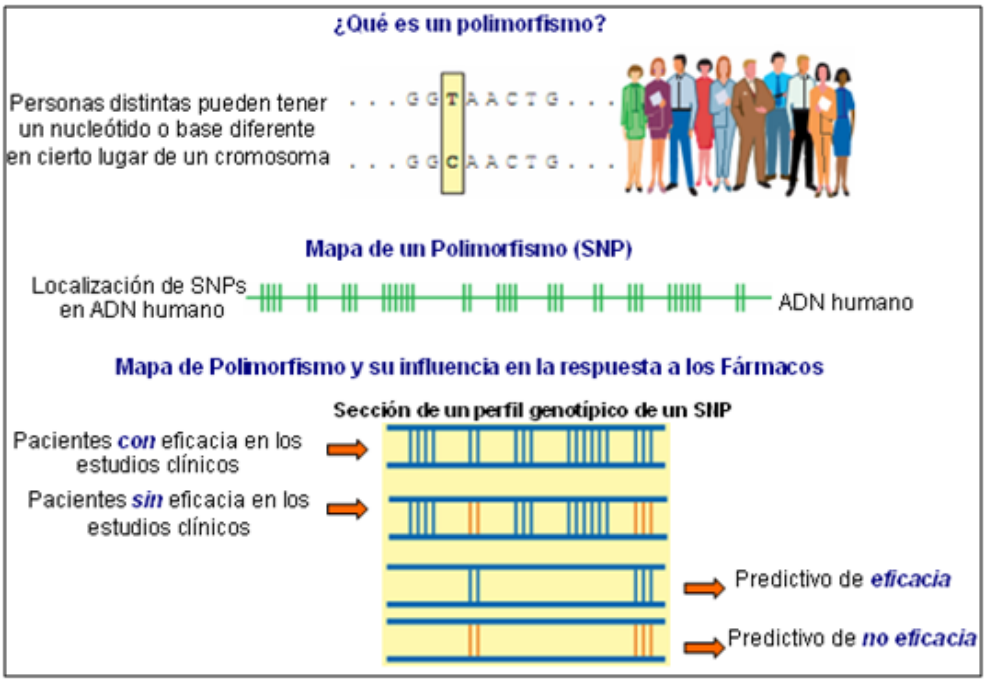
Existen varios motivos del punto farmacéutico para estudiar la ciclosporina, algunos de ellos son: primero debido a que tiene una variabilidad grande en los procesos farmacocinéticos y tiene un margen terapéutico estrecho que hace que se deba monitorizar y realizar un ajuste individualizado, es decir, caso a caso, de la dosis que se administra; segundo ya que su uso debe ser limitado debido a su nefrotoxicidad (Shu et al., 2015) dado que la afectación renal por tóxicos del medicamento presenta una inestabilidad del 5 % al 89 % y tercero porque es un sustrato de los polimorfismos donde en varios estudios se reportan notorias diferencias como por ejemplo que pacientes homocigotos para el genotipo CYP3A5*3 (están expuestas a altas concentraciones de ciclosporina), mientras que en otros estudios se reporta una correlación negativa de ciclosporina, existiendo diferencias en los resultados. En otros estudios realizado en la etnia, se ha reportado que el genotipo CYP3A5*3 está asociado a una alta concentración de ciclosporina en los chinos Han del norte de China y en otros estudios señalan que el genotipo CYP3A4*1G en el intron 10 demuestra que tiene una concentración más baja de ciclosporina que el genotipo CYP3A4*1/*1 en los chinos Han del este de China, ello nos indica que también los polimorfismos pueden afectar la administración de la dosis de ciclosporina en los pacientes de diferente etnia.

Un polimorfismo es un locus genético que está presente en dos o más alelos distintos, de forma que el alelo más raro tiene una frecuencia mayor o igual a 1 % (0,01) en la población general. El polimorfismo de nucleótido único (SNP) es la secuencia de ADN que consiste en la variación de una sola base. Son variaciones comunes de una sola base del ADN humano con una

frecuencia aproximadamente (Caratachea, 2007), donde distintas personas pueden tener secuencias diferentes o puede tratarse de la sustitución de unas de sus nucleótidos, por ejemplo, donde en un paciente hay una Citosina en otro paciente hay una Timina, como se muestra en la Figura (1.1), también puede ser más complicado y existir secuencia repetidas, en el caso de un paciente puede tener 20 copias y en otro paciente 18 copias. Siendo asociado a cambios sustanciales tanto en el metabolismo como en el efecto del fármaco.

Muchos de los fármacos, no son efectivos para un grupo de pacientes dentro de un mismo tratamiento. Es por esto que actualmente en algunos tratamientos, se están considerando los polimorfismos para predecir la respuesta clínica.

Figura 1.1. Descripción de un polimorfismo (Modificado de (Roses, 2000)).



Esta propiedad demuestra la importancia de estudiar este fármaco en este trabajo de investigación.

1.2.2. Motivación del punto de vista estadístico

Dentro de la literatura estadística se han planteado distintos modelos estadísticos, uno de ellos es el modelo lineal clásico utilizado frecuentemente para predecir. Otros modelos ya más avanzados son los MM y los MMS, ambos modelos utilizados también para predecir con datos más complejos. Algunas de las razones para utilizar los MM y MMS es que permiten analizar datos longitudinales o en clúster, además dan la posibilidad de trabajar con parte del modelo como efectos fijos y otra parte del modelo como efectos aleatorios y finalmente el hecho que facilitan el análisis de datos desbalanceados o desequilibrados.

Estos son algunos de los motivos más importantes para trabajar estos modelos estadísticos más avanzados, permitiendo en conjunto analizar las variables clínicas y las variables genéticas de los datos de concentración de ciclosporina en pacientes trasplantados chilenos.

Capítulo 2

Modelos Estadísticos

En este capítulo, nos introduce al estado del arte de los dos modelos estadísticos el MM y el MMS

2.1. Modelo Mixto (MM)

El MM o modelo multinivel permite trabajar con datos longitudinales o en clúster, donde ambas formas registran la relación entre una variable dependiente y diversas variables explicatorias, incluyendo una parte de efectos fijos asociado con uno o más covariables y otra parte de efectos aleatorios asociado con uno o más factores aleatorios.

Los parámetros de efectos fijos describen la relación de las covariables con la variable dependiente para toda una población, en cambio los parámetros de efectos aleatorios son específicos para cada grupos o individuos dentro de una población (West et al., 2014).

En la literatura científica podemos encontrar estudios de los modelos paramétricos de efectos mixtos (Laird y Ware 1982, Jones 1993, Vonesh y Chinchilli 1996, Verbeke y Molenberghs 2000, Diggle, Heagerty, Liang y Zeger

2002, y Demidenko 2004, entre otros). Estos desarrollos se hacen en general sobre la suposición de que los residuos sean distribuídos normalmente con varianza constante.

El **clúster** se define como datos agrupados cuya finalidad es dividir un conjunto de objetos en grupos, es decir, es cuando se mide la variable dependiente una vez para cada sujeto (la unidad de análisis) y se agrupa o se anida dentro de racimos de unidades, clasificando a un conjunto de individuos en grupos homogéneos que a menudo son independientes (West et al., 2014).

Los datos **longitudinales** se define como medidas repetidas en el tiempo, es decir, cuando se mide la variable dependiente (respuesta) varias veces para cada sujeto a través del tiempo, que generalmente están correlacionadas en un sujeto e independientes entre sujetos (West et al., 2014).

Los datos **Multinivel o jerárquicos** corresponden a los distintos niveles dentro de un modelo donde se puede trabajar en conjunto con el clúster y con datos longitudinales en dos o más niveles. Consideramos datos de dos niveles como se muestra en la tabla (2.1) (West et al., 2014):

Nivel 1 denota las observaciones detalladas de la base de datos donde se encuentra las covariables o variables exploratorias de efecto fijo y la variable dependiente que siempre se mide en este nivel.

Nivel 2 representa el siguiente nivel de la jerarquía y generalmente se refiere a grupos de unidades en conjuntos de datos longitudinales agrupados (grupos de clusters) y las covariables o variables exploratorias de efecto aleatorio.

Tabla 2.1. Ejemplo de estructura multinivel en dos niveles (Modificado de (West et al., 2014)).

Nivel 2		**Nivel 1**	
clúster	v. explicatorias(aleatorio)	v. dependiente	v. exploratorias(fijo)
1	tiempo(1)	C_0	20
⋮	⋮	⋮	⋮
n	tiempo(n)	C_n	n

2.1.1. Especificación del modelo

El MM propuesto por primera vez por Harville (1976) y Laird (1982), esta compuesto por la siguiente estructura general de modelo (West et al., 2014):

$$y_{ij} = X_{ij}^T \beta + Z_{ij}^T b_i + \epsilon_{ij}, \tag{2.1}$$

$$b_i \sim N(0, D) \quad \epsilon_i \sim N(0, R_i),$$

$$j = 1, 2, ..., n_i \quad i = 1, 2, ..., n,$$

donde y_{ij} denota la respuesta del la j-ésima observación del i-ésimo sujeto, X_{ij} y Z_{ij} son los vectores asociados a las covariables de efectos fijos y efectos aleatorios, los parámetros $\beta : (p \times 1)$ y $b_i : (q \times 1)$ generalmente se llaman los vectores de efectos fijos y los vectores de efectos aleatorios respectivamente y ϵ_{ij} corresponde al error de medición.

Recordemos que por definición, los efectos aleatorios son variables aleatorias. Suponemos que los q efectos aleatorios en el vector b_i siguen una distribución normal multivariada, con el vector de medias 0 y una matriz de varianza-covarianza indicada por D:

$$b_i \sim N(0, D) \tag{2.2}$$

Los elementos a lo largo de la diagonal principal de la matriz D representan las varianzas de cada efecto aleatorio en b_i y los elementos fuera de la diagonal representan las covarianzas entre dos efectos aleatorios. Debido a que hay q efectos aleatorios en el modelo asociado al i-ésimo sujeto, D es una matriz $q \times q$ que es simétrica y definida positiva. Los elementos de esta matriz se muestran como sigue:

$$D = Var(b_i) = \begin{pmatrix} Var(b_{1i}) & cov(b_{1i}; b_{2i}) & \cdots & cov(b_{1i}; b_{qi}) \\ cov(b_{1i}; b_{2i}) & Var(b_{2i}) & \cdots & cov(b_{2i}; b_{qi}) \\ \vdots & \vdots & \ddots & \vdots \\ cov(b_{1i}; b_{qi}) & cov(b_{2i}; b_{qi}) & \cdots & Var(b_{qi}) \end{pmatrix}$$

Los elementos (varianzas y covarianzas) de la matriz D se definen como funciones de un conjunto (normalmente) pequeño de parámetros de covarianza almacenados en un vector denotado θ_D, especificado en (2.4).

En contraste con el modelo lineal estándar, los residuos asociados con observaciones repetidas sobre el mismo sujeto en un modelo mixto pueden ser correlacionados. Se debe suponer que los residuos n_i en el vector ϵ_i para un sujeto dado, i, son variables aleatorias que siguen una distribución normal multivariada con un vector de medias 0 y una varianza-covarianza simétrica positiva definida Matriz R_i:

$$\epsilon_i \sim N(0, R_i) \tag{2.3}$$

También se asume que los residuos asociados con diferentes sujetos son

independientes entre sí. Además, se supone que los vectores de residuos, $\epsilon_1,\ldots,\epsilon_m$ y los efectos aleatorios, $b_1,\ldots,b_m$, son independientes entre sí. La forma general de la matriz R_i se puede representar de la siguiente forma:

$$R_i = Var(\epsilon_i) = \begin{pmatrix} Var(\epsilon_{1i}) & cov(\epsilon_{1i};\epsilon_{2i}) & \cdots & cov(\epsilon_{1i};\epsilon_{n_i i}) \\ cov(\epsilon_{1i};\epsilon_{2i}) & Var(\epsilon_{2i}) & \cdots & cov(\epsilon_{2i};\epsilon_{n_i i}) \\ \vdots & \vdots & \ddots & \vdots \\ cov(\epsilon_{1i};\epsilon_{n_i i}) & cov(\epsilon_{2i};\epsilon_{n_i i}) & \cdots & Var(\epsilon_{n_i i}) \end{pmatrix}$$

Los elementos (varianzas y covarianzas) de la matriz R_i se definen como funciones de un conjunto (normalmente) pequeño de parámetros de covarianza almacenados en un vector denotado θ_R, especificado en (2.5).

Estructuras de covarianza para la matriz D

Una matriz D sin restricciones adicionales sobre los valores de sus elementos (aparte de ser definida simetría y positiva) se denomina una estructura no estructurada (o general) D matriz. La simetría en la matriz $(q \times q)$ implica que el vector θ_D tiene $q \times (q+1)/2$ parámetros. La siguiente matriz es un ejemplo de una matriz D no estructurada, en el caso de un MM que tiene dos efectos aleatorios asociados con el i-ésimo sujeto.

$$D = Var(b_i) = \begin{pmatrix} \sigma^2_{b1} & \sigma_{b1,b2} \\ \sigma_{b1,b2} & \sigma^2_{b2} \end{pmatrix}$$

En este caso, el vector θ_D contiene tres parámetros:

$$\theta_D = \begin{pmatrix} \sigma^2_{b1} \\ \sigma_{b1,b2} \\ \sigma^2_{b2} \end{pmatrix} \tag{2.4}$$

También definimos otras estructuras más cuidadosamente para D imponiendo ciertas restricciones a la estructura de D. Una estructura muy comúnmente utilizada es la estructura de las componentes de la varianza, en la que cada efecto aleatorio en b_i tiene su propia varianza y todas las covarianzas en D se definen como cero. En general, el vector θ_D para la estructura de componentes de varianza requiere q parámetros de covarianza, definiendo las varianzas en la diagonal de la matriz D. Por ejemplo, en un modelo mixto que tiene dos efectos aleatorios asociados con el i-ésimo sujeto, una matriz de D de la componente de varianza tiene la siguiente forma:

$$D = Var(b_i) = \begin{pmatrix} \sigma_{b1}^2 & 0 \\ 0 & \sigma_{b2}^2 \end{pmatrix}$$

En este caso, el vector θ_D contiene dos parámetros:

$$\theta_D = \begin{pmatrix} \sigma_{b1}^2 \\ \sigma_{b2}^2 \end{pmatrix} \tag{2.5}$$

La matriz D no estructurada y estructuras de componentes de varianza para la matriz son las más utilizadas en la práctica entre otras. Por ejemplo, se podría permitir que los parámetros que representan las varianzas y covarianzas de los efectos aleatorios en el vector θ_D varían entre los diferentes subgrupos de casos (por ejemplo, varones y mujeres en un estudio longitudinal), si hubo una mayor varianza entre sujetos en los efectos seleccionados que se espera en un subgrupo comparado con otro (por ejemplo, los varones tienen más variabilidad en sus intercepciones).

