



15

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE CIENCIAS

ESTIMACION DE UN MODELO DE SUPERVIVENCIA A LA DURACION Y
TIEMPO DE MANIFESTACION DE LOS EVENTOS ADVERSOS
OBSERVADOS DENTRO DE UN ESTUDIO COMPARATIVO DE EFICACIA
FARMACOLOGICA.

T E S I S
QUE PARA OBTENER EL TITULO DE:
A C T U A R I O
P R E S E N T A :
FERNANDO CANSECO MELCHOR



DIRECTOR DE TESIS M. en C. CLAUDIA LARA PEREZ SOTO





Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



DRA. MARÍA DE LOURDES ESTEVA PERALTA
Jefa de la División de Estudios Profesionales de la
Facultad de Ciencias
Presente

Comunicamos a usted que hemos revisado el trabajo escrito: "Estimación de un modelo de supervivencia a la duración y tiempo de manifestación de los eventos adversos observados dentro de un estudio comparativo de eficacia farmacológica."

realizado por Fernando Canseco Melchor

con número de cuenta 08825688-1, quien cubrió los créditos de la carrera de: Actuarial

Dicho trabajo cuenta con nuestro voto aprobatorio.

Atentamente

Director de Tesis
 Propietario

M. en C. Claudia Lara Pérez Soto.

Claudia Lara P.S.

Propietario

Mat. Margarita Elvira Chávez Cano.

M. E. Chávez

Propietario

M. en C. José Salvador Zamora Muñoz.

Zamora Muñoz J.S.

Suplente

Dr. Alberto Contreras Cristán.

CC

Suplente

M. en C. Beatriz Eugenia Rodríguez Fernández.

Beatriz Rodríguez

Consejo Departamental de Matemáticas.

[Firma]
 FACULTAD DE CIENCIAS
 M. en C. José Antonio Flores Díaz.
 CONSEJO DEPARTAMENTAL
 DE
 MATEMÁTICAS

Estimación de un modelo de supervivencia a la duración y tiempo de manifestación de los eventos adversos observados dentro de un estudio comparativo de eficacia farmacológica.

Dedicado, con cariño, a mamá, papá y Graciela.

En profundo agradecimiento a mis maestros.

**Con especial agradecimiento a quienes asesoraron
este trabajo, particularmente al Dr. Jorge Olguín.**

ÍNDICE.

ÍNDICE.	1
INTRODUCCIÓN.	4
OBJETIVOS.	5
CAPÍTULO I : ANTECEDENTES.	6
1.1 Introducción.	6
1.2 El análisis de supervivencia dentro de la investigación farmacéutica.	6
1.3 Eventos adversos.	7
1.3.1 Eventos adversos dentro de la investigación farmacéutica.	9
1.3.2 Desarrollo de un ensayo clínico.	9
1.4 Sinopsis sobre el ensayo clínico base.	11
1.4.1 Antecedentes.	12
1.4.2 Diseño de la investigación clínica.	12
1.4.2.1 Criterios de selección de pacientes.	14
1.4.2.2 Tamaño de la población.	15
1.4.3 Desarrollo de la investigación clínica.	15
1.4.4 Registro y documentación de los eventos adversos dentro del ensayo clínico.	17
1.4.5 Criterios de evaluación de resultados del ensayo clínico.	17
1.5 El análisis de supervivencia como criterio de evaluación. ..	18
CAPÍTULO II : ANÁLISIS DE SUPERVIVENCIA.	19
2.1 Introducción.	19
2.2 Modelo probabilístico de supervivencia.	19
2.2.1 Conceptualización del tiempo de observación dentro de un estudio de supervivencia.	20
2.2.2 Función de supervivencia.	21
2.2.3 Función de distribución.	21
2.2.4 Función de densidad.	21
2.2.5 Función de riesgo.	22
2.2.6 Censura.	23
2.3 Estimación de la función de supervivencia.	26
2.3.1 Procedimientos no paramétricos.	26
2.3.1.1 Tabla de vida.	26
2.3.1.2 Función de supervivencia empírica.	28
2.3.1.3 Estimador de Kaplan-Meier.	29
2.3.2 Procedimientos paramétricos.	31
2.3.2.1 Modelo Weibull.	32

2.3.2.2 Modelo Exponencial.	33
2.3.2.3 Covariables.	34
2.3.2.4 Método de máxima verosimilitud para determinar estimadores.	35
2.3.2.4.1 Proceso de Newton-Raphson para determinar estimadores.	38
2.3.2.5 Intervalos de confianza.	39
2.4 Estimación y ajuste de un modelo paramétrico.	42
2.4.1 Transformaciones para métodos gráficos.	43
2.4.2 Pruebas de hipótesis para comparar la experiencia de supervivencia entre varias muestras.	45
2.4.2.1 Prueba de Mantel-Haenszel para dos muestras.	45
2.4.2.2 Generalización para dos muestras.	47
2.4.2.3 Generalización para más de dos muestras.	48
2.4.3 Proceso general de ajuste de un modelo paramétrico de supervivencia.	50
CAPÍTULO III : ESTIMACIÓN DE UN MODELO PARAMÉTRICO DE SUPERVIVENCIA.	51
3.1 Introducción.	51
3.2 Descripción de las bases de datos originales.	51
3.3 Base de datos de eventos adversos.	52
3.3.1 Integridad de la información.	57
3.3.2 Proceso de identificación de los eventos adversos emergentes.	57
3.3.3 Criterios de selección de registros para el ajuste de los modelos de supervivencia.	59
3.3.3.1 Modelo del tiempo de duración.	59
3.3.3.2 Modelo del tiempo anterior a la primera manifestación.	61
3.3.4 Determinación del tipo de censura.	63
3.3.5 Selección de los eventos adversos más frecuentes. ..	64
3.4 Estimación de la función de supervivencia por el proceso no paramétrico de Kaplan-Meier.	68
3.5 Estimación de la función de supervivencia aplicando los modelos paramétricos Exponencial y Weibull.	77
3.5.1 Aproximación por métodos gráficos.	77
3.5.2 Estimación de los modelos Exponencial y Weibull.	80
3.5.3 Criterios de selección de la función de supervivencia.	90
3.6 Proceso de prueba de hipótesis para comparar las experiencias de supervivencia entre tratamientos.	93
3.7 Resumen de procedimientos y resultados.	96
3.8 Conclusiones.	107
ANEXO 1. Formato para elaborar reportes resumidos de ensayos clínicos.	109

ANEXO 2. Desarrollo del ensayo clínico.	111
ANEXO 3. Formato utilizado para recabar información sobre los eventos adversos.	112
ANEXO 4 Bases de datos utilizadas en la construcción de los modelos de supervivencia que describen el tiempo de duración y el tiempo anterior a la primera manifestación de los eventos adversos emergentes por tipo de tratamiento.	113
BIBLIOGRAFÍA.	119

INTRODUCCIÓN.

Cuando se utiliza el término supervivencia, generalmente se le relaciona con temas relativos a mortalidad y en consecuencia a demografía. Afortunadamente esta palabra abarca ámbitos mucho más amplios. Por ejemplo, la evolución de diversos fenómenos que acontecen en el campo de la medicina, siguen los supuestos básicos que sustentan a esta área de la estadística, razón por la cual se prefiere aplicar este tipo de análisis como la herramienta adecuada para estudiarlos.

En este trabajo se analizará la duración y el tiempo anterior a la primera manifestación de los eventos adversos emergentes observados dentro de un ensayo clínico comparativo, enfocado al área farmacéutica. Para llevar a cabo lo anterior se aplicarán algunas de las técnicas que ofrece el análisis de supervivencia para construir un modelo matemático, basado en una estimación paramétrica, que describa razonablemente el comportamiento de ambos fenómenos con el paso del tiempo. Esta estimación se sustentará en la información disponible de las manifestaciones de los eventos adversos emergentes observados para cada uno de los tres tratamientos abordados dentro del ensayo clínico base. Cabe señalar que dicho ensayo se planteó con el objetivo de comparar la eficacia de un nuevo medicamento antihipercolesterolémico contra la de otros dos tratamientos conocidos. En este sentido, algunos de los resultados que se obtengan en este trabajo, pueden utilizarse también como punto de comparación. Debe aclararse que un evento adverso es cualquier suceso no deseable que le ocurra a un paciente durante el periodo de tratamiento médico, independientemente si existe algún tipo de relación entre uno y otro.

En el primer capítulo, se abordarán los argumentos que fundamentan la aplicación del análisis de supervivencia a la investigación farmacéutica, especialmente enfocados a la descripción del comportamiento de los eventos adversos emergentes. En este capítulo también se ha incluido un resumen del ensayo clínico que da origen a este trabajo.

El segundo capítulo presenta el marco teórico sobre el cual se sustentan los modelos de supervivencia que serán construidos. Los elementos de análisis que aquí se tratan, incluyen una exposición de los supuestos elementales del análisis de supervivencia y algunos métodos de estimación no paramétrica y paramétrica de la función de supervivencia.

En el tercer y último capítulo, se desarrollará todo el proceso de estimación de parámetros y validación de la función propuesta, bajo los modelos Weibull y Exponencial, para el tiempo de duración y el tiempo anterior a la primera manifestación de los eventos adversos emergentes en general, y para aquéllos observados dentro de los dos sistemas corporales con mayor incidencia de casos. También se explicará la labor de adecuación que fue necesario realizar sobre la información original, en virtud de que ésta no podía utilizarse directamente para llevar a cabo las diversas estimaciones que habrían de efectuarse.

Finalmente, los resultados alcanzados serán presentados por tipo de fenómeno y tratamiento, destacando los que corresponden al nuevo medicamento.

OBJETIVOS.

Construir un modelo paramétrico que describa el comportamiento del tiempo de duración y el tiempo anterior a la primera manifestación de los eventos adversos emergentes más frecuentes observados en el estudio farmacológico denominado "Estudio multinacional, multicéntrico, de asignación aleatoria, tratamiento doble ciego y doble falso del resultado de aplicar los tratamientos A, B y C (en donde C es una combinación de A y B) en la reducción de los niveles de colesterol LDL en pacientes con hipercolesterolemia primaria", a través de la estimación de la función de supervivencia.

Explicar la importancia de aplicar el análisis de supervivencia como herramienta de estudio y toma de decisiones en el campo de la investigación farmacológica.

Sustentar los modelos paramétricos desarrollados a través de la exposición de los argumentos estadísticos que los fundamentan.

Presentar las aportaciones y conclusiones, derivadas de este trabajo, en el contexto del estudio farmacológico que le da origen.

CAPÍTULO I : ANTECEDENTES.

1.1 Introducción.

Este capítulo tratará la importancia de aplicar el análisis de supervivencia a la investigación farmacéutica, mostrando que esta herramienta estadística es de gran utilidad para explicar algunos de los resultados que se obtienen dentro de un ensayo clínico. Será de especial atención la exposición que se hará sobre los eventos adversos; primero, en lo que respecta a sus características más generales y, después, en el contexto de una investigación médica específica. Asimismo, se sentarán algunas de las bases conceptuales que permitan alcanzar el objetivo central de este trabajo.

Cabe señalar, que para tener una mejor perspectiva del estudio base seleccionado, se ha incluido dentro de este capítulo, una breve síntesis del protocolo médico original, resaltando sus objetivos, los criterios de selección de los pacientes participantes, el diseño y desarrollo de la investigación, la forma en cómo se documentaron los eventos adversos y los criterios de evaluación de sus resultados.

1.2 El análisis de supervivencia dentro de la investigación farmacéutica.

Es posible asociar cualquier fenómeno de la naturaleza, o inducido por el hombre, con el tiempo; ya sea porque existe una relación directa, y por ello evidente, o bien porque se crea esa relación. En ambos casos nuestro interés siempre nos lleva a cuestionarnos sobre cómo un fenómeno evoluciona (aparece, se desarrolla y desaparece), es decir la forma en cómo se va manifestando con el paso del tiempo.

Es común encontrar, dentro de cualquier rama del conocimiento, una gran variedad de datos que dependen del tiempo de ocurrencia de un evento en particular. Para analizarlos, se recurre a estudios longitudinales, entre los cuales, el análisis de supervivencia es uno de ellos.

El análisis de supervivencia toma su nombre del proceso vida-muerte o mortalidad que opera sobre una población; su principal interés se centra en aquellas entidades que sobreviven o, más específicamente, en el tiempo que transcurre antes de que una entidad muera. No obstante lo anterior, el término supervivencia puede aplicarse a muchos otros fenómenos en diversas áreas. Algunos de ellos son: la duración de equipos y materiales sometidos a pruebas de resistencia o al uso cotidiano, los periodos de cura de enfermedades, el tiempo que demora en concluir una tarea una persona o una máquina, etc.

Dentro de la investigación farmacéutica, el análisis de supervivencia es relevante, ya que aporta elementos estadísticos de análisis en el proceso de estudio previo a la selección de mejores medicamentos. Su principal aplicación consiste en el desarrollo de modelos matemáticos que describan el comportamiento de ciertos fenómenos y expliquen algunos resultados que se presentan durante la fase experimental o de pruebas; por ejemplo: la forma en cómo se va a ir manifestando

la respuesta del paciente a un cierto medicamento, o el tiempo que tarda en aparecer algún evento no deseado después de iniciado un tratamiento, o la duración de este mismo evento en un periodo determinado o mientras se continúe ingiriendo un cierto fármaco, etc. Debe señalarse que en cualquiera de estos casos, u otros afines, queda de manifiesto un interés común: entender el desarrollo de un fenómeno aleatorio sujeto al tiempo.

Hoy en día, tanto laboratorios como empresas farmacéuticas tienen establecidas rigurosas normas de operación en lo que se refiere a sus procesos de investigación y comprobación de resultados. El diseño estadístico, como parte fundamental de ambos procesos, no sólo contribuye a una mejor planeación de la investigación en sí, también aportará elementos adicionales de análisis al momento de revisar y decidir sobre los resultados alcanzados; y es así como el análisis de supervivencia también plantea algunos criterios que permiten evaluar si un tratamiento o medicamento es efectivo y seguro.

1.3 Eventos adversos.

Dentro de la investigación farmacéutica se entiende por evento adverso cualquier suceso no deseado que experimente un paciente durante el desarrollo de un estudio clínico, sin prejuzgar si éste fue causado por el medicamento que se está investigando.

Los eventos adversos se clasifican según la relación que se establece entre el evento y el tratamiento:

- Eventos cuya ocurrencia no puede asociarse con el medicamento en investigación pero existe, al menos, una explicación alternativa que lleve a comprender el porqué de su manifestación.
- Eventos de relación remota; en donde podría establecerse algún tipo de relación temporal pero con carácter improbable, entre el tratamiento médico y el evento adverso ya que su ocurrencia puede ser atribuida a la presencia de alguna enfermedad concurrente o a la presencia de otras sustancias y elementos químicos.
- Eventos en donde es posible establecer una relación, lo cual significa que la ocurrencia del evento puede deberse a la administración de la sustancia o dosis que se investiga.
- Eventos con probable relación, la cual es aplicable sólo cuando, resultado de un adecuado seguimiento en la administración del medicamento, es poco probable atribuir su ocurrencia a otras causas diferentes al tratamiento mismo.
- Se dice que un evento adverso guarda una relación no accesible con la investigación cuando éste es el resultado de reacciones adversas sobre las cuales no se puede hacer juicio alguno ya que la información disponible es insuficiente o contradictoria y no puede ser complementada ni verificada, es decir, no hay explicación alguna sobre su ocurrencia.

Para establecer algún tipo de relación entre el evento y el tratamiento, deben aplicarse algunos criterios clínicos sobre la información recabada, al momento de elaborar los reportes sobre los eventos adversos. Los criterios básicos son :

- La duración del evento mientras el paciente participa en el estudio.
- Los efectos al discontinuar o continuar con el tratamiento.
- La evolución de la misma enfermedad u otras enfermedades propias de cada individuo participante.
- La presencia de otros medicamentos o tratamientos clínicos a los que estuviese sometido el paciente
- El estado de estrés físico, mental o ambos a la vez.
- Efectos farmacodinámicos¹ y farmacocinéticos² del medicamento, pues cada persona reacciona en forma distinta a un determinado agente químico y presenta diferentes tipos de respuesta a los procesos de absorción, distribución, asimilación y excreción de las sustancias activas de las medicinas.
- La intensidad de los eventos adversos, la cual puede presentarse en tres magnitudes distintas:
 1. Leve: generalmente se otorga este atributo a aquellos eventos de carácter transitorio y que no interfieren en el desempeño normal de las actividades del paciente.
 2. Moderado: se aplica a los eventos lo suficientemente intensos para interferir con las actividades cotidianas de los pacientes.
 3. Severo: es un evento adverso que incapacita y obstaculiza al paciente para realizar toda actividad que le es normal.

Otra forma de clasificar a los eventos adversos es de acuerdo con el momento en que aparecen con respecto al tiempo que dura un tratamiento médico. De acuerdo con este criterio, los eventos adversos pueden ser:

- De tratamiento emergente, sólo si comenzaron, por primera vez, después de iniciado el tratamiento; sin embargo, este concepto también es aplicable a aquellos eventos definidos en cualquiera de los siguientes dos casos:
 1. Eventos que reaparecieron después de iniciado el tratamiento; es decir, que sus primeras manifestaciones iniciaron y terminaron antes de comenzada la medicación para luego volverse a manifestar cuando el tratamiento ya estaba en curso.
 2. Los eventos que ya venían ocurriendo antes de comenzado el tratamiento, y que incrementaron su intensidad cuando éste dio inicio.
- De tratamiento no emergente sí:

¹ La farmacodinámica puede definirse como el estudio de los efectos bioquímicos y fisiológicos de las sustancias que componen los fármacos sobre los diferentes órganos y sistemas del cuerpo humano.

² Cuando se habla de los efectos farmacocinéticos de los medicamentos, se hace referencia a la forma en que éstos interactúan con los órganos y sistemas del cuerpo humano durante las diferentes fases de asimilación de los mismos.

-
1. Todas sus manifestaciones comenzaron y terminaron antes o después del periodo de tratamiento.
 2. Sus manifestaciones comenzaron antes de iniciado el tratamiento y no variaron en intensidad después de comenzado el mismo.
 3. Aquéllos que ya se venían manifestando antes de iniciado el tratamiento y cuya intensidad disminuyó cuando éste dio inicio.

Para contar con una mejor perspectiva sobre la importancia que tienen los eventos adversos dentro de un estudio clínico, a continuación se presenta la descripción del marco general en el que se desarrolla este tipo de estudios.

1.3.1 Eventos adversos dentro de la investigación farmacéutica.

Dentro de todo proceso de investigación farmacéutica, la fase de pruebas tiene una importancia fundamental. Los resultados que ahí se obtengan indicarán qué tan seguro y eficaz es un medicamento en el tratamiento de una determinada enfermedad. Por esta razón, los estudios que se planteen durante esta etapa buscan demostrar que los resultados de nuevos tratamientos son al menos semejantes a los resultados de tratamientos ya conocidos.

El término seguridad se asocia directamente con la presencia de los eventos adversos. Los estudios planteados con este objetivo buscan comparar los niveles de incidencia de estos eventos entre distintos tratamientos. Por otro lado, en los estudios de eficacia, se comparan los niveles de efectividad de los tratamientos, es decir, qué tanto contribuyen los fármacos que se están analizando a mitigar la enfermedad que se está combatiendo.

El principal objetivo de ambas pruebas es demostrar, con base en las evidencias recabadas a lo largo del ensayo, que el uso de un nuevo medicamento, un nuevo tratamiento o una nueva dosis es tan seguro, efectivo o, incluso mejor que el uso de otro u otros medicamentos anteriormente probados. De este modo, las pruebas que han de realizarse tendrán que validar este hecho a partir de la comparación de resultados.

Tanto las pruebas de seguridad como las de eficacia no son excluyentes, ambas están presentes en todo ensayo clínico y pueden ser planteadas en alguna de sus dos modalidades: aquéllas encaminadas a probar seguridad y eficacia ó eficacia y seguridad; el orden en el nombre del objetivo, indicará su propósito principal.

Aunque los eventos adversos tienen una relación directa con las pruebas de seguridad, en estudios donde el objetivo principal sea probar eficacia, también deberán registrarse y analizarse todos los eventos que lleguen a presentarse, ya que esto contribuye a reforzar el objetivo de seguridad.

A continuación, se explicará brevemente el proceso que debe seguir un ensayo clínico, cuya finalidad sea probar seguridad o eficacia como objetivos primarios.

1.3.2 Desarrollo de un ensayo clínico.

La finalidad de cualquier ensayo clínico es mostrar, bajo un contexto lógico y secuencial, un panorama de todo el estudio farmacológico por medio de la exposición de sus propósitos, el planteamiento de objetivos y parámetros, la descripción de las pruebas realizadas, los resultados obtenidos, la discusión de los

mismos y las conclusiones finales. El documento que se presenta para este tipo de estudios consta, básicamente, de las siguientes partes:

i) Introducción. Aquí se expone el objetivo primario de la investigación, el cual se plantea en función al tipo de estudio y pruebas que habrán de realizarse: seguridad o eficacia. El objetivo de seguridad se basa en la incidencia de los eventos adversos; por su parte el objetivo de eficacia se sustenta en la variable de respuesta planteada. En esta parte es usual incluir los antecedentes sobre el padecimiento y los tratamientos más comunes para combatirlo, tratando de no omitir información relevante.

ii) Objetivo del estudio. En esta sección se definen los objetivos que se pretenden alcanzar. El objetivo principal se establece en función de probar seguridad o eficacia según sea el caso; los objetivos secundarios se plantean de tal manera que contribuyan a reforzar al primero.

iii) Plan de investigación. Aquí se expone, en detalle, todo el plan de acción que se llevará a cabo. Los puntos esenciales que deben desarrollarse son los siguientes:

- a) *Plan y diseño.* En esta parte se presenta una descripción general de todo el ensayo explicando el tipo de diseño que se ha planteado para el mismo (abierto, ciego, doble ciego, doble falso, u otro), su duración, el ámbito territorial, el número de pacientes tratados así como las sustancias que serán comparadas.
- b) *Discusión sobre el diseño.* Se expondrán los argumentos para realizar la investigación, además de mencionar los puntos precisos de comparación que serán abordados.
- c) *Selección de la población.* Se expondrá la forma de reclutamiento de la población.
- d) *Tratamientos.* Se explica la forma de asignar a cada individuo en un grupo de estudio así como el tratamiento que va a ser aplicado en lo que respecta a su duración, dosis del medicamento, forma de administración y forma de aprovisionamiento.
- e) *Variables de seguridad y eficacia farmacológica.* En esta parte, se examinarán las variables de seguridad y eficacia definidas dentro de la investigación, a partir de la información recabada para los eventos adversos registrados o a través de la respuesta al tratamiento, según corresponda:
 - Los eventos adversos observados se documentarán en forma completa. Cada detalle se anota en el formato definido para tal efecto; algunos datos imprescindibles son: la fecha de inicio, duración, intensidad, tratamiento, la fecha de finalización y las acciones tomadas para controlarlo (si las hubo). También se determinará y registrará el tipo de relación del evento con el tratamiento en investigación. El levantamiento de esta información se realiza durante las visitas médicas de control. El procedimiento a seguir en caso de presentarse un evento adverso calificado como severo o grave, tenga o no relación con el ensayo, también se documentará en esta parte.
 - La variable primaria de eficacia reside en el nivel de respuesta que manifestaron los pacientes respecto al tratamiento, y es usual establecerlo por medio de un valor numérico que puede ser expresado

por sí mismo o en porcentaje. Este valor se basa, fundamentalmente, en los cambios favorables de los indicadores corporales alterados a consecuencia de la enfermedad padecida.

- iv) **Diseño del estudio.** En el diseño del estudio se describe el plan de análisis de los resultados del ensayo. En virtud de que estas investigaciones se plantean con el objeto de comparar seguridad o eficacia, se propone una prueba estadística adecuada la cual se utiliza como base para determinar las diferencias y similitudes entre todos los tratamientos para ambos tipos de objetivo.
- v) **Disposición de los pacientes.** Aquí se describe el grado de respuesta manifestado por los pacientes asignados en cada tratamiento.
- vi) **Evaluación de resultados.** En la parte final del ensayo se presenta la evaluación de los resultados de seguridad y eficacia con sustento en la evidencia estadística observada.
- La evaluación de seguridad se realiza a través de la comparación de la incidencia de los eventos adversos observados en cada tratamiento, tanto a nivel general como de forma particular según las variables demográficas previamente definidas, el tipo de evento observado, su intensidad o alguna otra característica de interés.
 - En la evaluación de eficacia se pretende demostrar que la respuesta del tratamiento experimental es al menos similar a la respuesta del tratamiento de control en virtud de que tanto uno como el otro contribuyeron a lograr una mejoría en los pacientes. Al igual que en el análisis de seguridad, la comparación de la respuesta de eficacia se hace tanto en forma general como particularizando en variables demográficas precisas.

Debe señalarse que no obstante que los criterios para la evaluación de seguridad y eficacia son distintos, los resultados de ambos guardan una relación muy estrecha pues todo ensayo clínico, no importando la forma en que halla sido planteado, pretende mostrar que un tratamiento farmacológico es confiable y tiene una buena respuesta.

El ANEXO 1 muestra una de las múltiples formas existentes para reportar, a manera de resumen, el desarrollo de cualquier ensayo clínico.

1.4 Sinopsis sobre el ensayo clínico base.

El análisis de supervivencia de los eventos adversos, sobre los cuales se desarrollará el presente trabajo, se desprende de la información del ensayo clínico conocido como "Estudio multinacional, multicéntrico, de asignación aleatoria, tratamiento doble ciego y doble falso del resultado de aplicar los tratamientos A, B y C (en donde C es una combinación de A y B) en la reducción de los niveles de colesterol LDL en pacientes con hipercolesterolemia primaria". La finalidad de este estudio se centró en probar eficacia.

A continuación, se presentará un resumen de los puntos relevantes del protocolo del estudio antes citado.

1.4.1 Antecedentes.

El colesterol es un compuesto químico indispensable para el funcionamiento normal de nuestro cuerpo además de ser un nutrimento; sin embargo, en cantidades excesivas es dañino. Por su origen, puede clasificarse en dos tipos: endógeno (se produce al interior del organismo, principalmente en el hígado donde también se almacena) y exógeno (se originan fuera del organismo y se adquiere por medio de los alimentos que se consumen).

Diversos estudios médicos han demostrado que el incremento en los niveles del colesterol (particularmente el relacionado con la lipoproteína LDL) es uno de los factores de mayor riesgo en el desarrollo de enfermedades cardiovasculares. Con la finalidad de disminuir este riesgo, se recomienda a los pacientes que presentan mediciones alteradas habituales que modifiquen sus hábitos alimenticios, además de someterse a un tratamiento farmacológico bajo vigilancia médica. Actualmente, uno de los métodos más utilizados en la reducción del colesterol endógeno es el que se basa en el uso de inhibidores de las enzimas encargadas de producirlo, particularmente la HMG-CoA reductasa (B-hidroximetilglutamaryl CoA reductasa) en su ciclo biosintético dentro del hígado.

El tratamiento identificado como A, se basa en un medicamento catalogado como uno de los más potentes en su tipo entre todos los medicamentos desarrollados para inhibir la función de la enzima HMG-CoA. Por su parte, el tratamiento B, relacionado con los fibratos, es de uso extendido en pacientes con hipercolesterolemia ya que ha demostrado ser eficaz además de bien aceptado por las personas que lo ingieren regularmente. El tratamiento C fue propuesto como un método alternativo dado que los resultados obtenidos en pruebas previas mostraron que hubo individuos que no respondieron ni al tratamiento A ni al tratamiento B por separado.

En este ensayo clínico se planteó comparar los resultados obtenidos de aplicar A con los resultados de B y C. En todo momento se consideró al tratamiento A como tratamiento de prueba.

Como antecedente se tiene que los eventos adversos reportados en las etapas que preceden esta investigación, fueron los siguientes: para el medicamento A se observó una elevación en los niveles de creatinina fosfoquinética y transaminasas; para el medicamento B se reportaron malestares gastrointestinales transitorios, y ocasionalmente salpullido, dolores musculares, acalambamientos musculares que pueden estar acompañados de un incremento en los niveles de creatinina fosfoquinética, incremento en las transaminasas hepáticas e incluso impotencia y alopecia en casos aislados.

1.4.2 Diseño de la investigación clínica.

En este punto, debe señalarse que este estudio se encuentra enmarcado en el contexto de una investigación mayor sobre los efectos y resultados del medicamento A en el tratamiento de la hipercolesterolemia primaria. Se definió hipercolesterolemia primaria como la afección que sufren las personas cuyos

valores en sus niveles de colesterol LDL³ y triglicéridos (TG⁴) son al menos de 4.12 mmol/L (160 mg/dL) y 3.97 mmol/L (350 mg/dL) respectivamente.

El objetivo primario del ensayo clínico base se centró en la evaluación de la eficacia del tratamiento A comparado con B y C. Esto se lograría observando los cambios en ciertos parámetros de medición de colesterol, lípidos y triglicéridos en la población tratada. La comparación en los niveles de seguridad fue planteada como objetivo secundario.

Con la finalidad de alcanzar el objetivo primario, se diseñó una prueba tendiente a comparar la reducción porcentual en los niveles de colesterol LDL y triglicéridos entre los pacientes seleccionados y sometidos a uno de los tres tratamientos. Estas pruebas se llevaron a cabo en diferentes centros hospitalarios de investigación (multicéntrico) ubicados en cuatro países (multinacional). La asignación de cada paciente en cada grupo de estudio se realizaría de manera aleatoria mientras que la distribución de los medicamentos se caracterizaría por estar planteada bajo un esquema doble ciego y doble falso (cada paciente recibió dos dosis iguales, una placebo y una activa, sin que él o su médico tratante tuviesen conocimiento sobre cuál de ellas contenía el medicamento).

La duración total de este proceso fue de 26 semanas continuas divididas en tres etapas:

1. Etapa de Reclutamiento.

El propósito de la Etapa de Reclutamiento fue eliminar todos los efectos residuales de cualquier tratamiento médico al que hubiese estado sometido el paciente antes de pasar a la siguiente fase.

En este ensayo, la Etapa de Reclutamiento tuvo una duración de 4 semanas. Durante este periodo, cada paciente visitó al médico en dos ocasiones, al principio y al final. En ambas visitas se verificó el cumplimiento de los criterios de inclusión en el estudio; y solo en caso de que el paciente hubiese cumplido con dichos criterios, pasaría a la siguiente etapa.

2. Primera Etapa.

Durante esta etapa, con una duración de 6 semanas y 4 visitas médicas (de la segunda a la quinta), se fijó como meta seleccionar, en forma aleatoria, sólo a aquellos pacientes susceptibles de participar en la siguiente fase de acuerdo con los resultados obtenidos en las medidas de los niveles colestéricos de cada uno.

³ LDL son las iniciales del nombre en inglés de la Lipoproteína de Baja Densidad (Low Density Lipoprotein). Una lipoproteína es una partícula soluble encargada de transportar el colesterol dentro del cuerpo a través de la sangre. Existen 5 tipos de lipoproteínas. Las LDL llevan el colesterol del hígado hacia los tejidos periféricos para depositarlo en las células que lo requieran. Estas lipoproteínas llegan a permanecer dentro del torrente sanguíneo durante varios días y son responsables de transportar casi tres cuartas partes del colesterol total. Debe señalarse que a mayor cantidad de colesterol intracelular, las LDL no podrán depositar su carga en las células, lo que originará un incremento del colesterol circulante, con los riesgos que esto conlleva.

⁴ TG es una abreviatura que denota a los lípidos conocidos como triglicéridos o triacilglicéridos. Los TG son el producto resultante de la reacción de ácidos grasos con el alcohol glicerol (en los seres vivos, los alcoholes más frecuentes son el glicerol y el colesterol). Los TG son abundantes en alimentos de origen animal. Cuando se consume un alimento rico en grasas, el organismo absorbe y empaqueta el colesterol junto con los TG dentro de la lipoproteína quilomión. Los quilomiónes circulan por los vasos linfáticos y la sangre, descargando los TG en los tejidos musculares (a donde aportan energía) o en los tejidos adiposos (donde se almacenan como reserva).

También se proporcionó a cada participante una dieta de estabilización adecuada y un tratamiento basado en la administración de medicamentos placebo⁵.

3. Segunda Etapa.

De entre aquellos sujetos que no fueron excluidos y que completaron la Primera Etapa en forma satisfactoria, se formaron tres grupos de tratamiento tomando como base criterios de asignación aleatoria para la inclusión de pacientes en cada uno. Cada grupo fue sometido a un tratamiento: A, B o C. El tratamiento se aplicó durante 16 semanas durante las cuales se verificaron 6 visitas médicas (incluyendo la quinta visita).

