

36
24.



UNIVERSIDAD NACIONAL AUTÓNOMA DE
MÉXICO

ESCUELA NACIONAL DE ESTUDIOS PROFESIONALES

"ACATLAN"

**"ESTUDIO DE CASOS Y CONTROLES DE
LEUCEMIA EN EL CENTRO MEDICO
NACIONAL SIGLO XXI UTILIZANDO
REGRESIÓN LOGÍSTICA"**

T E S I S
QUE PARA OBTENER EL TITULO DE:

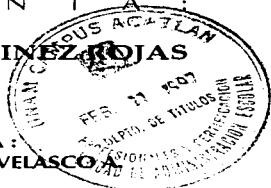
**LICENCIADA EN MATEMATICAS
APLICADAS Y COMPUTACIÓN**

P R E S E N T A :

CAROLINA MARTINEZ ROJAS

ASESORA :

DRA. SILVIA RUIZ VELASCO



FEBRERO 1997

**TESIS CON
FALLA DE ORIGEN**





Universidad Nacional
Autónoma de México



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.

Dedicada a:

Mis padres y hermanos por todo su apoyo y comprensión que me han proporcionado en las diferentes etapas de mi vida.

A la Dra. Silvia Ruíz Velasco Acosta por su tiempo, disposición e invaluable apoyo en la dirección de este trabajo.

A la unidad de Investigación Epidemiológica del Centro Médico Nacional Siglo XXI y en especial al Dr. Arturo Fajardo Gutiérrez por su valiosa colaboración para la obtención de los datos y apoyo en el proceso de desarrollo de la tesis.

Agradezco a:

Dios, por todo lo que me ha proporcionado

La universidad Nacional Autónoma de México
por mi formación académica,

A los profesores:

Ing. Beatriz Clavel Díaz,
M. en C. Eduardo Godoy Escoto,
Lic. Ma. Del Carmen Villar Patiño,
Ing. Reyes Laurencio Moncada,

por su valiosa aportación en la revisión
de esta tesis.

Al Instituto Mexicano del Petróleo por el programa
de apoyo a la titulación.

A la Gerencia de Análisis de Inversiones en Explotación
y especialmente al Ing. Rodrigo Pulido Maldonado y
equipo de trabajo por el apoyo y los recursos
proporcionados para la conclusión de este trabajo.

A todos mis compañeros y amigos.

CONTENIDO

CONTENIDO

INTRODUCCIÓN.....	1
1. CONCEPTOS BÁSICOS.....	3
1.1 Marco de referencia.....	3
1.2 Población y Tipos.....	5
1.3 Modelos.....	5
1.3.1 Modelo del triángulo epidemiológico.....	5
1.3.2 Modelo de la multitud de causas.....	6
1.3.3 Determinación de causalidad.....	6
1.3.4 Validez de una investigación epidemiológica.....	7
1.4 Escalas de medición.....	8
1.4.1 Medidas usadas en epidemiología.....	9
1.5 Tipos de estudios.....	12
1.5.1 Aleatorizado o experimental.....	13
1.5.2 Quasi-experimental.....	14
1.5.3 Observacional.....	14
1.6 Criterios para elegir un diseño.....	15
2. METODOLOGÍA PARA UN DISEÑO DE CASOS Y CONTROLES.....	16
2.1 Apreciación global y perspectiva histórica.....	16
2.2 Definición de casos.....	17
2.3 Definición de controles.....	18
2.3.1 Tipos de controles.....	18
2.3.2 Razón de controles para los casos.....	19
2.4 Características generales.....	19
2.4.1 Ventajas y desventajas.....	20
2.5 Datos que se pueden estimar a partir de un estudio de casos y controles.....	22
2.5.1 Medición de asociación entre enfermedad y factores de riesgo.....	24
2.5.1.1 Riesgo Relativo.....	24
2.5.1.2 Riesgo atribuible.....	25
2.5.1.3 Razón de momios.....	26
2.6 Consideraciones para el análisis.....	29
2.6.1 Confusión e Interacción.....	30
3. DIAGNÓSTICO DE LA INFORMACIÓN.....	31
3.1 Definición de variables.....	33
3.2 Tablas de frecuencias simples.....	34
3.3. Análisis bivariado para las variables de interés.....	36
4. REGRESIÓN LOGÍSTICA.....	77
4.1 Introducción al análisis multivariado.....	77
4.1.1 Origen.....	77
4.1.2 Recomendaciones para el estudio del análisis multivariado.....	77
4.1.3 Clasificación de los métodos.....	78
4.2 Modelación estadística.....	81
4.2.1 Modelos lineales.....	82
4.2.2 Ajuste de modelos lineales para datos binomiales.....	84
4.2.3 La transformación logística.....	85

4.3 Definición del modelo logístico.....	85
4.3.1 Ajuste del modelo logístico lineal para datos binomiales.....	87
4.3.2 Comparación de modelos logísticos lineales.....	90
4.3.3 No convergencia y sobreajuste.....	93
4.3.4 Consideraciones para seleccionar el modelo.....	94
4.4 Interpretación de los coeficientes de un modelo logístico lineal.....	94
4.4.1 Factor de exposición simple dicotómica.....	95
4.4.2 Factor de exposición politómica.....	96
4.4.3 Variable de exposición continua.....	97
4.4.4 Modelos con combinación de diferentes tipos.....	98
4.5 Regresión logística aplicada a un estudio de casos y controles.....	98
5. ANÁLISIS DE RESULTADOS.....	115
5.1 Interpretación de los resultados.....	115
5.2 Selección del modelo.....	118
CONCLUSIONES.....	120
APÉNDICE.....	122
BIBLIOGRAFÍA.....	169

INTRODUCCIÓN

INTRODUCCIÓN

El ser humano es una parte fundamental en el desarrollo de la sociedad, y la capacidad para poner en práctica sus facultades depende en gran parte de la salud, ésta sin embargo puede ser afectada por diversos factores tales como hábitos personales, factores socioeconómicos, efectos del medio ambiente o las mismas características físicas del individuo.

Para poder explicar el comportamiento de una enfermedad no solo se toma en cuenta la literatura médica sino que ha sido necesario realizar muchos estudios y establecer relaciones con otras disciplinas del conocimiento.