Estructuras de covarianza para la matriz R

La matriz de covarianza más simple para R_i es la estructura diagonal, en la cual se supone que los residuos asociados con las observaciones sobre un mismo sujeto no están correlacionados y tienen una varianza constante. La matriz diagonal R_i para cada sujeto i tiene la siguiente estructura:

$$R_i = Var(\varepsilon_i) = \sigma^2 I_{n1} \begin{pmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma^2 \end{pmatrix}$$

La estructura diagonal requiere un parámetro en θ_R, que define la varianza constante en cada punto del tiempo.

$$\theta_R = \begin{pmatrix} \sigma^2 \end{pmatrix}$$

También la estructura de simetría compuesta se utiliza frecuentemente para la matriz R_i. La forma general de esta estructura para cada sujeto i es la siguiente:

$$R_i = Var(\varepsilon_i) = \sigma^2 I_{n1} \begin{pmatrix} \sigma^2 + \sigma_1 & \sigma_1 & \cdots & \sigma_1 \\ \sigma_1 & \sigma^2 + \sigma_1 & \cdots & \sigma_1 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_1 & \sigma_1 & \cdots & \sigma^2 + \sigma_1 \end{pmatrix}$$

En la estructura de covarianza de simetría compuesta, hay dos parámetros en el θ_R vector que definen las varianzas y covarianzas en la matriz R_i:

$$\theta_R = \begin{pmatrix} \sigma^2 \\ \sigma_1 \end{pmatrix}$$

Obsérvar que se debe suponer que los residuos n_i asociados a los valores de respuesta observados para el i-ésimo sujeto tienen una covarianza constante, σ_1 y una varianza constante, $\sigma^2 + \sigma_1$, en la estructura de simetría compuesta. Esta estructura se utiliza a menudo cuando una admisión de correlación de residuos es igual aceptable (por ejemplo, ensayos repetidos bajo la misma condición en un experimento).

Modelo marginal del Modelo mixto (MM)

La matriz de especificación general de un modelo marginal para el sujeto i esta representado por la ecuación:

$$Y_i = X_i\beta + \epsilon_i^*, \tag{2.6}$$

donde

$$\epsilon_i^* \sim N(0, V_i^*)$$

En (2.4), la matriz de diseño X_i de $(n_i \times p)$ se construye de la misma forma que en un MM, así también de forma similar, β es un vector de efectos fijos, el vector ϵ_i^* representa un vector de errores marginales residuales. Los elementos de la matriz de varianza y covarianza marginal V_i^* de $(n_i \times n_i)$ se definen generalmente mediante un pequeño conjunto de parámetros de covarianza denominamos θ^*. Todas las estructuras utilizadas para la matriz

R_i descritas en la subsección 2.3.2 se pueden utilizar para especificar una estructura para V_i^*. También se permiten otras estructuras para V_i^*, como las mostradas en la sección 2.4. Obsérvese que toda la parte aleatoria del modelo marginal se describe solamente en términos de los residuos marginales ϵ_i^*. En contraste con el MM, el modelo marginal no implica los efectos aleatorios, b_i, por lo que no se pueden hacer inferencias sobre ellos y consecuentemente este modelo no es un MM.

Modelo marginal implícito del Modelo mixto (MM)

El MM introducido en (2.1) implica el siguiente modelo lineal marginal representado por la siguiente ecuación:

$$Y_i = X_i\beta + \epsilon_i^*, \tag{2.7}$$

donde

$$\epsilon_i^* \sim N(0, V_i)$$

y la matriz de varianza-covarianza, V_i se define como

$$V_i = Z_i D Z_i^{'} + R_i$$

El modelo marginal implícito define la distribución marginal del vector Y_i:

$$Y_i \sim N(X_i\beta; Z_i D Z_i^{'} + R_i) \tag{2.8}$$

La media marginal (o valor esperado) y la matriz de varianza marginal-covarianza de el vector Y_i es igual a

$$E(Y_i) = X_i\beta \tag{2.9}$$

y

$$Var(Y_i) = V_i = Z_i D Z_i' + R_i$$

Los elementos fuera de la diagonal en la matriz V_i $(n_i \times n_i)$ representan las covariancias marginales del vector Y_i. Estas covarianzas son en general diferentes de cero, lo que significa que en el caso de un conjunto de datos longitudinales, las observaciones repetidas sobre un individuo dado i están correlacionadas.

La distribución marginal especificada en (2.6), con media y varianza definidas en (2.7) es un punto relacionado con la estimación de verosimilitud en los MM planteado en la siguiente sección.

2.1.2. Estimación del modelo

En el MM, se estiman los parámetros de efecto fijo, β y los parámetros de covarianza, θ, es decir, θ_D y θ_R para las matrices D y R_i, respectivamente. En esta sección, se examinan las estimaciones máxima verosimilitud (MV) y máxima verosimilitud restringida (MVR), que son los métodos comúnmente utilizados para estimar estos parámetros.

2.1.2.1. Máxima verosimilitud (MV)

En general, la estimación de MV es un método para obtener estimaciones de los parámetros desconocidos, ya que son los valores de los parámetros que hacen que los valores observados de la variable dependiente sean más probables, dados los supuestos distributivos (West et al., 2014). En el MM la función de verosimilitud de β y θ hacen referencia a la distribución marginal de la variable dependiente Y_i definida en (2.6). La función de densidad de

probabilidad normal multivariante correspondiente es:

$$f(Y_i|\beta,\theta) = (2\pi)^{\frac{-n_i}{2}} det(V_i)^{\frac{-1}{2}} exp(-0{,}5 \times (Y_i - X_i\beta)' V_i^{-1}(Y_i - X_i\beta)) \quad (2.10)$$

Con base en la función de densidad de probabilidad en (2.8), y dada la información observada $Y_i = y_i$, la contribución de la función de verosimilitud para el i-ésimo sujeto se define como sigue:

$$L_i(\beta,\theta;y_i) = (2\pi)^{\frac{-n_i}{2}} det(V_i)^{\frac{-1}{2}} exp(-0{,}5 \times (y_i - X_i\beta)' V_i^{-1}(y_i - X_i\beta)) \quad (2.11)$$

La función de verosimilitud $L(\beta,\theta)$ definida en (2.9) para m individuos es:

$$L(\beta,\theta) = \Pi_i L_i(\beta,\theta) = \Pi_i (2\pi)^{\frac{-n_i}{2}} det(V_i)^{\frac{-1}{2}} exp(-0{,}5 \times (y_i - X_i\beta)' V_i^{-1}(y_i - X_i\beta)) \quad (2.12)$$

Usando (2.10), la función log-verosimilitud de (β,θ), se define como:

$$\begin{aligned} \ell(\beta,\theta) = lnL(\beta,\theta) = & -0{,}5n \times ln(2\pi) - 0{,}5 \times \sum_i ln(det(V_i)) \\ & - 0{,}5 \times \sum_i (y_i - X_i\beta)' V_i^{-1}(y_i - X_i\beta) \end{aligned} \quad (2.13)$$

2.1.2.2. Máxima verosimilitud restringida (MVR)

La estimación introducida en la década de los 70 por Patterson y Thompson como un método para estimar los componentes de la varianza en el contexto de los diseños de bloques incompletos o desequilibrados. La estimación MVR es una forma alternativa de estimar los parámetros de covarianza en θ preferentemente utilizada (West et al., 2014), ya que produce estimaciones imparciales de los parámetros de covarianza teniendo en cuenta la

pérdida de grados de libertad que resulta de la estimación de los efectos fijos en β.

Las estimaciones de θ se basan en la optimización de la siguiente función de Log-verosimilitud:

$$\ell_{MVR}(\theta) = -0{,}5 \times (n-p) \times ln(2\pi) - 0{,}5 \times \sum_i ln(det(V_i)) \\ - 0{,}5 \times \sum_i r_i' V_i^{-1} r_i - 0{,}5 \times \sum_i ln(det(X_i' V_i^{-1} X_i)), \tag{2.14}$$

donde

$$r_i = y_i - X_i \left(\left(\sum_i X_i' V_i^{-1} \right)^{-1} \sum_i X_i' V_i^{-1} y_i \right) \tag{2.15}$$

2.1.3. Criterio de información

El criterio de información es una forma de evaluar, comparar y seleccionar un modelo relativo que mejor se ajuste al conjunto de los datos, basado en su valor óptimo log-verosimilitud (West et al., 2014). Entre los criterios de comparación se encuentra el llamado criterio de información de Akaike (AIC).

$$AIC = -2 \times \ell(\hat{\beta}, \hat{\theta}) + 2p \tag{2.16}$$

El AIC penaliza el ajuste de un modelo para el número de parámetros que se estima. En (2.14) p representa el número total de parámetros que se están estimando en el modelo tanto para los efectos fijos como aleatorios.

2.1.4. Diagnóstico del modelo

Es importante llevar a cabo el diagnóstico de modelo para comprobar si las suposiciones distributivas para los residuos se han cumplido y si el ajuste del modelo es sensible a observaciones inusuales (outliers).

2.1.4.1. Diagnóstico residual

El diagnóstico residual se utiliza para decidir si existe o no un patrón específico en los residuos, el ejemplo más simple es decidir si un determinado conjunto de residuos contra los valores predichos presentan un patrón aleatorio o no, representados gráficamente para verificar los supuestos del modelo y para detectar observaciones atípicas y observaciones potencialmente influyentes (West et al., 2014). En general, los residuos deben ser evaluados para determinar la normalidad, la varianza constante y los valores atípicos.

2.1.4.2. Residuo condicional

Un residuo condicional es la diferencia entre el valor observado y el valor predicho condicional de la variable dependiente (West et al., 2014). Se expresa por medio de la ecuación:

$$\hat{\epsilon_i} = y_i - X_i\hat{\beta} - Z_i\hat{b_i} \tag{2.17}$$

En general, los residuos condicionales en su forma básica en (2.16) no son adecuados para verificar los supuestos del modelo y para detectar los valores atípicos.

Para aliviar los problemas con la interpretación de residuos condicionales que pueden tener desviaciones desiguales, consideramos escalar (es decir, dividir) los residuos por sus desviaciones estándar verdaderas o estimadas.

2.1.4.3. Diagnóstico para efectos aleatorios

Para el diagnóstico de los efectos aleatorios se debe considerar los predictores del vector de efectos aleatorios b_i, representado gráficamente por ejemplo, Histogramas, Cuantil-Cuantil (Q-Q plots), para investigar posibles valores atípicos que pueden justificar una investigación exhaustiva (West et al., 2014).

2.1.4.4. Coeficientes de correlación intraclase (CCI)

En general, el coeficiente de correlación intraclase (CCI) es una medida que describe la similitud (o homogeneidad) de las respuestas sobre la variable dependiente dentro de un grupo o una unidad de análisis (Cebolla, 2013).

$$\rho(Y_{ij}, Y_{i,j}) = \left[\frac{\sigma_0^2}{\sigma_0^2 + \sigma_\epsilon^2}\right] \tag{2.18}$$

Sus valores oscilan necesariamente entre 0 y 1, que debe ser interpretado como el grado de homogeneidad. Cuando $\rho = 1$ existe una dependencia máxima de los grupos, cuando $\rho = 0{,}5$ existe una dependencia media de los grupos y cuando $\rho = 0$ existe una mínima dependencia de los grupos.

2.2. Modelo Mixto Semiparamétrico (MMS)

Los MMS han sido discutidos y usados por varios autores para datos longitudinales (zeger y Diggle 1994; Zhang et al., 1998; Jacqmin-Gadda et al., 2002; Durban et al., 2005; entre otros.), donde generalmente utilizan los componentes paramétricos para modelar los efectos lineales y los componentes no paramétricos para modelar los efectos no lineales.

2.2.1. Especificación del modelo

El MMS es una extensión del MM y tiene como característica principal la inclusión de un componente no paramétrico asociado con alguna variable explicativa que no tiene relación lineal con la variable de respuesta (Ibacache-Pulgar et al., 2012). En este trabajo asumiremos que

$$Y_{ij} = X_{ij}^T\beta + Z_{ij}^T b_i + f(t_{ij}) + \epsilon_{ij}, \tag{2.19}$$

$$\epsilon_i \sim N(0, \sigma^2 I),$$

donde Y_{ij} denota la respuesta de la j-ésima observación del i-ésimo sujeto en el tiempo t_{ij}, X_{ij} y Z_{ij} son vectores $(p \times 1)$ y $(q \times 1)$ respectivamente, asociados a las covariables de efectos fijos y efectos aleatorios, β y b_i son vectores $(p \times 1)$ y $(q \times 1)$ asociados a los efectos fijos y los efectos aleatorios, f es una función arbitraria dos veces diferenciable que depende de los valores de la variable explicativa t y ϵ_{ij} corresponde al error aleatorio, para $i = 1, ..., n$ y $j = 1, ..., m_i$.