Esta etapa tuvo como objetivo recabar toda la información necesaria con la finalidad de probar la eficacia y seguridad de cada tratamiento.

En la figura 1.1, se muestra un esquema con las diferentes fases y tareas más relevantes de todo el ensayo clínico.

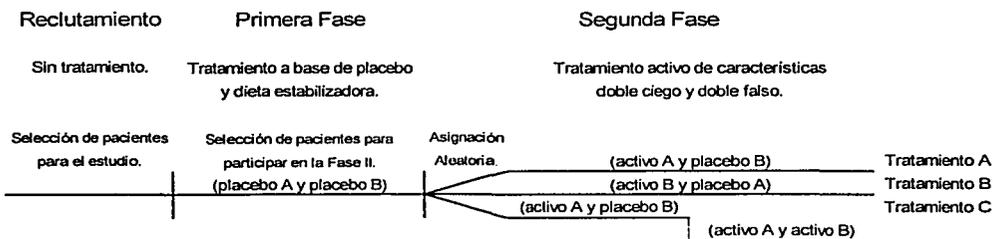


Figura 1.1 Fases y principales tareas dentro del ensayo clínico.

1.4.2.1 Criterios de selección de pacientes.

Los criterios primarios de inclusión aplicados fueron los siguientes:

- Hombres y mujeres, entre 18 y 75 años, cuyo diagnóstico haya sido hipercolesterolemia primaria.
- Podrían participar mujeres estériles.
- También podrían incluirse personas con un historial de enfermedad coronaria con 2 ó más factores de riesgo cardiovascular y niveles de colesterol LDL de, al menos, 3.35 mmol/L (130 mg/dL).

Entre los criterios de exclusión aplicados, encontramos los siguientes:

- Mujeres embarazadas o en periodo de lactancia.
- Pacientes con antecedentes de afecciones cardíacas durante los 6 meses anteriores al inicio del tratamiento.

⁵ Por placebo debe entenderse aquella sustancia que carece por si misma de acción terapéutica, pero puede inducir algún efecto curativo por sugestión.

-
- Personas con enfermedades o padecimientos crónicos (incluido el alcoholismo y el abuso de drogas), antecedentes de cáncer o antecedentes familiares de hipercolesterolemia.
 - Personas con antecedentes de desórdenes gastrointestinales.
 - Los que estuviesen sometidos a algún otro tratamiento contra la hipercolesterolemia o se encontrasen participando en algún otro ensayo en etapa de estudio.
 - Personas con hipersensibilidad a los inhibidores de la HMG-CoA o sus derivados.
 - Pacientes que hagan uso de algún medicamento contraindicado.

1.4.2.2 Tamaño de la población.

Para alcanzar el objetivo primario del ensayo clínico base, se estimó conveniente que el número total de sujetos requeridos para validar los resultados finales debería ser de alrededor de 378 personas (aproximadamente 126 participantes en cada uno de los tres grupos). Este tamaño de muestra se calculó con la finalidad de verificar igualdad entre los valores medios de colesterol de los pacientes de cada grupo después de concluido cada tratamiento, permitiendo una diferencia clínica máxima de hasta un 5%.

Por experiencias previas, se sabe que alrededor del 15% de los pacientes abandonan cualquier estudio de este tipo entre cada una de las visitas médicas de control. Este antecedente llevó a considerar que para esta investigación sería necesario contar con un mínimo de 444 pacientes al momento de iniciar la Segunda Etapa, garantizando con ello que cada grupo de tratamiento estaría integrado al menos por 148 personas. Para asegurar el número de personas requerido al momento de efectuarse la quinta visita médica, se decidió reclutar 665 individuos en total.

Como comentario adicional, vale la pena aclarar que para cualquier ensayo clínico la definición del tamaño de la población es muy importante, especialmente por razones de costo (en términos económicos y de tiempo). Este argumento debe motivar al investigador a determinar un tamaño de muestra adecuado que, además de considerar este importante factor, permita la realización de inferencias estadísticas significativas que garanticen resultados confiables para los fines del estudio. En virtud de lo anterior, debe señalarse que dentro de este trabajo se asumirá que el tamaño de población que se determinó para el estudio de eficacia farmacológica, garantiza la obtención de resultados confiables en la modelación de los tiempos de supervivencia para la duración y manifestación de los eventos adversos emergentes que se registraron.

1.4.3 Desarrollo de la investigación clínica.

Como ya ha sido explicado, el ensayo se dividió en tres partes.

La Fase de Reclutamiento dio inicio con la primera visita médica (visita de selección). Aquí se decidió si un paciente es apto para ser incluido en el estudio, aplicando en cada caso los criterios de selección establecidos. Sólo aquéllos que

hubiesen cumplido con esos criterios continuarían a la siguiente etapa. Cabe señalar que dentro de esta fase se procuró eliminar todo efecto previo de tratamiento hiperlipídico. Para lograrlo, se pidió a cada sujeto llevar una dieta especial además de descontinuar algún otro tratamiento que se hubiese comenzado previamente.

La Primera Fase estuvo caracterizada por su propósito de limpiar cualquier rastro de medicamentos o tratamientos anteriores. La administración de medicamentos placebo contribuiría a alcanzar este objetivo: cada persona ingirió una tableta placebo del medicamento A y B todas las noches durante 6 semanas.

En lo que respecta a las visitas médicas, estas tendrían como finalidad recoger información sobre la incidencia de eventos adversos, realizar pruebas de laboratorio y gabinete, y llevar registros precisos de control de algunos signos vitales.

Al inicio de la Segunda Fase, durante la quinta visita médica, se seleccionó a los pacientes que finalmente participarían en cada uno de los tratamientos. Los criterios más relevantes que se aplicaron fueron los siguientes:

- Pacientes cuyo nivel promedio de colesterol LDL obtenido de las mediciones recabadas durante la tercera y cuarta visita no hubiesen diferido en más de 12% respecto a 4.12 mmol/L (160 mg/dL). Al menos una de las medidas debería ser menor a 4.12 mmol/L o menor a 3.35 mmol/L (130 mg/dL).
- Pacientes con niveles de TG menores o iguales a 3.97 mmol/L (350 mg/dL) registrados también durante la tercera y cuarta visitas.
- Pacientes que hubiesen completado el tratamiento placebo administrado durante la Primera Fase en al menos un 80%.

Posterior a la selección, se llevó a cabo la asignación de pacientes en cada tratamiento de forma aleatoria.

La Segunda Etapa duró 16 semanas y se caracterizó por ser de tratamiento doble ciego y doble falso, teniendo como objetivo evaluar la eficacia y seguridad de cada tratamiento. Los pacientes asignados, divididos en tres grupos, fueron dotados con los siguientes medicamentos:

i) Grupo I

Se prescribió:

- una tableta placebo identificada con el medicamento B, y
- una tableta del medicamento A.

Ambas serían tomadas cada noche durante 16 semanas.

ii) Grupo II

Se prescribió:

- una tableta del medicamento B, y
- una tableta placebo identificada con el medicamento A.

Ambas se tomarían cada noche durante 16 semanas.

iii) Grupo III

Se prescribió:

-
- a)
- una tableta placebo identificada con el medicamento B, y
 - una tableta del medicamento A.

Ambas se tomarían cada noche durante las primeras 8 semanas;

- b)
- una tableta del medicamento B, y
 - una tableta del medicamento A.

Estas se tomarían durante las últimas 8 semanas.

Durante esta fase del estudio, se llevó un registro minucioso de todos los eventos adversos reportados.

El desarrollo completo de la investigación así como el programa de actividades en cada una de las visitas médicas se encuentra detallado dentro del ANEXO 2.

1.4.4 Registro y documentación de los eventos adversos dentro del ensayo clínico.

Los eventos adversos reportados durante el tiempo que duró el ensayo clínico fueron cuidadosamente registrados dentro de la Forma para Registro de Eventos Adversos (ver ANEXO 3) elaborada con esa finalidad.

Las anomalías reportadas en pruebas de laboratorio, consideradas como relevantes por el médico tratante, también se consideraron como eventos adversos.

La información recabada para cada uno de los eventos que se reportaron se muestra a continuación:

- 1) Fecha de inicio y término del evento.
- 2) Descripción del evento.
- 3) Duración del evento.
- 4) Severidad.
- 5) Acciones tomadas.
- 6) Resultado de dichas acciones.
- 7) Relación con el tratamiento.

Si el evento adverso fue calificado como serio, sería registrado en un documento especial, además de ser tomado en cuenta para su evaluación junto con los otros eventos.

1.4.5 Criterios de evaluación de resultados del ensayo clínico.

El objetivo de eficacia se evaluó comparando la variación porcentual entre los niveles medios de colesterol LDL, al principio y al final del tratamiento, para cada grupo. En este sentido, se planteó una prueba de hipótesis estadística para verificar la existencia de igualdad entre las variaciones medias del resultado de cada uno de los tres tratamientos.

El criterio de evaluación de seguridad se hizo comparando la incidencia de eventos adversos entre los tres grupos. Para facilitar esta comparación, se construyó una tabla donde se clasificó a cada evento por su tipo, frecuencia y sistema corporal al que afecta.

1.5 El análisis de supervivencia como criterio de evaluación.

Una de las formas más sencillas para clasificar las enfermedades es dividir las según la forma en que éstas se manifiestan; de este modo, se tienen enfermedades agudas y crónicas. Las primeras tienen un periodo breve de evolución y una alta intensidad; pueden ser sanadas, pero existe la posibilidad de volver a ser contraídas. Las segundas tienen un periodo largo y lento de evolución acompañado de una intensidad creciente, por lo cual también se les conoce como degenerativas. Estas enfermedades pueden ser controladas y las posibilidades de sanarlas disminuirán conforme su intensidad aumenta.

Dadas estas características, la exposición a los fármacos utilizados en el tratamiento de ambos tipos de dolencias es diferente. Para las enfermedades crónicas, la presencia de la sustancia activa es constante ya que su finalidad principal es controlar el padecimiento. En este sentido, algunos de los aspectos que deben ser mayormente cuidados dentro de la investigación y desarrollo de los medicamentos, particularmente en aquéllos de exposición continua, es la incidencia y, sobre todo, la duración de los eventos adversos (se busca que un gran número de pacientes tratados no manifiesten ningún evento adverso, y que la mayoría de los eventos que llegasen a reportarse sean de corta duración y tarden en aparecer).

En virtud de que la variable de interés en el estudio de los eventos adversos es el tiempo, el procedimiento estadístico adecuado para evaluar los resultados que se obtengan de su estudio es el análisis de supervivencia.

En las próximas páginas, se abordará el análisis de supervivencia como herramienta para estimar y ajustar un modelo matemático que describa el comportamiento del tiempo de duración y el tiempo anterior a la primera manifestación de los eventos adversos emergentes observados dentro del ensayo clínico detallado en las secciones precedentes.

CAPÍTULO II : ANÁLISIS DE SUPERVIVENCIA.

2.1 Introducción.

Este capítulo trata sobre los conceptos fundamentales del análisis de los fenómenos de supervivencia así como de aquellos procedimientos para la estimación y validación de un modelo matemático sobre un conjunto de datos. En primer lugar se explicará el término supervivencia y se definirán algunos de sus elementos para luego hablar de la función de supervivencia y los tipos de censura que serán tratados. Más adelante se abordará la estimación de un modelo por medio de la implementación del proceso no paramétrico de Kaplan–Meier y los modelos paramétricos Weibull y Exponencial, no sin antes presentar su forma de construcción o estimación de parámetros según corresponda. Finalmente, se hará referencia una de las formas para probar igualdad entre las experiencias de supervivencia de varias muestras aplicando la estadística de Mantel–Haeszel.

2.2 Modelo probabilístico de supervivencia.

Por supervivencia se entenderá el análisis de datos derivados de una variable aleatoria asociada con el tiempo. El objetivo fundamental de dicho análisis es describir la forma en que un determinado fenómeno impacta a una población a partir del estudio en que éste tarde en manifestarse para cada uno de los miembros que la integran; pudiendo también realizar esta caracterización sobre los elementos de una muestra aleatoria tomados de dicha población. En este sentido, el tiempo que transcurre hasta el momento en que el fenómeno se presenta, constituye el factor de mayor relevancia. Este tiempo se encuentra delimitado por dos eventos bien definidos, uno inicial y uno final. El evento de inicio, asociado con el momento de origen, se vincula con la acción de reclutamiento de las entidades o individuos que formarán parte de la muestra de acuerdo con criterios de selección determinados de antemano. El evento terminal, relacionado con el momento final, marca el instante en que el fenómeno de interés se manifiesta a través de la ocurrencia del evento que le es característico y es así como este evento indica el término del periodo de observación. Cabe señalar que cuando una entidad experimenta el evento final, el dato que se genera recibe el nombre de falla.

Por otro lado, debe entenderse como periodo de observación al lapso que transcurre entre el momento inicial y el final. En este periodo, cada entidad es sometida a diversos procesos cuya finalidad es registrar información adicional que contribuya a realizar una mejor descripción del fenómeno.

Una vez definidos estos elementos, se tiene una mejor comprensión de porqué el periodo de observación sigue un comportamiento aleatorio, ya que, aunque es posible conocer el momento en que da inicio, generalmente no es posible conocer el momento cronológico en que concluirá sin antes haberse verificado el evento final. Asimismo, tampoco es posible suponer que existe, de inicio, un modelo que explique la duración de los tiempos de observación, ya que aunque se parte del

supuesto de que todos los elementos que integran la muestra son similares no se tiene conocimiento a priori que algún factor inhiba o acelere la ocurrencia del evento final.

Cabe señalar que todos los elementos que intervienen en el análisis de supervivencia se describen a través del tiempo que antecede a la ocurrencia del evento final, aunque también es posible explicar la forma en que el fenómeno repercute en la población tomando como base las posibilidades que tenga cualquier entidad muestral de superar un determinado instante sin haber experimentado el evento final característico.

Son diversas las razones por las que resulta efectivo aplicar el análisis de supervivencia al estudio de ciertos fenómenos, algunas de ellas son:

- La principal variable que describe estos fenómenos es el tiempo de duración el cual siempre asume valores positivos.
- Estos fenómenos pueden ser observados incluso durante extensos periodos, pudiendo coexistir, dentro de una misma población, datos sobre tiempos de observación breves y largos.
- No toda la información de este tipo de fenómenos se refiere al tiempo preciso de duración (desde el momento exacto de inicio hasta el momento exacto de término definido por el evento final). Existen otros datos, conocidos como censuras, que provienen de entidades para las cuales existió un momento inicial y cuyo momento final se determina por sucesos diferentes al evento característico final.

Para efectos de este trabajo, se entenderá como fenómeno de supervivencia toda aquella manifestación que se apegue a los planteamientos expuestos en esta sección.

2.2.1 Conceptualización del tiempo de observación dentro de un estudio de supervivencia.

Cuando se plantea un estudio estadístico, generalmente su duración es limitada. En otros casos, por características propias del fenómeno, el tiempo de observación de algunos de sus integrantes es largo, entonces debe adecuarse el concepto de tiempo de observación al contexto del tiempo de duración de la investigación. De lo anterior, se entiende como tiempo de observación al tiempo calendario que permanece un determinado individuo o entidad de la muestra dentro de un lapso preestablecido para llevar a cabo las observaciones del fenómeno de supervivencia. El tiempo de supervivencia de cada miembro de la muestra equivaldrá al periodo en que cada uno de ellos permanece dentro del estudio medido a partir del momento de su inclusión hasta el momento en que se realice su último registro; ya sea porque experimentó la ocurrencia del evento definido como final (falla) o porque la entidad se censuró por alguna causa.

Una vez expuestas estas premisas, se procederá a detallar los puntos que fundamentan la teoría de supervivencia.

2.2.2 Función de supervivencia.

Al hablar de un modelo de supervivencia se hará referencia a la existencia de una determinada función de probabilidad asociada con una variable aleatoria T o tiempo de supervivencia, la cual por naturaleza propia tomará valores no negativos. Esta función cuantifica, a partir de un momento origen $t=0$, la probabilidad de que a un determinado tiempo t una cierta entidad no haya experimentado el evento que defina el momento final de observación; en otras palabras, la probabilidad de que la ocurrencia de dicho evento sea posterior a t . Se define entonces a $S(t)$ como la función de supervivencia:

$$S(t) = P[T > t]$$

Dado que $T \geq 0$ entonces:

- $S(0) = P[T > 0] = 1$
- $S(t)$ es no creciente, ya que en tanto $t \rightarrow \infty$ la probabilidad de ocurrencia del evento sobre la entidad en estudio aumenta, y la no ocurrencia del mismo disminuye.
- $\lim_{t \rightarrow \infty} S(t) = 0$

2.2.3 Función de distribución.

Se define $F(t)$ o función de distribución acumulada de T , como la probabilidad de que la variable aleatoria T tome valores menores o iguales a t . Esta función cuantificará la probabilidad de que el evento ocurra antes o al menos al tiempo t . Sea $F(t)$ función de distribución, entonces:

$$F(t) = P[T \leq t]$$

En donde:

- $F(t) = 1 - S(t)$
- $F(0) = 0$
- $\lim_{t \rightarrow \infty} F(t) = 1$

2.2.4 Función de densidad.

Se define $f(t)$ como la función de densidad de probabilidad de la variable aleatoria T .

La función de densidad, a diferencia de las funciones de distribución y supervivencia, reporta una medida instantánea de probabilidad sobre la ocurrencia del evento al tiempo t para la entidad en observación.

Lo anterior puede entenderse como la razón de cambio del valor de la función de distribución en el intervalo $t \leq T \leq t + \delta$ cuando $\delta \rightarrow 0$, dado que $T \geq 0$. Esto puede expresarse de la siguiente forma:

$$f(t) = \lim_{\delta \rightarrow 0} \frac{F(t + \delta) - F(t)}{\delta}$$

o bien:

$$f(t) = \frac{d}{dt} F(t)$$

De aquí es posible concluir que:

$$f(t) = -\frac{d}{dt} S(t)$$

2.2.5 Función de riesgo.

La función de riesgo $h(t)$ reporta la tasa instantánea de variación de la probabilidad de ocurrencia del evento final (probabilidad de falla) condicionada a que la entidad en estudio está presente al tiempo $T=t$. En otras palabras, es el valor límite del cociente de la probabilidad de falla en el intervalo $t \leq T \leq t + \delta$, condicionada a que $T \geq t$, entre el tamaño de dicho intervalo, cuando $\delta \rightarrow 0$.

$$h(t) = \lim_{\delta \rightarrow 0} \frac{P[t \leq T \leq t + \delta | T \geq t]}{\delta}$$

$$h(t) = \lim_{\delta \rightarrow 0} \frac{\frac{P[t \leq T \leq t + \delta]}{P[T \geq t]}}{\delta}$$

$$h(t) = \frac{1}{S(t)} \lim_{\delta \rightarrow 0} \frac{P[t \leq T \leq t + \delta]}{\delta}$$

$$h(t) = \frac{1}{S(t)} \lim_{\delta \rightarrow 0} \frac{F(t + \delta) - F(t)}{\delta}$$

$$h(t) = \frac{1}{S(t)} \frac{d}{dt} F(t)$$

$$h(t) = \frac{f(t)}{S(t)}$$

La forma más usual de expresar la función de supervivencia es a partir de la función de riesgo:

$$h(t) = \frac{f(t)}{S(t)}$$

$$h(t) = -\frac{1}{S(t)} \frac{d}{dt} S(t)$$

$$h(t) = -\frac{d}{dt} \ln S(t)$$

De aquí se concluye que:

$$\int h(t) dt = -\ln S(t)$$

$$S(t) = e^{-\int_0^t h(u) du}$$

Cabe señalar que la trascendencia que tiene la función de riesgo no radica en el hecho de que exista una estrecha relación entre ésta y la función de supervivencia, sino en que esta función describe la forma como incide el fenómeno en estudio con el paso del tiempo; en otras palabras, recoge la intensidad de que el evento final ocurra en cualquier momento posterior al origen para cualquiera de las entidades integrantes de la población. Lo anterior tiene como consecuencia que muchos de los modelos paramétricos utilizados para explicar el desarrollo de la supervivencia en una población, se basen en supuestos cimentados alrededor del comportamiento de $h(t)$.

2.2.6 Censura.

Una de las características más importantes de la información originada de un fenómeno de supervivencia es la presencia de datos censurados. Un dato censurado contiene información parcial de la variable aleatoria en virtud de que, por alguna razón, no fue posible registrar la ocurrencia del evento final para ciertas entidades en observación.

En términos prácticos, la censura se clasifica según la forma en que se trunca el desarrollo de la variable aleatoria así como a la movilidad o rigidez de los límites del mismo estudio. Por ejemplo, cuando se establece un periodo de seguimiento dentro del cual el inicio de las observaciones puede no ser el mismo para todas las entidades, pero el término estará definido por un tiempo máximo preestablecido, se tiene una clase de censura conocida como de tipo I.

Censura de tipo I:

Sea t_c un valor fijo; para una muestra t_1, \dots, t_n de la variable aleatoria T . Se define:

$$Y_i = \begin{cases} t_i & \text{si } t_i \leq t_c \\ t_c & \text{si } t_i > t_c \end{cases}$$

En este caso, el tiempo de ocurrencia del evento final es registrado exactamente en tanto se cumpla $t_i \leq t_c$. Gráficamente, la censura de tipo I puede apreciarse en la figura 2.1.

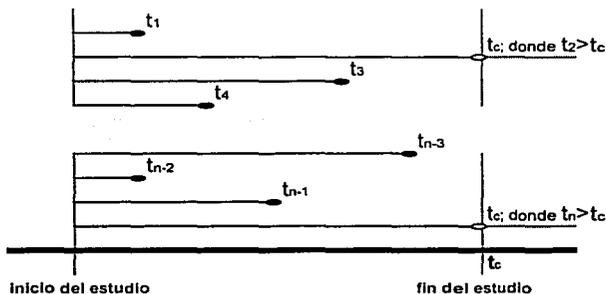


Figura 2.1 Censura de tipo I.

Si en lugar de fijar el momento de término del estudio para todas las entidades, se fija un número máximo de fallas, nos encontramos ante la censura de tipo II.

Censura de tipo II:

Sea r , un valor fijo el cual representa el número máximo de fallas que pueden permitirse dentro de una muestra de tamaño n de realizaciones de la variable aleatoria T , de tal manera que $r \leq n$. Sea $t_{(1)}, \dots, t_{(n)}$, las n realizaciones ordenadas de la variable aleatoria T , donde se cumple $t_{(1)} \leq \dots \leq t_{(n)}$. El seguimiento que se hace a las realizaciones de la variable cesará en el momento que se verifique la r -ésima falla; de tal manera que las $(n-r)$ observaciones restantes aportarán $t_{(r)}$ a la experiencia de supervivencia de la muestra. Gráficamente, este tipo de censura puede apreciarse en la figura 2.2.

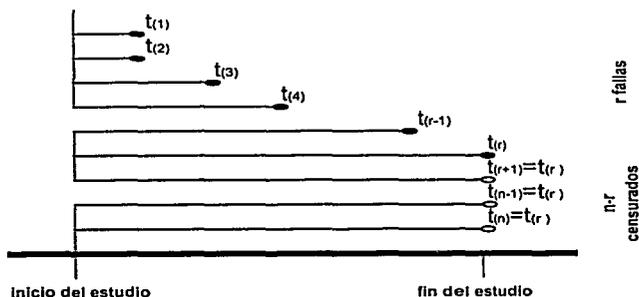


Figura 2.2 Censura de tipo II.

Por otro lado, si el proceso de censura no puede ser controlado a priori (como en los casos de censura de tipo I y II), tenemos otro caso de censura conocido como censura aleatoria.

Censura aleatoria:

Sean c_1, \dots, c_n un conjunto de variables aleatorias idénticamente distribuidas con función de densidad G ; c_i es el tiempo de censura asociado al i -ésimo individuo t_i . Las observaciones registradas para el fenómeno de supervivencia pueden ser expresadas como $(y_1, \delta_1), \dots, (y_n, \delta_n)$ donde:

$$y_i = \min(t_i, c_i)$$

$$\delta_i = \begin{cases} 1 & \text{si } t_i \leq c_i \text{ (} t_i \text{ no es censurado)} \\ 0 & \text{si } t_i > c_i \text{ (} t_i \text{ es censurado)} \end{cases}$$

Las formas de censura dentro de esta clasificación tienen uno de dos orígenes posibles:

- 1) Cuando no hay forma de continuar con el seguimiento de la variable aleatoria por alguna razón; por ejemplo, si se pierde la secuencia de la observación. En este caso se registra el último momento para el cual se tuvo algún dato cierto sobre la misma.
- 2) Cuando se alcanza el término del estudio sin que se hubiese verificado el evento final. Aquí se registra como dato de término el momento de fin del estudio.

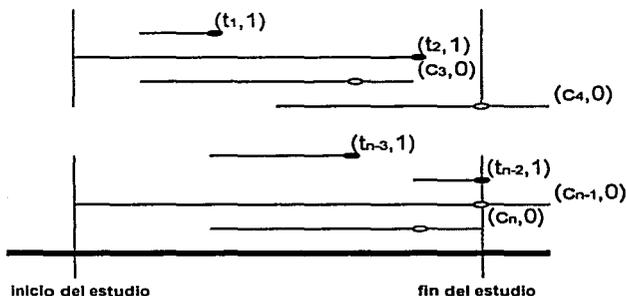


Figura 2.3 Censura aleatoria.

En la figura 2.3 se ilustra este tipo de censura.

Cabe señalar que en un fenómeno de supervivencia con censura aleatoria es usual asumir independencia entre las fallas y las censuras. Esto puede justificarse ya que durante el proceso de observación se tienen ingresos y salidas en forma aleatoria. Sin embargo, este supuesto no siempre se cumple.

Aunque existen otras clases de censura, los tres tipos básicos aquí tratados son los que se encuentran con mayor frecuencia en una gran variedad de fenómenos de supervivencia. Finalmente se dirá, que a estas tres clases de censura también se le conoce como censura por la derecha (se sabe donde inicia la observación pero no siempre se le puede dar un seguimiento completo a su término).

2.3 Estimación de la función de supervivencia.

2.3.1 Procedimientos no paramétricos.

Para poder explicar un fenómeno de supervivencia, puede construirse un modelo matemático que sintetice, de forma adecuada, las características de la variable aleatoria a la que se está dando seguimiento. Una manera de alcanzar este objetivo, es estimar la función de supervivencia a partir de la información muestral disponible.

La construcción de un modelo de supervivencia utilizando procedimientos no paramétricos se basa en que no es necesario plantear ningún supuesto sobre el comportamiento de los tiempos de supervivencia; esto significa que en principio no se establecerá modelo paramétrico alguno sobre la forma de la función de supervivencia, la función de distribución o la función de riesgo.

2.3.1.1 Tabla de vida.

Una primera aproximación de un estimador de $S(t)$ para una muestra aleatoria obtenida de un fenómeno de supervivencia es la tabla de vida. Con la finalidad de conocer esta técnica, a continuación se muestra su construcción intuitiva ante la presencia de censura de cualquier tipo. Se parte de dividir el tiempo de observación en $(n+1)$ intervalos $I_j = [t_{j-1}, t_j)$ donde $j=1, \dots, n+1$, y donde se cumple $t_0 < t_1 < \dots < t_{k-1} < t_k < t_{k+1} < \dots < t_n < t_{n+1}$, tal y como se muestra en la figura 2.4. Cabe señalar que el tamaño de los intervalos no necesariamente debe ser el mismo.

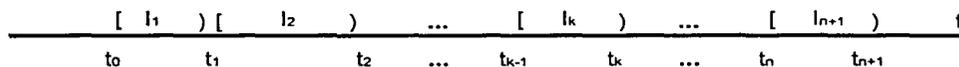


Figura 2.4 Intervalos integrantes de una partición sobre el tiempo de observación de un fenómeno de supervivencia.

Cada individuo de la muestra tiene asociado un tiempo de falla o de censura, de tal forma que solo se conoce el intervalo donde ocurrió la falla o la censura; en consecuencia, los datos básicos para la construcción del estimador son el número de fallas y censuras por intervalo. Ahora se supondrá que en el j -ésimo intervalo, I_j , existen d_j fallas y c_j censuras respectivamente, mientras que al inicio del mismo se encuentran presentes n_j individuos. De lo anterior puede definirse:

$$n_1 = n$$
$$n_j = n_{j-1} - d_j - c_j \quad \text{para } j=2, \dots, n+1$$

Nuestro interés recae en determinar la probabilidad de que un individuo sobreviva más allá del inicio de un intervalo I_k , lo que equivale a encontrar $S(t_k)$. Aprovechando el resultado $P[T > t_j | T > t_{j-1}]$, que es la probabilidad de sobrevivir a I_j dado que se sobrevivió a I_{j-1} , puede estimarse $S(t_k)$ como el producto de las

probabilidades condicionales de sobrevivir a cada intervalo anterior a I_k en forma individual:

$$S(t_k) = P[T > t_k] = P[T > t_1]P[T > t_2 | T > t_1] \dots P[T > t_k | T > t_{k-1}]$$

Para hallar el valor de $P[T > t_i | T > t_{i-1}]$, es necesario plantear algún supuesto sobre la forma en como se lleva a cabo el proceso de censura en cada intervalo. Por ejemplo, es común asumir que todas las censuras ocurren de manera uniforme a lo largo de I_j , lo que implica que cada individuo censurado permanece en riesgo, al menos, la mitad del intervalo. Entonces, el número promedio de entidades en riesgo en I_j es:

$$N_j = n_j - \frac{c_j}{2}$$

Bajo este supuesto, la probabilidad de falla en I_j dado que se sobrevivió a I_{j-1} es $\frac{d_j}{N_j}$; y en consecuencia, la probabilidad de sobrevivir a I_j dado que se sobrevivió a I_{j-1} es:

$$P[T > t_i | T > t_{i-1}] = 1 - \frac{d_j}{N_j} = \frac{N_j - d_j}{N_j}$$

Por lo tanto, el estimador de $S(t_k)$ está dado por la siguiente expresión:

$$\hat{S}(t_k) = \prod_{j=1}^k P[T > t_j | T > t_{j-1}] = \prod_{j=1}^k \left(\frac{N_j - d_j}{N_j} \right); \text{ para } k=1, \dots, n+1$$

Para presentar los componentes del cálculo y los resultados de este estimador, se utiliza un formato tabular el cual se muestra en la tabla 2.1.

Tabla 2.1 Formato utilizado para presentar los resultados de la tabla de vida. El ejemplo que se ilustra es válido para el k -ésimo intervalo.

Intervalo	Período de tiempo	d_j	c_j	n_j	N_j	$\frac{N_j - d_j}{N_j}$	$\hat{S}(t)$
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
I_k	$t_{k-1} - t_k$	d_k	c_k	n_k	$N_k = n_k - \frac{c_k}{2}$	$\frac{N_k - d_k}{N_k}$	$\prod_{j=1}^k \left(\frac{N_j - d_j}{N_j} \right)$
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

2.3.1.2 Función de supervivencia empírica.

Se define la función de distribución empírica sobre una muestra aleatoria no censurada, como:

$$\hat{F}(t) = \frac{\#t_i \leq t}{n}; \text{ para } t_1, \dots, t_n$$

donde $\#t_i$ es el número de entidades expuestas.

De la relación $F(t) = 1 - S(t)$, es posible hallar una expresión para la función de supervivencia empírica:

$$\begin{aligned}\hat{S}(t) &= 1 - \hat{F}(t) \\ \hat{S}(t) &= 1 - \left[\frac{\#t_i \leq t}{n} \right] \\ \hat{S}(t) &= \left[\frac{\#t_i > t}{n} \right]\end{aligned}$$

La finalidad de esta estimación es determinar los valores de $S(t)$ a partir del cálculo de los cocientes del número de entidades que no han experimentado el evento al tiempo t entre el número total de entidades expuestas.

La función de supervivencia, construida de esta forma, se caracteriza porque asume el valor 1 hasta antes de que ocurra la primera falla, y tomará el valor 0 después de que la última entidad experimente el evento en estudio. Lo anterior

lleva a asumir que \hat{S} es constante entre dos valores adyacentes de t donde se registre la ocurrencia de un evento además de tener una forma escalonada. Estas características se ilustran en la figura 2.5.

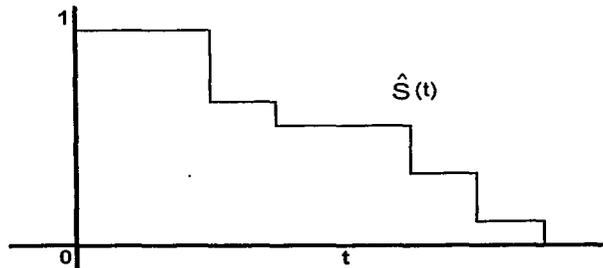


Figura 2.5 Función de supervivencia empírica.