En el área médica, los profesionales generalmente tienen datos auxiliares que sirven de apoyo para basar sus juicios clínicos, solo que para poder tomar una decisión, tienen que saber de donde provienen los datos y como fueron obtenidos, además deben conocer si las conclusiones basadas en ellos son estadísticamente válidas, en éste sentido la Estadística ha sido una disciplina fundamental en el estudio de la etiología de las enfermedades para comprender y evaluar cada vez mejor el fenómeno en estudio.

El presente trabajo es un estudio de casos y controles de leucemia desarrollado en la Unidad de Investigación en Epidemiología Clínica, del Hospital de Pediatría del Centro Médico Nacional Siglo XXI (IMSS).

El estudio de la leucemia comprende desde conocer sus mecanismos etiopatogénicos, hasta la rehabilitación en pacientes con secuelas, pasando por el diagnóstico y tratamiento adecuado y oportuno. En general no se sabe de un mecanismo único que origine esta enfermedad, sin embargo se conocen diferentes factores que se han encontrado asociados al desarrollo de la misma.

El objetivo principal de la investigación es determinar posibles factores de riesgo asociados al desarrollo de leucemia en niños residentes del Distrito Federal.

Tomando como punto de partida que se trata de un estudio comparativo observacional y de que la información base para el análisis de los datos se obtuvo a partir de un cuestionario, consideramos que la metodología adecuada a utilizar para poder explicar el comportamiento del fenómeno y alcanzar el objetivo es la siguiente:

En primer lugar se decide realizar un análisis descriptivo agrupando los datos mediante tablas de frecuencias simples para cada una de las variables en estudio pues creemos que esto nos permitirá conocer de manera general las características de la población y nos ayudará a sensibilizar cuáles son las variables que podrían ser de utilidad para la siguiente fase de análisis.

Como el interés principal es determinar si existe asociación entre la leucemia y los diferentes factores de riesgo que según la literatura médica se han mencionado como

posibles causas, se decide realizar un análisis bivariado bajo el supuesto de que la pareja de variables seleccionadas para cada prueba son independientes. Para corroborar este supuesto se considera conveniente utilizar tablas de contingencia debido a que es una herramienta muy útil que proporciona medidas sobre el grado de asociación que guardan las variables.

Dado que se trata de un estudio de casos y controles y puesto que los datos fueron colectados de manera retrospectiva, aunado a la característica de que el análisis bivariado solo permite el análisis de dos variables a la vez, consideramos que este análisis es importante, pero no suficiente para explicar el fenómeno, por lo tanto creemos que se requiere usar otra herramienta que nos permita obtener una representación matemática más general de la relación entre las variables. El análisis multivariado es en este sentido una herramienta más poderosa ya que permite el análisis simultáneo de más de dos variables.

En el entorno del análisis multivariado existe un gran número de métodos de análisis, sin embargo se considera que el método de regresión logística es el más adecuado para la última parte del análisis por lo siguiente:

- Permite explorar el efecto de muchos factores de riesgo.
- Es muy útil cuando la variable dependiente es dicotómica.
- Es particularmente apropiado para analizar datos dicotómicos que se han obtenido retrospectivamente.
- Permite controlar la presencia de variables de confusión y realizar interacciones.

Por lo tanto supondremos que si en el análisis bivariado se obtiene una aproximación clara de cuáles son las variables que están asociadas al fenómeno, con el método de regresión logística podremos realizar el análisis de todas las variables significativas al mismo tiempo, así como investigar si existe confusión o interacción entre las variables y de esta forma poder explicar cuáles son los factores de riesgo más importantes y proponer un modelo que se ajuste a los datos y nos acerque mas a la realidad.

El contenido general del trabajo puede resumirse de la siguiente forma:

En el capítulo uno se definen los conceptos tanto estadísticos como médicos principales para la comprensión del trabajo realizado. El capítulo dos contiene la parte teórica respecto a la metodología que se usa para llevar a cabo un estudio de casos y controles. En el tercer capítulo se definen las variables del estudio y se realiza la primera parte del análisis de los datos. El capítulo cuatro contiene la parte teórica respecto al método de regresión logística utilizado en el análisis final de los datos y finalmente en el capítulo cinco se hace el análisis de los resultados así como la selección del modelo.

CONCEPTOS BÁSICOS

1. CONCEPTOS BÁSICOS

1.1 Marco de referencia

Las estadísticas demográficas han sido muy útiles en el estudio de la etiología de las enfermedades, sin embargo existen muchas enfermedades poco comunes cuyo origen no se ha podido definir completamente.

En el mundo la primera causa de cáncer en niños es la leucemia, en Estados Unidos de Norteamérica (EUA) ocupa el 30% de todos los cánceres y la leucemia linfoblástica aguda el 85%, aunque en México no se conoce bien la frecuencia de ésta neoplasia, se sabe por reportes de algunos hospitales que la leucemia es la principal causa de atención en niños con cáncer¹.

Las muestras consideradas para el presente trabajo están formadas por dos grupos, una que corresponde a los casos con leucemia diagnosticada en cualquiera de sus variedades formada por 187 pacientes, y la otra corresponde a los controles con igual número de pacientes. Todos los niños incluidos en el estudio fueron seleccionados del Hospital de Pediatría del Centro Médico Nacional Siglo XXI y del Hospital General del Centro Médico la Raza, del Instituto Mexicano del Seguro Social.

La fuente de recopilación de datos fue un cuestionario codificado con las diferentes variables de estudio (ver apéndice). Para obtener la información se acudió al domicilio del niño en estudio y para algunos se obtuvo durante su internamiento en el hospital.

Los conceptos de casos, controles, así como las características del estudio se describen en los capítulos dos y tres, en este capítulo sólo se definen los conceptos básicos.

La **Estadística** es una ciencia que se ocupa, entre otras cosas, de recoger, organizar, analizar, interpretar y presentar la información que pueda ser expresada numéricamente. La **Bioestadística** es una rama de la estadística aplicada a las ciencias biológicas. La estadística es una parte vital de muchas investigaciones, desde la etapa inicial de un estudio (planeación), hasta la interpretación de resultados.

Un término general para cualquier característica o rasgo que puede medirse en un paciente o en otra unidad de estudio es la **variable**. Una variable se puede definir como un valor (medición), la estatura es un ejemplo de variable, se llama así porque entre una medición y otra sobre el mismo sujeto puede variar. Esta diferencia puede darse por variabilidad en los instrumentos que se usan y entre los que hacen la medición en cuanto a la forma de leer los instrumentos. Hay diferentes tipos de variables, sin embargo sólo se definen las que son de utilidad para el estudio.