Matricialmente, el modelo (2.19) puede ser escrito como

$$Y_i = X_i\beta + Z_i b_i + N_i \mathrm{f} + \epsilon_i, \tag{2.20}$$

donde $Y_i = (Y_{i1}, \ldots, Y_{im_i})^T$ es el vector de respuesta; X_i $(m_i \times p)$ es una matriz de diseño de efectos fijos con filas X_{ij}^T; Z_i $(m_i \times q)$ es una matriz de diseño de efectos aleatorios con filas Z_{ij}^T; N_i $(m_i \times r)$ es una matriz de incidencia con elementos dados por la función indicadora $I(t_{ij} = t_l^0)$, para $j = 1, \ldots, m_i$ y $l = 1, \ldots, r$, donde I es una matriz de identidad de dimensión $n = \sum_{i=1}^{m} m_i$; $b = (b_1^T, \ldots, b_m^T)^T$ es normal $(0, D(\phi))$, con $D(\phi) = diag(D, \ldots, D)$; f es un vector $(r \times 1)$ definido por $\mathrm{f} = (f(t_1^0), \ldots, f(t_r^0))^T$, donde t_j^0 son los valores ordenados y distintos de la variable explicativa t y ϵ_i es un vector de errores aleatorios $(m_i \times 1)$. Asumiremos que la distribución de Y e ϵ está dada por

$$\begin{cases} Y & \sim & N_n(X\beta + N\mathrm{f}; \sigma^2 I) \\ \epsilon & \sim & N_n(0; \sigma^2 I). \end{cases}$$

2.2.2. Estimación de los Parámetros

Sea $\theta = (\beta^T, \mathrm{f}^T, \sigma^2)^T$ el vector de parámetros que se debe estimar bajo el modelo MMS y sea $V = diag(V_1, \ldots, V_m)$, con $V_i = Z_i D Z_i^T + \sigma^2 I$. Entonces la función de log-verosimilitud (Zhang et al., 1998) para el modelo (2.20) es la siguiente:

$$\ell(\beta, \mathrm{f}; Y) = -\frac{1}{2} log|V| - \frac{1}{2}(Y - X\beta - N\mathrm{f})^T V^{-1}(Y - X\beta - N\mathrm{f}). \qquad (2.21)$$

Como $f(t)$ es una función no paramétrica, consideramos la siguiente función de log-verosimilitud penalizada (Green y Silverman, 1994):

$$\ell_p(\theta,\lambda) - \frac{\lambda}{2}\int_0^T [f''(t)]^2 dt = \ell_p(\theta,\lambda) - \frac{\lambda}{2} f^T K \mathrm{f}, \tag{2.22}$$

donde $\lambda > 0$ es un parámetro de suavizado que controla el balance entre la bondad del ajuste y el suavizado de la función estimada. Se asume que la función $f(t)$ pertenece a las funciones del espacio sobolev (Bourgain et al., 2000). La matriz K es una matriz de suavizamiento no negativa.

Asumiendo que el parámetro de suavizado λ y las componentes de varianza θ son fijos, el estimador de máxima verosimilitud penalizada (MVP) de β y f satisface la siguiente ecuación matricial:

$$\begin{bmatrix} X^T W X & X^T W N \\ N^T W X & N^T W N + \lambda K \end{bmatrix} \begin{bmatrix} \beta \\ \mathrm{f} \end{bmatrix} = \begin{bmatrix} X^T W Y \\ N^T W Y \end{bmatrix}, \tag{2.23}$$

donde $W = V^{-1} = diag(W_1, ..., W_m)$, con $W_i = V_i^{-1}$. La ecuación (2.23) tiene una solución única si y solo si la matriz $[X, NT]$ es de rango completo, donde $T = [1, t^0]$ y 1 es un vector de unos $(r \times 1)$.

Sea $W_x = W - WN(N^T W N + \lambda K)^{-1} N^T W$ y $W_f = W - WX(X^T W X)^{-1} X^T W$. Los estimadores para β y f están dadas por las siguientes ecuaciones:

$$\widehat{\beta} = (X^T W_x X)^{-1} X^T W_x Y \tag{2.24}$$

y

$$\widehat{\mathrm{f}} = (N^T W_f N + \lambda K)^{-1} N^T W_f Y. \tag{2.25}$$

Observe que las matrices $(X^T W_x X)^{-1}$ y $(N^T W_f N + \lambda K)^{-1}$ son definidas positivas dado que $[X, NT]$ es de rango completo.

La estimación de los efectos aleatorios específicos del sujeto b_i se pueden estimar de la siguiente manera:

$$\widehat{b_i} = DZ_i^T W_i(Y_i - X_i\widehat{\beta} - N_i\widehat{\mathrm{f}}). \tag{2.26}$$

2.2.3. Estimación del modelo usando un método mixto

De acuerdo a Zhang et al., (1998), el modelo mixto semiparamétrico (2.20) se puede escribir como un modelo mixto lineal modificado a través de una transformación lineal uno a uno según Green, (1987),

$$\mathrm{f} = T\delta + Ba, \tag{2.27}$$

donde δ y a son vectores con dimensiones 2 y $r-2$, $B = L(L^TL)^{-1}$ y L es una matriz de rango completo $r \times (r-2)$ que satisface $K = LL^T$ y $L^TT = 0$. Utilizando la igualdad $\mathrm{f}^TK\mathrm{f} = a^Ta$, la log-verosimilitud penalizada de la ecuación (2.22) se puede expresar como sigue:

$$\ell_p(\theta,\lambda) = -\frac{1}{2}log|V| - \frac{1}{2}(Y - X\beta - NT\delta - NBa)^TV^{-1}(Y - X\beta - NT\delta - NBa) - \frac{1}{2\tau}a^Ta, \tag{2.28}$$

donde $\tau = 1/\lambda$. Por lo tanto, cuando $\widehat{f}(t)$ es un spline cúbico, podemos escribir el modelo mixto semiparamétrico (2.20) como un modelo mixto lineal modificado

$$Y = X\beta + Zb + NT\delta + NBa + \epsilon, \tag{2.29}$$

donde $\beta_* = (\beta^T, \delta^T)^T$ son los coeficientes de regresión, $b_* = (a^T, b^T)^T$ son mutuamente independientes de los efectos aleatorios con una distribución $N(0, \tau I)$.

La MVP de $(\widehat{\beta}, \widehat{\mathrm{f}})$ corresponden a combinaciones lineales de los mejores predictores lineales no sesgados $\widehat{\beta}_*$ y $\widehat{a}$ bajo el modelo (2.29), con $\widehat{\mathrm{f}} = T\widehat{\delta} + B\widehat{a}$. Los estimadores del efecto aleatorio b_i también se pueden obtener como los predictores lineales no sesgados bajo el modelo (2.29).

Los errores estándar de $(\widehat{\beta}, \widehat{\mathrm{f}})$ y $\widehat{b_i}$ pueden ser calculados utilizando el enfoque frecuentista o bayesiano. El frecuentista calcula los errores estándar suponiendo que $f(t)$ es una verdadera función suave fija, mientras que los errores estándar bayesianos se calculan asumiendo un spline cúbico para f, cuya log-densidad es $-\lambda \mathrm{f}^T K \mathrm{f}/2$.

De acuerdo a Zhang et al., (1998), se puede estimar simultáneamente el parámetro de suavizamiento y los componentes de varianza utilizando máxima verosimilitud restringida (MVR), bajo la representación del modelo (2.29), tratando $\tau = 1/\lambda$ y θ como componentes de varianza.

Diagnóstico del modelo

Las técnicas de diagnósticos conocidas para los MM, también se puede extender para el caso de un MMS (Eubank, 1985). Además es posible analizar los residuos para la parte paramétrica y posteriormente para la parte no paramétrica, ya que una observación puede ser influyente en la parte paramétrica pero no afectar la parte no paramétrica o viceversa (Kim et al., 2002).

Para este trabajo se realizará el análisis del diagnóstico con las técnicas existentes del modelo mixto, pero separado en dos partes. Una para la parte paramétrica y otra para la parte semiparamétrica.

Capítulo 3

Metodología

En este capítulo se amplia la información acerca de los procedimientos que se utilizaron para la aplicación de los modelos desarrollados en el Capítulo 2.

3.1. Identificación de la base datos

Para este estudio se considero los datos de 65 pacientes trasplantados en el año 2015 en el Hospital San Juan de Dios de Santiago (Chile), que durante la estancia post-operatoria se les administró un medicamento inmunosupresor llamado ciclosporina para evitar el rechazo del órgano trasplantado.

El tratamiento de ciclosporina difiere por el tiempo de hospitalización que permaneció cada uno de los paciente, donde las dosis del medicamento fueron administradas, calculadas y ajustadas de acuerdo a los datos clínicos de cada paciente correspondientes a las variables explicativas del tipo clínico, además en cada uno de los pacientes la concentración de ciclosporina fue medida en la sangre, primero al inicio del tratamiento registrada en la variable C0 (ng/ml) y luego a las 2 horas registrada en la variable C2 (ng/ml), variando estas mediciónes entre 3 a 15 medidas por paciente, es decir, cada paciente tiene

medidas repetidas de concentración de ciclosporina en la sangre para C0 y C2, esto se debe, a la dependencia o necesidad del paciente hacia el medicamento, existiendo un desequilibro en los datos del estudio. También se registró en cada uno de los pacientes cinco polimorfismos correspondientes a un sustrato de la ciclosporina, variables explicativas del tipo genético no usadas para el calculo de dosis de ciclosporina, donde los polimorfismos CYP3A son los que afectan la actividad metabólica del medicamento a través del cuerpo y los polimorfismos MDR1 son una proteína de resistencia que actúa como transportadora del medicamento. A continuación las variables de la base de datos se muestran en la Tabla (3.1).

Tabla 3.1. Variables de la base de datos.

Variables		Descripción	Tipo
Respuesta	C0 (ng/ml)	concentración inicial de ciclosporina en la sangre	
	C2 (ng/ml)	concentración a las 2 horas de ciclosporina en la sangre	
Explicatorias	Sexo (H/M)	género de los pacientes	clínicas
	Peso (Kg)	peso de los pacientes	
	Edad (años)	edad de los pacientes	
	Tiempo (días)	tiempo del tratamiento	
	DosisC0 (mg)	dosis inicial de ciclosporina	
	DosisC2 (mg)	dosis a las 2 horas de ciclosporina	
	CYP3A4*3	polimorfismo asociado al metabolismo del medicamento	genéticas
	CYP3A4*22	polimorfismo asociado al metabolismo del medicamento	
	CYP3A5*B	polimorfismo asociado al metabolismo del medicamento	
	MDR1*3435	polimorfismo asociado al transporte del medicamento	
	MDR1*1236	polimorfismo asociado al transporte del medicamento	

3.2. Exploración de los datos

El primer proceso fue realizar una exploración de los datos de ciclosporina, con el fin de identificar características de los datos de los pacientes, para ello, se realizó un análisis o examen exploratorio de las variables clínicas y genéticas.

En las variables clínicas se obtuvo una evaluación general de los datos a

través de un análisis descriptivo, además se realizo una evaluación general e individual del comportamiento de la concentración de ciclosporina para C0 y C2 a través de unos diagramas de curvas, estos utilizados a modo de referencia para estudiar los modelos MM y MMS.

En las variables genéticas se comenzó con una tabla de frecuencia de los polimorfismos y posteriormente se obtuvo una evaluación general de los polimorfismos a través del análisis descriptivo por las condiciones: sexo, concentración de ciclosporina ($C0$) y ($C2$), edad, peso, dosisC0 y dosisC2.

Para este análisis exploratorio se utilizó el software Stata versión 14.

3.3. Análisis estadístico

Un segundo proceso fue analizar la base de datos de ciclosporina aplicando el MM con la fórmula (2.1), considerando el clúster, discutido anteriormente en el Capítulo 2.

Para modelar el MM se trabajara con cada variable de respuesta ($C0_{ij}$) y ($C2_{ij}$) de forma independiente.

- Aplicar el modelo vacío, no condicional del MM, es un paso preliminar muy útil en el análisis de datos multinivel, permite obtener una estimación de la varianza entre medidas repetidas de cada paciente y la varianza entre paciente. A partir de estas dos varianzas se puede calcular el CCI, esta medida corresponde al grado de información de la variabilidad en cada uno de los niveles. Este modelo se probó considerando las variables de respuestas ($C0_{ij}$) y ($C2_{ij}$) como efecto fijo y considerando el intercepto como efecto aleatorio, posteriormente se calculó el AIC de acuerdo a la fórmula (2.16) y el CCI de acuerdo a la

fórmula (2.18). La ecuación se expresa de la siguiente manera

$$C0_{ij} = \underbrace{\beta_0}_{efecto\ fijo} + \overbrace{b_0 + \varepsilon_{ij}}^{efecto\ aleatorio}$$

$$C2_{ij} = \underbrace{\beta_0}_{efecto\ fijo} + \overbrace{b_0 + \varepsilon_{ij}}^{efecto\ aleatorio}$$

- Aplicar el MM realizando un análisis univariante de las variables de respuesta. Se probó el modelo de forma individual considerando como efecto fijo ($C0_{ij}$) y ($C2_{ij}$) con cada una de las variables explicatorías y se considero como efecto aleatorio el intercepto y el tiempo (días). Ello con el fin de tomar en cuenta la variabilidad del tiempo, donde se selecciono las variables significativas con un valor-$p < 0{,}05$. Un ejemplo de la ecuación de los modelos univariante para cada variable de respuesta se expresa

$$C0_{ij} = \underbrace{\beta_0 + \beta_1 \times Sexo}_{efecto\ fijo} + \overbrace{b_0 + b_{1j} \times Tiempo + \varepsilon_{ij}}^{efecto\ aleatorio}$$

$$C2_{ij} = \underbrace{\beta_0 + \beta_1 \times Edad}_{efecto\ fijo} + \overbrace{b_0 + b_{1j} \times Tiempo + \varepsilon_{ij}}^{efecto\ aleatorio}$$

- Aplicar el MM a las variables seleccionadas del análisis univariante. Se probó el modelo considerando como efecto fijo ($C0_{ij}$) y ($C2_{ij}$) en

conjunto con cada una de las variables explicatorías estadísticamente significativo del análisis univariante y se considero como efecto aleatorio el intercepto y el tiempo (días), posteriormente se calculo el AIC de acuerdo a la fórmula (2.16).

De este proceso se obtuvo el MM (Modelo 1) para cada variable de respuesta.

- Probar si existe interacción entre las variables genéticas para $(C0_{ij})$ y $(C2_{ij})$, con el fin de observar si el efecto de una variable depende del nivel de otra variable (Jann, 2013). Se selecciona las variables significativas con un valor $p < 0{,}05$.

- Aplicar el MM obtenido anteriormente (Modelo 1), en conjunto con las variables genéticas seleccionadas de las interacciones. Se probó el modelo considerando como efecto fijo $(C0_{ij})$ y $(C2_{ij})$ en conjunto con las variables explicatorías y las interacciones y se considero como efecto aleatorio el intercepto y el tiempo (días). Este modelo se probó con cada una de las interacciones y en cada prueba se calculo el AIC de acuerdo a la fórmula (2.16) para seleccionar el mejor modelo.