Sin embargo, cuando la muestra tiene datos censurados, un mejor estimador lo proporciona el modelo de Kaplan–Meier.

2.3.1.3 Estimador de Kaplan–Meier.

El estimador no paramétrico más utilizado para la función de supervivencia, es el conocido con el nombre de Kaplan–Meier (K-M).

Este estimador, al igual que la tabla de vida, es el resultado del producto de las probabilidades individuales de sobrevivir después de un tiempo t_k dado que se sobrevivió inmediatamente antes de t_k . Su construcción también se basa en la implementación de intervalos, no necesariamente de igual tamaño, sobre el tiempo asociado a la información proveniente de un fenómeno de supervivencia. Como observación, debe señalarse que este estimador puede verse como el caso límite del proceso de construcción de la tabla de vida, cuando se propone un número de intervalos tendiente a infinito y cuyo tamaño tiende a cero.

Para determinar el estimador K-M, se recurrirá al proceso de estimación por máxima verosimilitud. Para ello, primero se propondrá la función de verosimilitud L , para luego encontrar el valor que la maximice.

Se parte de dividir el tiempo de observación en k -tiempos de falla, quienes cumplen $t_1 < \dots < t_k$; y donde al tiempo t_j se presentan d_j fallas y hay n_j entidades en riesgo. En el intervalo $I_j = [t_{j-1}, t_j)$, donde $j=1, \dots, k+1$, suponemos que se presentan c_j censuras, cuyo tiempo de censura es L_i^j , donde $i=1, \dots, c_j$; y además $t_0=0$ y $t_{k+1} \rightarrow \infty$. De lo dicho hasta este momento, es posible deducir la siguiente relación:

$$n_j = (d_j + c_j) + \dots + (d_k + c_k)$$

Ahora, se conviene ajustar los tiempos de censura infinitesimalmente a la derecha de tal manera que las censuras en t_j ocurren después de t_j .

Si la función de supervivencia para los tiempos de falla es $S(t)$, entonces la probabilidad de que una entidad falle en t_j puede escribirse como sigue:

$$S(t_j) - S(t_j + 0), \text{ donde } S(t_j + 0) = \lim_{x \rightarrow 0^+} S(t_j + x) \text{ para } j=1, \dots, k$$

La función de verosimilitud que se propone quedará expresada de la siguiente forma:

$$L = \prod_{j=1}^k \left[\left(\prod_{i=1}^{c_j} S(L_i^j) \right) [S(t_j) - S(t_j + 0)]^{d_j} \right] \prod_{i=1}^{c_{k+1}} S(L_i^{k+1})$$

donde cada componente tiene el siguiente significado:

$[S(t_j) - S(t_j + 0)]^{d_j}$ es la contribución de las fallas por intervalo,

$\prod_{i=1}^{c_j} S(L_i^j)$ es la contribución de las entidades que se censuraron durante el tiempo de observación; y,

$\prod_{i=1}^{c_{k+1}} S(L_i^{k+1})$ es la contribución de aquellos individuos que llegaron al final del tiempo de observación, y en consecuencia, se censuraron. Con la finalidad de maximizar L con respecto a $S(t)$, se harán las siguientes observaciones:

- $\hat{S}(t)$ debe ser discontinua en las t_{j_s} ya que de lo contrario $S(t_j)=S(t_j+0)$, y en consecuencia $L=0$.
- Por otro lado, al ser $\hat{S}(t)$ una función no creciente que cumple las características definidas en la sección 2.2.2 del presente capítulo, se tiene:

$$\begin{aligned}\hat{S}(t_1) &= \hat{S}(L^1) = 1, \text{ para } i=1, \dots, c_1 \\ \hat{S}(L^{j+1}) &= \hat{S}(t_j+0) = \hat{S}(t_{j+1}), \text{ para } j=1, \dots, k; i=1, \dots, c_{j+1}\end{aligned}$$

Entonces, la función de verosimilitud puede reescribirse de la siguiente forma:

$$L = \prod_{j=1}^k [S(t_j) - S(t_j+0)]^{d_j} S(t_j)^{c_j} S(t_{k+1})^{c_{k+1}}$$

Si $S(t_j+0) = P_j$, para $j=1, \dots, k$, y $P_0=1$, la función de verosimilitud que tendría que maximizarse es:

$$L_1 = \prod_{j=1}^k [P_{j-1} - P_j]^{d_j} P_j^{c_{j+1}}$$

con respecto a P_1, \dots, P_k .

Si ahora se define: $p_j = \frac{P_j}{P_{j-1}}$ y $q_j = 1 - p_j$, para $j=1, \dots, k$; la función L_1 puede reescribirse nuevamente para quedar como sigue:

$$\begin{aligned}L_1 &= \prod_{j=1}^k [p_1 p_2 \dots p_{j-1} - p_1 p_2 \dots p_j]^{d_j} (p_1 p_2 \dots p_j)^{c_{j+1}} \\ L_1 &= \prod_{j=1}^k [p_1 p_2 \dots p_{j-1} q_j]^{d_j} (p_1 p_2 \dots p_j)^{c_{j+1}}\end{aligned}$$

En este punto, debe observarse que si la última entidad es censurada, con tiempo de censura L^* , implicaría que $c_{k+1} > 0$ con lo que se tendría $\hat{S}(L^*) = \hat{S}(t_k+0)$, y en consecuencia $\hat{S}(t)$ queda indeterminado para $t > L^*$. En virtud de lo anterior, se considerará que $c_{k+1} = 0$.

Tomando en cuenta esta observación, L_1 puede expresarse como:

$$L_1 = \prod_{j=1}^k [p_1 p_2 \dots p_{j-1} q_j]^{d_j} (p_1 p_2 \dots p_j)^{c_j}$$

Con la finalidad de hallar una expresión más sencilla para L_1 , se aprovechará el resultado $n_j=(d_j+c_j)+\dots+(d_k+c_k)$. Después realizar una simplificación algebraica, se obtiene lo siguiente:

$$L_1 = \prod_{j=1}^k c_j^{d_j} p^{n_j-d_j}$$

Para maximizar L_1 con respecto a p_j fijo, se resuelve la ecuación resultante de igualar a cero la derivada del logaritmo natural de L_1 :

$$\frac{d\ln(L_1)}{dp} = 0$$

$$\hat{p}_j = \frac{n_j - d_j}{n_j}$$

Finalmente, para maximizar L con respecto a $S(t)$, se toma el resultado $\hat{S}(t_j+0) = \hat{S}(t_{j-1}+0) \left(\frac{n_j - d_j}{n_j}\right)$, con lo cual se obtiene el estimador $\hat{S}(t)$ que se venía buscando:

$$\hat{S}(t) = \prod_{j: t_j < t} \left(\frac{n_j - d_j}{n_j} \right)$$

Para concluir esta sección, se mencionará que el estimador de Kaplan–Meier tiene dos propiedades estadísticas importantes: es un estimador insesgado y consistente de $S(t)$. Para profundizar en estos temas, se sugiere consultar la bibliografía citada (Kalbfleisch, 1980).

2.3.2 Procedimientos paramétricos.

A diferencia de la estimación no paramétrica, en la estimación paramétrica se asumirá que los tiempos de supervivencia de la muestra manifiestan un comportamiento que puede ser explicado por medio de una función de densidad conocida con parámetros $\theta_1, \dots, \theta_n$.

En este caso, la tarea principal consistirá en determinar qué función se ajusta de mejor manera al fenómeno en estudio y, con base en la información disponible, estimar el valor de sus parámetros.

Como observación, debe señalarse que un modelo paramétrico puede ajustarse a partir de la evidencia que proporciona un ajuste no paramétrico previo.

En las secciones siguientes se tratarán uno de los modelos paramétrico más comunes (Weibull) para describir fenómenos de supervivencia relacionados con procesos biológicos, así como una de sus variantes considerada la más relevante (Exponencial). Para abordarlos, se plantearán algunos supuestos sobre el comportamiento de su función de riesgo.

2.3.2.1 Modelo Weibull.

La distribución Weibull, tuvo su origen de la observación y estudio de información de fenómenos enfocados a medir el tiempo aleatorio de falla de diversos materiales, componentes mecánicos y eléctricos cuando éstos eran sometidos a cierto esfuerzo.

El modelo Weibull se sustenta en expresar $h(t)$ como sigue:

$$h(t) = \lambda \gamma t^{\gamma-1}$$
$$t \geq 0$$
$$\lambda > 0$$
$$\gamma > 0$$

Este supuesto implica que la función de riesgo es monótona y puede establecerse una relación creciente o decreciente entre el tiempo y $h(t)$ según el valor que asuma el parámetro γ :

- Si $\gamma > 1$ se tiene una función creciente.
- Si $\gamma = 1$ se tiene una función constante en λ .
- Si $\gamma < 1$ se tiene una función decreciente.

Gráficamente, algunos ejemplos de la función de riesgo de este modelo, para diversos valores de λ y γ , se muestran en la figura 2.6.

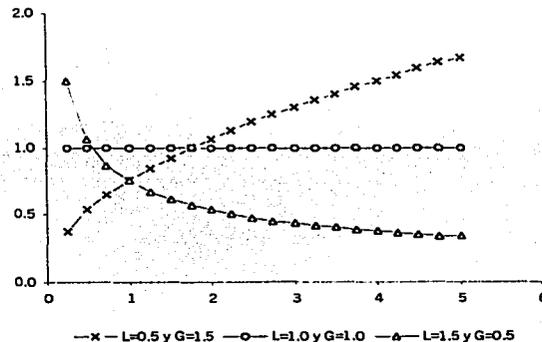


Figura 2.6 Función de riesgo del modelo Weibull para algunos valores de λ y γ .

La función de supervivencia está dada por:

$$S(t) = e^{-\int_0^t \lambda \gamma u^{\gamma-1} du}$$

$$S(t) = e^{-\lambda t^\gamma}$$

$$t \geq 0$$

En consecuencia, la función de densidad será:

$$f(t) = \lambda \gamma t^{\gamma-1} e^{-\lambda t^\gamma}$$

$$t \geq 0$$

cuya variable T se distribuye según el modelo Weibull con parámetros $\lambda > 0$ y $\gamma > 0$.

Como nota final de esta sección, se dirá que la distribución Weibull es de uso extendido en la modelación de procesos biológicos ya que, por lo general, describe adecuadamente el comportamiento de este tipo de fenómenos.

2.3.2.2 Modelo Exponencial.

El caso en el que $\gamma = 1$ para el modelo Weibull, merece especial atención. Este hecho implica asumir que los valores de la función de riesgo son constantes en cualquier momento, lo cual equivale a expresar:

$$h(t) = \lambda$$

$$t \geq 0$$

$$\lambda > 0$$

La forma gráfica de esta función, para algunos valores de λ , se muestran en la figura 2.7.

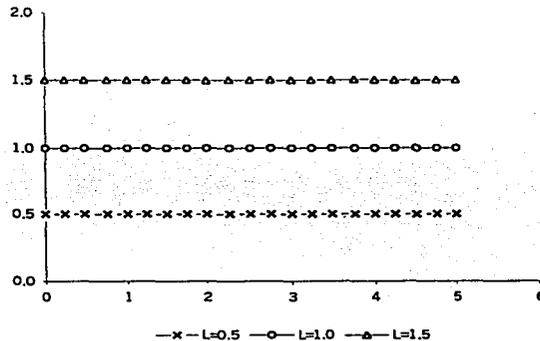


Figura 2.7 Función de riesgo del modelo Weibull cuando $\gamma = 1$, y para algunos valores de λ . A este modelo se le conoce como función Exponencial.

Por su parte, $S(t)$ queda definida de la siguiente forma:

$$S(t) = e^{-\int_0^t \lambda du}$$
$$S(t) = e^{-\lambda t}$$
$$t \geq 0$$

Partiendo de la equivalencia de $h(t)$ en términos de $f(t)$ y $S(t)$, es posible hallar una expresión para la función de densidad:

$$h(t) = \frac{f(t)}{S(t)}$$
$$h(t)S(t) = f(t)$$

Sustituyendo cada término por sus respectivos valores se obtiene $f(t)$:

$$f(t) = \lambda e^{-\lambda t}$$
$$t \geq 0$$

La relevancia de este caso radica en tres aspectos sobresalientes:

- $h(t)$ no depende de los valores que asuma la variable T , lo cual implica que el riesgo de ocurrencia del evento en estudio será el mismo en cualquier momento.
- La forma de $f(t)$ es equivalente a la de una función de densidad cuya variable T se distribuye de forma Exponencial con parámetro λ .
- Al ser un modelo univariado, se facilita su implementación e interpretación. Esto se abordará en secciones posteriores.

Finalmente, se mencionará que la función Exponencial proporciona un buen ajuste para describir el tiempo de vida útil de artículos manufacturados, además de ser de uso extendido en pruebas de confiabilidad.

2.3.2.3 Covariables.

Antes de explicar el método que se utilizará en el proceso de estimación del valor de los parámetros de las distribuciones propuestas, se hará una breve mención sobre la forma más general de construcción de un modelo paramétrico de supervivencia. Esta forma implica considerar la información complementaria que proporcionan otras variables las cuales ejercen cierta influencia sobre el comportamiento del fenómeno que se está analizando. A estas variables se les conoce con el nombre de covariables.

La presencia de covariables puede tratarse a través de la función de riesgo, siendo el Modelo de Regresión de Cox el más conocido para abordar este tema.

Este modelo se basa en el supuesto de que, a un tiempo $T=t$, existe una proporción entre la función de riesgo y la función de riesgo del fenómeno sin la influencia de sus covariables para cada integrante de una muestra aleatoria. Dicha proporción viene dada por una combinación lineal entre las diferentes covariables que intervienen. Lo anterior puede expresarse de la siguiente forma:

$$h(t) = e^{\beta'X} h_0(t)$$

donde $X=(x_1, \dots, x_p)'$ es la matriz de p -covariables, $\beta=(\beta_1, \dots, \beta_p)'$ es la matriz de sus respectivos coeficientes, y $h_0(t)$ es la función de riesgo "base". Por su parte, el producto $\beta'X = \beta_1 x_1 + \dots + \beta_p x_p$, indica la relación lineal más común que puede asumirse entre las p -variables.

En este punto, debe ser claro que para el ajuste de un modelo de supervivencia con covariables es necesario encontrar métodos apropiados que permitan estimar el valor de los coeficientes de las variables, así como formas adecuadas de interpretar estos resultados.

El manejo de covariables tiene importantes repercusiones; por ejemplo, la función de supervivencia quedaría expresada de la siguiente forma:

$$S(t) = e^{-\int_0^t e^{\beta'X} h_0(u) du}$$

Sin embargo, por las características propias de los fenómenos que se están modelando en este trabajo, el enfoque que prevalecerá será el de ausencia de covariables dentro el proceso de ajuste. No obstante lo anterior, y para tener mayores referencias sobre este tema, se sugiere consultar la bibliografía citada (Cantor, 1997 y Collet, 1994).

2.3.2.4 Método de máxima verosimilitud para determinar estimadores.

Existen diversos métodos para estimar el valor de un parámetro. Uno de ellos es el de máxima verosimilitud. La principal característica de este procedimiento es que selecciona como estimador el valor del parámetro que tiene la propiedad de maximizar el valor de la probabilidad de la muestra. Esto se consigue encontrando el valor del parámetro que maximiza la función de verosimilitud.

Los estimadores máximo verosímiles tienen algunas propiedades estadísticas deseables tales como ser suficientes (siempre que el estimador sea único) y eficientes (si es que el estimador existe); aunque, en la mayoría de los casos, estos estimadores son sesgados (Mood, 1974).

Por las ventajas expuestas, el procedimiento que se utilizará para estimar el valor de los parámetros de los modelos seleccionados será el de máxima verosimilitud.

El procedimiento general para su construcción es el siguiente: si se parte de una muestra aleatoria de n entidades con tiempos de supervivencia t_1, \dots, t_n no

censurados, cuya función de densidad para T es f(t), la función de verosimilitud está dada por:

$$L(\theta_1, \dots, \theta_n; \underline{t}_j) = \prod_{i=1}^n f(t_i)$$

Puede observarse que la función de verosimilitud depende de los parámetros desconocidos de la función de densidad. El estimador de máxima verosimilitud serán los valores de θ_i ; donde $i=1, \dots, n$; para los cuales la función de verosimilitud es máxima.

El caso más general es aquel en el cual se tienen datos censurados y no censurados. Si se cuenta con una muestra aleatoria de n observaciones, de las cuales m experimentaron el evento y (n-m) no lo experimentaron, la función de verosimilitud correspondiente estará compuesta por el producto de la probabilidad conjunta de las m entidades, que en los tiempos $t_1 < t_2 < \dots < t_m$ experimentaron el evento y las (n-m) entidades, que en los tiempos $t'_1 < t'_2 < \dots < t'_{n-m}$ no lo experimentaron. En la siguiente expresión se sintetiza la idea desarrollada anteriormente:

$$L(\theta_1, \dots, \theta_n; \underline{t}_j) = \prod_{i=1}^m f(t_i) \prod_{k=1}^{n-m} S(t'_k)$$

Para el primer término, es claro que aquellas entidades que se vieron afectadas por el evento de supervivencia (fallas) contribuirán a la función de verosimilitud con el valor que asuma la función de densidad evaluada en t_i . Por su parte, el segundo término, relacionado con los datos censurados, contribuirá con el valor de la función de supervivencia evaluada en t'_k .

Al simplificar esta expresión aplicando una función indicadora que diferencie los datos censurados de los no censurados, se tiene la forma general de la función de verosimilitud aplicable a muestras con censura de tipo I, II y aleatoria (Zamora, 1994):

$$\delta_i = \begin{cases} 1 & \text{si } t_i \text{ no es censurado} \\ 0 & \text{si } t_i \text{ es censurado} \end{cases}$$

$$L(\theta_1, \dots, \theta_n; \underline{t}_j) = \prod_{i=1}^n [f(t_i)]^{\delta_i} [S(t_i)]^{(1-\delta_i)}$$

Si ahora se aplica una transformación logarítmica, se facilita hallar el valor de los estimadores. Cabe señalar que se eligió esta transformación ya que garantiza que los valores que maximizan la función transformada son los mismos que maximizan la función original.

$$\ln[L(\theta_1, \dots, \theta_n; \underline{t}_j)] = \sum_{i=1}^n \{\delta_i [\ln(f(t_i))]\} + \sum_{i=1}^n \{(1-\delta_i) [\ln(S(t_i))]\}$$

Los estimadores por máxima verosimilitud para la función transformada, se obtienen al resolver el sistema de ecuaciones resultante de igualar a cero las derivadas parciales de cada estimador:

$$\begin{aligned} \frac{d}{d\theta_1} \text{Ln}[L(\theta_1, \dots, \theta_n; \underline{t}_i)] &= \frac{d}{d\theta_1} \left\{ \sum_{i=1}^n \{\delta_i [\text{Ln}(f(t_i))]\} + \sum_{i=1}^n \{[1 - \delta_i] [\text{Ln}(S(t_i))]\} \right\} = 0 \\ &\vdots \\ \frac{d}{d\theta_n} \text{Ln}[L(\theta_1, \dots, \theta_n; \underline{t}_i)] &= \frac{d}{d\theta_n} \left\{ \sum_{i=1}^n \{\delta_i [\text{Ln}(f(t_i))]\} + \sum_{i=1}^n \{[1 - \delta_i] [\text{Ln}(S(t_i))]\} \right\} = 0 \end{aligned}$$

Aunque el método de estimación del valor de los parámetros para las funciones propuestas en este trabajo es el mismo, el procedimiento de estimación es diferente entre una y otra. En el caso Exponencial, la expresión algebraica puede hallarse de la siguiente forma:

$$\begin{aligned} L(\lambda; \underline{t}_i) &= \prod_{i=1}^n (\lambda e^{-\lambda t_i})^{\delta_i} (e^{-\lambda t_i})^{(1-\delta_i)} \\ \text{Ln}[L(\lambda; \underline{t}_i)] &= \text{Ln} \left(\prod_{i=1}^n (\lambda^{\delta_i} e^{-\lambda t_i}) \right) \\ \text{Ln}[L(\lambda; \underline{t}_i)] &= \ln(\lambda) \sum_{i=1}^n \delta_i - \lambda \sum_{i=1}^n t_i \end{aligned}$$

Con m fallas dentro de la muestra (m datos no censurados):

$$\text{Ln}[L(\lambda; \underline{t}_i)] = \ln(\lambda)m - \lambda \sum_{i=1}^n t_i$$

Para maximizar la función logarítmica de verosimilitud se toma la derivada con respecto a λ y se iguala a cero:

$$\frac{d\text{Ln}[L(\lambda; \underline{t}_i)]}{d\lambda} = \frac{m}{\lambda} - \sum_{i=1}^n t_i = 0$$

Finalmente, el estimador del parámetro será:

$$\hat{\lambda} = \frac{m}{\sum_{i=1}^n t_i}$$

En el modelo Weibull los estimadores de λ y γ por máxima verosimilitud quedan como sigue:

$$L(\lambda, \gamma; \underline{t}_i) = \prod_{i=1}^n (\lambda \gamma t_i^{\gamma-1} e^{-\lambda t_i^\gamma})^{\delta_i} (e^{-\lambda t_i^\gamma})^{(1-\delta_i)}$$

$$\text{Ln}[L(\lambda, \gamma; \underline{t}_i)] = \text{Ln} \left(\prod_{i=1}^n (\lambda^{\delta_i} \gamma^{\delta_i} t_i^{(\gamma-1)\delta_i} e^{-\lambda t_i^\gamma}) \right)$$

$$\text{Ln}[L(\lambda, \gamma; \underline{t}_i)] = \text{Ln}(\lambda \gamma) \sum_{i=1}^n \delta_i + (\gamma - 1) \sum_{i=1}^n (\delta_i \text{Ln}(t_i)) - \lambda \sum_{i=1}^n t_i^\gamma$$

Con m fallas dentro de la muestra:

$$\text{Ln}[L(\lambda, \gamma; \underline{t}_i)] = \text{Ln}(\lambda \gamma)m + (\gamma - 1) \sum_{i=1}^n (\delta_i \text{Ln}(t_i)) - \lambda \sum_{i=1}^n t_i^\gamma$$

Finalmente, el sistema de ecuaciones es el siguiente:

$$\frac{d\text{Ln}[L(\lambda, \gamma; \underline{t}_i)]}{d\lambda} = \frac{m}{\lambda} - \sum_{i=1}^n t_i^\gamma = 0$$

$$\frac{d\text{Ln}[L(\lambda, \gamma; \underline{t}_i)]}{d\gamma} = \frac{m}{\gamma} + \sum_{i=1}^n ((\delta_i \text{Ln}(t_i)) - \lambda \sum_{i=1}^n t_i^\gamma \text{Ln}(t_i)) = 0$$

En la parte final del cálculo, se utilizará un procedimiento numérico conocido como Newton-Raphson ante la dificultad evidente para resolver este sistema de ecuaciones.

2.3.2.4.1 Proceso de Newton-Raphson para determinar estimadores.

Cuando no es sencillo resolver el sistema de ecuaciones que maximizan la función de verosimilitud, es usual recurrir a procedimientos numéricos para encontrar el valor de los estimadores a partir de la información muestral. Uno de estos procesos es el conocido como Newton-Raphson: sea $\mathbf{U}(\theta)$ matriz de $n \times 1$ que contiene la derivada de las funciones logarítmicas de verosimilitud con respecto a cada uno de los n -parámetros θ :

$$\mathbf{U}(\theta) = \left(\frac{d}{d\theta_1} \text{Ln}[L(\theta_1, \dots, \theta_n; \underline{t}_i)], \dots, \frac{d}{d\theta_n} \text{Ln}[L(\theta_1, \dots, \theta_n; \underline{t}_i)] \right)$$

Sea $\mathbf{I}(\theta)$ matriz de $n \times n$ con la segunda derivada negativa de las funciones logarítmicas de verosimilitud, de tal manera que en (j, k) se tenga:

$$-\frac{\partial^2}{\partial \theta_j \partial \theta_k} \text{Ln}[L(\theta_1, \dots, \theta_n; \underline{t}_i)] \equiv -\frac{\partial^2}{\partial \theta_j \partial \theta_k} \text{Ln}L(\theta)$$

$$I(\theta) = - \begin{pmatrix} \frac{\partial^2}{\partial \theta_1 \partial \theta_1} \text{LnL}(\theta) & \dots & \frac{\partial^2}{\partial \theta_1 \partial \theta_n} \text{LnL}(\theta) \\ \vdots & & \vdots \\ \frac{\partial^2}{\partial \theta_n \partial \theta_1} \text{LnL}(\theta) & \dots & \frac{\partial^2}{\partial \theta_n \partial \theta_n} \text{LnL}(\theta) \end{pmatrix}$$

Finalmente, el algoritmo para estimar el vector de parámetros θ en el ciclo $(s+1)$ del proceso se define como:

$$\hat{\theta}_{s+1} = \hat{\theta}_s + I^{-1}(\hat{\theta}_s) U(\hat{\theta}_s); \text{ para } s = 0, 1, \dots$$

Como ejemplo, a continuación se desarrollará esta expresión para el modelo Weibull, donde el valor de los parámetros será:

$$\begin{pmatrix} \hat{\lambda}_{s+1} \\ \hat{\gamma}_{s+1} \end{pmatrix} = \begin{pmatrix} \hat{\lambda}_s \\ \hat{\gamma}_s \end{pmatrix} + \begin{pmatrix} \frac{m}{\hat{\lambda}_s^2} & \sum_{i=1}^n t_i \hat{\gamma}_s (\ln(t_i)) \\ \sum_{i=1}^n t_i \hat{\gamma}_s (\ln(t_i)) & \frac{m}{\hat{\lambda}_s^2} + \lambda_s \sum_{i=1}^n t_i \hat{\gamma}_s (\ln(t_i))^2 \end{pmatrix}^{-1} \begin{pmatrix} \frac{m}{\hat{\lambda}_s} - \sum_{i=1}^n t_i \hat{\gamma}_s \\ \frac{m}{\hat{\lambda}_s} + \sum_{i=1}^n \delta_i \ln(t_i) + \lambda_s \sum_{i=1}^n t_i \hat{\gamma}_s (\ln(t_i)) \end{pmatrix}$$

Con la finalidad de asegurar una buena estimación, el algoritmo de Newton-Raphson debe dar inicio en un valor conveniente para $\hat{\theta}_0$ (por ejemplo, el de los estimadores obtenidos por métodos gráficos, tema que se comentará más adelante), y se detendrá cuando la diferencia entre los valores estimados en cada ciclo sea lo suficientemente pequeña a criterio de quien esté implementando el proceso, además de proporcionar el mejor ajuste.

Un resultado importante que debe mencionarse, es que la matriz de varianzas y covarianzas de los $\hat{\theta}_1, \dots, \hat{\theta}_n$ parámetros estimados por máxima verosimilitud, puede ser aproximada a través de la matriz inversa de las segundas derivadas negativas de las funciones logarítmicas de verosimilitud evaluadas en los valores calculados para el ciclo s (Collet, 1994):

$$\text{Var}(\hat{\theta}_s) \approx I^{-1}(\hat{\theta}_s)$$

Este resultado es particularmente útil al momento de proponer un intervalo de confianza sobre el valor de un parámetro específico.

2.3.2.5 Intervalos de confianza.

Una forma de aproximarse al valor real de los parámetros de una distribución muestral, es a través de los métodos de estimación puntual, entre los cuales se

encuentra el de máxima verosimilitud. Sin embargo, aunque los estimadores puntuales tengan buenas propiedades estadísticas e incluso el error que pueda cometerse durante el proceso de estimación sea mínimo, tienen la desventaja de que sólo aportan un único valor el cual, muy probablemente, no es el valor verdadero del parámetro que se está buscando. En este sentido, es conveniente proponer, junto con la estimación puntual $\hat{\theta}$ de un parámetro θ , un cierto intervalo que incluya a $\hat{\theta}$ con alguna medida de seguridad de que el parámetro verdadero θ está contenido en dicho intervalo. Para ello, debe encontrarse una estadística relacionada con el parámetro desconocido, pero cuya distribución no contenga a θ ; de este modo es posible construir un intervalo que acote el valor real del parámetro:

$$P[h_1(T) < \theta < h_2(T)] = 1 - \alpha$$

En este caso, $h_1(T)$ y $h_2(T)$ son funciones de la estadística T y $1 - \alpha$ es el coeficiente de confianza que proporcionará la medida de seguridad requerida. Con relación a los modelos que se están tratando, los intervalos de confianza se realizarán a través de estadísticas cuya distribución es asintótica a una distribución conocida.

Para construir un intervalo de confianza sobre el parámetro λ del modelo Exponencial, a partir de una muestra aleatoria con censura de tipo I o censura aleatoria, puede utilizarse la siguiente aproximación asintótica:

$$\frac{\hat{\lambda} - \lambda}{I(\hat{\lambda})^{-\frac{1}{2}}} \stackrel{a}{\sim} N(0, 1)$$

donde $I(\hat{\lambda})$ es la función de información o de Fisher dada por:

$$-E \left\{ \frac{d^2 \text{Ln}L(\lambda)}{d\lambda^2} \right\}$$

Para calcular $I(\hat{\lambda})$ se requieren los tiempos potenciales de censura L_i de todos los individuos de la muestra (Lawless, 1982 y Zamora, 1994); sin embargo, dado que esta información no siempre se conoce en la práctica se prefiere utilizar una aproximación asintótica equivalente:

$$\frac{\hat{\lambda} - \lambda}{I_0^{-\frac{1}{2}}} \stackrel{a}{\sim} N(0, 1)$$

donde $I_0 = - \frac{d^2 \text{Ln}L(\lambda)}{d\lambda^2} \Big|_{\hat{\lambda}}$.

Si se cuenta con una muestra aleatoria con m fallas, se tiene que $I_0 = \frac{m}{\hat{\lambda}^2}$, y en consecuencia:

$$\frac{\hat{\lambda} - \lambda}{\frac{\hat{\lambda}}{\sqrt{m}}} \stackrel{a}{\sim} N(0, 1)$$

converge asintóticamente a una distribución Normal estándar.

Ahora, cuando se tenga una muestra aleatoria de tamaño n , con censura de tipo II y m fallas, puede construirse un intervalo de confianza para λ a partir de la estadística (Miller, 1981):

$$2\lambda T \stackrel{a}{\sim} \chi^2_{2m}$$

donde $T = \sum_{i=1}^n t_i$. Cabe señalar que dentro de esta última expresión, $(n-m)$ entidades son censuradas y contribuyen a la suma en t_m fijo.

Recordando que el estimador por máxima verosimilitud de λ es $\hat{\lambda} = \frac{m}{\sum_{i=1}^n t_i}$, para cualquier tipo de censura, puede replantearse la estadística a quedar como sigue:

$$\frac{2m\lambda}{\hat{\lambda}} \stackrel{a}{\sim} \chi^2_{2m}$$

quién converge asintóticamente a una distribución Ji-cuadrada con $(2m)$ grados de libertad.

Por su parte, para construir intervalos de confianza sobre los parámetros del modelo Weibull, se recomienda utilizar una distribución alternativa equivalente en lugar de la distribución original. En este caso, se prefiere la función del Valor Extremo, cuya distribución es la siguiente:

$$f(x) = \frac{1}{b} e^{\left\{\frac{(x-u)}{b}\right\}} e^{\left\{-\left(\frac{x-u}{b}\right)\right\}}$$

con $u \in \mathbb{R}$ y $b > 0$.

Esta distribución se obtiene al aplicar las siguientes transformaciones al modelo Weibull:

$$t = e^x$$

$$\lambda = e\left(-\frac{u}{b}\right)$$

$$\gamma = \frac{1}{b}$$

La principal ventaja de trabajar con la distribución del Valor Extremo radica en el hecho de que u y b son parámetros de localización y escala respectivamente de una distribución simétrica, lo que facilita el manejo de la función.

Para encontrar estimadores de u y b se utiliza el proceso de máxima verosimilitud. Los intervalos de confianza para los parámetros se determinarán a partir de la distribución de los estimadores máximo verosímiles de u y b , que es asintótica a una distribución Normal bivariada (Zamora, 1994):

$$\left(\hat{u}, \hat{b}\right)^a \sim N((u, b), I^{-1}(u, b))$$

El término $I^{-1}(u, b)$ corresponde a la matriz inversa de la matriz de varianzas y covarianzas de Fisher. En la práctica, esta matriz puede estimarse por medio de $I_0(\hat{u}, \hat{b})$, que es la matriz de 2×2 de las segundas derivadas negativas de las

funciones logarítmicas de verosimilitud evaluadas en (\hat{u}, \hat{b}) .