V. Dependiente: es aquella para la cuál la distribución de las unidades en estudio en las diferentes categorías es meramente observada por el investigador.

¹Neglia JP. Robinson. Epidemiología de las leucemias agudas en la infancia. 1988; p.p. 727-745

V. Independiente: valor que puede ser o no controlado por el investigador.

Factor: variable en la cuál el investigador controla activamente la distribución de las unidades de estudio en las diferentes categorías.

Cuando una variable es observada como parte de un experimento o de un proceso de muestreo perfectamente definido, por lo general se le llama variable aleatoria.

Una distribución de probabilidad es un modelo para una variable aleatoria, que describe la forma en que la probabilidad está distribuida entre los valores posibles que la variable puede tener.

Los datos de todo tipo pero en especial los que se obtienen de observaciones sobre un elevado número de sujetos, deben organizarse de modo que colaboren a su propia interpretación. Un paso importante en la organización de los datos, es construir una distribución de frecuencia.

Dato cuantitativo o métrico: son mediciones usadas para identificar o describir sujetos u objetos por la cantidad o grado. Por ejemplo la edad y peso de una persona son ejemplos de datos cuantitativos.

Dato no cuantitativo (categórico): es un atributo, característica o propiedad categórica que puede ser usado para identificar o describir un objeto o sujeto. Por ejemplo: la ocupación o el consumo de algún producto (alto, medio, bajo).

Epidemiología: la epidemiología estudia la distribución de enfermedades en poblaciones humanas y los factores que determinan tal distribución.

Como objetivos principales en los estudios epidemiológicos pueden considerarse los siguientes:

- Describir el estado de salud de las poblaciones.
- Explicar la etiología de las enfermedades.
- Predecir la frecuencia de las enfermedades y estado de salud de las poblaciones.
- Controlar la distribución de las enfermedades.

Gran parte de la investigación epidemiológica gira en torno a la etiología de las enfermedades para conocer la historia natural de las mismas. La investigación etiológica consiste en la búsqueda de las causas de una enfermedad, las relaciones entre éstas y las magnitudes relativas de sus efectos en ella.

1.2 Población y tipos

Población: es un grupo de todas las unidades de estudio acerca de las cuáles una investigación en particular puede proporcionar información .

Población objetivo: es el grupo entero de unidades de estudio al cuál se tiene interés en aplicar las conclusiones.

Población real: es aquella de donde se obtuvo finalmente la muestra.

1.3 Modelos

Los modelos epidemiológicos tratan de explicar la ocurrencia de una enfermedad en términos de las interacciones entre factores que afectan el desarrollo de la misma. Los factores se han dividido en:

- ◆ **Factores del huésped o intrínsecos,** éstos factores afectan la susceptibilidad a la enfermedad (factores genéticos, de personalidad, clase social), por ejemplo.
- ◆ **Factores ambientales,** éstos influyen en la exposición y a veces afectan también indirectamente a la susceptibilidad, algunos de éstos factores pueden ser los agentes infecciosos, calor, presión atmosférica, agentes químicos, radiación etc..

Para poder establecer una relación causa-efecto se necesita como mínimo:

- Una secuencia temporal entre dos variables: la variable causa debe preceder en el tiempo a la variable efecto, es decir ésta no puede estar presente antes de la causa.
- El factor causa debe estar correlacionado con el factor efecto.
- La única explicación plausible del efecto observado tiene que ser el factor causa.

La condición suficiente y necesaria del principio de causalidad es expresada como sigue: Una variable A es causa de B si siempre que se da A acontece B, y nunca acontece B si previamente no se ha dado A. Sin embargo esta condición no se cumple siempre ya que la mayoría de las enfermedades se deben a más de una causa. En situaciones más complejas, la enfermedad X es causada por múltiples factores, digamos los factores A, B, C y D o el factor A puede causar más de una enfermedad. En otro ejemplo el factor A puede causar la enfermedad X sólo en presencia del factor B, de modo que si el factor A o B están presentes por separado, no puede observarse efecto causal alguno. Entre los modelos que se usan en estudios epidemiológicos se tienen los siguientes:

1.3.1 Modelo del triángulo epidemiológico

El modelo consta de tres componentes el huésped, el ambiente y el agente. En este modelo se establece que cada componente debe ser analizado y comprendido para poder captar y predecir las modalidades de una enfermedad. Un cambio en cualquier componente altera el equilibrio existente, aumentando o disminuyendo la frecuencia de la enfermedad.

1.3.2 Modelo de la multitud de causas

Aquí los efectos no dependen de las causas aisladas sino de cadenas de causas. Se considera que se puede evitar el desarrollo de una enfermedad cortando la cadena causal en diferentes puntos y por lo tanto no se requiere una comprensión completa de los mecanismos causales para adoptar medidas eficientes para la prevención y el control de una enfermedad.

Actualmente en el campo epidemiológico se utiliza *teoría de probabilidad y técnicas estadísticas* para probar y estimar a partir de observaciones, relaciones que se han hipotetizado como causales. En este campo no se utiliza el término factor causal sino "factor de riesgo", con el objeto de indicar que se trata de una variable que se "cree relacionada" con la probabilidad de que un individuo desarrolle la enfermedad. Al **factor de riesgo** se le define como el factor (agente) cuyo efecto en el resultado de interés está siendo estudiado. A diferencia del término "tratamiento", utilizado para afectar el factor resultado bajo consideración, se utiliza el término factor de riesgo cuando la exposición a dicho agente es accidental o incontrolable, o cuando se aplica para un propósito diferente al de afectar al factor resultado bajo estudio.

Para determinar cuál es el factor que se debe elegir como factor de riesgo, se debe considerar que esté asociado estadísticamente con la enfermedad; que la presencia del factor preceda a la ocurrencia de la enfermedad y finalmente que el efecto del factor (asociación) que se observa no sea completamente explicado por cualquier fuente de error.

1.3.3 Determinación de causalidad

Para determinar si un factor es causa de otro, denominado resultado, lo primero que se debe hacer es obtener un buen estimador² de la asociación que existe entre el supuesto factor causal y el factor respuesta. Después se debe determinar si la asociación observada es estadísticamente significativa, y de ser así, faltaría determinar si existe una relación causal.

A continuación se describen algunos criterios operacionales que ayudan a determinar si una relación es causal o no lo es.