  De este proceso se obtuvo el MM (Modelo final) para cada variable de respuesta.

- Aplicar el diagnóstico al MM (Modelo final) para $(C0_{ij})$ y $(C2_{ij})$, usando gráficos para observar el comportamiento de los residuos.

Un tercer proceso fue analizar la base de datos de ciclosporina usando el MMS con la fórmula (2.19), considerando los splines.

Para modelar el MMS se trabajara con cada variable de respuesta $(C0_{ij})$

y $(C2_{ij})$ de forma independiente.

- Aplicar splines, utilizando el MM (Modelo final) obtenido anteriormente. Para realizar el ajuste del modelo se utiliza la variable tiempo (días) como splines, con un valor de k_i, que va entre $i = 1, ..., 16$ nodos, para obtener el mejor ajuste del modelo en cada variable de respuesta y se calculo el AIC con la fórmula (2.14) en cada prueba para seleccionar el mejor modelo.

- Aplicar el diagnóstico al MMS, usando gráficos para observar el comportamiento de los residuos.

Para el análisis del MM y el MMS se utilizó el software Stata versión 14.

Capítulo 4

Resultados

A continuación en este capítulo se presentan los resultados obtenidos después de llevar a cabo la metodología planeada, inicialmente se exponen algunas características generales de la base de datos. Luego se muestran los resultados del análisis exploratorio realizado a la concentración de ciclosporina $C0$ y $C2$, y a los polimorfismos. Finalmente se presentan los resultados obtenidos de la aplicación de los modelos MM y MMS.

4.1. Base de datos

Los datos de ciclosporina son un ejemplo de un conjunto de datos agrupados de dos niveles y la concentración de ciclosporina en la sangre (unidad de análisis) anidados dentro de los pacientes.

Nuestro análisis utiliza un modelo de intercepto y tiempo aleatorio de dos niveles (lo que implica que las observaciones sobre un mismo paciente estén correlacionadas) y que el efecto fijo este asociado con el tratamiento.

Una parte de las 488 observaciones del conjunto de datos de ciclosporina se muestra en la Tabla (4.1), cada fila de datos representa los valores

individuales para un paciente trasplantado.

Tabla 4.1. Estructura de la base de datos.

Nivel 2				**Nivel 1**										
clúster	(aleatorio)	v. dependientes		v. explicatorias (fijo)										
Paciente	Tiempo	C0	C2	Sexo	Edad	Peso	Tiempo	DosisC0	DosisC2	CYP1	CYP2	CYP3	MDR1	MDR2
1	1	90	483	0	39	43.5	1	350	350	0	2	0	2	2
1	2	129	953	0	39	43.0	2	350	350	0	2	0	2	2
1	3	144	874	0	39	42.5	3	350	350	0	2	0	2	2
1	4	234	1087	0	39	42.0	4	325	300	0	2	0	2	2
1	5	180	1408	0	39	41.5	5	300	300	0	2	0	2	2
1	6	152	879	0	39	41.0	6	300	300	0	2	0	2	2
1	7	186	1058	0	39	40.0	7	275	250	0	2	0	2	2
1	8	129	1083	0	39	39.0	8	250	250	0	2	0	2	2
1	9	210	1774	0	39	37.5	9	200	200	0	2	0	2	2
1	10	165	1298	0	39	39.5	10	250	200	0	2	0	2	2
2	1	380	400	1	59	87.0	1	550	525	0	2	0	1	1
2	2	400	0	1	59	87.0	2	500	475	0	2	0	1	1
2	3	342	839	1	59	87.0	3	425	400	0	2	0	1	1
2	4	396	818	1	59	87.0	4	350	350	0	2	0	1	1
2	5	281	502	1	59	87.0	5	350	350	0	2	0	1	1
2	6	239	396	1	59	87.0	6	350	350	0	2	0	1	1
2	7	220	1031	1	59	89.2	7	350	350	0	2	0	1	1
2	8	405	1203	1	59	88.5	8	300	300	0	2	0	1	1
2	9	211	823	1	59	90.5	9	250	250	0	2	0	1	1
2	10	150	774	1	59	92.5	10	200	200	0	2	0	1	1
2	11	145	761	1	59	95.2	11	200	175	0	2	0	1	1
3	1	165	1229	1	55	91.0	1	550	550	0	2	0	1	2
3	2	284	1598	1	55	91.0	2	550	500	0	2	0	1	2
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮		
3	9	130	765	1	55	90.0	9	250	250	0	2	0	1	2
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮		
65	4	187	1107	0	41	58.0	4	250	250	0	2	0	1	2
65	5	109	598	0	41	63.5	5	200	200	0	2	0	1	2

CYP1=CYP3A4*3 CYP2=CYP3A5*B CYP3=CYP3A4*22 MDR1=MDR1*3435 MDR2=MDR1*1236

4.2. Análisis exploratorio

A continuación en esta sección se mostrarán los resultados obtenidos del análisis exploratorio realizado a las variables explicatorias de tipo clínicas y de tipo genéticas.

4.2.1. Análisis de las variables clínicas

En la primera Tabla (4.2), podemos observar un resumen general de las variables explicatorías de tipo clínicas.

Tabla 4.2. Distribución de las variables clínicas (N=65).

Variables	Media	D.E	Min.	Max.
Peso (Kg)	67.480	11.731	37.5	95.2
Edad (años)	51.154	12.548	25	74
DosisC0 (mg)	331.864	101.997	150	600
DosisC2 (mg)	328.846	100.470	150	550

Este resumen corresponde a la distribución de las variables clínicas de los 65 pacientes trasplantados, donde se aprecia que el peso general está entre 37.5 a 95.2 kilos, la edad general está entre 25 a 74 años, la dosis inicial dosisC0 está en un rango de 150 a 600 miligramos administrado en los pacientes y la dosis a las dos horas dosisC2 está en un rango de 150 a 550 miligramos administrado en los pacientes.

Posteriormente en la Tabla (4.3), observaremos un análisis descriptivo distribuido por la condición de género para la concentración de ciclosporina en la sangre de C0 y C2.

Tabla 4.3. Distribución según condición de género de la concentración de ciclosporina en la sangre.

	Sexo	Media	D.E	Min.	Max.	n	(%)
C0 (ng/ml)	Mujeres	185.54	84.06	31	812	29	(44.62 %)
	Hombres	211.23	116.59	33	799	36	(55.68 %)
C2 (ng/ml)	Mujeres	996.005	394.491	98	1998	29	(44.62 %)
	Hombres	960.293	345.727	237	1972	36	(55.68 %)

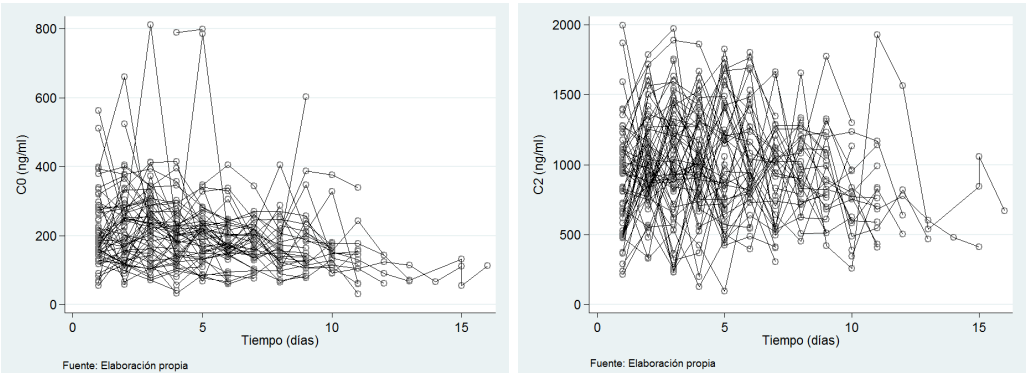
En este análisis descriptivo se aprecia que existen 29 mujeres y 36 hombres en la base de datos de los pacientes, donde la concentración de ciclosporina

en C0 el mayor promedio se da en los pacientes hombres con un valor promedio=211 y con una desviación estándar=117 correspondiente a la diferencia entre cada paciente hombre, en cambio en la concentración de ciclosporina en C2 el mayor promedio se da en las mujeres con un valor promedio=996 y con una desviación estándar=394 correspondiente a la diferencia entre cada paciente mujer.

4.2.1.1. Curvas de los datos de pacientes

A continuación en la Figura (4.1), se presentan las curvas generales de los pacientes trasplantados para C0 y C2.

Figura 4.1. Gráficos de spaghetti de la concentración de ciclosporina en la sangre de los pacientes trasplantados en el Hospital San Juan de Dios de Santiago, 2015.

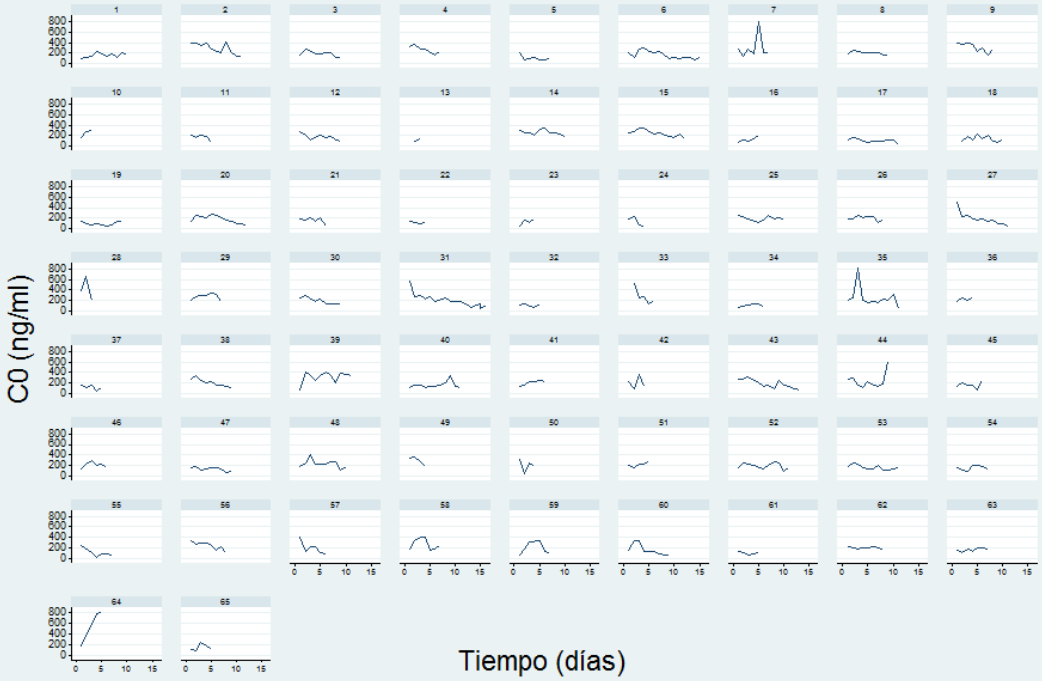


(a) C0 (b) C2

En estos gráficos de spaghetti se puede observar para C0 y C2, que hay una variabilidad a través del tiempo de la concentración de ciclosporina en la sangre en los pacientes trasplantados, en donde se percibe que no hay una tendencia lineal definida. Los puntos en el gráfico nos indican el nivel de ciclosporina en la sangre determinada por cada paciente en los distintos tiempos de medición, conectados con los segmentos de líneas rectas.

Posteriormente observaremos en la Figura (4.2), la tendencia individual de cada paciente trasplantado de la concentración de ciclosporina en la sangre para $C0$.

Figura 4.2. Gráfico de dispersión individual de la concentración de ciclosporina en la sangre para ($C0$) de los pacientes trasplantados en el Hospital San Juan de Dios de Santiago, 2015.



Fuente: Elaboración propia

Podemos observar en este gráfico, que la dispersión individual de la concentración de ciclosporina para C0 es variable en el tiempo, es decir, que no percibe una tendencia lineal en la mayoría de los 65 pacientes.

Luego en la Figura (4.3) observaremos la tendencia individual de cada paciente trasplantado de la concentración de ciclosporina en la sangre para $C2$.

Figura 4.3. Gráfico de dispersión individual de la concentración de ciclosporina en la sangre para ($C2$) de los pacientes trasplantados en el Hospital San Juan de Dios de Santiago, 2015.

Fuente:elaboración propia

Se aprecia en este gráfico, que la dispersión individual de la concentración de ciclosporina para C2 es más variable en el tiempo comparado con el gráfico de C0, es decir, que claramente no hay una tendencia lineal en los 65 pacientes.

4.2.2. Análisis de las variables genéticas

La primera Tabla (4.4), corresponde a la distribución de frecuencia de los cinco polimorfismos de la base de datos de ciclosporina.

Tabla 4.4. Distribución de frecuencia de los polimorfismos.

Polimorfismos	CYP3A4*3		CYP3A5*B		CYP3A4*22		MDR1*3435		MDR1*1236	
	n	%	n	%	n	%	n	%	n	%
0	54	83	3	5	63	97	19	29	15	23
1	11	17	20	31	2	3	31	48	36	55
2	-	-	42	65	-	-	15	23	14	22

0 = dos alelos salvajes; 1 = un alelo salvaje con un alelo mutante; 2 = dos alelos mutantes.

En esta distribución de frecuencias, se puede apreciar que para el polimorfismo CYP3A4*3, hay una mayor frecuencia con un 83 % correspondientes a 54 pacientes cuando existe la presencia de los dos alelos salvajes, además se observa que en este polimorfismo no hay presencia de los dos alelos mutantes. Para el polimorfismo CYP3A5*B, podemos observar que hay una mayor frecuencia con un 65 % correspondientes a 42 pacientes cuando existe la presencia de los dos alelos mutantes y una menor frecuencia con un 5 % asociados a 3 pacientes cuando existe la presencia de los dos alelos salvajes. Para el polimorfismo CYP3A4*22, se observa que hay una mayor frecuencia con un 97 % asociados a 63 pacientes cuando existe la presencia de los dos alelos salvajes, además se aprecia que en este polimorfismo no hay presencia de los dos alelos mutantes. Para el polimorfismo MDR1*3435, podemos observar que hay una mayor frecuencia con un 48 % correspondientes a 31 pacientes cuando existe la presencia de un alelo salvaje con un alelo mutante y una menor frecuencia con un 23 % asociados a 15 pacientes cuando existe la presencia de los dos alelos salvajes. Para el polimorfismo MDR1*1236, se observa que hay una mayor frecuencia con un 55 % correspondientes a 36 pacientes cuando existe la presencia de un alelo salvaje con un alelo mutante y una menor frecuencia con un 22 % asociados a 14 pacientes cuando existe la presencia de los dos alelos mutantes.