Por último se dirá, que los resultados arriba utilizados se derivan de la siguiente aproximación asintótica, cuando el tamaño de la muestra es grande ($n \rightarrow \infty$) ante la presencia de cualquier tipo de censura (Miller, 1981):

$$\left(\hat{\theta}_k\right)^a \sim N(\theta_k, I^{-1}(\theta_k))$$

A partir de lo anterior, es posible hallar intervalos de confianza directamente sobre los parámetros de la función original sin la necesidad de plantear una transformación.

2.4 Estimación y ajuste de un modelo paramétrico.

Un buen punto de partida para proponer algún tipo de ajuste sobre la forma de distribución real de la muestra (e incluso comparaciones entre varias muestras), es apoyarse en los métodos gráficos de ajuste ya que éstos aportan un elemento visual sobre la dispersión de los datos. Sus ventajas pueden resumirse en los siguientes puntos:

- 1) Proveen una representación puntual de la información.

-
- 2) Su presentación no requiere haber realizado supuesto alguno sobre la distribución real.
 - 3) Proporcionan una idea gráfica sobre la forma de la distribución real.
 - 4) Es un indicador eficaz de particularidades de los datos (outliers).
 - 5) Pueden ser implementados incluso en muestras censuradas.
 - 6) Al aplicar una transformación adecuada sobre los datos es posible obtener un gráfico con tendencia lineal; esto facilita el planteamiento de algún supuesto sobre la experiencia de supervivencia de la muestra.
 - 7) Pueden ser utilizados como base para ajustar una distribución específica a todo el conjunto de datos.

2.4.1 Transformaciones para métodos gráficos.

El mayor reto en el uso de los métodos gráficos radica en decidir si la forma de la curva típica de una determinada distribución ajusta adecuadamente a un conjunto de datos con sólo realizar una simple inspección visual. Esta decisión puede facilitarse si la apreciación se realiza sobre un conjunto de puntos que sigan la forma de una recta. Para lograr esto, se propone someter a la muestra a una transformación que produzca este efecto sobre los datos transformados. Si al graficar los nuevos puntos se obtiene aproximadamente una recta, puede decirse que existe una evidencia de que el modelo paramétrico seleccionado puede ser adecuado para explicar los datos originales; en caso contrario, debe optarse por otro modelo.

Para comprobar que el modelo Exponencial puede ajustarse a un conjunto de puntos, la transformación que se propone sobre la función de supervivencia se basa en el logaritmo natural:

$$S(t) = e^{-\lambda t}$$
$$-\text{Ln}S(t) = \lambda t$$

Si se desea comprobar que el modelo Weibull es quien mejor se ajusta a la muestra, la transformación propuesta es la siguiente:

$$S(t) = e^{-\lambda t^\gamma}$$
$$\text{Ln}(-\text{Ln}S(t)) = \text{Ln}\lambda + \gamma \text{Ln}(t)$$

Si al transformar los datos, y luego de graficar t contra $-\text{Ln}(\hat{S}(t))$, asumiendo que el estimador de $S(t)$ es $\hat{S}(t)$ de Kaplan-Meier, se obtiene una recta con pendiente positiva que cruza cerca del origen, puede decirse que la función Exponencial es adecuada. Este comportamiento puede apreciarse en la figura 2.8.

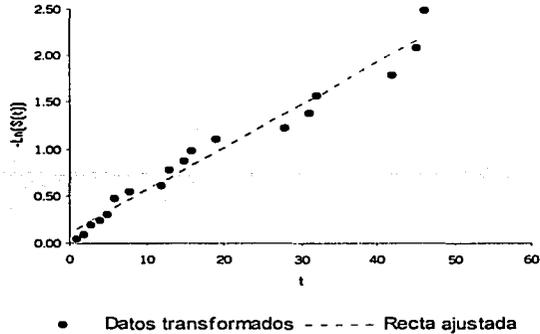


Figura 2.8 Conjunto de datos transformados que aproximan una recta. La tendencia de la misma sienta un precedente para suponer que el modelo Exponencial puede ser adecuado para describir los datos originales.

Por otro lado, si al graficar $\ln(t)$ contra $\ln(-\ln\hat{S}(t))$, asumiendo nuevamente que $\hat{S}(t)$ es el estimador de $S(t)$ por Kaplan–Meier, se obtiene una recta con pendiente γ que cruza el eje de las ordenadas en $\ln\lambda$, se tiene una evidencia de que el modelo Weibull es el indicado. El comportamiento gráfico se muestra en la figura 2.9

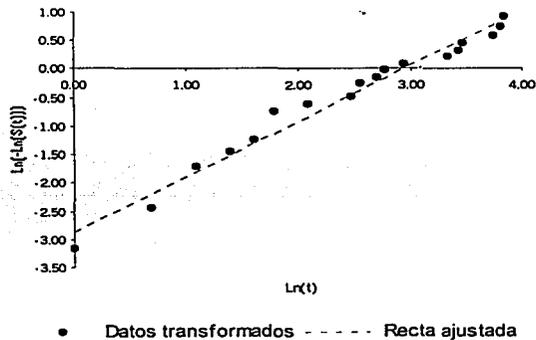


Figura 2.9 Conjunto de datos transformados que aproximan una recta. La forma de la misma sienta un precedente para suponer que el modelo Weibull puede ser adecuado para describir los datos originales.

Vale la pena observar que si se ajusta una recta al conjunto de puntos transformados, tanto el valor de la pendiente como el de la ordenada al origen pueden considerarse estimadores de los parámetros del modelo propuesto. Por ejemplo, para la distribución Weibull, la pendiente de la recta ajustada es un estimador de γ , mientras que la ordenada al origen es un estimador para $\ln\lambda$ (y a

partir de éste puede hallarse un estimador de λ). En el caso del modelo Exponencial, la pendiente de la recta ajustada es un estimador de λ .

Antes de concluir esta sección, debe señalarse que los parámetros estimados por métodos gráficos pueden utilizarse como valores iniciales en algunos procesos numéricos que permiten afinar nuestra estimación (p.j. Newton-Raphson), particularmente si el modelo cuenta con 2 ó más parámetros, lo que dificulta su cálculo directo (p.j. Weibull).

2.4.2 Pruebas de hipótesis para comparar la experiencia de supervivencia entre varias muestras.

Con la finalidad de comparar la experiencia de supervivencia entre varios grupos, se pueden utilizar algunos procedimientos no paramétricos entre los cuales se encuentra el conocido como Log-rank o prueba de Mantel-Haenszel (Collet, 1994 y Miller, 1981). En las secciones siguientes se abordará este tema.

2.4.2.1 Prueba de Mantel-Haenszel para dos muestras.

La formulación de esta prueba para dos muestras, parte de la construcción de una tabla donde se anotarán las características de los dos grupos a comparar en función de las entidades expuestas y las fallas observadas en un determinado momento. Los supuestos que se usarán son:

- Un primer grupo con t_{11}, \dots, t_{1n} observaciones y un segundo grupo con t_{21}, \dots, t_{2m} .
- Se asume que hay r fallas entre ambas muestras. Los tiempos de falla ordenados serán $t_{(1)}, \dots, t_{(r)}$ de tal forma que en $t_{(j)}$ ocurren d_{1j} fallas en el primer grupo y d_{2j} fallas en el segundo. Las fallas son independientes.
- Los expuestos en $t_{(j)}$ serán n_{1j} para el primer grupo y n_{2j} para el segundo.

Esto supuestos quedan resumidos dentro de la tabla 2.2, para cualquier $t_{(j)}$.

Tabla 2.2 Resumen de supuestos en la construcción de la estadística de prueba de Mantel-Haenszel para dos muestras al tiempo $t_{(j)}$.

$$\begin{aligned} d_j &= d_{1j} + d_{2j} && \text{fallas} \\ n_j &= n_{1j} + n_{2j} && \text{expuestos} \\ & \text{para } j=1, \dots, r \end{aligned}$$

Grupo	Fallas en $t_{(j)}$	Sobrevivientes en $t_{(j)}$	Expuestos justo antes de $t_{(j)}$
1	d_{1j}	$n_{1j} - d_{1j}$	n_{1j}
2	d_{2j}	$n_{2j} - d_{2j}$	n_{2j}
	$d_{1j} + d_{2j} = d_j$	$(n_{1j} - d_{1j}) + (n_{2j} - d_{2j}) =$ $(n_{1j} + n_{2j}) - (d_{1j} + d_{2j}) =$ $n_j - d_j$	$n_{1j} + n_{2j} = n_j$

En el planteamiento de la estadística de prueba se formulará la hipótesis de que no existen diferencias entre las experiencias de supervivencia de las muestras (hipótesis nula). Para verificar lo anterior se comparará el número de fallas observadas en ambos grupos contra el número de fallas esperadas, suponiendo la validez de la hipótesis planteada. Adicionalmente se asumirá que existe independencia entre las experiencias de supervivencia de ambas muestras, supuesto que se utilizará más adelante.

Si se fijan los valores de las fallas del grupo 1 en la tabla 2.2 se pueden reescribir las entradas de la misma tabla de tal manera que solo dependan de d_{1j} . Este replanteamiento se aprecia en la tabla 2.3.

Tabla 2.3 Replanteamiento de los supuestos en la construcción de la estadística de prueba de Mantel-Haenszel para dos muestras al tiempo $t_{(j)}$, después de fijar las fallas del primer grupo

Grupo	Fallas en $t_{(j)}$	Sobrevivientes en $t_{(j)}$	Expuestos justo antes de $t_{(j)}$
1	d_{1j}	$n_{1j} - d_{1j}$	n_{1j}
2	$d_j - d_{1j}$	$(n_j - n_{1j}) - (d_j - d_{1j})$	$n_j - n_{1j}$
	d_j	$n_j - d_j$	n_j

Dado que se están comparando dos grupos con n_j , d_j y n_{1j} fijos, es posible definir una nueva variable aleatoria D la cual recoge todas las posibles combinaciones de tener d_{1j} fallas exactamente al tiempo $t_{(j)}$, de entre d_j fallas sobre n_{1j} expuestos de un total de n_j posibles. Esta variable asume valores en el rango 0 al mínimo entre d_{1j} y n_{1j} , y sigue una distribución hipergeométrica:

$$P[D=d_{1j}] = \frac{\binom{d_j}{d_{1j}} \binom{n_j - d_j}{n_{1j} - d_{1j}}}{\binom{n_j}{n_{1j}}}$$

El valor esperado de las fallas del primer grupo en $t_{(j)}$ estará dado por la esperanza de D :

$$E(D) = \frac{n_{1j} d_j}{n_j}$$

Por su parte, la varianza está dada por:

$$\text{Var}(D) = \frac{n_{1j}(n_j - n_{1j})d_j(n_j - d_j)}{n_j^2(n_j - 1)}$$

Bajo la hipótesis nula, la comparación entre las fallas realmente observadas y las fallas esperadas en $t_{(j)}$ puede expresarse como la diferencia de ambas:

$$d_{1j} - E(D)$$

$$d_{1j} - \frac{n_{1j}d_j}{n_j}$$

El conjunto total de diferencias sobre las r fallas será:

$$U = \sum_{j=1}^r (d_{1j} - E(D))$$

$$U = \sum_{j=1}^r d_{1j} - \sum_{j=1}^r E(D)$$

Finalmente, la estadística de Mantel-Haezsel tiene la siguiente forma:

$$MH = \frac{\sum_{j=1}^r (d_{1j} - E(D))}{\sqrt{\sum_{j=1}^r \text{Var}(D)}} = \frac{U}{\sqrt{V}}$$

donde:

$$V = \text{Var}(U) = \sum_{j=1}^r \text{Var}(D) = \sum_{j=1}^r \frac{n_{1j}(n_j - n_{1j})d_j(n_j - d_j)}{n_j^2(n_j - 1)}; \text{ por independencia de las } d_{1j}\text{'s.}$$

Cabe señalar que:

$$\frac{U}{\sqrt{V}} \stackrel{a}{\sim} N(0, 1)$$

converge asintóticamente a una distribución Normal con media cero y varianza 1; sin embargo, algunos autores prefieren utilizar la expresión:

$$\frac{U^2}{V} \stackrel{a}{\sim} \chi^2$$

la cual converge asintóticamente a una distribución Ji-cuadrada con un grado de libertad. Ambas expresiones facilitan hallar valores en tablas.

2.4.2.2 Generalización para dos muestras.

La generalización de la prueba para dos muestras al tiempo $t_{(j)}$ puede expresarse como:

$$w[d_{1j} - E(D)]$$

donde w actúa como un ponderador para las diferencias entre las fallas observadas y esperadas. La estadística U estará definida de la siguiente manera:

$$U = \sum_{j=1}^r w(d_{1j} - E(D)) = \sum_{j=1}^r (wd_{1j} - wE(D))$$

Utilizando el resultado $\text{Var}(aX) = a^2 \text{Var}(X)$, se encuentra una expresión para V:

$$V = \text{Var}(U) = \text{Var}(wD) = w^2 \text{Var}(D)$$

Asumiendo independencia entre las d_{1j} s, se tiene que la estadística de prueba es como sigue:

$$\frac{U}{\sqrt{V}} = \frac{\sum_{j=1}^r (wd_{1j} - wE(D))}{\sqrt{\sum_{j=1}^r w^2 \text{Var}(D)}} \stackrel{a}{\sim} N(0, 1) \quad \text{ó} \quad \frac{U^2}{V} \stackrel{a}{\sim} \chi^2$$

Si $w=1$, se obtiene la prueba de Mantel-Haenszel. Por otro lado, si $w=n_j$ la estadística resultante se le conoce como prueba de Wilcoxon.

2.4.2.3 Generalización para más de dos muestras.

En forma análoga a la idea desarrollada para dos muestras, la construcción de la estadística de prueba para más de dos muestras también se fundamenta en comparar el total de las diferencias entre las fallas observadas (d_{kj}) para cada uno de los grupos, con sus respectivos valores esperados ($E(D_k)$). La generalización de supuestos puede apreciarse en la tabla 2.4

Tabla 2.4 Resumen de supuestos en la construcción de la estadística de prueba de Mantel-Haenszel para g-muestras al tiempo $t_{(t)}$.

g-muestras	1	2	...	g	
Fallas en $t_{(t)}$	d_{1j}	d_{2j}	...	d_{gj}	d_j
Sobrevivientes en $t_{(t)}$	$n_{1j} - d_{1j}$	$n_{2j} - d_{2j}$...	$n_{gj} - d_{gj}$	$n_j - d_j$
Expuestos justo antes de $t_{(t)}$	n_{1j}	n_{2j}	...	n_{gj}	n_j

La estadística de prueba quedará expresada en forma matricial:

$$U'V^{-1}U \stackrel{a}{\sim} \chi^2_{(g-1)} \dots\dots\dots(1)$$

El producto resultante converge asintóticamente a una distribución Ji-cuadrada con $(g-1)$ grados de libertad, donde:

$$D = \begin{pmatrix} w \sum_{j=1}^r d_{1j} \\ \vdots \\ w \sum_{j=1}^r d_{g-1j} \end{pmatrix};$$

$$E(D) = \begin{pmatrix} wE(D_1) \\ \vdots \\ wE(D_{g-1}) \end{pmatrix};$$

$$U = (D - E(D));$$

$$V = \text{Var}(D) = \begin{pmatrix} w^2 \text{Var}(D_1) & \cdots & w^2 \text{Cov}(D_1, D_k) & \cdots & w^2 \text{Cov}(D_1, D_{g-1}) \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ w^2 \text{Cov}(D_k, D_1) & \cdots & w^2 \text{Var}(D_k) & \cdots & w^2 \text{Cov}(D_k, D_{g-1}) \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ w^2 \text{Cov}(D_{g-1}, D_1) & \cdots & w^2 \text{Cov}(D_{g-1}, D_k) & \cdots & w^2 \text{Var}(D_{g-1}) \end{pmatrix}.$$

A manera de ejemplo, a continuación se muestra el desarrollo de la ecuación (1) de esta sección para tres muestras, y con $w=1$:

$$\left(\begin{matrix} \sum_{j=1}^r (d_{1j} - \frac{n_{1j}d_j}{n_j}) & \sum_{j=1}^r (d_{2j} - \frac{n_{2j}d_j}{n_j}) \\ \sum_{j=1}^r \frac{n_{1j}(n_j - n_{1j})d_j(n_j - d_j)}{n_j^2(n_j - 1)} & \sum_{j=1}^r \frac{n_{1j}(-n_{2j})d_j(n_j - d_j)}{n_j^2(n_j - 1)} \\ \sum_{j=1}^r \frac{n_{2j}(-n_{1j})d_j(n_j - d_j)}{n_j^2(n_j - 1)} & \sum_{j=1}^r \frac{n_{2j}(n_j - n_{2j})d_j(n_j - d_j)}{n_j^2(n_j - 1)} \end{matrix} \right)^{-1} \begin{pmatrix} \sum_{j=1}^r (d_{1j} - \frac{n_{1j}d_j}{n_j}) \\ \sum_{j=1}^r (d_{2j} - \frac{n_{2j}d_j}{n_j}) \end{pmatrix} = p$$

En este caso, p es el valor de la estadística de prueba de Mantel-Haenszel.

2.4.3 Proceso general de ajuste de un modelo paramétrico de supervivencia.

El proceso general de ajuste de un modelo paramétrico de supervivencia se sujetará a los siguientes pasos:

- 1) Se parte de una muestra aleatoria t_1, \dots, t_n , censurada.
- 2) Se asume que la muestra sigue una determinada distribución, definida a partir de la evidencia encontrada del análisis resultante de aplicar métodos gráficos y no paramétricos de ajuste.
- 3) Se estiman los valores $\hat{\theta}_n$ de los parámetros de la distribución que se ha supuesto.
- 4) En caso de contar con k-muestras, puede formularse una prueba de hipótesis con la finalidad de verificar si entre ellas existe una experiencia de supervivencia común. La hipótesis nula a probar se plantea en los siguientes términos:

$H_0: S(t_1) = \dots = S(t_k)$
contra
$H_1: S(t_i) \neq S(t_j)$ al menos para algún $i \neq j$

CAPÍTULO III: ESTIMACIÓN DE UN MODELO PARAMÉTRICO DE SUPERVIVENCIA.

3.1 Introducción.

Este capítulo trata sobre el proceso de estimación de un modelo paramétrico de supervivencia que describe y permite comparar algunos aspectos relacionados con el desarrollo del tiempo de duración y el tiempo anterior a la primera manifestación de los eventos adversos emergentes registrados durante el ensayo clínico base, para cada uno de los tratamientos médicos que están siendo analizados. Primero, se explicará la forma como se organizó la información de la base de datos original para, posteriormente, poder identificar la que sería de utilidad en la construcción de los modelos. En esta parte, se detallarán los criterios que se aplicaron para identificar los registros correspondientes a los eventos adversos emergentes, así como el tipo de censura presente en cada muestra. Luego, se abordará el proceso que hubo de realizarse en la construcción de un modelo de supervivencia para cada uno de los dos fenómenos que se están estudiando y los eventos adversos observados, tanto en su forma general como para los dos sistemas corporales con mayor incidencia de casos, y por tipo de tratamiento. Cabe señalar que estos ajustes se realizaron a través de la estimación de la función de supervivencia bajo los modelos Weibull y Exponencial. Para alcanzar este objetivo, se recurrió a procesos estadísticos de estimación no paramétrica y paramétrica, que fueron tratados en el capítulo anterior.

En la siguiente parte, se justificará la razón por la cual fue necesario desarrollar una prueba de hipótesis adicional, cuyo objetivo sería verificar igualdad entre las experiencias de supervivencia de cada uno de los tratamientos.

Al final del capítulo, se presentará un resumen con los resultados obtenidos así como las conclusiones y aportaciones que se desprenden del presente estudio.

Por último, debe señalarse que la mayor parte del análisis, así como la construcción de los modelos, se llevó a cabo utilizando el software estadístico S.A.S. (Statistical Analysis System) para Windows en su octava versión.

3.2 Descripción de las bases de datos originales.

Los datos recabados durante las diversas visitas médicas, se encuentran organizados por temas y están concentrados en 25 bases de datos. Los campos de cada base pueden clasificarse en dos tipos: aquéllos con información general, la cual permite identificar a cada paciente participante, y los que contienen información específica sobre el tema que se aborda dentro de la propia base de datos. Los números y nombres asignados a cada base se describen en la tabla 3.1.

Tabla 3.1 Bases de datos, por número y nombre, con información recabada durante las visitas médicas efectuadas durante la realización del ensayo clínico.

No. De Base	Nombre de la base de datos
1	Datos generales e información demográfica.
2 y 3	Antecedentes familiares y personales de hiperlipidemia.
4, 5, 6, 7 y 11	Signos vitales, exploración física, historia clínica y factores de riesgo.
8	Medicación concomitante.
9 y 10	Evaluación cardíaca.
12	Evaluación dietética
21 y 13	Evaluación de lípidos y cumplimiento de dieta.
20 y 14	Resultados de laboratorio.
15	Cumplimiento del tratamiento.
16	Hábitos de ingesta de alcohol.
17	Eventos adversos.
18	Fin del estudio/ retiro prematuro.
19	Hematología y química sanguínea.
22	Urianálisis.
23	Criterios de exclusión.
24	Criterios de asignación aleatoria.
25	Criterios de inclusión.

3.3 Base de datos de eventos adversos.

La base de datos donde se registró la información relativa a los eventos adversos (base No. 17), está integrada con información de 665 pacientes reclutados inicialmente. De este total, sólo 375 calificaron para participar dentro de la Segunda Etapa del ensayo clínico, donde habrían de someterse a un tratamiento médico específico después de haber sido asignados a uno de los tres grupos de estudio en forma aleatoria. La distribución de pacientes por tratamiento, se resume en la tabla 3.2.

Tabla 3.2 Resultado de la asignación aleatoria de pacientes por cada tratamiento.

Total	Pacientes sin tratamiento	Pacientes con tratamiento	
		A	B
		128	126
		121	
665	290	375	

Los eventos adversos observados, incluso aquellos de pacientes no incluidos dentro de algún tratamiento, fueron registrados en la base de datos de forma individual, es decir un registro por cada evento. Sin embargo, si un mismo evento se presentó en forma intermitente, con dos o más manifestaciones separadas por uno o más días entre sí, quedaría registrado tantas veces como manifestaciones se hubiesen reportado, anotando en cada registro la información específica de cada manifestación. Por otro lado, si un paciente declaró no padecer evento adverso alguno, sólo quedarían anotados sus datos personales en la base una sola vez dentro de un único registro. En resumen, si un paciente reportó uno o más eventos adversos, aparecerían tantos registros como manifestaciones de cada evento hubiesen sido reportadas. En este sentido, la base de datos de

eventos adversos puede analizarse a partir de la información de las diversas manifestaciones de cada evento.

Dentro del total de pacientes asignados a un tratamiento, 207 no experimentaron evento alguno mientras que 168 experimentaron al menos un evento adverso. Los eventos adversos reportados por estos últimos fue de 316; los registros totales (manifestaciones) de estos eventos fue de 346.

La distribución de eventos y registros de los pacientes asignados a un tratamiento, se muestra en la tabla 3.3.

Tabla 3.3 Distribución de pacientes asignados a un tratamiento, según manifestaron haber experimentado algún evento adverso. Total de eventos adversos por tratamiento. Total de registros de eventos adversos por tratamiento.

Tratamiento	Pacientes			Eventos adversos	Registros de eventos adversos
	Total	Sin eventos adversos	Con al menos un evento adverso		
A	124	67	57	114	125
B	130	71	59	97	105
C	121	69	52	105	116
	375	207	168	316	346

Como dato informativo, se mencionará que de los 290 pacientes no incluidos en algún tratamiento; 217 no reportaron eventos, mientras que 73 reportaron al menos un evento adverso. Los registros de eventos para esta última población fue de 116 en total.

Para alcanzar el objetivo planteado en este trabajo, nuestro interés se centrará en los registros (manifestaciones) de eventos adversos emergentes. Cabe recordar que esta clasificación incluye sólo aquellos eventos adversos que:

- 1) Comenzaron por primera vez después de iniciado el tratamiento (figura 3.1).

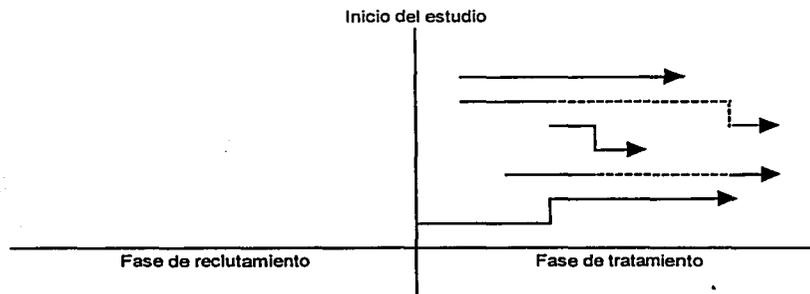


Figura 3.1 Eventos adversos que comenzaron después de iniciado el tratamiento.

2) Reaparecieron una vez iniciado el tratamiento (figura 3.2), ó

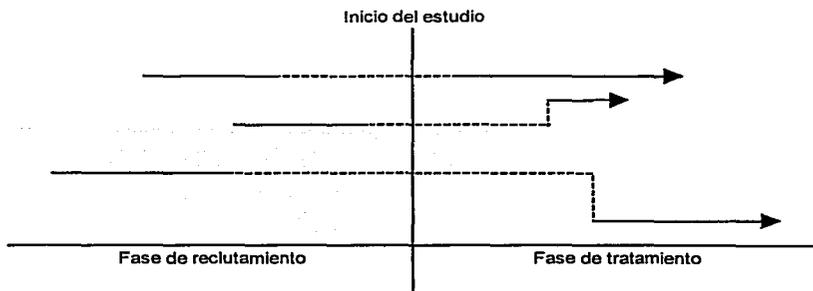


Figura 3.2 Eventos adversos que reaparecieron después de iniciado el tratamiento.

3) Incrementaron su intensidad una vez comenzado el tratamiento (figura 3.3).

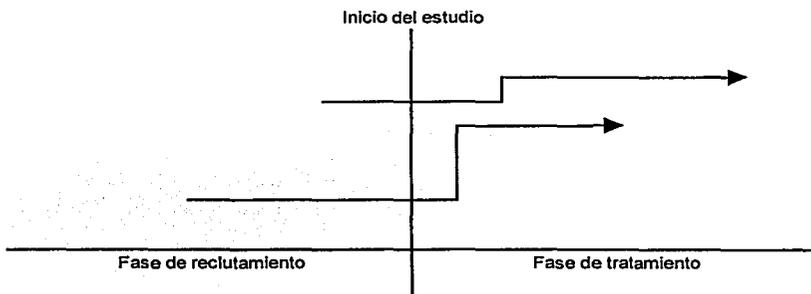


Figura 3.3 Eventos adversos que incrementaron su intensidad después de iniciado el tratamiento.

Antes de explicar el proceso de análisis y organización de la información para llevar a cabo el ajuste de los modelos de supervivencia, se procederá a describir la estructura de la base de datos de los eventos adversos.

Esta base está compuesta por 42 variables numeradas en forma ascendente. Todas ellas tienen como finalidad caracterizar la información de los eventos adversos dentro del contexto del estudio clínico. Entre las primeras ocho, las que mayormente se mencionarán en éste trabajo son:

La variable 3 (PATNR), donde se registra el número aleatorio que identifica a cada paciente cuando éste es asignado a un tratamiento.

La variable 4 (FOLGEKT), proporciona información sobre la posición que ocupa cada manifestación de los eventos adversos registrados para cada paciente; por ejemplo, si una persona no reportó ningún evento o reportó un evento con una sola manifestación, el valor de este campo es 1; por el contrario, si hubiese reportado un solo evento con más de dos manifestaciones, la primer manifestación

registrada asume el valor 1 y la segunda el valor 2. En resumen, si un mismo paciente reportó más de un evento con más de una manifestación por evento, el valor del campo FOLGEKT indicará el orden en que cada manifestación fue registrada.

En las variables 6 (TREATMGR), 7 (CENTER) y 8 (CRF_NO) se anotó el tratamiento asignado, la clave del centro hospitalario donde fue reclutado el paciente y su número de expediente, respectivamente.

De las variables 9 a la 42, se recabó información específica sobre cada una de las manifestaciones de cada evento adverso reportado. El registro de los eventos adversos, se hizo por medio del nombre del evento anotado por el médico en el reporte correspondiente. Este dato se codificó en términos de diccionarios especializados y diseñados para este propósito; en este caso se eligió el "Coding Symbols for Thesaurus of Adverse Reaction Terms" (COSTART), cuyo objetivo es crear un lenguaje común dentro de la industria farmacéutica.

En la tabla 3.4, puede apreciarse un resumen de toda esta información, ordenada por el número de la variable. También se ha incluido el nombre de la misma, tal y como aparece dentro de la propia base de datos, así como una breve explicación de su significado y contenido.

Tabla 3.4 Descripción de las variables 9 a 42 de la base de datos de Eventos Adversos.

No. de variable	Nombre de la variable en la base de datos	Nombre	Explicación
9	EVENT	Evento	Se reporta si el registro corresponde a una manifestación de un evento adverso o no corresponde a una manifestación de un evento adverso, donde: 1 no hay evento adverso 2 hay un primer evento adverso 3 hay más de un primer evento adverso
10	EVENT1	Evento adverso específico	Clave del evento adverso específico.
11	SERIOUS	Gravedad del evento	Se indica la seriedad de la manifestación del evento adverso tomando como punto de referencia la vida del paciente.
12	DATEVSDD	Día de inicio	Día de inicio de la manifestación del evento (dd).
13	DATEVSM	Mes de inicio	Mes de inicio de la manifestación del evento (mm).
14	DATEVSY	Año de inicio	Año de inicio de la manifestación del evento (yy).
15	AE_SDT	Fecha de inicio	Fecha de inicio de la manifestación del evento (dd/mm/yy)
16	AE_SPH	Fase donde inició la manifestación del evento adverso	Fase del estudio donde inició la manifestación del evento adverso, donde: 1 antes de iniciado el tratamiento 2 durante el periodo de tratamiento 3 después de finalizado el tratamiento 0.5 fase no definida
17	AE_SRS	Días relativos entre la fecha de inicio de la manifestación del evento y la fecha de inicio del tratamiento.	Si $AE_SDT < SM_SDT$ entonces: $AE_SRS = AE_SDT - SM_SDT$ Si $AE_SDT \geq SM_SDT$ entonces: $AE_SRS = AE_SDT - SM_SDT + 1$ donde SM_SDT es la fecha de inicio del tratamiento ⁶ .
18	AE_SRE	Días relativos entre la fecha de inicio de la manifestación del evento y la fecha de fin del tratamiento	$AE_SRE = AE_SDT - SM_EDT - 1$ Donde SM_EDT es la fecha de fin del tratamiento ⁷ .
19	DATEVEDD	Día de fin	Día de fin de la manifestación del evento (dd).
20	DATEVEMM	Mes de fin	Mes de fin de la manifestación del evento (mm).
21	DATEVEYY	Año de fin	Año de fin de la manifestación del evento (yy).
22	AE_EDT	Fecha de fin	Fecha de fin de la manifestación del evento (dd/mm/yy).
23	AE_EPH	Fase donde finalizó la manifestación del evento	Fase del estudio donde finalizó la manifestación del evento adverso, donde: 1 antes de iniciado el tratamiento 2 durante el periodo de tratamiento 3 después de finalizado el tratamiento 0.5 fase no definida
24	AE_ERS	Días relativos entre la fecha de fin de la manifestación del evento y la fecha de inicio del tratamiento	$AE_ERS = AE_EDT - SM_SDT + 1$ donde SM_SDT es la fecha de inicio del tratamiento.
25	AE_ERE	Días relativos entre la fecha de fin de la manifestación del evento y la fecha de fin del tratamiento	$AE_ERE = AE_EDT - SM_EDT$ donde SM_EDT es la fecha de fin del tratamiento.
26	INTENS	Intensidad	Clave de intensidad de la manifestación del evento adverso.
27	TODRUG	Relación	Clave de relación de la manifestación del evento con el tratamiento.
28 a 30	MEASUR1 MEASUR2 MEASUR3	al Acciones tomadas	Acciones tomadas para disminuir o eliminar los efectos del evento adverso.
31 a 33	OUTCOME1 OUTCOME2 OUTCOME3	al Resultados	Resultado de las consecuencias del evento adverso.
34 a 39	DETAIL1 al DETAIL6	al Detalles	Detalles relativos al resultado reportado en los campos OUTCOME 1, 2 y 3.
40	COST_TE	Término COSTART	Clasificación del evento adverso de acuerdo con el código COSTART.
41	BODY_SYS	Sistema corporal	Sistema del cuerpo que afecta el evento adverso según la clasificación COSTART
42	COSTART		Clave del código COSTART para clasificar eventos adversos.