- Fuerza de asociación (a mayor asociación, mayor es la evidencia de causalidad).
- Claridad en la secuencia de eventos.
- Existencia de una relación dosis-respuesta.
- Consistencia de los resultados.
- Coherencia en los resultados (se fortalece una interpretación causal si los resultados obtenidos no entran en conflicto con el conocimiento existente, sobre la historia natural de la enfermedad o con otros hechos ya aceptados).

²El estimador es una regla que nos dice como determinar a partir de cualquier muestra un valor numérico para estimar cierto parámetro de la población.

1.3.4 Validez de una investigación epidemiológica

La validez de un estudio se refiere a la presencia o ausencia de sesgos originados en el diseño o análisis, pueden hacer que esté distorsionada la estimación del efecto que el supuesto factor causal tiene en el factor resultado. El estimador de la medida de un efecto puede estar afectado por dos tipos diferentes de errores:

- **Errores aleatorios:** se originan por la naturaleza aleatoria del proceso de muestreo y producen una diferencia entre el estimador de la medida del efecto $\hat{\theta}$, y la medida del efecto que realmente se está estimando θ_0 . Esta diferencia depende del tamaño de la muestra, de las características del estimador que se esté utilizando y de la variabilidad de las características de interés en la población bajo estudio.
- **Errores sistemáticos (sesgos):** se originan por problemas metodológicos en el diseño o análisis del estudio. Estos errores dan origen a la diferencia entre lo que el estimador está realmente estimando θ_0 y la verdadera medida del efecto de interés θ .

El error total entre el estimador de la medida del efecto y la medida del efecto, se define como:

$$(\hat{\theta} - \theta) = (\hat{\theta} - \theta_0) + (\theta_0 - \theta)$$

errores aleatorios errores sistemáticos

En los estudios epidemiológicos se considera más importante controlar la validez que la precisión; es mejor un estudio sin sesgos aunque restringido en alcances a un estudio más general pero con problemas de validez.

En un estudio epidemiológico se dice que hay validez interna cuando no hay sesgo en el estimador, de tal forma que los resultados pueden extenderse a la población objetivo.

Los sesgos se pueden originar por:

- 1) Selección de sujetos para el estudio (sesgos de selección).
- 2) Medición de las variables de interés (sesgos de información).
- 3) Presencia de factores extraños (sesgos de confusión).

Los sesgos de selección son distorsiones en el estimador como consecuencia de la forma en la que se seleccionan los sujetos de la población. Estos sesgos pueden surgir por defectos en el diseño del estudio, por ejemplo:

- Errores por falta de información.
- Errores por supervivencia selectiva (muerte antes de finalizar el estudio).
- Errores en la selección de grupos que van a ser comparados.

Los sesgos de información surgen por errores en la medición de las variables de interés, puede ser que se utilice un instrumento de medición defectuoso, por ejemplo un cuestionario, una entrevista o bien que se lleve a cabo un diagnóstico diferente entre los grupos que se desean comparar.

Cuando se habla de factores extraños se hace referencia a "factores de confusión", en un estudio de causa y efecto se dice que los dos factores son confusos cuando no es posible determinar cuál factor causa el efecto que se está investigando.

Estrategias para controlar los factores de confusión a nivel de diseño:

- Se puede restringir el estudio a sujetos que tengan un mismo valor en los factores de confusión.
- Realizar apareamiento o acoplamiento (el objetivo de esto es formar los grupos de riesgo o tratamiento de tal forma que la distribución de los factores de confusión sea la misma en cada uno de ellos). El apareamiento se lleva a cabo buscando sujetos que tengan valores similares en los factores de confusión que se quieren controlar.

Estrategias para controlar los factores de confusión a nivel de análisis:

- Estratificación: es un procedimiento de ajuste utilizado cuando los factores de confusión son categóricos. Por ejemplo, se puede dar el caso en que se toma una muestra aleatoria simple de individuos en una comunidad y sucede que no hay mujeres en la muestra, el sexo podría ser un factor importante en lo que se está estudiando y, si éste es el caso, podría haberse obtenido una muestra muy deficiente. Para evitar esa posibilidad, se toma una muestra estratificada obtenida separando las unidades de estudio de la población en grupos no sobrepuestos, llamados estratos, y entonces seleccionar la muestra de cada estrato, en este caso la población es dividida en estratos según sexo, de esta forma se puede asegurar que cada sexo está representado en proporción a su distribución en la población. De igual modo podría estratificarse a la población en base a edad, estrato socioeconómico etc..
- Utilización de modelos matemáticos: se utilizan modelos matemáticos para explicar la relación entre el factor respuesta y los factores de riesgo y de confusión, obteniendo estimadores de la medida del efecto del factor de riesgo, ajustados por la presencia de factores de confusión. Entre los modelos más usados en epidemiología se tienen el modelo logístico y el análisis de covarianza.

1.4 Escalas de medición

Las escalas de medición pueden dividirse básicamente en: escala cuantitativa, escala nominal y escala ordinal. La escala cuantitativa permite la medición de todos los valores fraccionarios posibles dentro de un intervalo; permite obtener medidas como 169,7 cm, 33 grados en el caso de la temperatura. En estos casos la característica que se está midiendo es *cuantitativa* y a un conjunto de tales mediciones se le conoce como datos continuos, estatura, peso, presión arterial son ejemplos de características cuantitativas que se miden

generalmente con escalas de intervalos. El número de integrantes en una familia es también un dato cuantitativo, pero no se mide en escala de intervalos, en el caso de que sólo se permiten datos enteros se les llama datos discretos.

Quando se clasifican las mediciones en categorías como "alto", "medio" o "bajo" por ejemplo, la característica es cualitativa y éstas mediciones dan origen a **datos categóricos** que *consisten de cantidades o número de individuos en cada categoría*. Hay dos tipos de datos categóricos, dependiendo de si es natural o no la secuencia en que pueden ordenarse éstos. En el ejemplo anterior hay un orden natural; es decir medio está entre bajo y alto, en este caso la escala de medición se llama ordinal. Si no hay orden natural, la escala se llama nominal y aquí las categorías solo tienen nombres y no hay implicada secuencia alguna, por ejemplo el color de cabello (castaño, rubio, negro).

1.4.1 Medidas usadas en epidemiología

Medidas de frecuencia.