Luego la Tabla(4.5), representa un análisis descriptivo de los polimorfismos distribuidos por las condiciones: sexo, concentración de ciclosporina ($C0$) y ($C2$), edad, peso, dosisC0 y dosisC2.

Tabla 4.5. Análisis descriptivo de los polimorfismos.

VARIABLES	CYP3A4*3				CYP3A5*B						CYP3A4*22				MDR1*3435						MDR1*1236					
	0		1		0		1		2		0		1		0		1		2		0		1		2	
Sexo(M/H)	25	29	4	7	2	1	9	11	18	24	28	35	1	1	9	10	15	16	5	10	8	7	14	22	7	7
C0(media/D.E)	196	95	225	145	208	103	211	121	194	95	201	106	183	53	199	91	205	115	193	98	185	98	206	107	203	104
C2(media/D.E)	982	374	942	327	862	434	959	373	991	359	973	367	1041	371	1024	403	954	355	961	343	944	392	973	349	1012	385
Edad(media/D.E)	51	13	56	10	62	2	52	13	51	12	51	13	58	7	50	14	52	11	54	14	50	12	52	13	53	12
Peso(media/D.E)	68	12	64	10	62	6	68	11	68	12	67	12	69	15	65	11	69	12	68	12	67	12	68	11	68	13
DosisC0(media/D.E)	350	109	342	107	352	117	351	111	346	108	348	109	343	107	353	109	340	110	360	105	352	108	344	109	355	112
DosisC2(media/D.E)	331	101	316	95	360	115	330	104	326	98	329	101	322	77	337	105	320	100	338	93	336	103	322	97	339	106

0 = dos alelos salvajes 1 = un alelo salvaje con un alelo mutante 2 = dos alelos mutantes.

En este análisis descriptivo podemos observar para la variable sexo, que la mayor frecuencia se encuentra en el polimorfismo CYP3A4*22 en los hombres (n=35) cuando hay dos alelos salvajes y la menor frecuencia se encuentra en el polimorfismo CYP3A4*22 en los hombres (n=1) y mujeres (n=1) cuando hay un alelo salvaje con un alelo mutante y en el polimorfismo CYP3A5*B en los hombres (n=1) cuando hay dos alelos salvajes.

Para la variable C0, se observa que hay un promedio mayor en el polimorfismo CYP3A4*3 con un valor de 225 ng/ml cuando hay un alelo salvaje con un alelo mutante y un promedio menor en el polimorfismo CYP3A4*22 con un valor de 183 ng/ml cuando hay un alelo salvaje con un alelo mutante.

Para la variable C2, se observa que hay un promedio mayor en el polimorfismo CYP3A4*22 con un valor de 1041 ng/ml cuando hay un alelo salvaje con un alelo mutante y un promedio menor en el polimorfismo CYP3A5*B con un valor de 862 ng/ml cuando hay dos alelos salvajes.

Para la variable edad, se observa que la edad promedio mayor esta en el polimorfismo CYP3A5*B con un valor de 62 años cuando hay dos alelos salvajes y la edad promedio menor esta en los polimorfismos MDR1*3435 y MDR1*1236 con un valor de 50 años cuando hay dos alelos salvajes.

Para la variable peso, se observa que el peso promedio mayor esta en los polimorfismos CYP3A4*22 y MDR1*3435 con un valor promedio=69 kilos cuando hay un alelo salvaje con un alelo mutante y el peso promedio menor esta en el polimorfismo CYP3A5*B con un valor promedio=62 kilos cuando hay presencia de los dos alelos salvajes.

Para la variable dosisC0, se observa que hay un promedio mayor en el polimorfismo MDR1*3435 con un valor de 360 mg cuando hay dos alelos mutantes y un promedio menor en el polimorfismo MDR1*3435 con un valor de 340 mg cuando hay un alelo salvaje con un alelo mutante.

Para la variable dosisC2, se observa que hay un promedio mayor en el polimorfismo CYP3A5*B con un valor de 360 mg cuando hay dos alelos salvajes y un promedio menor en el polimorfismo CYP3A4*3 con un valor de 316 mg cuando hay un alelo salvaje con un alelo mutante.

4.3. Análisis de los Modelos

En esta sección observaremos los resultados obtenidos de la aplicación de los modelos: el MM y el MMS, utilizando los datos de ciclosporina para C0 y C2.

4.3.1. Aplicación Modelo Mixto (MM)

En esta parte del trabajo se expondrán los resultados de las distintas etapas de la aplicación del MM.

4.3.1.1. Modelo vacío

Un primer análisis del MM se realiza con el modelo vacío. Recordemos que este análisis no incluye ninguna variable explicatoria. En la Tabla (4.6) podemos apreciar los parámetros obtenidos del modelo vacío para C0 y C2.

Tabla 4.6. Modelo vacío del Modelo mixto de la concentración de ciclosporina en la sangre para C0 y C2 aplicado en los pacientes trasplantados en el Hospital San Juan de Dios, 2015.

	Modelo vacío C0				Modelo vacío C2		
Efecto fijo	Coeficiente β	D.E	Valor p	Efecto fijo	Coeficiente β	D.E	Valor p
constante	201.440	8.919	<0.001**	constante	983.685	24.554	<0.001**
Efecto aleatorio	Parámetro		CCI	Efecto aleatorio	Parámetro		CCI
var (σ)	3984.372		34(%)	var (σ)	22514.330		17(%)
var (ε)	7800.996		66(%)	var (ε)	113196.100		83(%)
AIC	5856.273			AIC	7045.316		

**valor significativo

En el resultado de estos modelos vacíos podemos observar en C0 que el valor promedio de concentración de ciclosporina en la sangre es de 201 ng/ml y el grado de dependencia de los pacientes con el coeficiente de correlación intraclase (CCI) es 0.34, esto nos informa que el 34 % de la variabilidad entre las medias es atribuible a diferencias entre paciente, mientras que el 66 % restante de la variabilidad es atribuible a la evolución del transcurso del tiempo dentro de cada paciente.

Para C2 el valor promedio de concentración de ciclosporina en la sangre es de 984 ng/ml y el grado de dependencia de los pacientes es de CCI=0.17, esto nos informa que el 17 % de la variabilidad entre las medias es atribuible

a diferencias entre paciente, mientras que el 83 % restante de la variabilidad es atribuible a la evolución del transcurso del tiempo dentro de cada paciente.

4.3.1.2. Análisis univariante del Modelo Mixto (MM)

Para determinar que variables explicatorias aportan para medir concentración de ciclosporina en la sangre en C0 y C2, se realiza el análisis univariante del MM, este resultado se muestra en la Tabla (4.7). Recordemos que este análisis se desarrollado de forma individual.

Tabla 4.7. Análisis univariante del Modelo mixto de la concentración de ciclosporina en la sangre para C0 y C2 en los pacientes trasplantados en el Hospital San Juan de Dios, 2015.

		Univariante C0				Univariante C2		
Variable	Categoría	Coeficiente β	D.E	Valor p	Categoría	Coeficiente β	D.E	Valor p
Sexo	constante	187.650	13.945		constante	1003.563	36.260	
		44.854	18.782	0.017**		-49.148	48.077	0.307
Edad	constante	67.039	36.941		constante	949.363	102.421	
		2.911	0.703	<0.001**		0.561	1.930	0.771
Peso	constante	144.840	49.380		constante	1391.327	128.993	
		1.018	0.736	0.167		-6.213	1.870	<0.001**
Tiempo	constante	241.479	12.520		constante	1039.851	41.684	
		-8.471	1.837	<0.001**		-12.287	6.345	0.053**
DosisC0	constante	96.288	19.179					
		0.322	0.047	<0.001**				
DosisC2					constante	0.401	0.185	
						850.812	62.450	<0.001**
CYP3A4*3	constante	202.077	10.379		constante	984.266	26.194	
	1	36.596	25.735	0.160	1	-39.618	64.769	0.541
CYP3A5*B	constante	228.050	46.120		constante	918.172	122.968	
	1	-0.830	49.240	0.987	1	44.500	129.966	0.732
	2	-27.987	47.650	0.557	2	72.026	126.636	0.570
CYP3A4*22	constante	210.861	9.821		constante	974.254	24.419	
	1	-25.776	55.930	0.645	1	89.302	134.943	0.508
MDR1*3435	constante	202.446	18.011		constante	1033.652	44.453	
	1	19.462	22.892	0.395	1	-84.877	55.918	0.129
	2	-3.386	26.978	0.900	2	-80.131	65.802	0.223
MDR1*1236	constante	187.920	20.140		constante	958.323	51.065	
	1	31.070	23.990	0.190	1	12.109	60.342	0.841
	2	27.800	28.980	0.340	2	58.118	72.714	0.424

1=un alelo salvaje con un alelo mutante 2 = dos alelos mutantes

**valor significativo

En este análisis univariante se observa que en C0 las variables explicatorias Sexo, Edad, Tiempo y DosisC0 fueron estadísticamente significativas, por lo cual son un aporte para medir la concentración de ciclosporina en C0.

Para C2 las variables explicatorias Peso, Tiempo y DosisC2 fueron estadísticamente significativas, por consiguiente son un aporte para medir la concentracion de ciclosporina en C2.

Observemos que en ambos modelos solo las variables que obtuvieron un resultado estadísticamente significativo son variables clínicas, ninguna de las variables genéticas en este análisis individual obtuvo un resultado significativo.

4.3.1.3. Análisis del Modelo Mixto (Modelo 1)

Se prueba el MM para C0 y C2 con las variables explicatorias seleccionadas del análisis univariante, el resultado de los parámetros obtenidos se observa en la Tabla (4.8).

Tabla 4.8. Modelo mixto con las variables explicatorias seleccionadas del análisis univariante, aplicado en C0 y C2 en los pacientes trasplantados (Modelo 1).

	Modelo 1 de C0				Modelo 1 de C2		
Efecto fijo	Coeficiente β	D.E	Valor p	Efecto fijo	Coeficiente β	D.E	Valor p
Sexo	5.380	18.611	0.773	Peso	-6.903	2.048	<0.001**
Edad	2.770	0.711	<0.001**	Tiempo	8.191	10.550	0.437
DosisC0	0.411	0.080	<0.001**	DosisC2	0.591	0.299	0.048**
Tiempo	5.033	3.385	0.137	Constante	1212.112	158.170	<0.001**
Constante	-98.134	51.434	0.056**				
Efecto aleatorio	Parámetro			Efecto aleatorio	Parámetro		
var(tiempo)	118.000			var(tiempo)	705.270		
var(constante)	3561.120			var(constante)	60304.010		
cov(tiempo,constante)	-154.108			cov(tiempo,constante)	-5946.382		
var(error)	5735.091			var(error)	105429.3		
AIC	5760.935			AIC	7010.051		

**valor significativo

Se observa en el MM (Modelo 1) de C0 que las variables explicatorias Edad y DosisC0 obtuvieron un resultado estadísticamente significativas, es decir que estas variables son un aporte para medir la concentración de ciclosporina en la sangre en C0.

En el MM (Modelo 1) de C2 las variables explicatorias Peso y DosisC2 obtuvieron un resultado estadísticamente significativas, es decir que estas variables son un aporte para medir la concentración de ciclosporina en la sangre en C2.

Observemos que en ambos modelos la cantidad de variables explicatorias que obtuvieron un resultado significativo disminuyó en comparación con la cantidad de variables explicatorias del análisis univariante.

4.3.1.4. Interacciones

Recordemos que el análisis anterior del MM no incluyó variables genéticas, dado que estas variables no fueron estadísticamente significativas en el análisis univarante, por esta razón es necesario evaluar si existe interacción entre las variables genéticas.

A continuación en la Tabla (4.9), se presenta el resultado de dos interacciones de los polimorfismos para la concentración de ciclosporina en la sangre en C0. Una interacción ocurre entre los polimorfismos CYP3A4*3 y CYP3A5*B asociados al metabolismo del medicamento y la otra interacción ocurre entre los polimorfismos MDR1*3435 y MDR1*1236 asociados al transporte del medicamento.

Tabla 4.9. Interacción de los polimorfismos en C0.

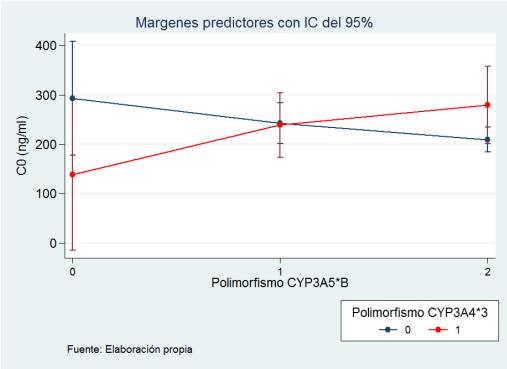
Polimorfismos		Categoría		Coeficiente β	D.E	Valor p
CYP3A4*3	CYP3A5*B	constante		293.258	58.924	
		1	1	151.494	105.199	0.150
		1	2	225.328	106.188	0.034**
MDR1*3435	MDR1*1236	constante		149.084	29.734	
		1	1	-19.373	48.241	0.688
		1	2	-233.989	69.918	0.001**
		2	1	-174.935	84.887	0.039**
		2	2	-284.431	98.847	0.004**

1 = un alelo salvaje con un alelo mutante 2 = dos alelos mutantes
**valor significativo

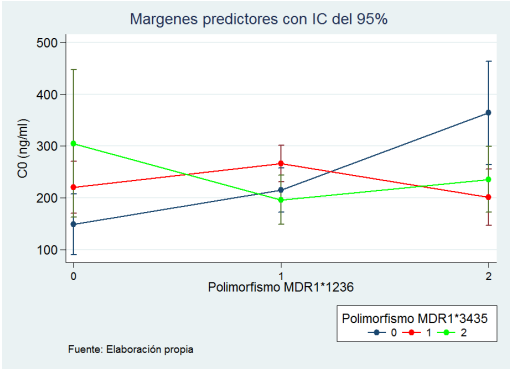
Se puede apreciar que dentro de las interacciones asociadas para el grupo de citocromo, la interacción que fue significativa es aquella que consideraba un alelo salvaje con un alelo mutante para el polimorfismo CYP3A4*3 y dos alelos mutantes para el polimorfismos CYP3A5*B. Además se observa que hay interacciones de forma simultanea asociadas al grupo de enzima, una interacción ocurre cuando considera un alelo salvaje con un alelo mutante para el polimorfismo MDR1*3435 y dos alelos mutantes para el polimorfismo MDR1*1236, otra interacción es cuando considera dos alelos salvajes para el polimorfismo MDR1*3435 y un alelo salvaje con un alelo mutante para el polimorfismo MDR1*1236 y por último una interacción cuando considera dos alelos mutantes para el polimorfismo MDR1*3435 y dos alelos mutantes para el polimorfismo MDR1*1236.