⁶ Se definió el campo SM_SDT para anotar el resultado del cálculo de la fecha de inicio del tratamiento. Este campo no se encuentra en la base de datos original.

⁷ Se definió el campo SM_EDT para anotar el resultado del cálculo de la fecha de fin del tratamiento. Este campo no se encuentra en la base de datos original.

3.3.1 Integridad de la información.

El primer paso dentro del proceso de análisis correspondió a validar la integridad de la información de todos los registros relacionados con un tratamiento. Esto implicó verificar que los datos de cada registro estuviesen completos. En este sentido, se encontró que para algunos de ellos, relacionados con un evento adverso, la fecha de inicio (AE_SDT), o de término (AE_EDT), o ambas fechas estaban incompletas (p.j. existía el mes y año de inicio o fin del evento, pero no existía el día). En otros, el dato no definido fue la fase de inicio del evento (AE_SPH=0.5), o de fin del evento (AE_EPH=0.5); e incluso se detectaron casos con fechas incompletas y fases no definidas para un mismo registro.

Por su parte, para los registros no relacionados con un evento, el dato mayormente faltante fue la fecha de inicio o de fin del tratamiento.

En virtud de lo anterior, se decidió continuar con el proceso de análisis sin considerar 23 registros cuya información estaba incompleta. La distribución de estos registros, por tratamiento, puede verse en la tabla 3.5.

Tabla 3.5 Distribución de registros de pacientes asignados a un tratamiento, clasificados en registros completos e incompletos.

Tratamiento	Registros con EA	Completos	Incompletos
A	125	114	11
B	105	100	5
C	116	114	2
	346	328	18

Tratamiento	Registros sin EA	Completos	Incompletos
A	67	65	2
B	71	70	1
C	69	67	2
	207	202	5

3.3.2 Proceso de identificación de los eventos adversos emergentes.

A partir de la información de la fase de inicio (AE_SPH) y término (AE_EPH) de los registros relacionados con un evento adverso, es posible plantear una clasificación respecto a sí cada evento puede ser considerado como emergente o no emergente. Según este planteamiento, se pueden identificar 6 casos:

- 1) El evento comenzó y terminó en la fase I.
- 2) El evento comenzó en la fase I y terminó en la fase II.
- 3) El evento comenzó en la fase I y terminó en la fase III.
- 4) El evento comenzó y terminó en la fase II.
- 5) El evento comenzó en la fase II y terminó en la fase III.
- 6) El evento comenzó y terminó en la fase III.

En la figura 3.4, se ilustran los 6 casos citados arriba.

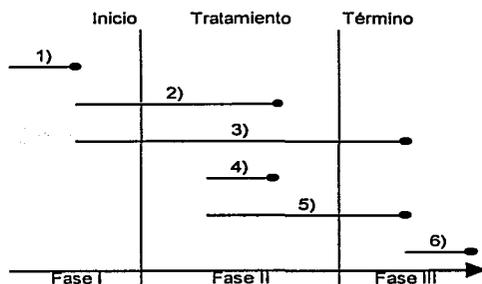


Figura 3.4 Clasificación de eventos adversos según la fase del tratamiento donde iniciaron y terminaron.

Los caso 1) y 6), no pueden considerarse eventos adversos emergentes en virtud de que no se apegan a la definición, pues se desarrollaron completamente fuera de la fase de tratamiento. Por el contrario, los caso 4) y 5) son claramente eventos adversos emergentes, ya que aparecieron por primera vez durante el periodo de tratamiento, independientemente del número de manifestaciones reportadas o de la intensidad de cada una de ellas. En lo que respecta a los casos 2) y 3), podrían ser considerados como eventos adversos emergentes sólo en dos situaciones:

- a) si la intensidad reportada después de iniciado el tratamiento fue mayor a la reportada antes de que éste hubiese dado comienzo para una misma manifestación, o
- b) si reapareció una o más veces después de que el tratamiento dio inicio, aunque no halla variado la intensidad en cada nueva manifestación.

Cabe aclarar que si un evento adverso apareció por primera vez antes de iniciado el tratamiento y desapareció después de iniciado el tratamiento para luego reaparecer, no sería considerado como de tratamiento emergente, salvo que hubiese incrementado su intensidad después de comenzado el tratamiento para la misma manifestación.

Aplicando estos criterios, fue posible agrupar los eventos adversos registrados en emergentes y no emergentes, según se expresa en la tabla 3.6.

Tabla 3.6 Clasificación de eventos adversos en emergentes y no emergentes por tratamiento.

Tratamiento	Registros	Eventos	EA emergentes	EA no emergentes
A	114	103	67	36
B	100	92	65	27
C	114	103	75	28
	328	298	207	91

Una vez identificados los eventos emergentes, se procedió a clasificar cada uno de los registros según la fase donde dio inicio. Los resultados se encuentran resumidos en la tabla 3.7.

Tabla 3.7 Clasificación de registros de eventos adversos emergentes y no emergentes según la fase del tratamiento donde iniciaron y terminaron, por tipo de tratamiento.

Clasificación del total de registros por tipo de evento según la fase de inicio y término (ver figura 3.4).

Tratamiento	EA Emergentes	1	2	3	4	5	6	Total
A	67	2	0	0	61	12	0	75
B	65	4	0	0	61	8	0	73
C	75	3	0	0	72	7	2	84
	207	9	0	0	194	27	2	232

Tratamiento	EA No Emergentes	1	2	3	4	5	6	Total
A	36	17	13	2	1	1	5	39
B	27	15	7	1	0	0	4	27
C	28	12	13	0	0	0	5	30
	91	44	33	3	1	1	14	96

Tratamiento	EA por tratamiento	1	2	3	4	5	6	Total
A	103	19	13	2	62	13	5	114
B	92	19	7	1	61	8	4	100
C	103	15	13	0	72	7	7	114
	298	53	33	3	195	28	16	328

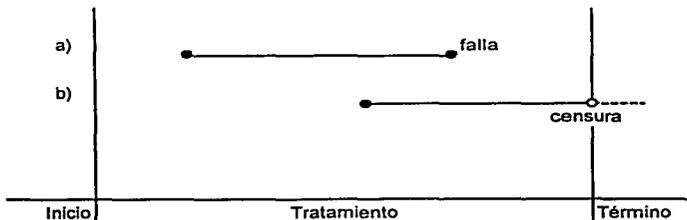
3.3.3 Criterios de selección de registros para el ajuste de los modelos de supervivencia.

3.3.3.1 Modelo del tiempo de duración.

Después de identificar los registros de eventos emergentes, se procedió a seleccionar aquéllos susceptibles de ser considerados dentro de cada uno de los modelos de supervivencia que serían construidos.

Para el ajuste del tiempo de duración, se eligieron los registros correspondientes a las manifestaciones de eventos adversos emergentes que dieron inicio durante la fase de tratamiento, tomando como referencia el valor del campo de la fase de inicio del evento, AE_SPH=2. Debe notarse que el suceso que origina la observación del fenómeno duración, es la manifestación explícita de una queja

(evento adverso) por parte de un paciente sometido a un tratamiento. Por otro lado, el suceso que determina el momento en que deja de observarse el fenómeno duración, es la manifestación explícita, por parte del mismo paciente, de haber dejado de sufrir la molestia que lo aquejaba. Gráficamente, este comportamiento se ilustra en la figura 3.5.

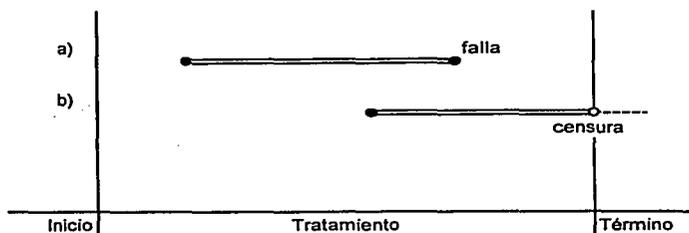


- a) EAE con una manifestación, la cual concluye dentro del periodo de estudio (falla).
 b) EAE con una manifestación, la cual se prolongó hasta el término del estudio (censura)

Figura 3.5 Registros de eventos adversos emergentes elegibles para el ajuste del tiempo de duración. Caso general.

Sin embargo, se encontraron algunos casos especiales cuya solución debe ser comentada:

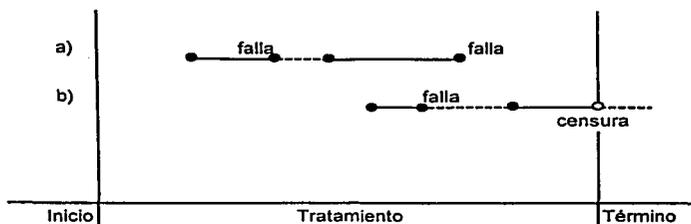
- Si para una misma manifestación, se reportaron dos o más registros iguales (registros repetidos), se seleccionó solo uno de ellos. Esta particularidad se ilustra en la figura 3.6.



- a) EAE con 2 manifestaciones iguales. Se elige solo una para el ajuste (falla).
 b) EAE con 2 manifestaciones iguales. Se elige solo una para el ajuste (censura).

Figura 3.6 Registros de eventos adversos emergentes elegibles para el ajuste del tiempo de duración. La doble línea indica que dentro de la base existen dos registros iguales para una misma manifestación; en este caso solo se selecciona un solo registro.

- Si un evento tuvo dos o más manifestaciones iniciadas dentro de la fase de tratamiento (AE_SPH=2), todos estos registros se considerarían dentro del ajuste. Esta situación queda ilustrada en la figura 3.7.



a) EAE con 2 manifestaciones diferentes. Se eligen ambas para el ajuste (fallas).

b) EAE con 2 manifestaciones diferentes: la primera es una falla y la segunda un dato censurado. Ambas se eligen para el ajuste.

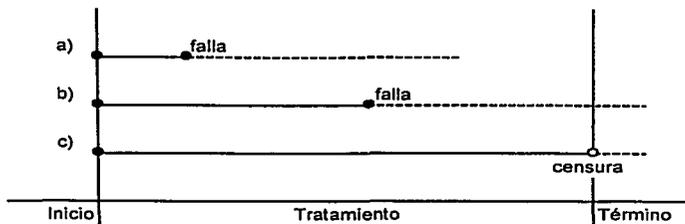
Figura 3.7 Registros de eventos adversos emergentes elegibles para el ajuste del tiempo de duración. La línea punteada que une dos o más registros, indica que se trata de un mismo evento con dos o más manifestaciones. Si estas manifestaciones comenzaron dentro de la fase de tratamiento, sus registros son elegibles para el ajuste de duración.

3.3.3.2 Modelo del tiempo anterior a la primera manifestación.

En lo que respecta a los registros elegibles para ajustar el modelo del tiempo anterior a la primera manifestación, se aplicaron dos criterios de selección de datos:

- 1) Se tomó la información sobre el tiempo transcurrido desde el comienzo del tratamiento hasta el inicio de la primera manifestación de cada evento adverso emergente ocurrido dentro del periodo de tratamiento. Para identificar el primer registro del evento, se usó como referencia el valor de la variable AE_SPH=2 y la fecha de inicio de la manifestación, AE_SDT.
- 2) Se tomó la información del tiempo que duró el tratamiento de aquellos pacientes que entraron a la Segunda Etapa del estudio y que declararon no haber padecido evento adverso alguno.

Cabe hacer notar que para este modelo, el suceso que da inicio a la observación del fenómeno es el comienzo del propio tratamiento; en este sentido, todos los pacientes que fueron seleccionados para participar en la Segunda Fase debían ser observados, independientemente si manifestaron o no eventos adversos. Por otro lado, el suceso que determina el fin de la observación del fenómeno, es la declaración explícita del paciente de haber comenzado a padecer una cierta molestia (evento adverso). En la figura 3.8 se ilustra este caso general.

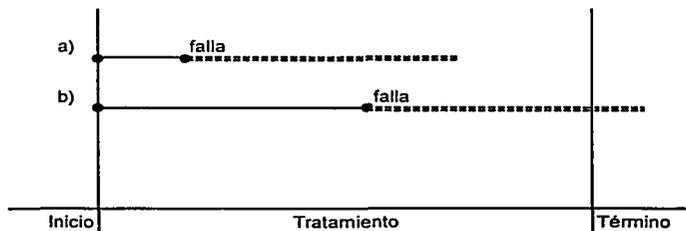


a) y b) Tiempo medido desde el inicio del tratamiento hasta la aparición de un EAE. Ambos casos se toman en cuenta para el ajuste (fallas).
 c) Tiempo medido desde el inicio del tratamiento hasta el término del mismo. Este paciente no presentó EA alguno, sin embargo esta información se toma en cuenta dentro del ajuste pues se trata de un dato censurado.

Figura 3.8 Registros de eventos adversos emergentes elegibles para el ajuste del tiempo anterior a la aparición de la primera manifestación. Caso general.

En la construcción del modelo para el tiempo anterior a la primera manifestación también se presentó una situación especial al momento de seleccionar los datos:

- Si para la primera manifestación, se hubiesen reportado dos o más registros iguales, solo se tomaría como referencia uno solo de los registros, según se muestra en la figura 3.9.



a) y b) Tiempo medido desde el inicio del tratamiento hasta la aparición de un EAE. En este caso el primer registro del evento está repetido. Ambos casos se toman en cuenta para el ajuste (fallas).

Figura 3.9 Registros de eventos adversos emergentes elegibles para el ajuste del tiempo anterior a la primera manifestación. La doble línea indica que dentro de la base existen dos registros iguales; en este caso solo se toma uno de ellos como referencia.

Una vez aplicados estos criterios de selección de información, se elaboró la tabla 3.8 donde se aprecia el total de registros de la base original que en principio serían utilizados para desarrollar ambos ajustes antes de identificar el tipo de censura operante en cada uno de los fenómenos que se están modelando.

Tabla 3.8 Clasificación de registros de eventos adversos emergentes para los ajustes de duración y manifestación después de aplicar los criterios generales de selección para la construcción de los modelos de supervivencia, por tipo de tratamiento.

Registros por tipo de evento según la fase de inicio y término (ver figura 3.4).

Tratamiento	EA Emergentes	1	2	3	4	5	6	Total
A	67	2	0	0	60	10	0	72
B	65	4	0	0	61	8	0	73
C	75	3	0	0	71	7	2	83
	207	9	0	0	192	25	2	228

Tratamiento	Registros para Duración
A	70
B	69
C	78
	217

Tratamiento	Registros para la Primera Manifestación	Registros relacionados con un EA	Registros no relacionados con un EA
A	133	68	65
B	135	65	70
C	142	75	67
	410	208	202

3.3.4 Determinación del tipo de censura.

Como explicado en la sección 2.2.6 del capítulo II; un dato censurado recoge parcialmente información sobre la variable aleatoria que se está estudiando, básicamente porque para dicho dato, la ocurrencia del evento final no pudo verificarse dentro del periodo de observación por alguna razón. En este sentido, algunos de los registros seleccionados se ajustan a esta definición.

En el caso del modelo para el tiempo de duración, un dato se consideró como censurado cuando la fase y fecha de inicio de dicha manifestación sí eran conocidas, pero la fase y fecha de término, no estaban definidas (AE_EDT=. y AE_EPH=.)⁸. En consecuencia, se sustituyó la fecha de fin de la manifestación del evento por la fecha de término del estudio y se asumió que la fase de término de la manifestación es AE_EPH=2.

En lo que respecta al modelo para el tiempo anterior a la primera manifestación, un dato fue catalogado como censurado cuando, a pesar de que el paciente se sometió a un cierto tratamiento, no reportó la ocurrencia de evento adverso alguno durante todo el periodo. En este caso, la fecha de término de la observación se asoció con la fecha de término del estudio.

⁸ A diferencia de un registro incompleto; donde la fecha de fin de la manifestación (AE_EDT) es parcialmente conocida (existe el mes y año, pero no el día) y la fase de fin de la manifestación no está definida (AE_EPH=0.5); para los registros censurados no se conoce ningún dato asociado ni con la fecha de fin de la manifestación ni con la fase de término de la misma (AE_EDT=. y AE_EPH=.).

Cabe señalar que en ambos casos, la censura identificada corresponde al tipo I, donde el valor fijo t_c coincide con la fecha de fin del tratamiento.

En la tabla 3.9 se muestra un resumen de los datos censurados y no censurados por tipo de ajuste y tratamiento después de aplicar los criterios antes señalados.

Tabla 3.9 Clasificación de registros de eventos adversos emergentes en censurados y no censurados por tipo de ajuste y tratamiento.

Tratamiento	Registros para duración.				Registros para la primera manifestación.			
	Totales	No censurados	Censurados	% de datos censurados	Totales	No censurados	Censurados	% de datos censurados
A	70	56	14	20.00%	133	68	65	48.87%
B	69	61	8	11.59%	135	65	70	51.85%
C	78	68	10	12.82%	142	75	67	47.18%
	217	185	32		410	208	202	

Por último, se mencionará que la tabla 3.9 refleja la distribución de los datos por tipo de tratamiento que se utilizarán como base en el ajuste de los respectivos modelos de supervivencia. El ajuste a este nivel, es válido para todas las manifestaciones de los eventos emergentes registrados durante el ensayo clínico.

3.3.5 Selección de los eventos adverso más frecuentes.

Se ha visto que el análisis de supervivencia es una herramienta eficaz que permite comprender mejor el comportamiento y desarrollo de los eventos adverso con el paso del tiempo. En este sentido, si nos referimos a enfermedades que requieren una exposición constante y prolongada a las sustancias activas (p.ej. la hiperlipidemia), el tiempo que tardan en manifestarse y la duración de los eventos adversos que lleguen a presentarse se convierten en dos aspectos que deben cuidarse con especial atención, particularmente en aquellos casos frecuentes o en aquellos sistemas corporales con mayor incidencia de casos. Con base en este argumento, se procedió a desarrollar modelos de supervivencia para el tiempo de duración y el tiempo anterior a la primera manifestación para los registros de manifestaciones de los dos sistemas corporales con mayor número de reportes de eventos adversos emergentes, por tipo de tratamiento. Cabe señalar que el ajuste a este nivel para el tiempo de manifestación, no incluye los registros de pacientes que no presentaron evento adverso alguno en virtud de que no es posible asociar estos registros a ningún sistema específico. Así mismo, se aclara que se tomó la decisión de no desarrollar ajustes por tipo de evento, ya que en todos los casos el número de registros asociados no fue lo suficientemente grande (menos de 15 registros para el caso con mayor información), y que es lo que comúnmente se hace.

El resultado de la clasificación de registros de eventos por sistema corporal se presenta en la tabla 3.10. La variable de la base de datos utilizada para llevar a cabo esta clasificación fue BODY_SYS (sistema corporal).

Tabla 3.10 Clasificación de registros de eventos adversos emergentes por sistema corporal, tipo de ajuste y tratamiento.

Sistema Corporal	Registros para duración.					Registros para manifestación. (solo de aquellos relacionados con un sistema corporal)				
	A	B	C	Total	%	A	B	C	Total	%
El cuerpo como un todo	24	23	15	62	28.57%	22	23	14	59	28.37%
Cardiovascular	5	8	2	15	6.91%	5	7	2	14	6.73%
Digestivo	14	15	26	55	25.35%	14	15	25	54	25.96%
Endócrino	0	1	0	1	0.46%	0	1	0	1	0.48%
Hemático y linfático	0	0	1	1	0.46%	0	0	1	1	0.48%
Metabólico/ nutricional	9	1	5	15	6.91%	9	1	5	15	7.21%
Musculosquelético	3	3	6	12	5.53%	3	2	6	11	5.29%
Nervioso	5	5	7	17	7.83%	5	5	7	17	8.17%
Respiratorio	2	3	10	15	6.91%	2	3	9	14	6.73%
Piel y anexos	4	3	0	7	3.23%	4	2	0	6	2.88%
Sentidos	1	2	1	4	1.84%	1	2	1	4	1.92%
Urogenital	3	5	5	13	5.99%	3	4	5	12	5.77%
	70	69	78	217		68	65	75	208	

Los sistemas elegidos para realizar el ajuste final fueron: "el cuerpo como un todo"⁹ y el digestivo. El resumen de los registros censurados y no censurados, por sistema, tratamiento y tipo de ajuste, se observa en la tabla 3.11.

Tabla 3.11 Clasificación de registros de eventos adversos emergentes, en censurados y no censurados, para los sistemas "el cuerpo como un todo" y digestivo, por tipo de ajuste y tratamiento.

Tratamiento	"El cuerpo como un todo"					
	Duración			Manifestación		
	No Censurado	Censurado	%	No Censurado	Censurado	%
A	22	2	9.09%	22	0	0.00%
B	20	3	15.00%	23	0	0.00%
C	13	2	15.38%	14	0	0.00%
	55	7		59	0	

Tratamiento	Digestivo					
	Duración			Manifestación		
	No Censurado	Censurado	%	No Censurado	Censurado	%
A	12	2	16.67%	14	0	0.00%
B	14	1	7.14%	15	0	0.00%
C	22	4	18.18%	25	0	0.00%
	48	7		54	0	

⁹ "El cuerpo como un todo", es la traducción más adecuada que se hizo del término en inglés "body as a whole", tomado de la base de datos original. Este sistema corporal engloba todos aquellos eventos adversos que se vinculan con dos o más sistemas simultáneamente, y cuya manifestación afecta a todo el cuerpo.

Al retomar los antecedentes del ensayo clínico, se recuerda que ya se esperaba la presencia de algunos eventos relacionados con los tratamientos (ver sección 1.4.1 del capítulo I). De entre estos eventos, la elevación en los niveles de transaminasas hepáticas (pruebas hepáticas anormales) además de otros malestares gastrointestinales, están englobados dentro de los eventos del sistema digestivo. En lo que respecta a los dolores y acalambramientos musculares, se asociaron como parte de los eventos del sistema "el cuerpo como un todo". Cabe señalar que el incremento en los niveles de creatinina fosfoquinética, evento adverso especialmente esperado, no figura dentro de ninguno de los dos sistemas seleccionados¹⁰. En las tablas 3.12a y 3.12b, se muestra una distribución detallada de los registros asociados con cada evento adverso por sistema, tratamiento y tipo de ajuste.

Tabla 3.12a Clasificación de eventos adversos emergentes del sistema "el cuerpo como un todo" por tipo de ajuste y tratamiento.

Registros para el sistema "el cuerpo como un todo"						
Evento adverso específico	Duración			Manifestación		
	A	B	C	A	B	C
Vientre o abdomen distendido	2	0	0	2	0	0
Dolor abdominal	5	4	4	4	4	4
Heridas por accidente	3	4	1	3	4	1
Astenia	1	3	2	1	3	2
Dolor de espalda	2	1	0	1	1	0
Dolor en el pecho	2	1	0	2	1	0
Dolor interno de pecho	1	0	0	1	0	0
Fiebre	0	2	1	0	2	1
Gripe y su sintomatología	2	2	2	2	2	2
Dolor de cabeza	3	4	3	3	4	2
Niveles hormonales alterados	1	0	0	1	0	0
Dolor en las piernas	1	2	1	1	2	1
Dolor (malestar generalizado)	0	0	1	0	0	1
Procedimientos quirúrgicos	1	0	0	1	0	0
	24	23	15	22	23	14

¹⁰ El incremento en los niveles de creatinina fosfoquinética se clasificó como parte del sistema metabólico/nutricional. En este sistema, se contabilizaron 15 registros, tanto para el ajuste del tiempo de duración como para el ajuste del tiempo de manifestación. La distribución de los mismos se aprecia en la tabla 3.10. Para el evento en particular, se detectaron 8 registros cuya distribución queda como sigue: 4 para el tratamiento A, 1 para el tratamiento B y 3 para el tratamiento C.

Tabla 3.12b Clasificación de eventos adversos emergentes del sistema digestivo por tipo de ajuste y tratamiento.

Registros para el sistema Digestivo						
Evento adverso específico	Duración			Manifestación		
	A	B	C	A	B	C
Anorexia	0	0	1	0	0	1
Colelitiasis	0	1	0	0	1	0
Colitis	2	0	0	2	0	0
Constipación	0	2	3	0	2	3
Diarrea	2	3	4	2	3	4
Boca seca	0	2	1	0	2	1
Dispepsia	2	2	6	2	2	6
Enteritis	0	1	1	0	1	1
Flatulencia	1	2	1	1	2	1
Gastritis	0	1	0	0	1	0
Pruebas hepáticas anormales.	3	0	2	3	0	2
Nauseas	3	1	3	3	1	3
Úlcera péptica	0	0	0	0	0	0
Sed	0	0	0	0	0	0
Vómito	1	0	4	1	0	3
	14	15	26	14	15	25

A partir de este punto, todas las estimaciones que se realizaron se basaron en la información que fue seleccionada sobre las manifestaciones de los eventos adversos emergentes, según los criterios señalados en las secciones precedentes. En este sentido, la variable aleatoria T, que se usará en adelante, representa el tiempo entre el inicio y el término de las manifestaciones de los eventos adversos emergentes, en el caso del ajuste para el tiempo de duración; y el tiempo entre el inicio del tratamiento e inicio de la primera manifestación de un evento adverso emergente, para el ajuste del tiempo anterior a la primera manifestación.

En el ANEXO 4 se muestra esta información final depurada.

Cabe señalar que el proceso de ajuste de los modelos de supervivencia, se abordó asumiendo que no hay covariables que influyan el desarrollo de los fenómenos que se están estudiando. Lo anterior queda validado por el diseño mismo del ensayo clínico, el cual no permite influencia alguna de variables externas ya que, desde un principio, se controlaron todos los factores que pudiesen haber afectado su desarrollo (ver secciones 1.4 y siguientes del capítulo I). En virtud de lo anterior, la variable T a la que se ha hecho referencia explica, por sí misma, la duración o el tiempo de manifestación de los eventos adversos emergentes del ensayo clínico ya mencionado.

3.4 Estimación de la función de supervivencia por el proceso no paramétrico de Kaplan–Meier.

El siguiente paso del proceso de ajuste, corresponde a la estimación de la función de supervivencia por medio del estimador no paramétrico de Kaplan–Meier. La justificación matemática de este modelo se trató en la sección 2.3.1.3 del capítulo II.

Debe recordarse que el estimador K-M, proporciona una buena aproximación sobre la forma de la distribución real además de servir como base para desarrollar un ajuste paramétrico posterior.

La estimación se llevó a cabo en dos etapas; en la primera, se aplicó a los 6 conjuntos de datos generales por tipo de ajuste (tiempo de duración y tiempo anterior a la primera manifestación) y tipo de tratamiento (A, B y C). Los resultados se han resumido en las tablas 3.13a y 3.13b; los gráficos respectivos se aprecian en las figuras 3.10a y 3.10b.

En la segunda etapa, se aplicó a los 12 conjuntos de datos cuyos criterios de agrupación fueron: tipo de ajuste (tiempo de duración y tiempo anterior a la primera manifestación), sistema corporal (“el cuerpo como un todo” y digestivo) y tratamiento (A, B y C). Los resultados alcanzados pueden verse en las tablas 3.14 a 3.17. También se incluyeron los valores de cada variable por tipo de ajuste, así como el resultado de la función de supervivencia empírica obtenida. Por su parte, los gráficos correspondientes a cada función ajustada se pueden observar en las figuras 3.11 a 3.14.

Tabla 3.13a Estimación del modelo K-M para el ajuste del tiempo de duración de los EA Emergentes por tipo de tratamiento.

Ajuste del modelo Kaplan-Meier para el tiempo de duración de los EA Emergentes.					
A		B		C	
t	$\hat{s}(t)$	t	$\hat{s}(t)$	t	$\hat{s}(t)$
0	1.0000	0	1.0000	0	1.0000
1	0.9429	1	0.8696	1	0.9487
2	0.9000	2	0.8551	2	0.8974
3	0.8143	3	0.7681	3	0.7436
4	0.7857	4	0.7246	4	0.6923
5	0.7429	5	0.6812	5	0.6795
6	0.6857	6	0.6377	7	0.6667
7	0.6714	7	0.6232	8	0.5882
8	0.6429	8	0.5507	9	0.5752
9	0.6286	9	0.5362	10	0.5621
12	0.6143	10	0.5072	12	0.5359
13	0.5857	11	0.4928	14	0.4967
15	0.5714	14	0.4348	15	0.4706
16	0.5571	15	0.4203	16	0.4575
18	0.5429	24	0.4058	20	0.4314
19	0.5286	25	0.3913	21	0.4183
24	0.5000	27	0.3478	24	0.4048
28	0.4853	30	0.3327	27	0.3643
29	0.4383	32	0.2852	28	0.3373
30	0.4227	33	0.2693	29	0.3104
31	0.3444	34	0.2376	31	0.2834
32	0.3287	35	0.2218	32	0.2699
36	0.3131	36	0.2060	35	0.2415
42	0.2966	44	0.1888	36	0.1989
45	0.2801	46	0.1716	40	0.1846
46	0.2472	49	0.1373	51	0.1693
48	0.2307	56	0.1144	52	0.1539
57	0.2097	59	0.0687	57	0.1197
58	0.1864			65	0.1026
62	0.1631			92	0.0821
85	0.1305				

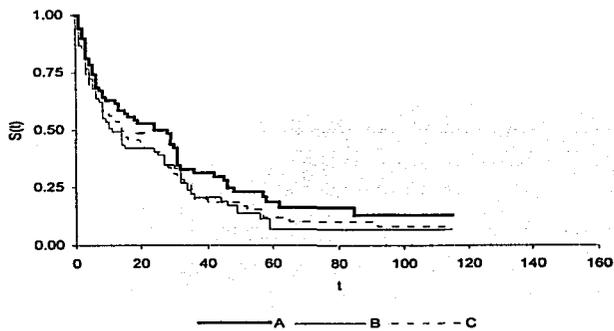


Figura 3.10a Gráfica del ajuste del modelo K-M para el tiempo de duración de los EA Emergentes por tipo de tratamiento.

Tabla 3.13b Estimación del modelo K-M para el ajuste del tiempo anterior a la primera manifestación de los EA Emergentes por tipo de tratamiento.

Ajuste del modelo Kaplan-Meier para el tiempo anterior a la primera manifestación de los EA emergentes.					
A		B		C	
t	$\hat{s}(t)$	t	$\hat{s}(t)$	t	$\hat{s}(t)$
0	1.0000	0	1.0000	0	1.0000
2	0.9924	3	0.9852	3	0.9859
4	0.9848	4	0.9778	4	0.9577
5	0.9773	6	0.9704	7	0.9507
9	0.9697	9	0.9629	8	0.9437
11	0.9621	12	0.9554	10	0.9366
15	0.9545	14	0.9480	14	0.9155
18	0.9470	16	0.9405	17	0.9085
22	0.9394	17	0.9330	21	0.9014
23	0.9242	18	0.9256	22	0.8732
25	0.9167	24	0.9181	24	0.8662
27	0.9014	25	0.9107	26	0.8521
29	0.8475	26	0.9032	29	0.7958
30	0.8398	27	0.8957	30	0.7887
32	0.8321	28	0.8883	33	0.7817
34	0.8166	29	0.8733	38	0.7746
36	0.8089	31	0.8584	47	0.7535
38	0.8011	34	0.8435	48	0.7465
40	0.7855	37	0.8285	50	0.6972
42	0.7778	41	0.8136	54	0.6830
43	0.7622	43	0.8062	56	0.6758
46	0.7466	46	0.7987	57	0.6474
49	0.7311	51	0.7912	61	0.6403
51	0.7233	53	0.7688	64	0.6332
52	0.7078	55	0.7614	67	0.6260
53	0.7000	56	0.7166	70	0.6189
56	0.6922	57	0.6942	72	0.6047
57	0.6300	58	0.6867	76	0.5976
58	0.6144	59	0.6718	77	0.5834
63	0.6066	68	0.6569	80	0.5762
68	0.5989	70	0.6494	81	0.5691
71	0.5911	80	0.6419	82	0.5620
72	0.5833	82	0.6041	85	0.5478
75	0.5678	85	0.5814	87	0.5407
82	0.5600	99	0.5588	92	0.5193
83	0.5444	102	0.5512	93	0.5122
85	0.5211	104	0.5437	95	0.5051
86	0.5133	107	0.5207	97	0.4980
87	0.4978	110	0.5129	102	0.4909
89	0.4822			111	0.4753
105	0.4744			113	0.4570

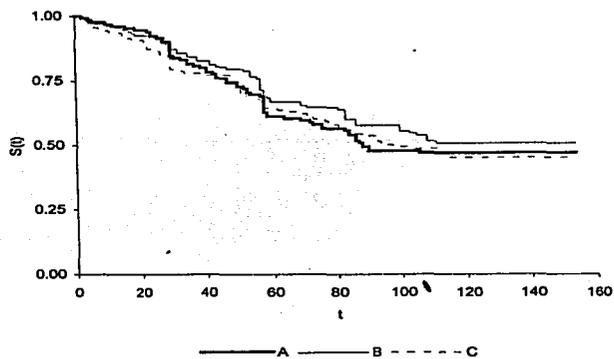


Figura 3.10b Gráfica del ajuste del modelo K-M para el tiempo anterior a la primera manifestación de los EA Emergentes por tipo de tratamiento.