Se utilizan para caracterizar la ocurrencia de un factor respuesta dicotómico, éstas medidas son importantes tanto en investigaciones descriptivas como etiológicas. Pueden expresar el grado relativo de morbilidad y mortalidad.

Existen dos tipos de medidas de morbilidad:

- *Medidas de incidencia:* son medidas de frecuencia basadas en casos nuevos que ocurren durante un lapso de tiempo. Para ajustar la incidencia al tamaño de la población que se está observando, por lo general se mide como la razón del número de casos nuevos que ocurren en un tiempo dado al número de individuos inicialmente en riesgo de desarrollar la enfermedad.
- *Medidas de prevalencia:* Por lo general se mide como la razón del número de casos en un momento dado en el tiempo, al número de individuos en la población que interesa en ese momento en el tiempo.

La figura 1 muestra la diferencia entre incidencia y prevalencia. Supóngase que cada segmento de línea representa un caso de enfermedad desde el inicio (principio del segmento lineal) hasta que la enfermedad ha llegado a su final (fin del segmento lineal), además; considérese que 1.000 personas están en riesgo en un momento cualquiera del día. La incidencia para el día 1 es 4 casos por cada 1.000 personas (4 segmentos lineales nuevos) y para el día 2 es de 2 casos por 1.000 personas (2 segmentos lineales nuevos). La prevalencia al final del día 1 es de 4 por 1.000 (existen 4 segmentos lineales) y al fin del día 2 es de 6 (existen 6 segmentos lineales). Es obvio que dos enfermedades pueden tener una incidencia idéntica y una podría tener una prevalencia mayor si su duración (tiempo desde el inicio de la enfermedad hasta el fin de su curso) es grande.

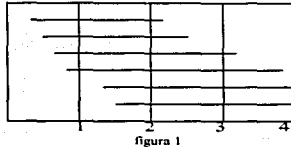


figura 1

(seis casos de una enfermedad representada al transcurso del tiempo por segmentos lineales)

Hay dos medidas de incidencia:

1. Riesgo.
2. Tasa.

El riesgo es la probabilidad de que un individuo no enfermo desarrolle una enfermedad específica, en un período de tiempo determinado. Se han desarrollado varios métodos para medir la cantidad relativa de enfermedad que hay en diferentes poblaciones.

Riesgo relativo

$RR = \frac{\text{Incidencia de la enfermedad en el grupo expuesto}}{\text{Incidencia de la enfermedad en el grupo no expuesto}}$

Razón³ de momios⁴

$OR = \frac{\text{Momio en favor de la enfermedad en el grupo expuesto}}{\text{Momio en favor de la enfermedad en el grupo no expuesto}}$

Riesgo atribuible

Se define como la incidencia de la enfermedad en un grupo expuesto menos la incidencia de la enfermedad en un grupo no expuesto.

El concepto de tasa es similar al de proporción⁵, excepto que se usa un multiplicador (por ejemplo 1000, 10000 o 100000) y se calcula por un espacio específico de tiempo. El multiplicador recibe el nombre de base.

³ La razón es una medida de la rapidez de cambio de un fenómeno en general por unidad de tiempo.

⁴ El momio en favor de un evento se define como la frecuencia con que el evento ocurre dividida entre la frecuencia con la que no sucede.

⁵ Es el número de observaciones con una característica dada dividida entre el número total de observaciones.

$$tasa = \left(\frac{a}{a+b} \right) * base$$

La tasa de incidencia es el potencial instantáneo para el cambio en el estado de enfermedad por unidad de tiempo, al tiempo t , relativo al tamaño de la población no enferma al tiempo t .

Para determinar como medir la incidencia, se debe definir que medida se va a usar (riesgo o tasa) y cómo se va a estimar esa medida, el criterio para tomar una decisión depende del objetivo del estudio. Si el objetivo es predecir un cambio individual en el estado de salud con base en ciertas características, se deberá conocer el "riesgo" de desarrollar la enfermedad.

Por otro lado si el objetivo es investigar la etiología de las enfermedades, tomando en cuenta los efectos de algunos factores, la forma de medir la incidencia depende del tipo de enfermedad y el tiempo durante el cuál se observa la ocurrencia de nuevos casos.

La prevalencia mide la cantidad de enfermedad en una población en un momento dado en el tiempo. Debido a que la prevalencia está en función de la duración de una enfermedad, la mayor aplicación se da en la planeación de servicios al cuidado de salud para esa enfermedad. Para medir la prevalencia se puede utilizar la prevalencia puntual y la prevalencia en un período.

La prevalencia puntual es la probabilidad al tiempo t , de que un individuo en una población, padezca la enfermedad que se está estudiando, el estimador que se usa se define en la siguiente expresión.

$$\hat{p}_t = \frac{C_t}{N_t} \text{ donde:}$$

N_t es el tamaño de la población en estudio al tiempo t .

$C_t = N_t - N_s$ es el número de casos con prevalencia al tiempo t .

N_s es el número de individuos sanos al tiempo t .

con la prevalencia puntual se puede calcular el cociente entre la probabilidad de estar enfermo al tiempo t y la probabilidad de no estarlo.

La prevalencia en un período es la probabilidad de que un individuo en una población esté enfermo, en cualquier tiempo dentro de un período dado (t_0, t) de duración $\Delta_t = (t - t_0)$. La prevalencia se estima como:

$$P(t_0, t) = \frac{C(t_0, t)}{N} = \frac{C_0 + I}{N}$$

donde $C(t_0, t)$ es el número de personas que tenían la enfermedad en cualquier tiempo durante el período de seguimiento (t_0, t) . Incluye a los enfermos en t_0 y a los casos incidentes (I) , detectados durante el período.

Medidas de mortalidad

Las medidas de mortalidad son medidas de frecuencia análogas a las de incidencia, sólo que aquí el interés es la muerte. Una medida para expresar la frecuencia de muerte es la función de riesgo del tiempo de supervivencia T . Se puede definir como la probabilidad "instantánea" de falla en un tiempo particular, dado que hay supervivencia hasta ese momento.

Medidas del efecto del factor de riesgo

Son medidas que sirven para evaluar el efecto que el factor de riesgo tiene en el factor respuesta, con la finalidad de investigar la etiología, tratamiento y prevención de enfermedades.