Para apreciar de mejor forma las interacciones entre los polimorfismos se visualiza con un gráfico para confirmar la existencia de interacción. Esto se realiza usando la concentración de ciclosporina en la sangre de C0, como se observa en la Figura (4.4).

Figura 4.4. Gráfico de interacción entre los polimorfismos en C0.



(a) CYP3A4*3 y CYP3A5*B



(b) MDR1*3435 y MDR1*1236

Podemos observar en el gráfico (a) que existe una mayor concentración de ciclosporina en la sangre cuando hay dos alelos salvajes en el polimorfismo CYP3A4*3, pero en la medida que aumentan los alelos mutantes del polimorfismo CYP3A5*B, la concentración de ciclosporina en la sangre disminuye y existe una menor concentración de ciclosporina en la sangre cuando hay un alelo salvaje con un alelo mutante en el polimorfismo CYP3A4*3, pero en la medida que aumentan los alelos mutantes del polimorfismo CYP3A5*B la concentración de ciclosporina en la sangre aumenta. La interacción entre estos dos polimorfismos ocurre cuando aumentan los alelos mutantes.

Podemos observar en el gráfico (b) que existe una mayor concentración de ciclosporina en la sangre cuando hay dos alelos mutantes en el polimorfismo MDR1*3435, pero en la medida que disminuyen los alelos mutantes del polimorfismo MDR1*1236, la concentración de ciclosporina en la sangre disminuye, pero si aumentan los alelos mutantes del polimorfismo MDR1*1236 la concentración de ciclosporina en la sangre aumenta y existe una menor concentración de ciclosporina en la sangre cuando hay dos alelos salvajes en el polimorfismo MDR1*3435, pero en la medida que aumentan los alelos mu-

tantes del polimorfismo MDR1*1236 la concentración de ciclosporina en la sangre aumenta. Las interacciones entre estos dos polimorfismos ocurre en la medida que aumenten o disminuyan los alelos mutantes.

Posteriormente en la Tabla (4.10), se presenta el resultado de una interacción entre los polimorfismos para la concentración de ciclosporina en la sangre en C2. Esta interacción ocurre entre los polimorfismos MDR1*3435 y MDR1*1236 asociados al transporte del medicamento.

Tabla 4.10. Interacción entre los polimorfismos en C2.

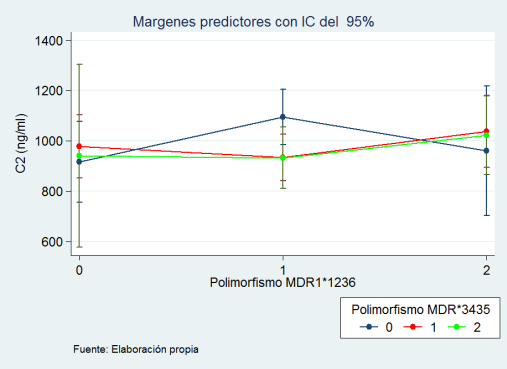
Polimorfismos		Categoría		Coeficiente β	D.E	Valor p
MDR1*3435	MDR1*1236	constante		916.072	81.940	
		1	1	-223.207	127.132	0.070**
		1	2	-13.318	182.650	0.942
		2	1	-186.705	219.7	0.395
		2	2	36.266	254.618	0.887

1 = un alelo salvaje con un alelo mutante 2 = dos alelos mutantes
**valor significativo

Se observa que en este grupo de enzimas, es posible que ocurra una interacción cuando considera un alelo salvaje con un alelo mutante para el polimorfismo MDR1*3435 y un alelo salvaje con un alelo mutante para el polimorfismo MDR1*1236 con un valor-p cercano a 0.05, es por esta razón es que no se descarta que esta interacción pueda ser estadísticamente significativa.

Para confirmar de mejor forma ésta interacción entre estos polimorfismos se visualizará en un gráfico. Esto se realiza usando la concentración de ciclosporina en la sangre en C2, como se muestra en la Figura (4.5).

Figura 4.5. Gráfico de interacción entre los polimorfismos en C2.



En este gráfico podemos observar que existe una mayor concentración de ciclosporina en la sangre cuando hay un alelo salvaje con un alelo mutante en el polimorfismo MDR1*3435, pero en la medida que aumentan los alelos mutantes del polimorfismo MDR1*1236, la concentración de ciclosporina en la sangre aumenta y existe una menor concentración de ciclosporina en la sangre cuando hay dos alelos salvajes en el polimorfismo MDR1*3435, pero en la medida que aumentan los alelos mutantes del polimorfismo MDR1*1236 la concentración de ciclosporina en la sangre disminuye. Las interacciones entre estos dos polimorfismos ocurre en la medida que aumenten o disminuyan los alelos mutantes.

Estos resultados de las interacciones en C0 y C2, se usarán para el análisis posterior del MM.

4.3.1.5. Análisis del Modelo Mixto (Modelo Final)

En el análisis del MM con las interacciones se obtuvieron los siguientes parámetros, como se muestra en la Tabla (4.11). Recordemos que este modelo utilizo las variables explicatorias significativas del análisis univariante más las interacciones de los polimorfismos para cada concentración de ciclosporina.

Tabla 4.11. Modelo mixto con interacción, aplicado en C0 y C2 en los pacientes trasplantados en el Hospital San Juan de Dios, 2015 (Modelo final).

	Modelo final de C0				Modelo final de C2		
Efecto fijo	Coeficiente β	D.E	Valor p	Efecto fijo	Coeficiente β	D.E	Valor p
Sexo	-38.222	19.079	0.045**	Peso	-6.073	2.238	0.007**
Edad	3.065	0.666	<0.001**	Tiempo	2.354	11.514	0.838
DosisC0	0.502	0.080	<0.001**	DosisC2	0.443	0.327	0.175
Tiempo	8.861	3.713	0.017**	MDR			
CYP				1 1	-220.951	134.530	0.101
1 1	193.619	86.371	0.025**	1 2	-80.987	188.628	0.668
1 2	284.539	92.304	0.002**	2 1	-171.293	234.115	0.464
MDR				2 2	-96.671	267.053	0.717
1 1	8.803	42.381	0.835	Constante	1151.466	179.441	<0.001**
1 2	-218.911	66.786	<0.001**				
2 1	-158.888	74.723	0.033**				
2 2	-261.338	86.772	0.003**				
Constante	-157.196	78.111	0.044**				
Efecto aleatorio	Parámetro			Efecto aleatorio	Parámetro		
var(tiempo)	206.937			var(tiempo)	708.861		
var(constante)	980.659			var(constante)	65568.4		
cov(tiempo,constante)	88.955			cov(tiempo,constante)	-6289.822		
var(error)	5497.070			var(error)	105530.8		
AIC	5641.666			AIC	6933.917		

CYP=interacción polimorfismos CYP3A4*3 y CYP3A5*B

MDR=interacción polimorfismos MDR1*3435 y MDR1*1236

1 = un alelo salvaje con un alelo mutante 2 = dos alelos mutantes

**valor significativo

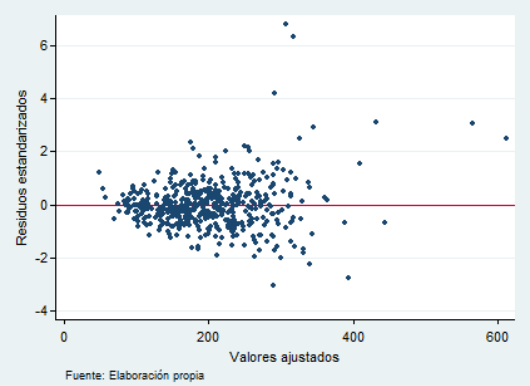
Se observa que el MM (Modelo final) de C0 las variables explicatorias Sexo, Edad, Tiempo, DosisC0 y las interacciones con los polimorfismos obtuvieron un resultado estadísticamente significativas, es decir que estas variables son un aporte para medir la concentración de ciclosporina en la sangre en C0.

En el MM (Modelo final) de C2 solo la variable explicatoria Peso obtuvo un resultado estadísticamente significativo, es decir que esta variable es un aporte para medir la concentración de ciclosporina en la sangre en C2. Con este resultado el modelo final de C2 no es muy favorable para los objetivos de investigación, por consiguiente no se llevará a acabo el diagnóstico del modelo, ya que el modelo se usará más tarde para modelar con el MMS.

4.3.1.6. Diagnóstico Modelo Mixto (MM)

El diagnóstico de varianza constante se puede observar en la Figura (4.6), donde verificaremos los supuestos distributivos para C0.

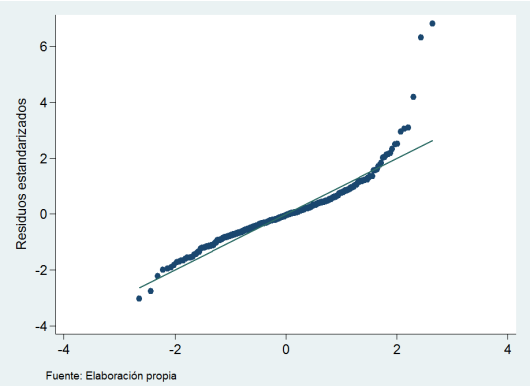
Figura 4.6. Gráfico de residuos v/s valores ajustados del Modelo final de C0.



En esta prueba de varianza se puede apreciar que los residuos presentan una estructura de tipo embudo con respecto a los valores predichos condicionales del modelo, es decir, que no hay una tendencia clara dentro de los residuos, por lo tanto tenemos una varianza que no es constante.

En la Figura (4.7) podemos observar la tendencia de los residuos estandarizados bajo una gráfica q-q plot.

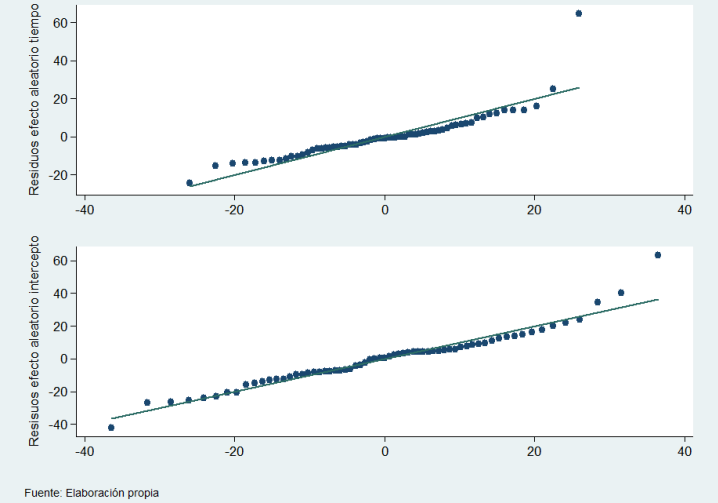
Figura 4.7. Gráfico Q-Q plot de los residuos estandarizados del Modelo final de C0.



Se aprecia en esta gráfica que la forma de los residuos presentan una distribución con colas mas largas, es decir, que los residuos se desvía de una distribución normal.

Finalmente en la Figura (4.8) podemos observar la tendencia de los residuos de los efectos aleatorios del tiempo y el intercepto.

Figura 4.8. Gráfico Q-Q normal de los residuos aleatorios del Modelo final usando el Modelo mixto.



En el gráfico del panel superior correspondiente a los residuos aleatorios del tiempo, se aprecia que existen indicios de un valor atípico positivo, en términos del tiempo (residuos aleatorios mayores de 25) asociado con los pacientes. En el gráfico del panel inferior correspondiente a los residuos aleatorios del intercepto, podemos observar que pueden existir dos valores atípicos positivos, en términos del intercepto (residuos aleatorios mayores de 35) asociado con los pacientes.

Este análisis de diagnóstico se observar que el modelo final de C0 no cumple con una distribución de normalidad, es por esa razón que necesita un ajuste de los datos.

4.3.2. Aplicación del Modelo Mixto Semipamétrico (MMS)

Para realizar el ajuste de los modelos en C0 y C2 se utilizarán los splines. Se recuerda que los splines son una función de suavizamiento que están unidos en puntos (nodos).

Para modelar el MMS en la concentración de ciclosporina en la sangre en C0 y C2, se utilizan los modelos finales obtenido del MM y la variable tiempo (días) como splines probando con distintos valores para K. El resultado de este modelo se puede apreciar en la Tabla (4.12).

Tabla 4.12. Modelo mixto semiparamétrico con splines, aplicado en C0 y C2 en los pacientes trasplantados en el Hospital San Juan de Dios, 2015.