Tabla 3.14 Estimación del modelo K-M para el ajuste del tiempo de duración y el sistema "el cuerpo como un todo", por tipo de tratamiento.

Tiempo de duración para los EAE del sistema "el cuerpo como un todo".						
A			B		C	
t	$\hat{S}(t)$	t	$\hat{S}(t)$	t	$\hat{S}(t)$	
0	0	1.0000	0	1.0000	0	1.0000
1	1	0.9583	3		1	0.9333
2	2	0.9167	3		3	
3	3		3		3	0.8000
4	3	0.8333	3	0.8261	4	0.7333
5	4	0.7917	4		10	0.6667
6	5	0.7500	4	0.7391	12	0.6000
7	6		5	0.6957	14	0.5333
8	6		6		16	0.4667
9	6	0.6250	6	0.6087	24	0.4000
10	8	0.5833	7	0.5652	27	0.3333
11	12	0.5417	8		31	0.2667
12	13		8		51	0.2000
13	13	0.4583	8		52	0.1333
14	15	0.4167	8	0.3913	55*	
15	16	0.3750	10	0.3478	115*	
16	19	0.3333	14	0.3043		
17	28	0.2917	25	0.2609		
18	31	0.2500	30	0.2174		
19	32	0.2083	30*			
20	42	0.1667	32	0.1630		
21	45	0.1250	34	0.1087		
22	46	0.0833	42*			
23	73*		71*			
24	92*					

*indica dato censurado.

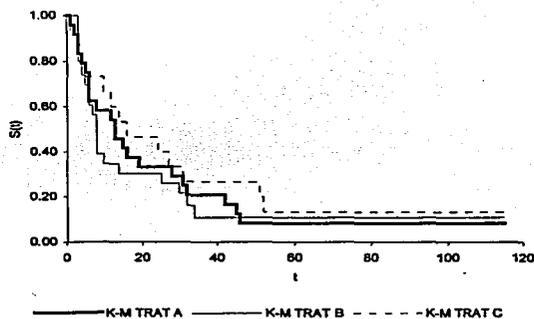


Figura 3.11 Gráfica del ajuste del modelo K-M para el tiempo de duración y el sistema "el cuerpo como un todo", por tipo de tratamiento.

Tabla 3.15 Estimación del modelo K-M para el ajuste del tiempo de duración y el sistema digestivo, por tipo de tratamiento.

Tiempo de duración para los EAE del sistema digestivo.							
		A		B		C	
	t	$\hat{S}(t)$	t	$\hat{S}(t)$	t	$\hat{S}(t)$	
0	0	1.0000	0	1.0000	0	1.0000	
1	2		2	0.9333	1	0.9615	
2	2	0.8571	6	0.8667	2		
3	3		8	0.8000	2		
4	3		14		2	0.8462	
5	3	0.6429	14	0.6667	3		
6	4	0.5714	24	0.6000	3		
7	18	0.5000	27	0.5333	3		
8	24		32		3		
9	24	0.3571	32	0.4000	3		
10	28*		33	0.3333	3		
11	31	0.2679	36	0.2667	3	0.5769	
12	36	0.1786	49	0.2000	7*		
13	85	0.0893	53*		8		
14	109*		59		8	0.4945	
15			59	0.0000	15	0.4533	
16					20		
17					20	0.3709	
18					27	0.3297	
19					28	0.2885	
20					29	0.2473	
21					34*		
22					36		
23					36		
24					36	0.0989	
25					49*		
26					83*		

*indica dato censurado.

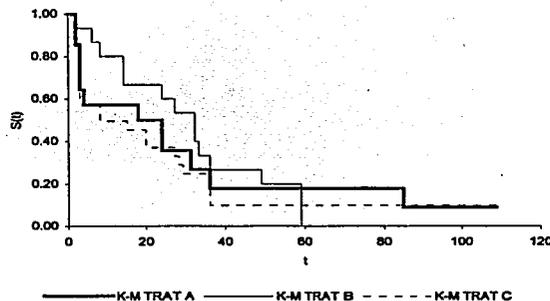


Figura 3.12 Gráfica del ajuste del modelo K-M para el tiempo de duración y el sistema digestivo, por tipo de tratamiento.

Tabla 3.16 Estimación del modelo K-M para el ajuste del tiempo anterior a la primera manifestación y el sistema "el cuerpo como un todo", por tipo de tratamiento.

Tiempo de manifestación para los EAE del sistema "el cuerpo como un todo".						
A			B		C	
t	$\hat{S}(t)$	t	$\hat{S}(t)$	t	$\hat{S}(t)$	
0	0	1.0000	0	1.0000	0	1.0000
1	9	0.9545	6	0.9565	3	0.9286
2	11	0.9091	16	0.9130	4	0.8571
3	15	0.8636	18	0.8696	22	0.7857
4	18	0.8182	24	0.8261	29	0.7143
5	23	0.7727	28	0.7826	30	0.6429
6	29		31	0.7391	50	0.5714
7	29		37	0.6957	54	0.5000
8	29	0.6364	41	0.6522	61	0.4286
9	30	0.5909	53		70	0.3571
10	32	0.5455	53	0.5652	76	0.2857
11	36	0.5000	56		85	
12	40	0.4545	56		85	0.1429
13	42	0.4091	56	0.4348	92	0.0714
14	52	0.3636	59	0.3913	95	0.0000
15	53	0.3182	68	0.3478		
16	57	0.2727	70	0.3043		
17	71	0.2273	80	0.2609		
18	75		82			
19	75	0.1364	82	0.1739		
20	85	0.0909	85	0.1304		
21	89	0.0455	99	0.0870		
22	105	0.0000	107	0.0435		
23			110	0.0000		

*indica dato censurado.

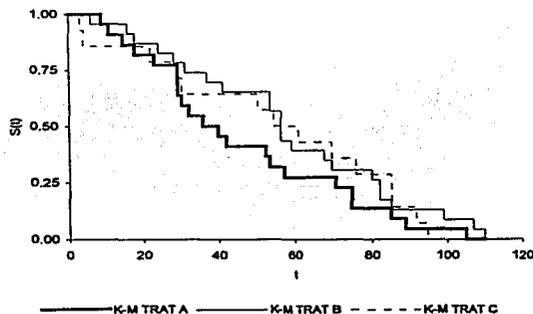


Figura 3.13 Gráfica del ajuste del modelo K-M para el tiempo anterior a la primera manifestación y el sistema "el cuerpo como un todo", por tipo de tratamiento.

Tabla 3.17 Estimación del modelo K-M para el ajuste del tiempo anterior a la primera manifestación y el sistema digestivo, por tipo de tratamiento.

Tiempo de manifestación para los EAE del sistema digestivo.							
A			B		C		
t	$\hat{s}(t)$	t	$\hat{s}(t)$	t	$\hat{s}(t)$		
0	0	0	1.0000	0	1.0000	0	1.0000
1	4	0.9286	3	0.9333	3	0.9600	
2	5	0.8571	4	0.8667	4		
3	23	0.7857	9	0.8000	4		
4	27	0.7143	12	0.7333	4	0.8400	
5	29		34	0.6667	21	0.8000	
6	29	0.5714	43	0.6000	24	0.7600	
7	38	0.5000	46	0.5333	29		
8	43		51	0.4667	29		
9	43	0.3571	53	0.4000	29		
10	49		59	0.3333	29		
11	49	0.2143	82		29	0.5600	
12	57	0.1429	82	0.2000	33	0.5200	
13	82	0.0714	85	0.1333	38	0.4800	
14	86	0.0000	99		50	0.4400	
15			99	0.0000	54	0.4000	
16					57	0.3600	
17					64	0.3200	
18					67	0.2800	
19					72	0.2400	
20					77	0.2000	
21					92	0.1600	
22					97	0.1200	
23					102	0.0800	
24					111	0.0400	
25					113	0.0000	

*indica dato censurado.

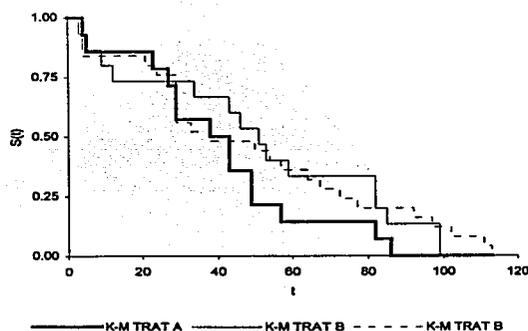


Figura 3.14 Gráfica del ajuste del modelo K-M para el tiempo anterior a la primera manifestación y el sistema digestivo, por tipo de tratamiento.

3.5 Estimación de la función de supervivencia aplicando los modelos paramétricos Exponencial y Weibull.

Continuando con el proceso de ajuste, se estimó la función de supervivencia $S(t)$, implementando los modelos paramétricos Exponencial y Weibull para el tiempo de duración y el tiempo anterior a la primera manifestación de los eventos adversos emergentes. Primero, se estimó el valor de los parámetros para cada modelo seleccionado utilizando una aproximación por métodos gráficos. Posteriormente, se mejoró esta estimación aplicando el método de máxima verosimilitud.

3.5.1 Aproximación por métodos gráficos.

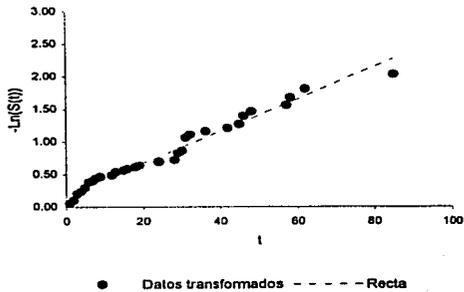
Como se mencionó anteriormente, los métodos gráficos proporcionan una evidencia objetiva sobre la forma paramétrica de la información que se está analizando, además de una estimación burda del valor de los parámetros del modelo que halla sido seleccionado.

Con la finalidad de verificar si las distribuciones elegidas son adecuadas para describir el tiempo de duración y el tiempo anterior a la primera manifestación de los eventos adversos, se aplicó a cada conjunto de datos la transformación asociada con cada modelo (ver sección 2.4.1 del capítulo II). Para realizar lo

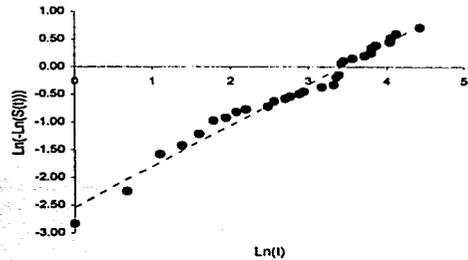
anterior, se asumió que el estimador de $S(t)$ es $\hat{S}(t)$ obtenido por el proceso de Kaplan-Meier. Se graficó t contra $\text{Ln}(\hat{S}(t))$, y $\text{Ln}(t)$ contra $\text{Ln}(-\text{Ln}\hat{S}(t))$ para compara la forma de los datos transformados con la forma típica de la recta que caracteriza a cada una de las funciones seleccionadas. Finalmente, se estimaron los valores de los parámetros de cada modelo a partir de los valores de la pendiente y la ordenada al origen de una recta ajustada sobre los puntos transformados. Para ilustrar este proceso, a continuación se muestran las figuras 3.15 y 3.16 donde es posible apreciar la tendencia gráfica de los datos después de aplicar las respectivas transformaciones; tanto en el ajuste general del tiempo de duración como del tiempo anterior a la primera manifestación. También se ha incluido el valor estimado de los parámetros. Por otro lado, las gráficas correspondientes por tipo de fenómeno, sistema y tratamiento se han omitido. En su lugar solo se muestra la tabla 3.18 donde aparecen los parámetros que fueron estimados.

Concluido este proceso, pudo verificarse que ambos modelos describen razonablemente bien los fenómenos que se están estudiando. Posteriormente se explicará la forma cómo se eligió uno de ellos para cada sistema y tipo de tratamiento.

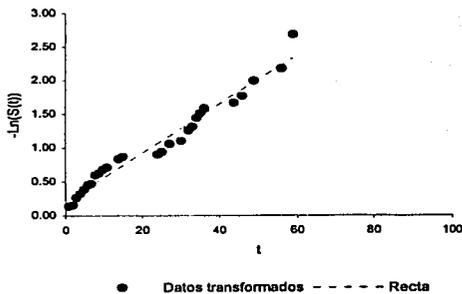
Por último se mencionará que, en el caso del modelo Weibull, los parámetros encontrados por métodos gráficos se utilizaron como valores de arranque en la implementación del proceso iterativo Newton-Raphson.



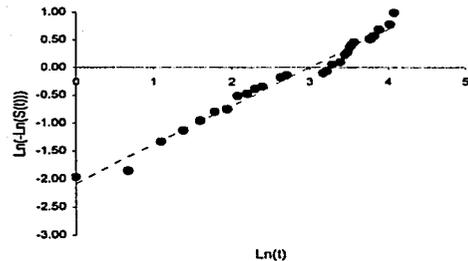
A: a) Parámetros: $\hat{\lambda} = 0.0247$



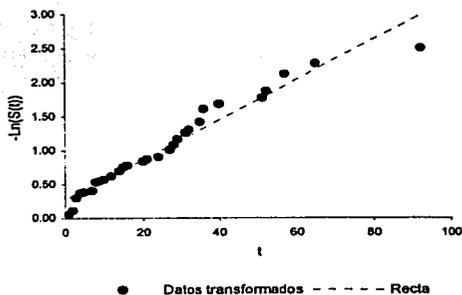
b) Parámetros: $\hat{\lambda} = 0.0784$ y $\hat{\gamma} = 0.7412$



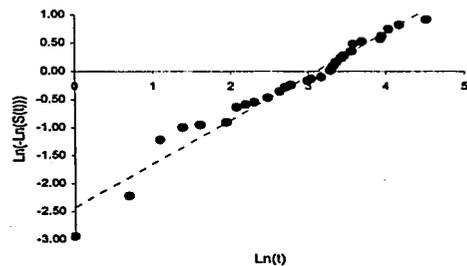
B: a) Parámetros: $\hat{\lambda} = 0.0362$



b) Parámetros: $\hat{\lambda} = 0.1242$ y $\hat{\gamma} = 0.6929$

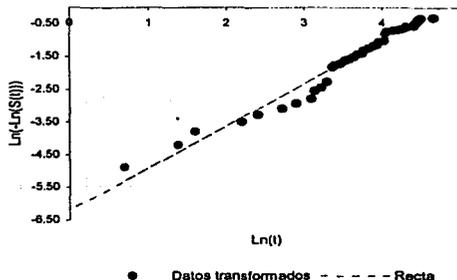
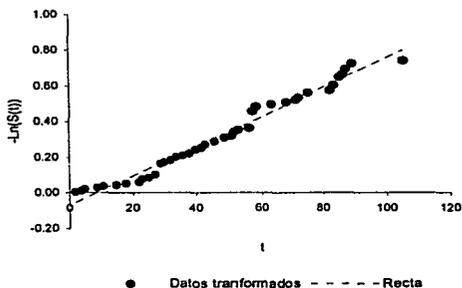


C: a) Parámetros: $\hat{\lambda} = 0.0295$



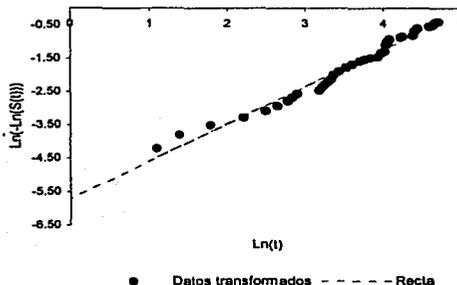
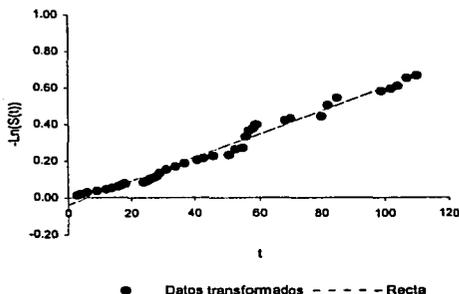
b) Parámetros: $\hat{\lambda} = 0.0874$ y $\hat{\gamma} = 0.7819$

Figura 3.15 Gráficas de los datos transformados para el ajuste del tiempo de duración de los EA Emergentes de los tratamientos A, B y C según los modelos Exponencial (a) y Weibull (b); así como las respectivas rectas ajustadas y los valores estimados de sus parámetros.



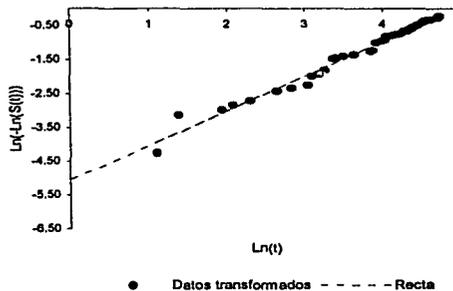
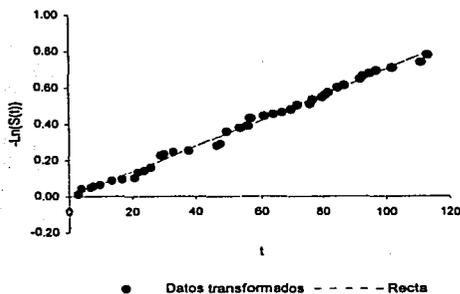
A: a) Parámetros: $\hat{\lambda} = 0.0084$

b) Parámetros: $\hat{\lambda} = 0.0021$ y $\hat{\gamma} = 1.2841$



B: a) Parámetros: $\hat{\lambda} = 0.0064$

b) Parámetros: $\hat{\lambda} = 0.0033$ y $\hat{\gamma} = 1.1167$



C: a) Parámetros: $\hat{\lambda} = 0.0071$

b) Parámetros: $\hat{\lambda} = 0.0064$ y $\hat{\gamma} = 1.0198$

Figura 3.16 Gráficas de los datos transformados para el ajuste del tiempo anterior a la primera manifestación de los EA Emergentes de los tratamientos A, B y C, según los modelos Exponencial (a) y Weibull (b); así como las respectivas rectas ajustadas y los valores estimados de sus parámetros.

Tabla 3.18 Valores estimados, por métodos gráficos, de los parámetros de los modelos Exponencial y Weibull por tipo de ajuste, sistema y tipo de tratamiento (A, B y C).

		A		B		C	
		Exponencial	Weibull	Exponencial	Weibull	Exponencial	Weibull
Tiempo de duración para los EAE del sistema "el cuerpo como un todo".							
Parámetros	$\hat{\lambda}$	0.0458	0.0568	0.0513	0.102	0.0338	0.0808
	$\hat{\gamma}$		0.9754		0.860		0.7876
Tiempo anterior a la primera manifestación de los EAE del sistema "el cuerpo como un todo".							
Parámetros	$\hat{\lambda}$	0.0327	0.0016	0.0282	0.001	0.0229	0.0285
	$\hat{\gamma}$		1.6649		1.557		0.8517
Tiempo de duración para los EAE del sistema digestivo.							
Parámetros	$\hat{\lambda}$	0.0262	0.1645	0.0335	0.028	0.0472	0.0841
	$\hat{\gamma}$		0.6055		1.004		0.8650
Tiempo anterior a la primera manifestación de los EAE del sistema digestivo.							
Parámetros	$\hat{\lambda}$	0.0349	0.0161	0.0200	0.033	0.0252	0.0177
	$\hat{\gamma}$		1.0890		0.824		1.0210

3.5.2 Estimación de los modelos Exponencial y Weibull.

Tomando los resultados desarrollados en la sección 2.3.2.4 del capítulo II, se calcularon los valores de los parámetros máximo verosímiles para las funciones seleccionadas. En el caso Weibull, se aplicó el algoritmo de Newton-Raphson para encontrar una aproximación numérica de λ y γ , iniciando este proceso en los

valores $\hat{\lambda}$ y $\hat{\gamma}$ estimados por métodos gráficos (ver la sección 2.3.2.4.1 del capítulo II y la sección 3.5.1 del capítulo III).

Los valores encontrados se aprecian en las tablas 3.19 y 3.20 en el caso general; y 3.21 a 3.24 en los casos por sistema corporal. Cabe señalar que para el modelo Weibull, los valores obtenidos son el resultado de aplicar el método numérico citado durante 10 ciclos.

Por otro lado, las gráficas donde se comparan las funciones Exponencial y Weibull resultantes contra el modelo Kaplan-Meier, puede verse en las figuras 3.17a, 3.17b y 3.17c para el tiempo de duración; y 3.18a, 3.18b y 3.18c. para el tiempo de manifestación. Por su parte, en las figuras 3.19 a 3.22; tanto a), como b) y c); se muestran las gráficas comparativas por tipo de ajuste y sistema para cada uno de los tres tratamientos.

Tabla 3.19 Valores de los estimadores de los modelos Exponencial y Weibull, obtenidos por máxima verosimilitud para el ajuste del tiempo de duración de los EA Emergentes.

		Tiempo de duración para los EAE					
		A		B		C	
Parámetros	$\hat{\lambda}$	Exponencial	Weibull	Exponencial	Weibull	Exponencial	Weibull
		$\hat{\gamma}$	0.0283	0.0601	0.0419	0.0809	0.0362
			0.7966		0.8099		0.7953

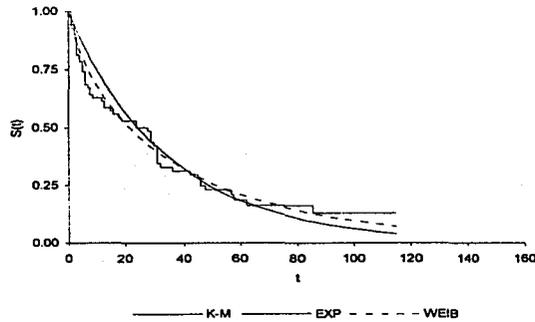


Figura 3.17a Gráfica de $\hat{S}(t)$ para el ajuste del tiempo de duración de los EA Emergentes y el tratamiento A. Los modelos que están siendo comparados son: K-M, Exp(0.0283) y Weib(0.0601, 0.7966).

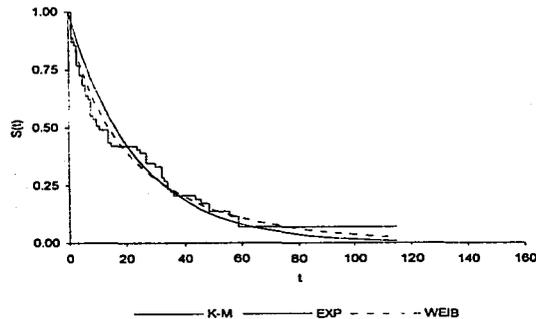


Figura 3.17b Gráfica de $\hat{S}(t)$ para el ajuste del tiempo de duración de los EA Emergentes y el tratamiento B. Los modelos que están siendo comparados son: K-M, Exp(0.0419) y Weib(0.0809, 0.8099).

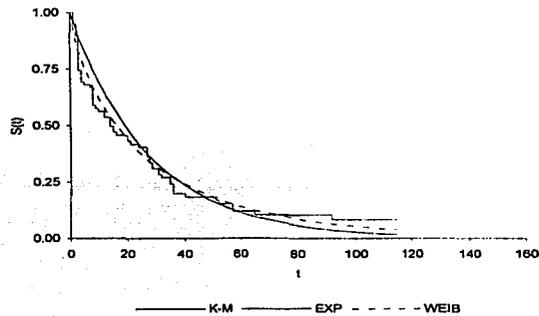


Figura 3.17c Gráfica de $\hat{S}(t)$ para el ajuste del tiempo de duración de los EA Emergentes y el tratamiento C. Los modelos que están siendo comparados son: K-M, Exp(0.0362) y Weib(0.0760, 0.7953).

Tabla 3.20 Valores de los estimadores de los modelos Exponencial y Weibull, obtenidos por máxima verosimilitud para el ajuste del tiempo anterior a la primera manifestación de los EA Emergentes.

		Tiempo anterior a la primera manifestación de los EAE					
		A		B		C	
		Exponencial	Weibull	Exponencial	Weibull	Exponencial	Weibull
Parámetros	$\hat{\lambda}$	0.0066	0.0027	0.0057	0.0025	0.0066	0.0054
	$\hat{\gamma}$		1.1936		1.1842		1.0447

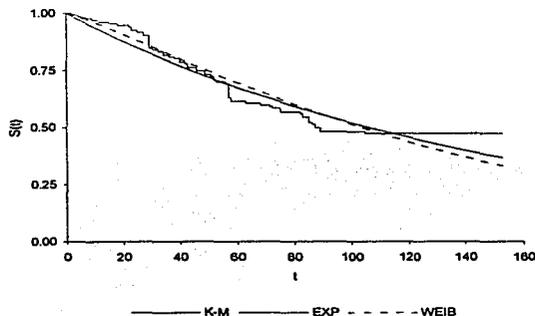


Figura 3.18a Gráfica de $\hat{S}(t)$ para el ajuste del tiempo anterior a la primera manifestación de los EA Emergentes y el tratamiento A. Los modelos que están siendo comparados son: K-M, Exp(0.0066) y Weib(0.0027, 1.1936).

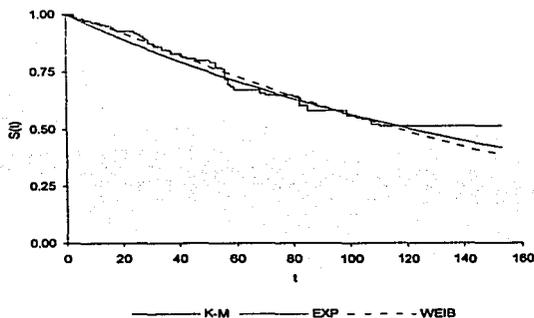


Figura 3.18b Gráfica de $\hat{S}(t)$ para el ajuste del tiempo anterior a la primera manifestación de los EA Emergentes y el tratamiento B. Los modelos que están siendo comparados son: K-M, Exp(0.0057) y Weib(0.0025, 1.1842).

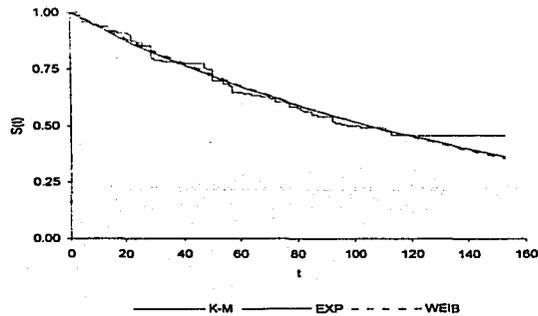


Figura 3.18c Gráfica de $\hat{S}(t)$ para el ajuste del tiempo anterior a la primera manifestación de los EA Emergentes y el tratamiento C. Los modelos que están siendo comparados son: K-M, Exp(0.0066) y Weib(0.0054, 1.0447).

Tabla 3.21 Valores de los estimadores de los modelos Exponencial y Weibull, obtenidos por máxima verosimilitud para el ajuste del tiempo de duración, el sistema "el cuerpo como un todo" y los tratamientos A, B y C.

Tiempo de duración para los EAE del sistema "el cuerpo como un todo".							
		A		B		C	
Parámetros		Exponencial	Weibull	Exponencial	Weibull	Exponencial	Weibull
	$\hat{\lambda}$	0.0422	0.0653	0.0549	0.0711	0.0311	0.0631
	$\hat{\gamma}$		0.8756		0.9189		0.8107

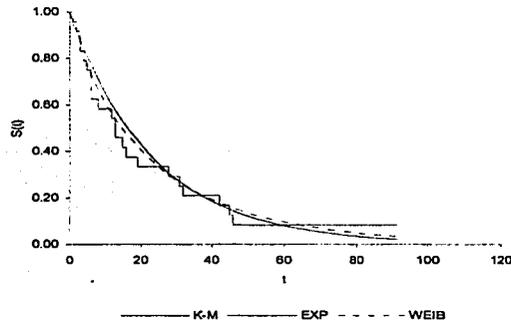


Figura 3.19a Gráfica de $\hat{S}(t)$ para el ajuste del tiempo de duración, el sistema "el cuerpo como un todo" y el tratamiento A. Los modelos que están siendo comparados son: K-M, Exp(0.0422) y Weib(0.0653, 0.8756).

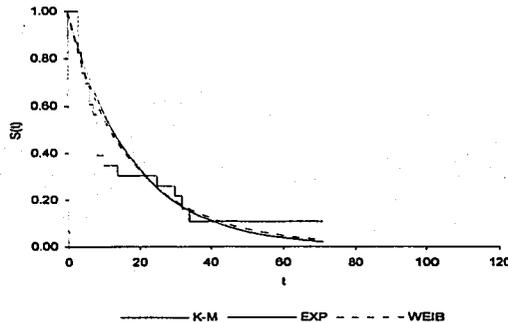


Figura 3.19b Gráfica de $\hat{S}(t)$ para el ajuste del tiempo de duración, el sistema "el cuerpo como un todo" y el tratamiento B. Los modelos que están siendo comparados son: K-M, Exp(0.0549) y Weib(0.0711, 0.9189).

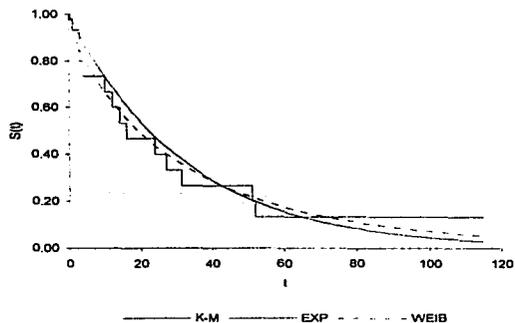


Figura 3.19c Gráfica de $\hat{S}(t)$ para el ajuste del tiempo de duración, el sistema "el cuerpo como un todo" y el tratamiento C. Los modelos que están siendo comparados son: K-M, Exp(0.0311) y Weib(0.0631,0.8107).

Tabla 3.22 Valores de los estimadores de los modelos Exponencial y Weibull, obtenidos por máxima verosimilitud para el ajuste del tiempo de duración, el sistema digestivo y los tratamientos A, B y C.

Tiempo de duración para los EAE del sistema digestivo.							
		A		B		C	
		Exponencial	Weibull	Exponencial	Weibull	Exponencial	Weibull
Parámetros	$\hat{\lambda}$	0.0323	0.0911	0.0313	0.0059	0.0474	0.0950
	$\hat{\gamma}$		0.7232		1.4558		0.7909

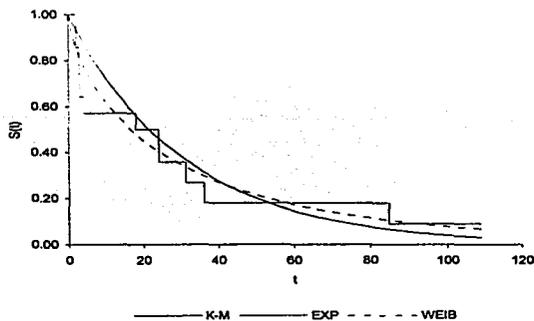


Figura 3.20a Gráfica de $\hat{S}(t)$ para el ajuste del tiempo de duración, el sistema digestivo y el tratamiento A. Los modelos que están siendo comparados son: K-M, Exp(0.0323) y Weib(0.0911,0.7232).

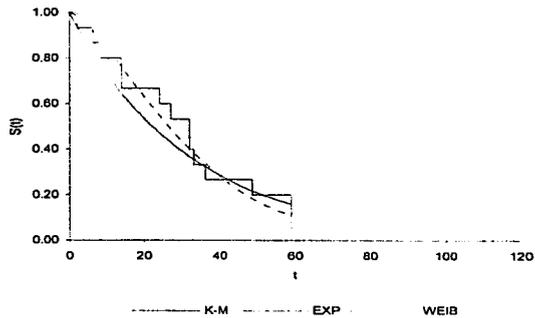


Figura 3.20b Gráfica de $S(t)$ para el ajuste del tiempo de duración, el sistema digestivo y el tratamiento B. Los modelos que están siendo comparados son: K-M, Exp(0.0313) y Weib(0.0059,1.4558).