En función de las escalas con las que se miden los factores de riesgo y respuesta, se describen algunos casos:

1. Riesgo categórico y respuesta dicotómica: en este caso la escala de medición de la variable riesgo es categórica y de la variable respuesta también pero solo tiene dos categorías. Las medidas del efecto en este caso involucran una comparación directa de medidas de frecuencia para diferentes categorías del factor de riesgo. Se compara la medida de frecuencia de cada categoría con la medida de frecuencia de una categoría específica. A las medidas del efecto para este caso, se les puede dividir en medidas de razón (razón de tasas de incidencia, riesgo relativo por ejemplo) y medidas de diferencia (diferencia de tasas de incidencia, riesgo atribuible por ejemplo).
2. Riesgo numérico y respuesta dicotómica: para este caso la variable riesgo es numérica y la variable respuesta es categórica con dos categorías (éxito o fracaso). En este caso se utiliza un modelo conocido como logístico lineal, para explicar la relación entre el factor riesgo y respuesta.
3. Riesgo dicotómico y respuesta continua: en este caso la variable riesgo es categórica y la variable respuesta es numérica. En este caso como medida del efecto puede emplearse la diferencia de medias de la variable respuesta entre los grupos de exposición.

1.5 Tipos de estudios

Un estudio puede ser descriptivo o comparativo. En un estudio descriptivo se tiene una sola población, se lleva a cabo cuando se conoce poco acerca de la ocurrencia o la historia

natural de una enfermedad. Son útiles para explorar asociaciones entre factores de riesgo potenciales y una enfermedad, cuando se tienen conocimientos limitados.

En un estudio comparativo se tienen dos o más poblaciones que se comparan respecto a ciertas características. La finalidad es identificar los factores de riesgo para las enfermedades, la estimación de sus efectos en la misma y el planteamiento de posibles estrategias de intervención.

Los estudios se pueden dividir en:

- Aleatorizado o Experimental.
- Quasi-experimental.
- Observacional.

1.5.1 Aleatorizado o experimental

Un estudio aleatorizado es aquél en el cual las categorías del factor bajo estudio (tratamiento) se asignan aleatoriamente a los sujetos bajo investigación. La distribución aleatoria de las unidades de estudio a los grupos de tratamiento es un seguro automático contra posibles factores de confusión además se pueden hacer inferencias estadísticas formales.

Existen varios tipos de diseños experimentales, algunos son:

1. Un diseño totalmente aleatorio: es aquel en el que cada unidad de estudio tiene la misma probabilidad de ser asignado a cualquier grupo de tratamiento bajo consideración.
2. Un diseño por bloques aleatorios: es aquel en el que las unidades de estudio son separadas primero en bloques (estratos) no traslapados y después asignados al azar a los grupos de tratamiento separadamente dentro de cada bloque.
3. Diseño cruzado: en este tipo de diseño el objetivo es que la unidad de estudio (paciente) sea su propio control. Por ejemplo si se están investigando dos tratamientos, A y B, un grupo de pacientes podría asignarse al tratamiento A por 2 semanas y después al B por el mismo tiempo. Un segundo grupo se asigna al mismo tratamiento en sentido contrario primero (B y después A). En este diseño la existencia de un efecto residual del tratamiento dado en el primer período puede afectar los resultados encontrados en el segundo.
4. Pruebas o ensayos clínicos: son estudios experimentales que comprenden a seres humanos como unidades de estudio.

1.5.2 Quasi-experimental

Un estudio Quasi-experimental es aquel en el que el investigador asigna las categorías del factor de tratamiento a los sujetos bajo investigación pero no aleatoriamente sino por conveniencia de los propios sujetos.

1.5.3 Observacional

Un estudio observacional es aquel en el que el investigador no asigna las categorías del factor de riesgo a los sujetos bajo investigación sino que simplemente observa la categoría que posee cada sujeto. Dependiendo de si la información que se requiere se obtiene en un sólo punto en el tiempo o en varios puntos, o si se hace o no un seguimiento a través del tiempo de los sujetos, un estudio puede clasificarse en:

- **Transversal:** en este tipo de estudio la información que se requiere se obtiene en un sólo punto hipotético en el tiempo y por lo mismo no puede haber un seguimiento de los sujetos bajo investigación. Generalmente se planean para responder a: ¿qué es lo que está pasando en este momento?.
- **Longitudinal:** en este caso se obtiene la información que se requiere, sobre los mismos aspectos, en dos o más puntos hipotéticos en el tiempo, o bien se hace un seguimiento de los objetos bajo investigación.

De acuerdo al tiempo de ocurrencia del factor de riesgo o tratamiento y del factor respuesta relativo al período en el que se lleva a cabo la investigación, un estudio puede ser:

- **Prospectivo:** en un estudio prospectivo se parte de la medición del factor de riesgo o tratamiento en la población sin enfermedad y después de cierto tiempo se mide el valor del factor respuesta.
- **Retrospectivo:** en un estudio retrospectivo se parte de la medición del factor respuesta y posteriormente se obtiene la medición sobre la exposición previa al factor de estudio, es decir; se procede del efecto a la causa potencial, (el factor de riesgo y el factor respuesta ocurren antes del inicio de la investigación).
- **No direccionado:** en este caso la medición sobre el factor de riesgo y sobre el factor respuesta se refieren al mismo punto en el tiempo.

Dos diseños básicos usados en epidemiología para investigar la etiología de una enfermedad son:

ESTUDIO COHORTE

En este diseño se conoce la información sobre el factor de riesgo al principio del período de seguimiento; es comparativo, longitudinal y por lo general se trabaja con casos de incidencia. Los estudios cohorte responden a la pregunta ¿que pasará? y por lo tanto, su

dirección en el tiempo es hacia delante o prospectivo (figura 2). Durante la investigación se observan a los grupos (expuestos y no expuestos) por cierto tiempo para constatar el efecto de éstas características definidas.

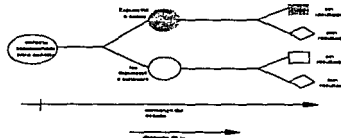


figura 2

ESTUDIO DE CASOS Y CONTROLES

El enfoque de este estudio es identificar una muestra de pacientes con la enfermedad que interesa "casos" y una segunda muestra de personas que no tienen la enfermedad "controles" y después comparar la frecuencia de las causas probables de la enfermedad en los dos grupos. En un estudio caso-control, los individuos son seleccionados en base a la presencia o ausencia de patología y se determina la frecuencia de los factores causales posibles en sus historias clínicas. Un estudio de casos y controles es un estudio retrospectivo o hacia atrás y que puede trabajar con casos de incidencia o prevalencia, se llevan a cabo para contrastar alguna hipótesis etiológica específica, sin embargo cuando no se tiene clara una hipótesis etiológica específica, se hace para explorar los antecedentes de las personas afectadas y no afectadas, con base en varias hipótesis etiológicas posibles.