	MMS de C0				MMS de C2		
Efecto fijo	Coeficiente β	D.E	Valor p	Efecto fijo	Coeficiente β	D.E	Valor p
Sexo	-50.127	19.871	0.012**	Peso	-7.249	2.141	<0.001**
Edad	2.961	0.690	<0.001**	DosisC2	1.240	0.346	<0.001**
DosisC0	0.604	0.083	<0.001**	Tiempo1	155.812	33.852	<0.001**
Tiempo1	65.973	14.770	<0.001**	Tiempo2	-1584.66	484.583	<0.001**
Tiempo2	-965.006	297.082	<0.001**	Tiempo3	1820.798	601.403	0.002**
Tiempo3	1173.437	372.728	0.002**	MDR			
Tiempo4	-1958.632	866.512	0.024**	1 1	-252.041	130.962	0.054**
Tiempo5	2426.459	1130.547	0.032**	1 2	-193.614	178.016	0.277
CYP				2 1	-183.360	233.550	0.432
1 1	231.216	90.001	0.010**	2 2	-175.168	258.442	0.498
1 2	332.445	96.373	<0.001**	Constante	548.342	206.268	0.008**
MDR							
1 2	-184.007	70.619	0.009**				
2 1	-158.597	77.440	0.041**				
2 2	-242.071	90.651	0.008**				
Constante	-255.151	83.223	0.002**				
Efecto aleatorio	Parámetro			Efecto aleatorio	Parámetro		
var(tiempo)	156.680			var(tiempo)	972.369		
var(constante)	1159.864			var(constante)	88428.21		
cov(tiempo,constante)	182.071			cov(tiempo,constante)	-9085.523		
var(error)	5277.151			var(error)	96656.12		
AIC	5579.743			AIC	6880.008		

CYP=interacción polimorfismos CYP3A4*3 y CYP3A5*B

MDR=interacción polimorfismos MDR1*3435 y MDR1*1236

1 = un alelo salvaje con un alelo mutante 2 = dos alelos mutantes

**valor significativo

Se observa en el MMS de C0 con K=6, que las variables explicatorias Sexo, Edad, Tiempo, DosisC0 y las interacciones con los polimorfismos obtuvieron un resultado estadísticamente significativas, es decir que estas variables son un aporte para medir la concentración de ciclosporina en la sangre en C0, también podemos observar que el valor del AIC es menor que el valor del modelo final del MM.

En el MMS de C2 con K=4, que las variables explicatorias Peso, Tiempo, DosisC2 y la interacción del polimorfismo obtuvieron un resultado estadísticamente significativas, cuando hay un alelo salvaje con un alelo mutante del polimorfismo MDR1*3435 con un alelo salvaje con un alelo mutante del polimorfismo MDR1*1236, es decir que estas variables son un aporte para medir la concentración de ciclosporina en la sangre en C2, con un valor de AIC más bajo que el modelo final del MM.

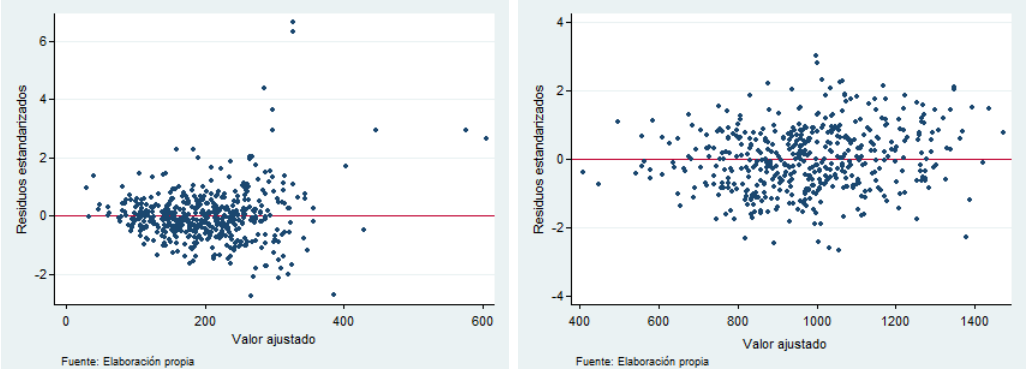
4.3.2.1. Diagnóstico Modelo Mixto Semiparamétrico (MMS)

Recordemos que el diagnóstico del MMS se ejecuta en dos partes. Un diagnóstico para la parte paramétrica para ambas concentraciones C0 y C2 y un diagnóstico para la parte semiparamétrica también para ambas concentraciones C0 y C2.

4.3.2.2. Diagnóstico parte paramétrica.

Se comienza con el análisis del diagnóstico de la varianza, la Figura (4.9) nos muestra el resultado para C0 y C2.

Figura 4.9. Gráficos de residuos v/s valores ajustados del Modelo mixto semiparamétrico.



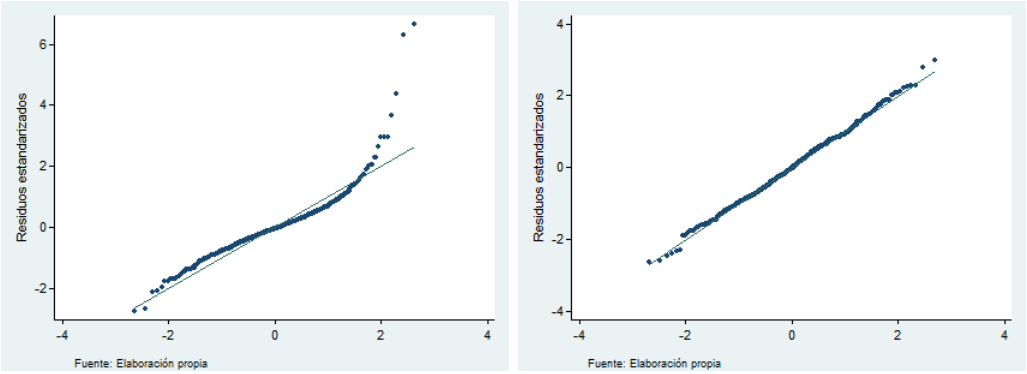
(a) Varianza C0 (b) Varianza C2

En ésta prueba de varianza del modelo mixto semiparamétrico, podemos observar en la varianza C0 que los residuos están distribuidos un poco más aleatorios comparado con la varianza del MM, pero se percibe que hay datos alejados de la linea, es decir presenta datos atípicos, podemos suponer que aún la varianza no es contante.

En la varianza C2 se puede observar que los residuos están distribuidos más aleatoriamente en forma de bandas sobre y bajo la linea recta, por lo tanto podemos suponer que la varianza es constante.

Posteriormente se presentan los gráficos q-q plot, donde observaremos la tendencia distributiva del modelo agrupado en cuantiles, para la concentración de ciclosporina en la sangre de C0 y C2, como se muestra en la Figura (4.10)

Figura 4.10. Gráficos Q-Q plot de los residuos del Modelo mixto semiparamétrico.

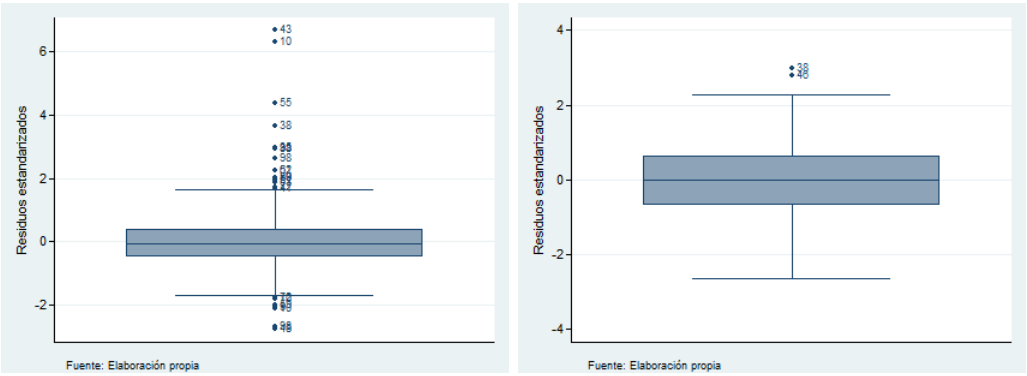


(a) q-q plot C0 (b) q-q plot C2

Se observa que en el gráfico q-q plot para C0, hay una cola más largas en el extremo superior de los residuos, que puede tener origen en que existan datos atípicos, con este resultado podemos suponer que C0 no sigue una distribución normal. En el q-q plot para C2, podemos observar que los residuos están sobre la linea recta solo se visualiza algunos puntos fuera de linea, podemos suponer que tiene una distribución normal.

Para comprobar estos datos atípicos, observaremos los gráficos de caja para C0 y C2, como se muestra en la Figura (4.11).

Figura 4.11. Gráficos de caja de los datos atípicos del Modelo mixto semiparamétrico.



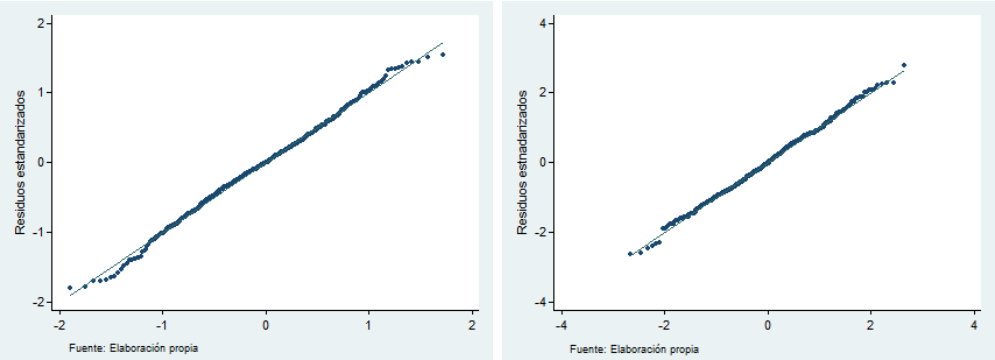
(a) datos atípicos de C0 (b) datos atípicos de C2

Podemos observar en estos gráficos de caja que para C0 hay muchos datos atípicos en el extremo superior y en el extremo inferior, por lo tanto estos outliers necesitarán ser evaluados.

Para C2 se observa que hay solo dos datos atípicos en el extremo superior, estos datos también serán evaluados.

A continuación se presentan los gráficos q-q plot sin datos atípicos, como se muestra en la Figura (4.12).

Figura 4.12. Gráficos Q-Q plot de los residuos sin datos atípicos del Modelo mixto semiparamétrico.

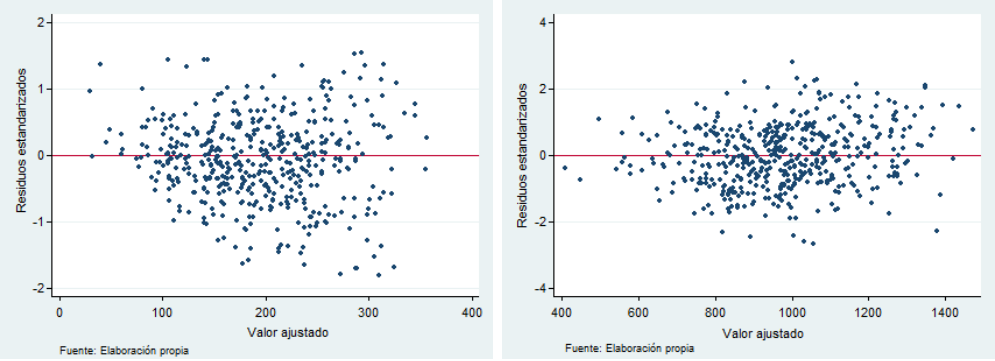


(a) q-q plot ajustado C0 (b) q-q plot ajustado C2

Después de eliminar los datos atípicos influyentes, se puede apreciar que en los gráficos q-q plot sin datos atípicos, para C0 y C2 la mayoría de los puntos están sobre la linea recta logrando mejorar el resultado anterior de los q-q plot, habiendo linealidad.

Luego es necesario volver a revisar el comportamiento de la varianza para C0 y C2, esto se aprecia en la Figura (4.13).

Figura 4.13. Gráficos de residuos ajustados del Modelo mixto semiparamétrico.



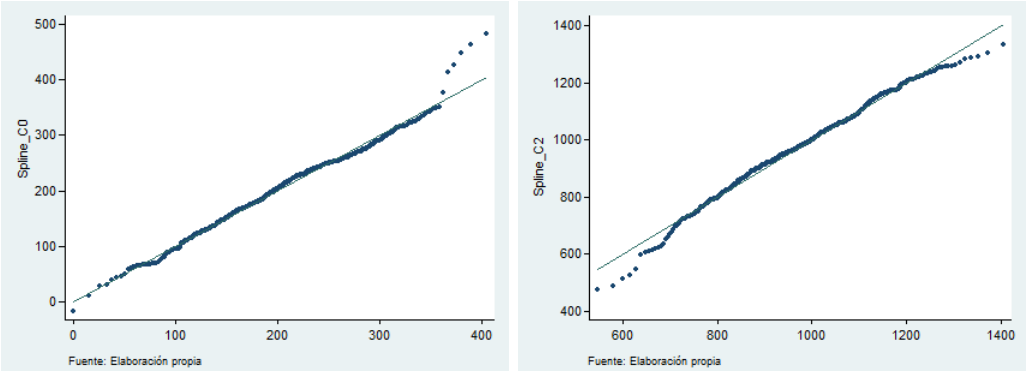
(a) Varianza ajustada C0 (b) Varianza ajustada C2

En estos gráficos de varianza ajustada, podemos observar que en ambos gráficos los puntos están distribuidos aleatoriamente en forma de bandas sobre y bajo la linea recta, esto nos confirma que la varianza en C0 y C2 es constante.

4.3.2.3. Diagnóstico parte semiparamétrica

Se comienza observando la tendencia distributiva de los splines, esto se visualiza en la Figura (4.14) para C0 y C2.

Figura 4.14. Gráficos Q-Q plot de los splines del Modelo mixto semiparamétrico.

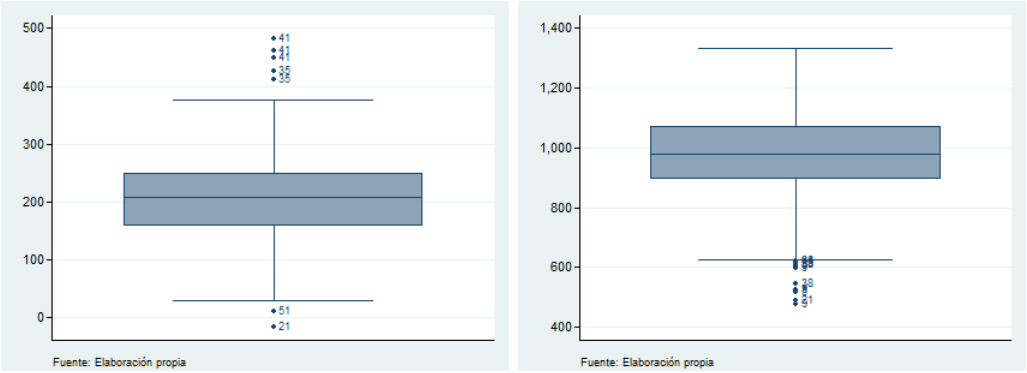


(a) q-q plot spline C0 (b) q-q plot spline C2

Podemos observar en ambos gráficos que existen puntos fuera de la linea recta, C0 tiene una cola mas larga en la parte superior y C2 tiene puntos en el extremo superior e inferior fuera de linea recta, por consiguiente es necesario evaluar si estos datos atípicos son influyentes.