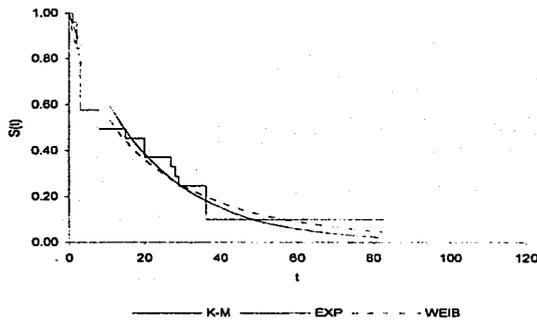


Figura 3.20c Gráfica de $S(t)$ para el ajuste del tiempo de duración, el sistema digestivo y el tratamiento C. Los modelos que están siendo comparados son: K-M, Exp(0.0474) y Weib(0.0950,0.7909).

Tabla 3.23 Valores de los estimadores de los modelos Exponencial y Weibull, obtenidos por máxima verosimilitud para el ajuste del tiempo anterior a la primera manifestación, el sistema "el cuerpo como un todo" y los tratamientos A, B y C.

Tiempo anterior a la primera manifestación de los EAE del sistema "el cuerpo como un todo":							
Parámetros		A		B		C	
		Exponencial	Weibull	Exponencial	Weibull	Exponencial	Weibull
	$\hat{\lambda}$	0.0219	0.0008	0.0175	0.0002	0.0185	0.0018
	$\hat{\gamma}$		1.7961		2.0614		1.5512

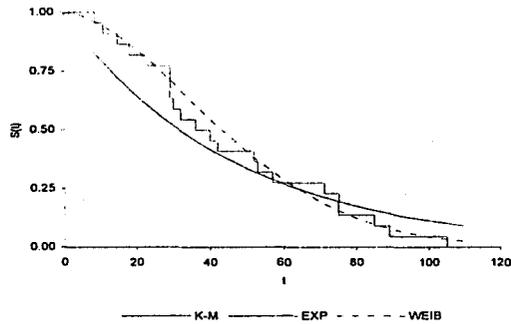


Figura 3.21a Gráfica de $\hat{S}(t)$ para el ajuste del tiempo de manifestación, el sistema "el cuerpo como un todo" y el tratamiento A. Los modelos que están siendo comparados son: K-M, Exp(0.0219) y Weib(0.0008,1.7961).

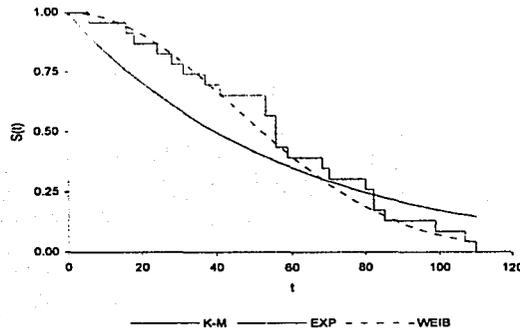


Figura 3.21b Gráfica de $\hat{S}(t)$ para el ajuste del tiempo de manifestación, el sistema "el cuerpo como un todo" y el tratamiento B. Los modelos que están siendo comparados son: K-M, Exp(0.0175) y Weib(0.0002,2.0614).

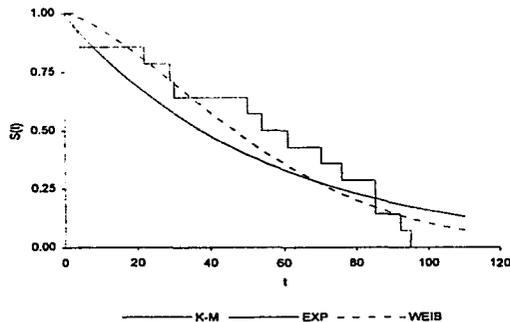


Figura 3.21c Gráfica de $\hat{S}(t)$ para el ajuste del tiempo de manifestación, el sistema "el cuerpo como un todo" y el tratamiento C. Los modelos que están siendo comparados son: K-M, Exp(0.0185) y Weib(0.0018,1.5512).

Tabla 3.24 Valores de los estimadores de los modelos Exponencial y Weibull, obtenidos por máxima verosimilitud para el ajuste del tiempo anterior a la primera manifestación, el sistema digestivo y los tratamientos A, B y C.

Tiempo anterior a la primera manifestación de los EAE del sistema digestivo.							
Parámetros		A		B		C	
		Exponencial	Weibull	Exponencial	Weibull	Exponencial	Weibull
	$\hat{\lambda}$	0.0248	0.0016	0.0197	0.0045	0.0203	0.0051
	$\hat{\gamma}$		1.6889		1.3511		1.3300

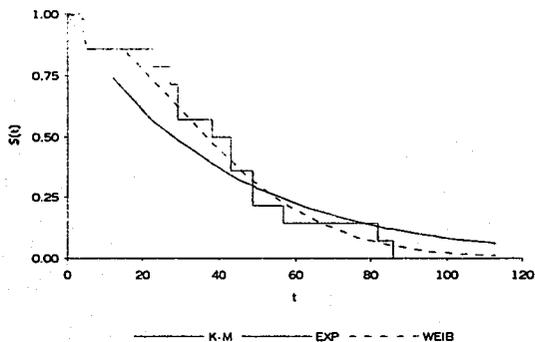


Figura 3.22a Gráfica de $\hat{S}(t)$ para el ajuste del tiempo de manifestación, el sistema digestivo y el tratamiento A. Los modelos que están siendo comparados son: K-M, Exp(0.0248) y Weib(0.0016,1.6889).

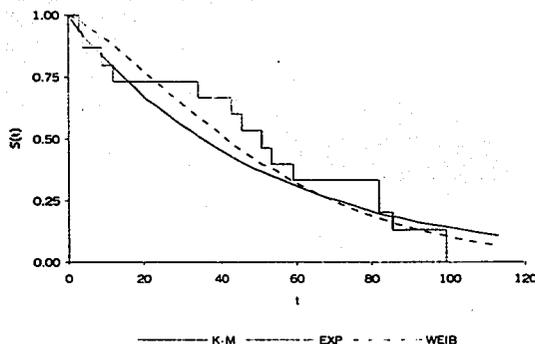


Figura 3.22b Gráfica de $\hat{S}(t)$ para el ajuste del tiempo de manifestación, el sistema digestivo y el tratamiento B. Los modelos que están siendo comparados son: K-M, Exp(0.0197) y Weib(0.0045,1.3511).

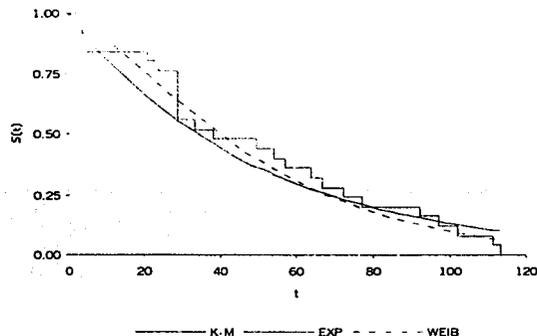


Figura 3.22c Gráfica de $\hat{S}(t)$ para el ajuste del tiempo de manifestación, el sistema digestivo y el tratamiento C. Los modelos que están siendo comparados son: K-M, Exp(0.0203) y Weib(0.0051, 1.3300).

3.5.3 Criterios de selección de la función de supervivencia.

Concluido el proceso de estimación de parámetros, y después de graficar las funciones obtenidas contra el estimador no paramétrico de Kaplan-Meier, puede decirse que, tras una inspección visual, la función Weibull describe adecuadamente la experiencia de supervivencia de los fenómenos que se están estudiando, sin embargo el ajuste a través de la función Exponencial ofrece una buena alternativa en algunos casos.

En el capítulo II, se mencionaron algunas propiedades y aplicaciones prácticas de ambas funciones. Es claro que el modelo Weibull tiene la ventaja de ser muy versátil, ya que sus dos parámetros le permiten asumir múltiples formas lo cual favorece al ajuste mismo, sin embargo su implementación es complicada en virtud de que los procesos para estimar dichos parámetros no son directos, y en ocasiones, tampoco son sencillos. Por su parte, el modelo Exponencial no es tan versátil como el Weibull, pero su estimación e implementación es directa y simple, además de que puede utilizarse para describir e interpretar una gran variedad de fenómenos con facilidad.

Con base en estos hechos, y siguiendo el principio de parsimonia, se consideró conveniente plantear un procedimiento formal para decidir si la función Exponencial es adecuada para describir la experiencia de supervivencia de los fenómenos tiempo de duración y tiempo anterior a la primera manifestación de los eventos adversos emergentes, asumiendo que la distribución Weibull lo hace.

Debe recordarse que la distribución Exponencial puede tratarse como un caso especial de la distribución Weibull cuando $\gamma = 1$; en este sentido, si se construye un intervalo de confianza para γ , y este intervalo abarca el valor 1, puede decirse que el fenómeno que se está analizando también puede modelarse a través de la función Exponencial. En virtud de lo anterior, solo se calcularon los límites de los intervalos de confianza para γ ; por tipo de ajuste, sistema y tratamiento; y en

aquellos casos donde dicho intervalo incluyera al valor 1, se consideraría que la función Exponencial es adecuada para describir la experiencia de supervivencia que se está estudiando.

El intervalo de confianza para γ , a un nivel de significancia α , es el siguiente:

$$(\hat{\gamma} - z_{1-\frac{\alpha}{2}} \text{E.S.} < \gamma < \hat{\gamma} + z_{1-\frac{\alpha}{2}} \text{E.S.}) = 1 - \alpha$$

donde:

$\hat{\gamma}$ es el valor del parámetro estimado por el método de máxima verosimilitud; $z_{1-\frac{\alpha}{2}}$ es el valor del percentil $1 - \frac{\alpha}{2}$ de una distribución Normal estándar; y E.S. es el valor estimado del error estándar de γ , aproximado por medio del valor de la varianza de γ , obtenida a partir de la matriz inversa de las segundas derivadas negativas de las funciones logarítmicas de verosimilitud evaluadas en los valores estimados de los parámetros λ y γ dentro del ciclo (s+1) del proceso Newton-Raphson donde se detuvo la secuencia iterativa de cálculo del valor de los parámetros ($I^{-1}(\hat{\lambda}_s, \hat{\gamma}_s)$ donde s=9).

Si se considera un nivel de significancia $\alpha = 0.05$, los límites de los intervalos así como el modelo paramétrico seleccionado para describir el fenómeno de supervivencia, quedan resumidos en las tablas 3.25a y 3.25b para los ajustes generales, y 3.26 para los casos por sistema y tratamiento.

Tabla 3.25a Límites de los intervalos de confianza para γ del modelo Weibull, estimados para el tiempo de duración de los EA Emergentes, por tipo de tratamiento. Función elegida para describir el fenómeno en estudio.

Tratamiento	Parámetro $\hat{\gamma}$	Límite inferior	Límite superior	Función
Tiempo de duración de los EAE				
A	0.7966	0.6228	0.9704	Weibull
B	0.8099	0.6440	0.9757	Weibull
C	0.7953	0.6449	0.9458	Weibull

Tabla 3.25b Límites de los intervalos de confianza para γ del modelo Weibull, estimados para el tiempo anterior a la primera manifestación de los EA Emergentes, por tipo de tratamiento. Función elegida para describir el fenómeno en estudio.

Tratamiento	Parámetro $\hat{\gamma}$	Límite inferior	Límite superior	Función
Tiempo anterior a la primera manifestación de los EAE				
A	1.1936	0.9375	1.4497	Exponencial
B	1.1842	0.9188	1.4495	Exponencial
C	1.0447	0.8286	1.2607	Exponencial

Tabla 3.26 Límites de los intervalos de confianza para γ del modelo Weibull, estimados sobre los ajustes de los modelos de supervivencia del tiempo de duración y el tiempo anterior a la primera manifestación de los EA Emergentes, por sistema y tipo de tratamiento. Función elegida para describir el fenómeno en estudio.

Tratamiento	Parámetro $\hat{\gamma}$	Límite inferior	Límite superior	Función
Tiempo de duración/ "el cuerpo como un todo"				
A	0.8756	0.5881	1.1631	Exponencial
B	0.9189	0.6087	1.2290	Exponencial
C	0.8107	0.4550	1.1664	Exponencial
Tiempo de duración/ digestivo				
A	0.7232	0.3971	1.0494	Exponencial
B	1.4558	0.8113	2.1004	Exponencial
C	0.7909	0.5266	1.0552	Exponencial
Tiempo de manifestación/ "el cuerpo como un todo"				
A	1.7961	1.2050	2.3871	Weibull
B	2.0614	1.3685	2.7542	Weibull
C	1.5512	0.8345	2.2680	Weibull
Tiempo de manifestación/ digestivo				
A	1.6889	0.9641	2.4136	Exponencial
B	1.3511	0.7572	1.9450	Exponencial
C	1.3300	0.8959	1.7641	Exponencial

En los casos donde el intervalo construido contiene el valor 1, se prefirió el modelo Exponencial por las ventajas descritas al principio de esta sección, salvo en el ajuste del tiempo de manifestación y el sistema "el cuerpo como un todo" para el tratamiento C, donde se escogió el modelo Weibull con la finalidad de abordar un solo tipo de función para un mismo fenómeno y un mismo sistema.

3.6 Proceso de prueba de hipótesis para comparar las experiencias de supervivencia entre tratamientos.

Para comprobar si existe algún tipo de relación entre las experiencias de supervivencia de cada uno de los tres tratamientos, tanto a nivel general como por sistema corporal, se planteó una prueba de hipótesis basada en la estadística de Mantel-Haezsel. La hipótesis a comprobar queda definida en los siguientes términos:

$H_0: S(t_A)=S(t_B)=S(t_C)$
contra
$H_1: S(t_i) \neq S(t_j)$ al menos para algún $i \neq j$ donde $i=A,B,C$ y $j=A,B,C$

Puesto que la comparación habría de realizarse entre tres grupos diferentes, se aplicó la forma matricial de la estadística de prueba, tal y como fuera definida dentro de la sección 2.4.2.3 de capítulo II.

Los resultados de la prueba, por tipo de ajuste y tratamiento se muestra en las tablas 3.27a y 3.27b; mientras que los resultados por sistema y tratamiento, se aprecian en las tablas 3.28 y 3.29. En cada una de ellas, se han incluido las entradas de la matriz U (diferencia entre las fallas observadas y las fallas esperadas) y las entradas completas de la matriz de varianzas y covarianzas V ; no obstante que para la estadística de prueba, solamente hubo de considerarse las primeras cuatro entradas (varianza de A y B, y covarianza entre A y B).

Por último se mencionará, a manera de conclusión, que los resultados de las pruebas de hipótesis aplicadas, indican que no hay evidencias suficientes para suponer que existen diferencias significativas entre las experiencias de supervivencia de los tres tratamientos, por sistema y tipo de ajuste.

Tabla 3.27a Resultado de la prueba de hipótesis para comparar las experiencias de supervivencia correspondientes al ajuste del tiempo de duración de los EA Emergentes entre los tratamientos A, B y C.

Tiempo de duración				
U		V		
		A	B	C
A	-10.457	A 40.1022	-18.1506	-21.9516
B	7.089	B -18.1506	35.7989	-17.6483
C	3.368	C -21.9516	-17.6483	39.5999

Mantel-Haeszel		
X^2	G. de L.	$P > X^2_2$
2.9278	2	0.2313

No hay evidencias suficientes para rechazar H_0 .

Tabla 3.27b Resultado de la prueba de hipótesis para comparar las experiencias de supervivencia correspondientes al ajuste del tiempo anterior a la primera manifestación de los EA Emergentes entre los tratamientos A, B y C.

Tiempo anterior a la primera manifestación				
U		V		
		A	B	C
A	2.5426	A 44.1843	-22.1868	-21.9975
B	-6.5969	B -22.1868	46.2575	-24.0707
C	4.0543	C -21.9975	-24.0707	46.0682

Mantel-Haeszel		
X^2	G. de L.	$P > X^2_2$
0.9523	2	0.6212

No hay evidencias suficientes para rechazar H_0 .

Tabla 3.28 Resultado de la prueba de hipótesis para comparar las experiencias de supervivencia correspondientes al ajuste del tiempo de duración, por tipo de sistema, y para los tratamientos A, B y C.

Duración / el cuerpo como un todo				
U		V		
		A	B	C
A	0.2225	A 12.4203	-6.3837	-6.0366
B	2.8308	B -6.3837	10.8884	-4.5047
C	-3.0533	C -6.0366	-4.5047	10.5413

Mantel-Haeszel		
X ²	G. de L.	P>X ² ₂
1.1442	2	0.5643

No hay evidencias suficientes para rechazar H₀.

Duración / digestivo				
U		V		
		A	B	C
A	-0.6615	A 7.6767	-3.6194	-4.0573
B	-2.6068	B -3.6194	9.3705	-5.7511
C	3.2683	C -4.0573	-5.7511	9.8084

Mantel-Haeszel		
X ²	G. de L.	P>X ² ₂
1.1685	2	0.5575

No hay evidencias suficientes para rechazar H₀.

Tabla 3.29 Resultado de la prueba de hipótesis para comparar las experiencias de supervivencia correspondientes al ajuste del tiempo anterior a la primera manifestación, por tipo de sistema, y para los tratamientos A, B y C.

Manifestación / el cuerpo como un todo				
U		V		
		A	B	C
A	4.6916	A 11.4040	-7.2487	-4.1553
B	-4.0746	B -7.2487	13.2211	-5.9724
C	-0.6171	C -4.1553	-5.9724	10.1277

Mantel-Haeszel		
X ²	G. de L.	P>X ² ₂
2.0687	2	0.3555

No hay evidencias suficientes para rechazar H₀.

Manifestación / digestivo				
U		V		
		A	B	C
A	3.8900	A 7.3477	-2.8885	-4.4592
B	-1.0008	B -2.8885	10.2130	-7.3245
C	-2.8892	C -4.4592	-7.3245	11.7837

Mantel-Haeszel		
X ²	G. de L.	P>X ² ₂
2.0901	2	0.3517

No hay evidencias suficientes para rechazar H₀.

3.7 Resumen de procedimientos y resultados.

El proceso de ajuste que se realizó, tuvo como finalidad implementar un modelo paramétrico de supervivencia que describiera el desarrollo del tiempo de duración y del tiempo anterior a la primera manifestación de los eventos adversos emergentes reportados en cada uno de los tres tratamientos comparados dentro del estudio de eficacia seleccionado para tal efecto. Este proceso se realizó en dos etapas: durante la primera, se revisó, caracterizó y seleccionó la información necesaria de la base de datos de eventos adversos para llevar a cabo el ajuste; en la segunda, se construyó el propio modelo con base en la información elegida y la herramienta estadística discutida dentro del capítulo II. En virtud de lo antes expuesto, se alcanzaron los siguientes resultados:

Se encontró que el análisis de supervivencia es una técnica estadística adecuada para estudiar fenómenos aleatorios que incluyen información completa e información incompleta sobre las realizaciones de una misma variable aleatoria.

En el campo de la investigación médica, pueden hallarse frecuentemente fenómenos de este tipo, para los cuales no siempre es posible dar seguimiento puntual al desarrollo de cada realización de la variable dentro de los límites establecidos para su propia observación. Un ejemplo, lo constituye el estudio de ciertas características de las manifestaciones de los eventos adversos: el tiempo que duran y el tiempo que tardan en aparecer por primera vez, dentro del contexto de un estudio farmacológico. En ambos casos, la variable principal es el tiempo que debe transcurrir desde un cierto momento hasta la ocurrencia de un cierto evento, del cual no siempre se tiene un dato exacto y cierto sobre su realización, lo que conlleva a manejar datos con información completa (fallas) e información parcial (censuras) al mismo tiempo.

Para describir la duración y el tiempo de manifestación de los eventos adversos emergentes observados en cada uno de los tres tratamientos del ensayo clínico base utilizado para realizar este trabajo, se escogieron dos modelos paramétricos con la finalidad de ajustar la función de supervivencia $S(t)$ a los datos seleccionados. Los modelos empleados fueron el Weibull y el Exponencial; ambos utilizados comúnmente para describir fenómenos relacionados con procesos biológicos. Cada uno se ajustó a todas las manifestaciones de los eventos adversos emergentes, válidas para el tiempo de duración y tiempo anterior a la primera aparición, por tipo de tratamiento; y a las manifestaciones de los eventos observados en los dos sistemas corporales con mayor número de casos, también por tratamiento. Este proceso requirió, como primer paso, la construcción de un modelo no paramétrico utilizando el estimador Kaplan-Meier, quien permitió visualizar la experiencia de supervivencia de los fenómenos en estudio a través de la forma gráfica de la función de supervivencia empírica. El estimador de K-M también se utilizó como punto de partida para implementar los modelos paramétricos citados. Para ello, se aplicaron métodos gráficos de ajuste utilizando ciertas transformaciones sobre la información disponible con la finalidad de validar los modelos seleccionados y encontrar una estimación burda del valor de sus parámetros. Después de aplicar las respectivas transformaciones, y luego de graficar los datos resultantes, se confirmó que ambas funciones sí eran adecuadas para describir la duración y el tiempo de manifestación de los eventos adversos.

Así mismo, se calculó el valor de sus parámetros, los cuales serían utilizados como punto de inicio en la implementación de métodos de cálculo más formales.

Un siguiente paso en la construcción del modelo de supervivencia, consistió en hallar el valor de los parámetros de las distribuciones por medio de un procedimiento que asegurase algunas propiedades estadísticas deseables para ellos. El método aplicado fue el de máxima verosimilitud para datos censurados.

Para la función Exponencial, el cálculo del valor de su único parámetro fue directo; sin embargo en la función Weibull fue necesario utilizar un método numérico de aproximación para sus dos parámetros; en este caso se escogió el procedimiento iterativo de Newton-Raphson, tomando como base los valores encontrados por métodos gráficos.

Al graficar las funciones resultantes contra el estimador de Kaplan-Meier, pudo confirmarse que ambos modelos describen satisfactoriamente la experiencia de supervivencia de los fenómenos en estudio, especialmente la distribución Weibull; sin embargo, y con la finalidad optar por un modelo sencillo, se planteó la construcción de un intervalo de confianza alrededor del valor de $\hat{\gamma}$ de esta última para que, en caso de que el intervalo abarcase el valor 1, se tuviese un elemento objetivo de decisión para preferir la distribución Exponencial. Derivado de lo anterior, y después de encontrar los límites de los intervalos, se optó por el modelo Exponencial para caracterizar ambos fenómenos de supervivencia, exceptuando el ajuste general del tiempo de duración y el ajuste del tiempo de manifestación del sistema "el cuerpo como un todo", donde prevaleció el modelo Weibull.

Sobre el ajuste del tiempo de duración de las manifestaciones de los eventos adversos emergentes en general, se encontró que la distribución Weibull es adecuada para describir dicho fenómeno. La función de supervivencia estimada $\hat{S}(t)$ es monótona decreciente y tiende a cero con cierta "rapidez" (entre el día 0 y 20 del tratamiento, donde $\hat{S}(t) > 0.4$), asumiendo valores cercanos a este número después de 120 días. Por otro lado, al analizar $\hat{h}(t)$, se encontró que también se trata de una función monótona decreciente, cuyos valores se indeterminan cuando t converge a cero por la izquierda; pero cuando $t \rightarrow \infty$, este proceso se torna "lento". Cabe señalar que este comportamiento puede interpretarse en el sentido de que la mayor parte de las manifestaciones registradas son breves, aunque existen algunas (las menos) de larga duración, particularmente aquellas que aparecieron poco después de iniciado un tratamiento y no habían concluido al momento de su término, dando como resultado datos censurados los cuales ocasionan que la función de supervivencia estimada presente una "cola larga" asintótica a cero. Así mismo, la forma de la función de riesgo indica que el patrón que sigue el proceso de conclusión de las manifestaciones de los eventos es muy intenso durante los días posteriores al inicio de cualquier tratamiento, pero tiende a disminuir progresivamente con el paso del tiempo. Como observación adicional, se mencionará que las experiencias de supervivencia de B y C son semejantes, especialmente durante los primeros 20 días; mientras que en el caso de A, cuya experiencia de supervivencia es ligeramente diferente, el proceso de terminación

de los eventos adversos resulta ser el menos eficaz ($\hat{h}_A(t) < \hat{h}_B(t)$ y $\hat{h}_A(t) < \hat{h}_C(t)$). Las figuras 3.23a y 3.23b muestran los comportamientos descritos arriba.

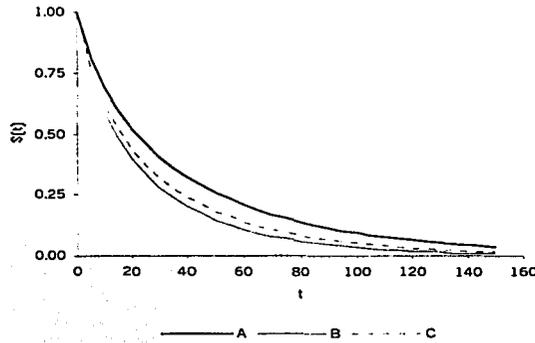


Figura 3.23a Gráficas de $\hat{S}(t)$ para el ajuste del tiempo de duración de los eventos adversos emergentes por tipo de tratamiento: A: Weib(0.0601,0.7966); B: Weib(0.0809,0.8099) y C: Weib(0.0760,0.7953).

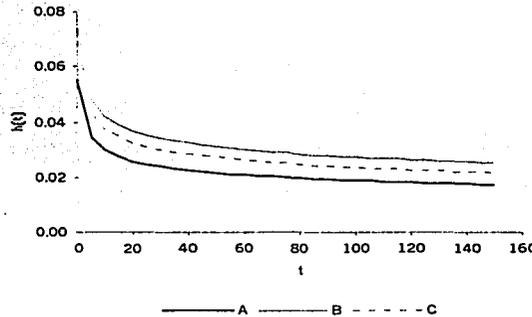


Figura 3.23b Gráficas de $\hat{h}(t)$ para el ajuste del tiempo de duración de los eventos adversos emergentes por tipo de tratamiento: A: Weib(0.0601,0.7966); B: Weib(0.0809,0.8099) y C: Weib(0.0760,0.7953).

Respecto al tiempo de duración de los eventos adversos emergentes del sistema "el cuerpo como un todo", puede apreciarse que $\hat{S}(t)$, descrito a través del modelo Exponencial, muestra un comportamiento semejante al del caso general, aunque aquí $\hat{S}(t) > 0.4$ entre el día 0 y 15 de tratamiento, lo cual indica que los eventos de este sistema duran menos que el promedio general. Por otro lado, $\hat{h}(t)$ es constante en todo momento y no depende de los días bajo tratamiento. Así mismo, el proceso de culminación de los eventos de A no es el mejor, ya que $\hat{h}_B(t) > \hat{h}_A(t)$. En las figuras 3.24a y 3.24b, puede observarse el comportamiento antes referido.

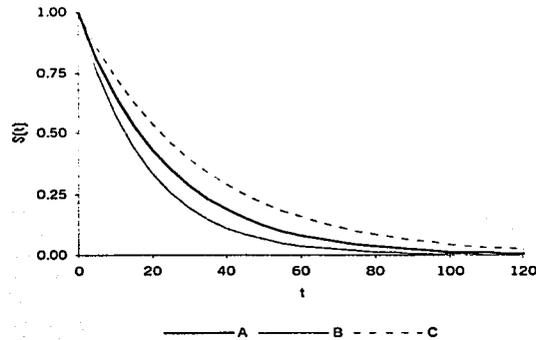


Figura 3.24a Gráficas de $\hat{S}(t)$ para el ajuste del tiempo de duración de los eventos adversos emergentes del sistema "el cuerpo como un todo" por tipo de tratamiento: A: Exp(0.0422); B: Exp(0.0549) y C: Exp(0.0311).

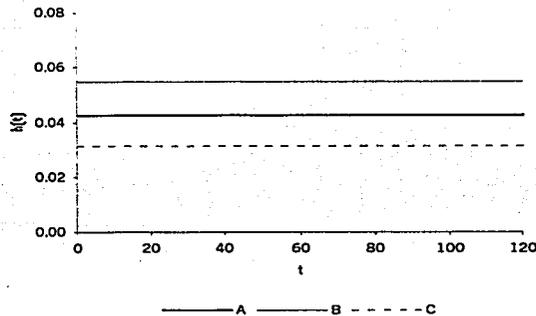


Figura 3.24b Gráficas de $\hat{h}(t)$ para el ajuste del tiempo de duración de los eventos adversos emergentes del sistema "el cuerpo como un todo" por tipo de tratamiento: A: Exp(0.0422); B: Exp(0.0549) y C: Exp(0.0311).

Para modelar el tiempo de duración de los eventos adversos emergentes del sistema digestivo se utilizó la distribución Exponencial. Aquí, la función de supervivencia estimada sigue un comportamiento parecido al del caso general, mientras que $\hat{h}(t)$ es constante (como en el sistema "el cuerpo como un todo"). También puede apreciarse que las experiencias de supervivencia de los tratamientos A y B son muy parecidas, mientras que los procesos de terminación de sus eventos son igualmente eficientes. Las figuras 3.25a y 3.25b permiten apreciar lo antes señalado.

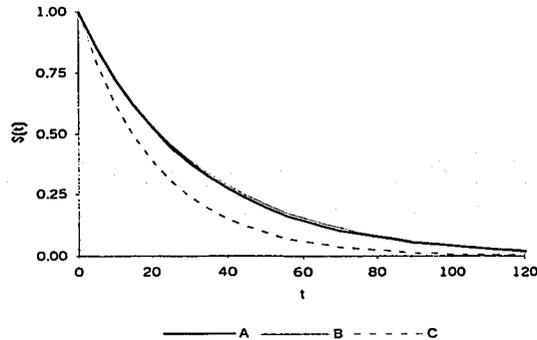


Figura 3.25a Gráficas de $\hat{S}(t)$ para el ajuste del tiempo de duración de los eventos adversos emergentes del sistema digestivo por tipo de tratamiento: A: Exp(0.0323); B: Exp(0.0313) y C: Exp(0.0474).

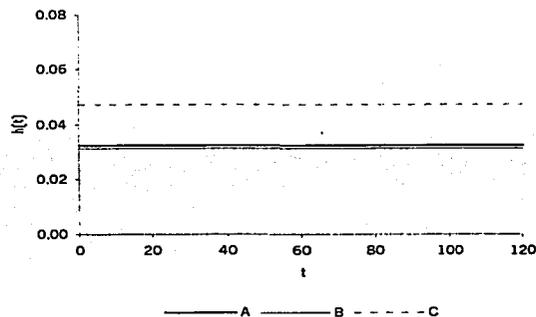


Figura 3.25b Gráficas de $\hat{h}(t)$ para el ajuste del tiempo de duración de los eventos adversos emergentes del sistema digestivo por tipo de tratamiento: A: Exp(0.0323); B: Exp(0.0313) y C: Exp(0.0474)

En lo que respecta al ajuste del tiempo anterior a la primera manifestación de los eventos adversos emergentes, se eligió la función Exponencial para modelar el caso general y el del sistema digestivo; mientras que en el caso del sistema "el cuerpo como un todo", se utilizó la función Weibull.

Para la descripción en general del tiempo anterior a la primera manifestación, la forma de $\hat{S}(t)$ indica que se trata de una función monótona decreciente que converge "lentamente" a cero (entre el día 0 y 100 del tratamiento, $\hat{S}(t) > 0.5$). En este punto, debe señalarse que para este mismo ajuste el gráfico del estimador de Kaplan-Meier muestra que para valores de $t > 105$, el comportamiento de la función de supervivencia tiende a ser asintótico a $y = 0.50$ (ver figura 3.10b). Esta observación puede interpretarse como que del total de pacientes sometidos a cualquiera de los tres tratamientos, alrededor de la mitad manifestaron no haber padecido evento adverso alguno durante el tiempo que duró el estudio. Ahora, en aquellos pacientes que sí reportaron la ocurrencia de un evento, se tiene que la primera manifestación tardó varios días en ocurrir, habiendo incluso eventos que aparecieron por primera vez uno o dos días antes de que concluyera el

tratamiento. En lo que respecta a la función de riesgo $\hat{h}(t)$, se asumió que es constante para cualquier valor de t , lo que implica que el proceso de aparición de cualquier evento adverso no depende de los días que un paciente permanece en tratamiento sino de otros factores que podrían estar relacionados con efectos farmacocinéticos y farmacodinámicos.

Por último, puede observarse que la experiencia de supervivencia de los tratamientos A y C es igual, por lo que puede afirmarse que el proceso de aparición de los eventos adversos en ambos tratamientos es equivalente, aunque comparado con el del tratamiento B, resulta ser el menos favorable (debe recordarse que, en el contexto del ensayo clínico base, se esperaría que el tratamiento A fuese aquel que presentase más pacientes sin eventos, y en aquellos que sí los presentasen, estos tardasen más día en aparecer). En las figuras 3.26a y 3.26b es posible visualizar los patrones de comportamiento descritos.