1.6 Criterios para elegir un diseño

La elección de un diseño depende de varios factores:

- Objetivo de la investigación.
- Características de la enfermedad que se estudia.
- Disponibilidad de la información.
- Disponibilidad de tiempo y recursos.
- Consideraciones éticas.

En los estudios epidemiológicos la naturaleza de la enfermedad que se estudia es un factor determinante en la elección de un diseño. Si la enfermedad es poco frecuente o con largos períodos de latencia, sería mejor utilizar un estudio de casos y controles en vez de un cohorte, si por el contrario la enfermedad es muy frecuente y con períodos cortos de latencia conviene utilizar un estudio cohorte. El nivel de conocimiento existente acerca de la enfermedad en estudio determinará a su vez los objetivos de la investigación, puede ser que se tenga en mente una hipótesis etiológica específica o que sólo se quiera explorar asociaciones entre múltiples factores de riesgo o tratamiento y la enfermedad.

METODOLOGÍA PARA UN DISEÑO DE CASOS Y CONTROLES

2. METODOLOGÍA PARA UN DISEÑO DE CASOS Y CONTROLES

2.1 Apreciación global y perspectiva histórica

Los orígenes del método caso-control tiene sus raíces en la epidemiología misma. El desarrollo de algunas aproximaciones comienza en 1830 por P. C. Luis en París, sin embargo no es sorprendente que la investigación médica en el siglo XIX fue clave para el establecimiento del campo epidemiológico. En Londres fue formada la Sociedad Epidemiológica⁶, en parte, como respuesta a la epidemia del cólera; las primeras investigaciones consistieron de una descripción de la enfermedad de interés, una discusión de los casos conocidos y la cronología de desarrollo, solo algunas indagaciones calcularon tasas y compararon éstas en los subgrupos.

El desarrollo de la epidemiología durante la segunda mitad del siglo XIX fue dominado por la revolución bacteriológica. El paradigma antiguo que había considerado como antecedente importante los factores sociales en la enfermedad, fue remplazada por la teoría del microbio.

Si las bacterias fueron agentes de la enfermedad, entonces un epidemiologista necesita solamente trazar el punto de contacto de la infección para determinar la etiología de la enfermedad.

Un gran número de desarrollos en la primera mitad del siglo XX contribuyeron a la evolución del método caso-control. Estos incluyeron el desarrollo de métodos en Sociología y dentro de la comunidad médica, los casos clínicos serios fueron una aproximación estándar para la investigación de enfermedad. La inclusión de estadísticos en la salud pública cambió notablemente el firmamento epidemiológico. El establecimiento de la Bioestadística como una disciplina central en escuelas de salud pública fue otro foco para cerrar la interacción entre epidemiólogos y estadísticos.

En E.U. el investigador Dorn⁷ simuló muchos de los pensamientos que tomaron lugar durante los 50's, y algunas de las mejores innovaciones estadísticas facilitaron la difusión del estudio de caso-control. Un ejemplo es la idea de Cornfield⁸, para usar la razón de momios y estimar el riesgo relativo en una investigación de caso-control. El desarrollo subsecuente del método caso-control fue resultado de la contribución de muchos epidemiólogos, estadísticos y sociólogos.

En la actualidad el método caso-control se usa en una variedad de aplicaciones que van más allá de la investigación etiológica y están enfocadas a resolver un gran número de problemas dentro del sector salud y la medicina.

⁶ Thompson. *Statistical Analysis of Case-Control Studies*. Epidemiologic Reviews, Vol.16, p.p. 9

⁷ Thompson. *Statistical Analysis of Case-Control Studies*. Epidemiologic Reviews, Vol.16, p.p. 12-14

⁸ Kleinbaum DG. *Epidemiologic Research*, 1982, p.p. 4-6

2.2 Definición de casos

El grupo formado por sujetos que representan un resultado particular (enfermedad, muerte), se les conoce como **casos**. Algunos autores conceptualizan el problema de definición de casos en términos de *sensibilidad*, *especificidad* y sus efectos de validación.

Para ilustrar los conceptos de sensibilidad y especificidad se tomarán como base los datos de la tabla 1, considerando el registro de hospital como la fuente más segura de información acerca de si la paciente fue o no expuesta a los rayos X durante el embarazo. La *sensibilidad* es el grado en que los pacientes que verdaderamente manifiestan una característica son clasificados como tales; en este ejemplo, la sensibilidad de la aseveración de la madre de que había sido expuesta a los rayos X es 24/37, ó 65 %. *Especificidad* es el grado en que los pacientes que no manifiestan una característica son clasificados en forma correcta; en este ejemplo es 31/38 u 82 %.

Historia de hospital	Total	Aseveración del paciente		
		Con rayos X	Sin rayos X	No sabían
Con rayos X	37	24	10	3
Sin rayos X	38	2	31	5
Total	75	26	41	8

Tabla 1

La misma información, proporcionada por la determinación de la sensibilidad y de la especificidad, se obtiene calculando los valores predictivos (falsos positivo y falso negativo) de los datos provenientes de lo que se considera la fuente menos segura.

El valor falso positivo es la proporción de personas que en realidad poseen una característica manifestada por ellas, en este caso 24/26, ó 92 %. El valor falso negativo es la proporción de personas que verdaderamente no tienen una característica que han mencionado no tener, en este ejemplo es 31/41, ó 76%.

Cuando se desea evaluar pruebas prediagnósticas destinadas a descubrir enfermedades en poblaciones, es preferible considerar la validez en términos de sensibilidad y de especificidad y no del valor predictivo. Sin embargo, cuando se considera la validez de la información obtenida en los estudios de casos y controles hay poco que escoger entre los dos métodos, sobre todo cuando se basan en datos de entrevistas.

Para definir los casos se necesita:

- Describir de manera clara y precisa el problema a investigar, así como el criterio a utilizar.
- Determinar las fuentes de donde se obtendrán los casos, pueden ser:

1. Todos o una muestra de preferencia aleatoria de las personas que presentan el resultado, atendidas en una determinada institución durante un intervalo de tiempo. El marco de muestreo pueden ser las hojas de admisión hospitalaria o las listas de diagnóstico de egreso del paciente.
 2. Todos o una muestra de las personas que presentan el resultado de una población más general, (municipio, estado); en cierto período de tiempo.
- Definir si se trabaja con casos de incidencia o de prevalencia. Si el objetivo del estudio es la etiología de la enfermedad es preferible trabajar con casos nuevos, diagnosticados en un período de tiempo determinado.