Observemos los datos atípicos para C0 y C2, en la Figura (4.15).

Figura 4.15. Gráficos de caja de los datos atípicos de los splines del Modelo mixto semiparamétrico.

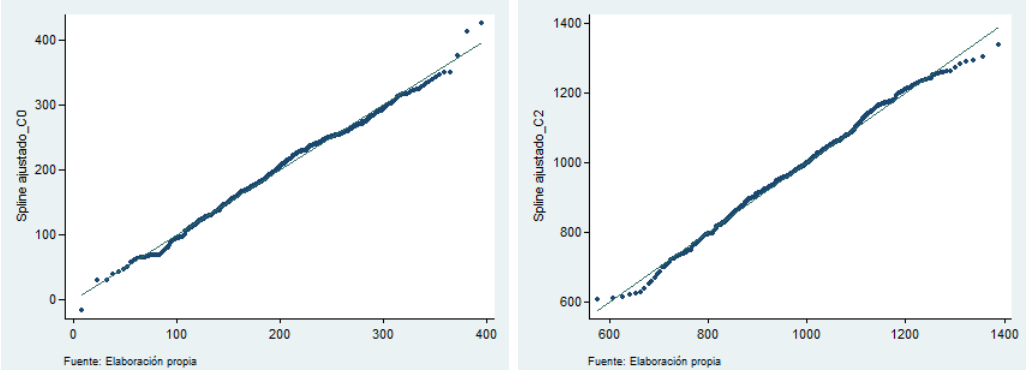


(a) datos atípicos del spline C0 (b) datos atípicos del spline C2

Se observa que en ambos gráficos de caja en los splines existen datos atípicos, en los splines C0 hay datos atípicos en ambos extremos superior e inferior y en los splines C2 hay datos atípicos en el extremo inferior. Estos datos serán evaluados y ajustados para observar el posterior comportamiento de los datos.

Después de evaluar y ajustar los datos atípicos se observa nuevamente los gráficos q-q plot para C0 y C2, estos gráficos se muestra a continuación en la Figura (4.16).

Figura 4.16. Gráficos Q-Q plot de los splines sin datos atípicos del Modelo mixto semiparamétrico.



(a) q-q plot spline ajustado C0 (b) q-q plot spline ajustado C2

Podemos observar que ambos gráficos la mayoría de puntos están sobre la linea recta, mejorando respecto del q-q plot anterior de los splines, habiendo linealidad en los datos. Estos resultados confirman que los splines tienen una distribución normal.

En estas dos partes del análisis del diagnóstico del MMS, podemos observar que para ambas concentraciones de ciclosporina en la sangre de C0 y C2 los datos logran mejorar respecto al MM.

4.3.3. Comparación de criterio de información (AIC)

En la Tabla (4.13), se muestra una comparación de los resultados del AIC y la cantidad de variables significativas por cada tipo de modelo.

Tabla 4.13. Comparación de Modelos.

	C0		C2	
Tipo de Modelo	AIC	**	AIC	**
Mixto vacío	5856	0	7045	0
Mixto 1	5761	2	7010	2
Mixto final	5642	9	6934	1
Mixto semiparamétrico	5580	13	6880	6

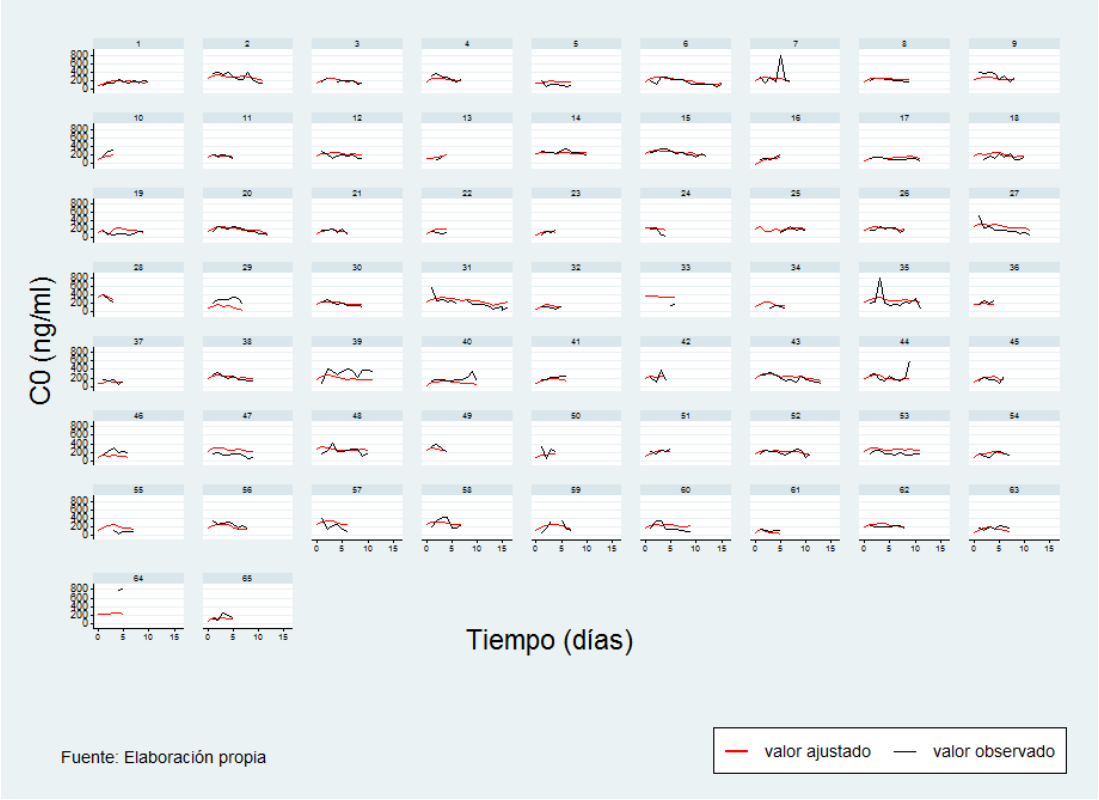
**cantidad de variables significativas por modelo.

Podemos observar que en esta tabla de comparación de los modelos, en C0 y C2 el modelo vacío tiene un mayor valor de AIC y en los siguientes modelos el valor de AIC empieza a disminuir, pero el modelo semiparamétrico es el que logra un menor valor de AIC y una mayor cantidad de variables significativas incluyendo variables clínicas como genéticas.

4.3.3.1. Curvas ajustadas con spline

Para complementar el análisis de los datos, se realizará el diagrama de las curvas individuales del comportamiento de los datos de los pacientes incluyendo el ajuste de las curvas con los splines, este análisis nos ayudara a observar si nuestro modelo semiparamétrico logra mejorar los datos de los pacientes, para esto primero observaremos un diagrama de las curvas para C0 como se muestra en la Figura (4.17), posteriormente observaremos un diagrama de las curvas para C2 como se muestra en la Figura (4.18).

Figura 4.17. Gráfico de dispersión individual con ajuste Spline del Modelo mixto semiparamétrico, aplicado a la concentración de ciclosporina en la sangre ($C0$) de los pacientes trasplantados en el Hospital San Juan de Dios de Santiago, 2015.



Fuente: Elaboración propia

Podemos observar la dispersión individual de los 65 pacientes de la concentración de ciclosporina en C0, donde las curvas observadas están representadas con lineas de color negro y las curvas ajustadas del suavizamiento de los splines del MMS con lineas de color rojo. Este ajuste logra mejorar la tendencia del comportamiento de los datos de los pacientes.

Figura 4.18. Gráfico de dispersión individual con ajuste Spline del Modelo mixto semiparamétrico, aplicado a la concentración de ciclosporina en la sangre ($C2$) de los pacientes trasplantados en el Hospital San Juan de Dios de Santiago, 2015.

Fuente: elaboración propia

Podemos observar la dispersión individual de los 65 pacientes de la concentración de ciclosporina en C2, donde las curvas observadas están representadas con lineas de color negro y las curvas ajustadas del suavizamiento de los splines del MMS con lineas de color rojo. Este ajuste logra mejorar la tendencia del comportamiento de los datos de los pacientes.

Capítulo 5

Conclusiones

Bajo el modelo vacío del modelo mixto, podemos concluir que este modelo fue nuestro análisis referencial en el cual obtuvimos el grado de dependencia de los pacientes, es este caso los pacientes no son homogéneos internamente y las observaciones son independientes.

Bajo el modelo mixto para C0, podemos concluir que en el modelo final las variables Sexo, Edad, DosisC0, Tiempo y las interacciones entre los polimorfismos CYP3A4*3 con CYP3A5*B y los polimorfismos MDR1*3435 con MDR1*1236 obtuvieron parámetros estadísticamente significativos con un valor AIC=5642, es decir que estas variables son un aporte para medir la concentración de ciclosporina en la sangre, pero la aplicacion del diagnóstico del modelo no logra cumplir con la normalidad. En el modelo mixto para C2, podemos concluir que en modelo final solo la variable Peso obtuvo un parámetro estadísticamente significativo con un valor AIC=6934, este resultado no es muy bueno para medir concentración de ciclosporina. En este modelo C0 tiene mejores resultados en la modelación, ya que hay más variables clínicas y genéticas que son significativas.

Bajo el modelo mixto semiparamétrico para C0, podemos concluir que las variables Sexo, Edad, DosisC0, Tiempo1, Tiempo2, Tiempo3, Tiempo4, Tiempo5 y las interacciones entre los polimorfismos CYP3A4*3 con CYP3A5*B y los polimorfismos MDR1*3435 con MDR1*1236 obtuvieron parámetros estadísticamente significativos con un valor AIC=5580 más bajo que el AIC del modelo mixto, es decir que todas las variables son un aporte para medir la concentración de ciclosporina en la sangre y en la aplicación del diagnostico del modelo logró obtener una distribución normal. En el modelo mixto semiparamétrico para C2, podemos concluir que las variables Peso, DosisC2, Tiempo1, Tiempo2, Tiempo3 y la interacción entre el polimorfismo MDR1*3435 con MDR1*1236 obtuvieron parámetros estadísticamnete significativos con un valor AIC=6880 más bajo que el AIC del modelo mixto, es decir que estas variables son un aporte para medir la concentración de ciclosporina en la sangre, aumentando notoriamente la cantidad de variables que aportan al modelo. En este modelo C0 obtiene mejores resultados en la modelación, ya que hay más variables clínicas y genéticas que son significativas.

Finalmente podemos concluir que el ajuste realizado a la variable tiempo en ambas concentraciones de ciclosporina en la sangre C0 y C2, logra mejorar la cantidad de variables que aportan para medir dosis de ciclosporina, por consiguiente un mejor modelo para estimar dosis de ciclosporina es el modelo mixto semiparamétrico cuando la concentración de ciclosporina en la sangre es C0, porque incluye mayor cantidad de variables clínicas y genéticas con un menor valor-p estadísticamente significativo y obtiene un menor valor AIC.

Bibliografía

Bourgain, J., Brezis, H., y Mironescu, P. (2000). Lifting in sobolev spaces. *Journal d'analyse mathématique*, 80(1):37–86.

Caratachea, M. A. C. (2007). Polimorfismos genéticos: Importancia y aplicaciones. *Revista del Instituto Nacional de Enfermedades Respiratorias*, 20(3):213–221.

Cebolla, H. (2013). Introducción al análisis multinivel. *colección cuadernos metodológicos*, (49).

de Rada, B. S. D. (2008). Monitorización y efectos secundarios de los inmunosupresores en el trasplante. In *Anales del Sistema Sanitario de Navarra*, volume 29, pages 207–218.

Eubank, R. (1985). Diagnostics for smoothing splines. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 332–341.

Green, P. J. (1987). Penalized likelihood for general semi-parametric regression models. *International Statistical Review/Revue Internationale de Statistique*, pages 245–259.

Green, P. J. y Silverman, B. W. (1994). *Nonparametric regression and generalized linear models: a roughness penalty approach*. CRC Press.

Ibacache-Pulgar, G., Paula, G. A., y Galea, M. (2012). Influence diagnostics for elliptical semiparametric mixed models. *Statistical Modelling*, 12(2):165–193.

Jann, B. (2013). Predictive margins and marginal effects in stata.

Kim, C., Park, B. U., y Kim, W. (2002). Influence diagnostics in semiparametric regression models. *Statistics & probability letters*, 60(1):49–58.

Oltra, B. P., RUIXO, J. P., TORRES, N. J., y MATEU, L. P. (2002). *Modelado farmacocinético de ciclosporina en pacientes con trasplante renal*. Universitat de València, Facultat de Farmàcia.

Roses, A. D. (2000). Pharmacogenetics and the practice of medicine. *Nature*, 405(6788):857–865.

Seoane, J. (2014). ¿ modelos mixtos (lineales)? una introducción para el usuario temeroso.

Shu, W.-y., Li, J.-l., Wang, X.-d., y Huang, M. (2015). Pharmacogenomics and personalized medicine: a review focused on their application in the chinese population. *Acta Pharmacologica Sinica*, 36(5):535–543.

West, B. T., Galecki, A. T., y Welch, K. B. (2014). *Linear mixed models*. CRC Press.

Wu, H. y Zhang, J.-T. (2006). *Nonparametric regression methods for longitudinal data analysis: mixed-effects modeling approaches*, volume 515. John Wiley & Sons.

Zhang, D., Lin, X., Raz, J., y Sowers, M. (1998). Semiparametric stochastic mixed models for longitudinal data. *Journal of the American Statistical Association*, 93(442):710–719.