Por otro lado, y a diferencia del modelo general, la forma de la función de supervivencia por sistema corporal presenta notables diferencias en virtud de que en este caso se excluyó del ajuste la información de los pacientes que no presentaron ningún tipo de evento, ya que estos datos no pueden asociarse con ningún sistema en particular.

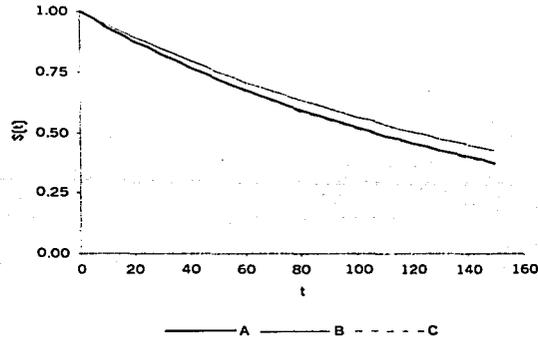


Figura 3.26a Gráficas de $\hat{S}(t)$ para el ajuste del tiempo anterior a la primera manifestación de los eventos adversos emergentes por tipo de tratamiento: A: Exp(0.0066); B: Exp(0.0057) y C: Exp(0.0066).

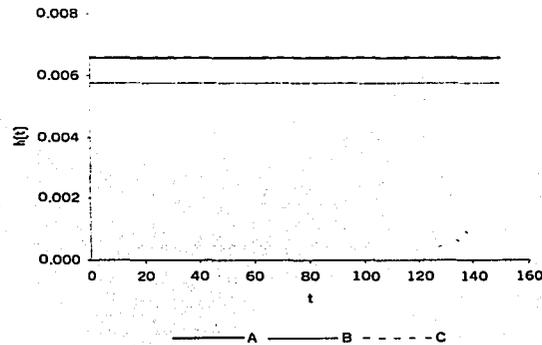


Figura 3.26b Gráficas de $\hat{h}(t)$ para el ajuste del tiempo anterior a la primera manifestación de los eventos adversos emergentes por tipo de tratamiento: A: Exp(0.0066); B: Exp(0.0057) y C: Exp(0.0066).

Se decidió describir el tiempo anterior a la primera manifestación de los eventos adversos del sistema "el cuerpo como un todo" a través del modelo Weibull en virtud de que el valor 1 queda fuera de los límites del intervalo de confianza planteado para el parámetro γ . $\hat{S}(t)$ es decreciente y toma valores cercanos a cero para $t > 120$. Por su parte, la función de riesgo es creciente, lo que indica que el proceso de aparición de un evento adverso para este sistema corporal se intensifica con el paso del tiempo. En las figuras 3.27a y 3.27b puede apreciarse este comportamiento.

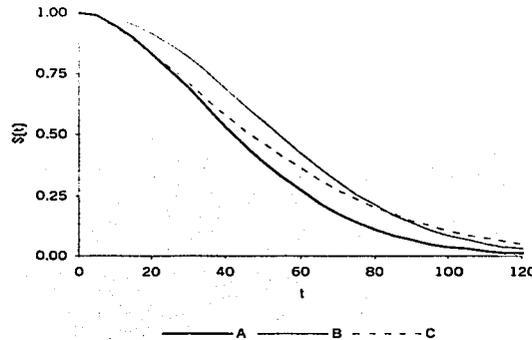


Figura 3.27a Gráficas de $\hat{S}(t)$ para el ajuste del tiempo anterior a la primera manifestación de los eventos adversos emergentes del sistema "el cuerpo como un todo" por tipo de tratamiento: A: Weib(0.0008,1.7961); B: Weib(0.0002,2.0614) y C: Weib(0.0018,1.5512).

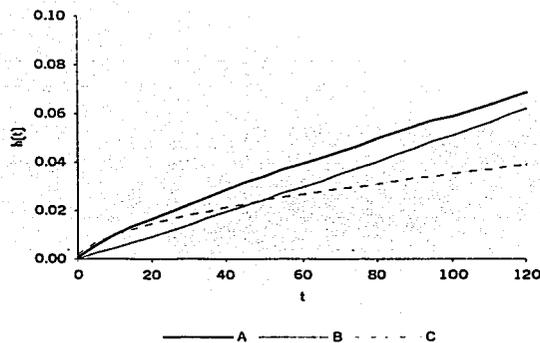


Figura 3.27b Gráficas de $\hat{h}(t)$ para el ajuste del tiempo anterior a la primera manifestación de los eventos adversos emergentes del sistema "el cuerpo como un todo" por tipo de tratamiento: A: Weib(0.0008,1.7961); B: Weib(0.0002,2.0614) y C: Weib(0.0018,1.5512).

Por su parte, para modelar el tiempo de anterior a la primera manifestación de los eventos adversos emergentes del sistema digestivo se utilizó la distribución Exponencial. La correspondiente función de supervivencia estimada, también sigue un comportamiento decreciente como en el sistema "el cuerpo como un todo", aunque aquí la función de riesgo es constante. Debe observarse también que los tratamientos B y C siguen un patrón de supervivencia similar. Las figuras 3.28a y 3.28b permiten apreciar lo antes señalado.

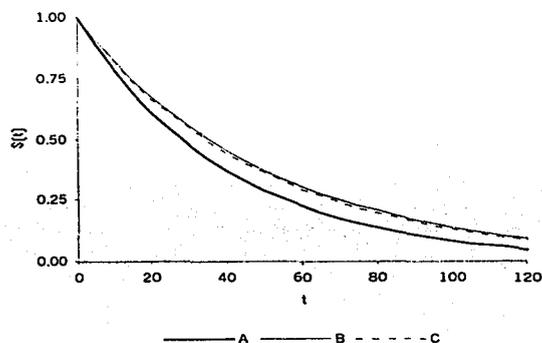


Figura 3.28a Gráficas de $\hat{S}(t)$ para el ajuste del tiempo anterior a la primera manifestación de los eventos adversos emergentes del sistema digestivo por tipo de tratamiento: A: Exp(0.0248); B: Exp(0.0197) y C: Exp(0.0203).

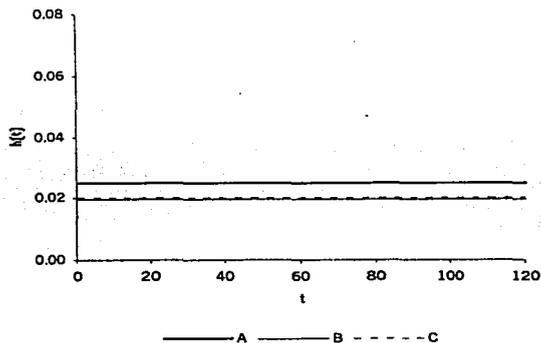


Figura 3.28a Gráficas de $\hat{h}(t)$ para el ajuste del tiempo anterior a la primera manifestación de los eventos adversos emergentes del sistema digestivo por tipo de tratamiento: A: Exp(0.0248); B: Exp(0.0197) y C: Exp(0.0203).

Un resultado importante dentro del análisis de los modelos ajustados, es el cálculo de algún valor de tendencia central de la variable. Este dato proporciona información sobre el número medio de días que duran las manifestaciones de los eventos adversos, en el caso del ajuste del tiempo de duración; o bien, el número promedio de días que deben transcurrir antes de que aparezca un evento adverso por primera vez después de iniciado el tratamiento, en el caso del ajuste del tiempo anterior a la primera manifestación. Asimismo, este dato puede utilizarse como un punto de comparación de la eficacia entre los tratamientos. Es claro que el valor de tendencia central ideal es el valor esperado, $E(T)$; sin embargo, cuando su cálculo es complicado y la función de probabilidad no es simétrica, como en las funciones que se han venido trabajando, se prefiere utilizar el valor del tiempo

medio de supervivencia (mediana) dado por t para el cual $\hat{S}(t)=0.50$ (Collet, 1994). En la tabla 3.30 se puede observar el valor del tiempo medio de supervivencia de los ajustes realizados.

En el caso del modelo general de duración, son más los días que demoran en desaparecer los eventos adversos del tratamiento A respecto a B y C; mientras que en el modelo del tiempo anterior a la primera manifestación, el tratamiento C es quien presenta los eventos adversos que tardan menos días en aparecer. Para los ajustes por sistema corporal y tratamiento en duración, el valor del tiempo medio de supervivencia del tratamiento A no resulta ser el más bajo: en el caso del sistema "el cuerpo como un todo" es el tratamiento B quien reporta los eventos menos largos, mientras que en el sistema digestivo es el tratamiento C. Por otro lado, para el tiempo anterior a la primera manifestación, es el tratamiento B donde los eventos tardan más tiempo en aparecer para ambos sistemas.

Todos estos datos indican que el tratamiento A no es significativamente superior a B o C; sin embargo, solo el resultado de la prueba de Mantel-Haenszel podría confirmar esta apreciación.

Tabla 3.30 Tiempo medio de supervivencia por tipo de ajuste, sistema corporal y tratamiento. Las cantidades están expresadas en días.

Tiempo de duración de los EAE			
Sistema/Trat.	A	B	C
General	21.54	14.18	16.11
"El cuerpo como un todo"	16.43	12.63	22.29
Digestivo	21.46	22.15	14.62

Tiempo anterior a la primera manifestación de los EAE			
Sistema/Trat.	A	B	C
General	105.72	121.00	104.93
"El cuerpo como un todo"	42.03	53.94	46.62
Digestivo	27.92	35.17	34.16

Finalmente, para validar igualdad entre las experiencias de supervivencia de los tres tratamientos, se propuso una prueba de hipótesis aplicando la estadística de Mantel-Haenszel. El resultado de ésta demostró, que a diferencia de lo que se hubiese esperado, las experiencias de supervivencia para la duración y tiempo de manifestación de los eventos adversos emergentes entre los tres tratamientos no son diferentes. En este punto, vale la pena observar que el resultado de la prueba de hipótesis puede tomarse como un indicio para suponer que un mismo modelo paramétrico (tanto en distribución como en valor de sus parámetros) puede describir la experiencia de supervivencia de los tres tratamientos; aunque, para efectos de este trabajo solo se procuró utilizar un mismo tipo de distribución para los casos de los eventos adversos en general y por sistema corporal, dejando de lado buscar la coincidencia en el valor de sus parámetros, ya que esto último va más allá de los objetivos planteados en un inicio.

De todo lo antes expuesto, puede decirse que, en términos generales y para los dos sistemas corporales con mayor incidencia de casos, la experiencia de supervivencia de la duración y del tiempo anterior a la primera manifestación de los eventos adversos emergentes observados en el tratamiento A, es semejante a la de B y C, en el contexto del ensayo clínico comparativo que se ha venido analizando.

3.8 Conclusiones.

Se ha visto que el análisis de supervivencia proporciona elementos para modelar el comportamiento de ciertos fenómenos que pueden vincularse con el paso del tiempo y, de esta manera estudiarlos y comprenderlos mejor; no obstante que el término supervivencia se siga asociando con la mortalidad, es claro que sus aplicaciones y métodos de análisis van más allá de la sola comprensión y modelación de patrones que permitan explicar la dualidad contingente vida-muerte. A lo largo de este trabajo queda de manifiesto que el análisis de supervivencia también puede aplicarse al estudio de fenómenos que ocurren en el ámbito de la investigación médica, y en particular, dentro de la industria farmacéutica en el proceso de investigación de la seguridad y eficacia de los medicamentos. En este contexto, se ha mostrado que esta herramienta estadística resulta de gran utilidad durante la etapa de pruebas farmacológicas, ya que proporciona los elementos necesarios para la implementación de modelos matemáticos que permitan describir el comportamiento de algunos fenómenos, tales como los eventos adversos, logrando con ello establecer criterios estadísticos de decisión que contribuyan a la selección de mejores medicinas. En virtud de lo anterior, se concluye que las posibilidades de aplicación que tiene el análisis de supervivencia abarcan otras disciplinas y áreas del conocimiento muy diferentes a las concepciones clásicas.

El proceso de ajuste, expuesto en detalle en páginas anteriores, se concretó a desarrollar un modelo para la función de supervivencia, con base en la información de la duración y el tiempo transcurrido antes de la primera manifestación de los eventos adversos emergentes observados dentro de un ensayo clínico comparativo entre tres tratamientos anti-hipercolesterolemicos. Los resultados alcanzados se presentaron en dos niveles de detalle: para todos las manifestaciones de los eventos adversos emergentes por tipo de ajuste, y para las manifestaciones de cada uno de los dos sistemas corporales con mayor incidencia de casos por tipo de tratamiento. Previo a la construcción del modelo, fue necesario revisar el antecedente clínico que dio origen a este trabajo. Esta tarea se realizó no solo con el objeto de contar con un panorama general de todo el estudio farmacológico, sino sobre todo con la finalidad de entender las motivaciones de los diversos profesionales que contribuyeron en el diseño del ensayo, y de este modo contar con mayores elementos de juicio al momento de presentar los resultados concretos de este ejercicio estadístico. De lo anterior se desprende que una de las actividades más importantes que fueron implementadas, consistió en el análisis del protocolo clínico, lo cual resultó enriquecedor tanto en el desarrollo del propio ajuste como en una mejor comprensión de los conceptos médicos y bioestadísticos alrededor de las pruebas que se hacen a los medicamentos.

Debe mencionarse que otra parte fundamental de este trabajo lo constituyó la construcción y adecuación de la base de datos con la información necesaria para la estimación de los modelos de supervivencia. Esta labor requirió, primero, tener plena claridad en los objetivos por alcanzarse, para luego poder analizar las características de los datos y así, estar en posibilidades de discernir aquellos útiles para el ajuste. De aquí se concluye que la calidad de los resultados que se logren

depende de dos factores estrechamente vinculados: la capacidad en el manejo de la información disponible y la calidad de los datos.

Con respecto a las funciones seleccionadas, debe recordarse que, desde el principio de este trabajo se propuso utilizar distribuciones adecuadas para describir fenómenos biológicos. El ajuste implicó integrar diversas herramientas estadísticas, tanto de estimación de parámetros como de validación de hipótesis, en un proceso encaminado a la resolución de un problema real y concreto: modelar la experiencia de supervivencia de la duración y el tiempo previo a la primera aparición de los eventos adversos. Como resultado de este proceso, se concluye que las funciones seleccionadas describen satisfactoriamente los tiempos de supervivencia de los eventos adversos emergentes: el modelo Weibull para la duración; y el modelo Exponencial para la primera manifestación; ambos para el caso general; mientras que en los ajustes por sistema corporal y tratamiento, se prefirió la función Exponencial excepto para un sistema corporal específico. Así mismo, debe destacarse que los elementos estadísticos abordados, fueron revisados y discutidos antes de ser utilizados, garantizando con ello una buena sustentación técnica de todo el trabajo realizado. Finalmente, se concluye que, gracias a una adecuada estructuración de los elementos de análisis planteados, es posible obtener resultados útiles, no solo para los fines de la propia investigación; sino sobre todo, se contribuye a enriquecer la formación profesional, lo cual constituye el único resultado permanente.

ANEXO 1. Formato para elaborar reportes resumidos de ensayos clínicos.

Nombre de la compañía		Reporte Tabular del Ensayo.		(Para uso exclusivo de la autoridad competente)
Nombre comercial del medicamento				
Nombre de la sustancia activa		Página	Número	
No. de referencia		Página	de	
Fecha de reporte:	Número:	Periodo del estudio (mes/ año):		
Criterios de evaluación:				
Seguridad/ Eficacia (primarios):				
Seguridad/ Eficacia (secundarios):				
Otros:				
Diseño estadístico				

Nombre de la compañía		Reporte Tabular del Ensayo.		(Para uso exclusivo de la autoridad competente)
Nombre comercial del medicamento				
Nombre de la sustancia activa		Página	Número	
No. de referencia		Página de		No. de modificación
Fecha de reporte	Número	Periodo del estudio (mes/ año):		
Resultados sobre seguridad				
Resultados sobre eficacia				
Otros resultados:				
Conclusiones:				

ANEXO 2. Desarrollo del ensayo clínico.

Etapas del ensayo	Reclutamiento								
	Fase I				Fase II				
	1	2	3	4	5	6	7	8	9
Visita médica de control	1	2	3	4	5	6	7	8	9
Semana	-10	-6	-4	-2	0	4	8	12	16
Actividad									
Verificación de criterios de inclusión	X								
Verificación de criterios de exclusión	X	X	X	X	X				
Asignación aleatoria al grupo de tratamiento					X				
Se recaba el consentimiento del paciente	X								
Información demográfica	X								
Antecedentes familiares de colesterol y enfermedades del corazón									
Antecedentes osteoarticulares del paciente	X								
Antecedentes cardiovasculares del paciente	X								X
Hábitos de consumo de tabaco	X								
Verificación de riesgo cardiovascular	X								
Tratamientos concomitantes	X	X	X	X	X	X	X	X	X
Enfermedades concomitantes	X	X	X	X	X	X	X	X	X
Exploración física completa	X				X				X
Historia médica	X	X	X	X	X	X	X	X	X
Signos vitales, masa corporal y nivel de actividad	X	X	X	X	X	X	X	X	X
Hábitos de consumo de alcohol	X				X		X		X
Electrocardiograma	X				X				X
Prescripción de dieta	X								
Evaluación del cumplimiento de la dieta		X	X	X	X		X		X
Distribución del medicamento		X	X	X	X	X	X	X	
Verificación del cumplimiento del tratamiento			X	X	X	X	X	X	X
Registro de eventos adversos		X	X	X	X	X	X	X	X
Se evalúa si el paciente continúa en el estudio o no		X	X	X	X				
Anotaciones finales									X

ANEXO 3. Formato utilizado para recabar información sobre los eventos adversos.

ENSAYO NO.	EVENTOS ADVERSOS					INICIALES DEL PACIENTE	NUMERO DE CENTRO / CODIGO	
<p>¿Se presentó algún evento adverso? No <input type="radio"/> Si <input type="radio"/> En caso afirmativo, favor de llenar este reporte marcando la opción más adecuada en cada caso.</p>								
Eventos Adversos	Serie 1.No 2.SI*	Razón para designarse como serie (es factible más de una respuesta) 1.Fallecimiento 2.Puso en peligro la vida del paciente 3.Hospitalización súbita o prolongación de la hospitalización ya existente 4.Anormalidad congénita/defecto de nacimiento 5.Persistencia o agravamiento de incapacidad o invalidez 6.Evento relevante medicamento	Gravedad	Fecha Inicio/Fin	Intensidad	Relación con el tratamiento	Acción(es) tomadas en relación al tratamiento (es factible más de una respuesta)	Resultado del evento
				dd mm aa	1.Levé 2.Moderado 3.Severo	1.Ninguna 2.Remota 3.Posible 4.Probable 5.No accesible	1.Ninguna 2.Reducción en la dosis del medicamento 3.Se discontinuó el medicamento y luego se reintrodujo 4.El medicamento se discontinuó en forma permanente 5.Se indicó otro medicamento o tratamiento, favor de especificarlo en la forma de registro para tratamiento concomitantes 6.Otras (especificar) 7.Insuficiente o se prolongó la hospitalización 8.Se requirió o se prolongó la hospitalización	1.Resultado 2.Hubo una mejoría 3.Sin cambio 4.Empeoró 5. 6.Fallecimiento 7.No hubo seguimiento adecuado
1.	<input type="radio"/> <input type="radio"/>	<input type="radio"/>		dd mm aa	<input type="radio"/> <input type="radio"/> <input type="radio"/>	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2.	<input type="radio"/> <input type="radio"/>	<input type="radio"/>		dd mm aa	<input type="radio"/> <input type="radio"/> <input type="radio"/>	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3.	<input type="radio"/> <input type="radio"/>	<input type="radio"/>		dd mm aa	<input type="radio"/> <input type="radio"/> <input type="radio"/>	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4.	<input type="radio"/> <input type="radio"/>	<input type="radio"/>		dd mm aa	<input type="radio"/> <input type="radio"/> <input type="radio"/>	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5.	<input type="radio"/> <input type="radio"/>	<input type="radio"/>		dd mm aa	<input type="radio"/> <input type="radio"/> <input type="radio"/>	<input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/>	<input type="radio"/>	<input type="radio"/>
* Si la respuesta fue afirmativa, llene una hoja complementaria por cada evento adverso serio reportado y envíe ambas al médico/investigador responsable del ensayo en forma inmediata.								
Otros detalles sobre los eventos adversos: _____								
Fecha		Nombre, firma y dirección del médico/investigador que reporta.						
dd mm aa								

ANEXO 4. Bases de datos utilizadas en la construcción de los modelos de supervivencia que describen el tiempo de duración y el tiempo anterior a la primera manifestación de los eventos adversos emergentes por tipo de tratamiento.

a) Bases de datos para el ajuste del modelo del tiempo de duración de las manifestaciones de los EAE, por sistema y tratamiento.

DURACION/ EL CUERPO COMO UN TODO					
A		B		C	
Dato	t	Dato	t	Dato	t
1	1	1	3	1	1
2	2	2	3	2	3
3	3	3	3	3	3
4	3	4	3	4	4
5	4	5	4	5	10
6	5	6	4	6	12
7	6	7	5	7	14
8	6	8	6	8	16
9	6	9	6	9	24
10	8	10	7	10	27
11	12	11	8	11	31
12	13	12	8	12	51
13	13	13	8	13	52
14	15	14	8	14	55*
15	16	15	10	15	115*
16	19	16	14		
17	28	17	25		
18	31	18	30		
19	32	19	30*		
20	42	20	32		
21	45	21	34		
22	46	22	42*		
23	73*	23	71*		
24	92*				

DURACION/ DIGESTIVO					
A		B		C	
Dato	t	Dato	t	Dato	t
1	2	1	2	1	1
2	2	2	6	2	2
3	3	3	8	3	2
4	3	4	14	4	2
5	3	5	14	5	3
6	4	6	24	6	3
7	18	7	27	7	3
8	24	8	32	8	3
9	24	9	32	9	3
10	28*	10	33	10	3
11	31	11	36	11	3
12	36	12	49	12	7*
13	85	13	53*	13	8
14	109*	14	59	14	8
		15	59	15	15
				16	20
				17	20
				18	27
				19	28
				20	29
				21	34*
				22	36
				23	36
				24	36
				25	49*
				26	83*

* indica que el dato es censurado.

b) Bases de datos para el ajuste del modelo del tiempo anterior a la primera manifestación de los EAE, por sistema y tratamiento.

MANIFESTACION/ EL CUERPO COMO UN TODO					
A		B		C	
Dato	t	Dato	t	Dato	t
1	9	1	6	1	3
2	11	2	16	2	4
3	15	3	18	3	22
4	18	4	24	4	29
5	23	5	28	5	30
6	29	6	31	6	50
7	29	7	37	7	54
8	29	8	41	8	61
9	30	9	53	9	70
10	32	10	53	10	76
11	36	11	56	11	85
12	40	12	56	12	85
13	42	13	56	13	92
14	52	14	59	14	95
15	53	15	68		
16	57	16	70		
17	71	17	80		
18	75	18	82		
19	75	19	82		
20	85	20	85		
21	89	21	99		
22	105	22	107		
		23	110		

MANIFESTACION/ DIGESTIVO					
A		B		C	
Dato	t	Dato	t	Dato	t
1	4	1	3	1	3
2	5	2	4	2	4
3	23	3	9	3	4
4	27	4	12	4	4
5	29	5	34	5	21
6	29	6	43	6	24
7	38	7	46	7	29
8	43	8	51	8	29
9	43	9	53	9	29
10	49	10	59	10	29
11	49	11	82	11	29
12	57	12	82	12	33
13	82	13	85	13	38
14	86	14	99	14	50
		15	99	15	54
				16	57
				17	64
				18	67
				19	72
				20	77
				21	92
				22	97
				23	102
				24	111
				25	113

* indica que el dato es censurado.

c) Bases de datos para el ajuste de los modelo generales del tiempo de duración y el tiempo anterior a la primera manifestación de los EAE, por tratamiento.

DURACION					
A		B		C	
Dato	t	Dato	t	Dato	t
1	1	1	1	1	1
2	1	2	1	2	1
3	1	3	1	3	1
4	1	4	1	4	1
5	2	5	1	5	2
6	2	6	1	6	2
7	2	7	1	7	2
8	3	8	1	8	2
9	3	9	1	9	3
10	3	10	2	10	3
11	3	11	3	11	3
12	3	12	3	12	3
13	3	13	3	13	3
14	4	14	3	14	3
15	4	15	3	15	3
16	5	16	3	16	3
17	5	17	4	17	3
18	5	18	4	18	3
19	6	19	4	19	3
20	6	20	5	20	3
21	6	21	5	21	4
22	6	22	5	22	4
23	7	23	6	23	4
24	8	24	6	24	4
25	8	25	6	25	5
26	9	26	7	26	7
27	12	27	8	27	7*
28	13	28	8	28	8
29	13	29	8	29	8
30	15	30	8	30	8
31	16	31	8	31	8
32	18	32	9	32	8
33	19	33	10	33	8
34	24	34	10	34	9
35	24	35	11	35	10
36	27*	36	14	36	12
37	28	37	14	37	12
38	28*	38	14	38	14
39	28*	39	14	39	14

MANIFESTACION					
A		B		C	
Dato	t	Dato	t	Dato	t
1	1*	1	3	1	3
2	2	2	3	2	3
3	4	3	4	3	4
4	5	4	6	4	4
5	9	5	8*	5	4
6	11	6	9	6	4
7	15	7	12	7	7
8	18	8	14	8	8
9	22	9	16	9	10
10	23	10	17	10	14
11	23	11	18	11	14
12	25	12	24	12	14
13	26*	13	25	13	17
14	27	14	26	14	21
15	27	15	27	15	22
16	27*	16	28	16	22
17	29	17	29	17	22
18	29	18	29	18	22
19	29	19	31	19	24
20	29	20	31	20	26
21	29	21	34	21	26
22	29	22	34	22	29
23	29	23	37	23	29
24	30	24	37	24	29
25	32	25	41	25	29
26	34	26	41	26	29
27	34	27	43	27	29
28	34*	28	46	28	29
29	36	29	51	29	29
30	38	30	53	30	30
31	40	31	53	31	33
32	40	32	53	32	38
33	42	33	55	33	47
34	43	34	56	34	47
35	43	35	56	35	47
36	46	36	56	36	48
37	46	37	56	37	50
38	49	38	56	38	50
39	49	39	56	39	50

DURACION					
A		B		C	
Dato	t	Dato	t	Dato	t
40	29	40	15	40	14
41	29	41	24	41	15
42	29	42	25	42	15
43	30	43	27	43	16
44	31	44	27	44	20
45	31	45	27	45	20
46	31	46	27*	46	21
47	31	47	30	47	22*
48	31	48	30*	48	24
49	32	49	32	49	27
50	36	50	32	50	27
51	41*	51	32	51	27
52	42	52	33	52	28
53	45	53	34	53	28
54	46	54	34	54	29
55	46	55	35	55	29
56	48	56	36	56	31
57	55*	57	42*	57	31
58	55*	58	44	58	32
59	55*	59	46	59	34*
60	57	60	49	60	35
61	57*	61	49	61	35
62	58	62	53*	62	36
63	62	63	53*	63	36
64	66*	64	56	64	36
65	73*	65	59	65	40
66	85	66	59	66	49*
67	85*	67	71*	67	51
68	92*	68	75*	68	52
69	93*	69	83*	69	55*
70	109*			70	57
				71	57
				72	65
				73	83*
				74	92
				75	98*
				76	98*
				77	98*
				78	115*

MANIFESTACION					
A		B		C	
Dato	t	Dato	t	Dato	t
40	51	40	57	40	50
41	52	41	57	41	50
42	52	42	57	42	50
43	53	43	58	43	50
44	56	44	59	44	51*
45	57	45	59	45	54
46	57	46	68	46	54
47	57	47	68	47	56
48	57	48	70	48	57
49	57	49	71*	49	57
50	57	50	80	50	57
51	57	51	82	51	57
52	57	52	82	52	61
53	58	53	82	53	64
54	58	54	82	54	67
55	63	55	82	55	70
56	68	56	85	56	72
57	71	57	85	57	72
58	72	58	85	58	76
59	75	59	99	59	77
60	75	60	99	60	77
61	82	61	99	61	80
62	83	62	102	62	81
63	83	63	104	63	82
64	85	64	105*	64	85
65	85	65	107	65	85
66	85	66	107	66	87
67	86	67	107	67	92
68	87	68	109*	68	92
69	87	69	110	69	92
70	89	70	111*	70	93
71	89	71	111*	71	95
72	105	72	111*	72	97
73	107*	73	111*	73	102
74	108*	74	111*	74	104*
75	109*	75	111*	75	107*
76	110*	76	111*	76	109*
77	110*	77	111*	77	110*
78	111*	78	111*	78	110*
79	111*	79	111*	79	110*
80	111*	80	111*	80	111

DURACION					
A		B		C	
Dato	t	Dato	t	Dato	t

MANIFESTACION					
A		B		C	
Dato	t	Dato	t	Dato	t
121	118*	121	115*	121	113*
122	118*	122	115*	122	114*
123	118*	123	115*	123	114*
124	118*	124	118*	124	115*
125	119*	125	118*	125	116*
126	119*	126	118*	126	117*
127	122*	127	119*	127	118*
128	126*	128	121*	128	118*
129	126*	129	125*	129	118*
130	130*	130	126*	130	118*
131	132*	131	130*	131	118*
132	132*	132	132*	132	118*
133	138*	133	133*	133	119*
		134	134*	134	121*
		135	146*	135	126*
				136	132*
				137	132*
				138	134*
				139	139*
				140	139*
				141	140*
				142	153*

* indica que el dato es censurado.

BIBLIOGRAFÍA.

1. ALLISON, Paul D. **"Survival analysis using the SAS® System: a practical guide"**. Cary, North Carolina, E.U.A., SAS Institute Inc., 1995 (c 1995), 292p.
2. CANAVOS, George C. **"Probabilidad y estadística. Aplicaciones y métodos"**. México, Ed. McGraw-Hill/Interamericana de México, 1991 (c 1988), 651 p.
3. CANTOR, Alan. **"Extending SAS® survival analysis techniques for medical research"**. Cary, North Carolina, E.U.A., SAS Institute Inc., 1997 (c 1997), 208p.
4. COLLETT, D. **"Modelling survival data in medical research"**. Bungay, Suffolk, Reino Unido, Chapman & Hall, (c 1994), 347p.
5. D'AGOSTINO, Ralph B. y Michael A. Stephens. **"Goodness-of-fit techniques"**. E.U.A., Marcel Dekker Inc., (c 1986), 560p.
6. EVERITT, Brian S. y Geoff Der. **"A handbook of statistical analyses using SAS"**. Osney Mead, Oxford, Reino Unido, Chapman & Hall, 1996 (c 1996), 158p.
7. KALBFLEISCH, J.D. y R.L. Prentice. **"The statistical analysis of failure time data"**. E.U.A., John Wiley & Sons Inc., (c 1980), 321p.
8. LAWLESS, J.F. **"Statistical models and methods for lifetime data"**. E.U.A., John Wiley & Sons Inc., (c 1982), 580p.
9. LONDON, Dick. **"Survival models and their estimation"**. 2ª Ed., Winsted and Avon, Connecticut, E.U.A., Actex Publications, (c 1988), 326p.
10. MILLER, Rupert G., Jr. **"Survival analysis"**. E.U.A., John Wiley & Sons Inc., (c 1981), 238p.

-
11. MOOD, Alexander M., Franklin A. Graybill y Duane C. Boes. **"Introduction to the theory of statistics"**. 3ª Ed., E.U.A., McGraw-Hill Inc., (c 1974), 564p.
 12. TUDELA, Victoria. **"El Colesterol: lo bueno y lo malo"**. 2ª Ed., México, Ed. Fondo de Cultura Económica, Primera reimpresión 2001 (c 2000), (La Ciencia para Todos / 140), p. 73
 13. ZAMORA, José Salvador. **"Análisis Estadístico de Tiempos de Falla Univariados"**. Tesis para obtener el título de Actuario. México D.F., Enero de 1994, U.N.A.M.
 14. Protocolo del Estudio Clínico: **"Estudio multinacional, multicéntrico, de asignación aleatoria, tratamiento doble ciego y doble falso del resultado de aplicar los tratamientos A, B y C (en donde C es una combinación de A y B) en la reducción de los niveles de colesterol LDL en pacientes con hipercolesterolemia primaria"**. Versión: 2.3, Fecha: 15 de octubre de 1996. Bayer de México, S.A. de C.V.