2.3 Definición de controles

El grupo formado por sujetos que no presentan el resultado específico (enfermedad, muerte), se conoce como "controles" o testigos.

El propósito del grupo control es determinar la tasa o riesgo de exposición que tendrían los casos, suponiendo que no hay asociación entre el factor de riesgo y la enfermedad. Para elegir el grupo de controles se deben tomar en cuenta los siguientes aspectos:

- Considerar que la información para los controles se pueda obtener de manera similar a los casos.
- Decidir si se realizará un apareamiento como herramienta para el control de factores de confusión.
- Cuidar que los controles se tomen de una población similar a la que originó los casos.

2.3.1 Tipos de controles

Algunos de los mas comunes son:

- Los miembros de la institución (hospital) de donde se obtuvieron los casos, pero que no presenten el resultado.
- La población de un área determinada (es adecuada siempre y cuando los casos sean representativos de los que ocurren en esta población).
- Los familiares de los casos.
- Muerte certificada.

De éstos, el mas usado es el primero (controles hospitalarios), debido a que se pueden seleccionar controles con personas que registren patrones similares al de los casos, y que provengan del mismo estudio base.

Los controles pueden seleccionarse por muestreo aleatorio simple, aleatorio estratificado o una selección que garantice un apareamiento para controlar los factores de confusión.

2.3.2 Razón de controles para los casos

Dentro de las consideraciones estadísticas para definir la razón de controles para los casos, algunos autores sugieren que cuando el número de sujetos disponibles en un grupo (casos) es limitado, un incremento en el otro grupo (control) aumenta la potencia del estudio. El incremento de controles a casos conduce a una ganancia en potencia alcanzando hasta una razón 4 a 1.

2.4 Características generales

La característica central de un diseño de casos y controles es la comparación de dos grupos, uno con una característica específica y otro sin la característica. La frecuencia de los factores sospechosos hipotetizados del individuo relacionado con la característica es comparado en los dos grupos. Como en otros métodos de investigación epidemiológica es necesario empezar definiendo y delineando el problema bajo consideración. La definición del problema dictará los pasos a seguir en el estudio, particularmente la selección de casos y controles.

El contexto de solución de un problema radica en dos tipos de acciones:

- la necesidad de prevenir y tratar los casos o resolver el problema en cuestión y
- la necesidad de prevenir la ocurrencia futura de casos.

El método de caso-control ayuda a proporcionar la información necesaria para emprender ambas acciones.

En un estudio de casos y controles, una muestra de casos es seleccionada de la población de individuos quienes tienen la enfermedad a ser estudiada y una muestra de controles es seleccionada de individuos que no la tienen. La información acerca de cuáles factores están asociados con la enfermedad es obtenida retrospectivamente para cada persona en el estudio. Por tal razón el estudio es conocido como retrospectivo.

Para obtener un grupo de controles que sea comparable con los casos, los primeros pueden ser apareados con los últimos respecto a variables de confusión. Así por ejemplo, cuando un caso particular entra a un estudio, una o más personas del mismo sexo, raza, edad y condición socioeconómica semejantes, son identificados e introducidos al grupo control. A menudo, el propósito de un estudio de caso-control es aprender lo suficiente acerca de las causas de una enfermedad de interés para estrechar los límites de posibilidades.

Usualmente en este tipo de estudios se pretende contrastar una hipótesis etiológica concreta, sin embargo, en las ocasiones en que no se tiene clara una hipótesis específica entonces se

realizan para explorar los antecedentes de las personas afectadas y no afectadas por la enfermedad, en base a varias hipótesis plausibles.

La figura 3 ilustra de manera general el proceso del estudio, el cuadrado representa a un primer grupo de individuos que vienen siendo los casos, mientras que el otro grupo representa los controles, los antecedentes de los casos y controles se analizan por un período previo para detectar la presencia (área sombreada) o ausencia (área clara) de características o factores de riesgo predisponentes.

En los diseños de caso-control, la naturaleza de la indagación es de retroceso como lo indica la flecha que apunta de regreso para mostrar la dirección del proceso de investigación.

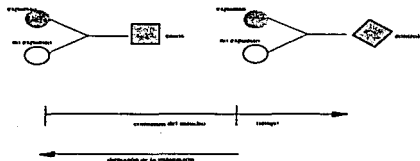


figura 3

La principal diferencia en el enfoque que se le da al estudio (prospectivo o retrospectivo), es que en un estudio cohorte (prospectivo) se inicia con un factor de riesgo o exposición y busca las consecuencias, un estudio de caso-control (retrospectivo) toma la consecuencia como punto de partida e investiga los precursores o factores de riesgo.

2.4.1 Ventajas y desventajas

Ventajas.

Las dos principales ventajas en este tipo de estudios están relacionadas con la eficiencia e informatividad. Se considera eficiente porque permite tratar con resultados raros. El método es también muy informativo. Permite probar los efectos e interacciones de un gran número de factores relacionados con los resultados obtenidos. En términos generales las ventajas se pueden resumir de la siguiente manera:

Un diseño de caso-control:

- Es apropiado para observar enfermedades poco comunes.

- Es ideal para investigadores que requieren obtener algunos datos preliminares para escribir un protocolo de un estudio más completo y prolongado.
- Permite analizar la influencia de varios factores de riesgo mediante un solo estudio.

Un diseño cohorte:

- Permite describir la experiencia subsecuente a la exposición (evaluación de la enfermedad, tasas de progresión etc.).
- Se puede controlar más la calidad en la medición de las variables.
- Permite explorar la presencia de otros factores de riesgo.
- El nivel del factor de estudio se observa antes de que se detecte la enfermedad, por lo tanto se puede asegurar que la causa hipotetizada precedió a la ocurrencia de la enfermedad.

Un diseño transversal:

- Es el que mejor determina la situación en un momento dado de una enfermedad y sirve para evaluar procedimientos de diagnóstico.
- Es útil para planear servicios y programas de salud.

Desventajas

Para un diseño de caso-control se pueden mencionar las siguientes:

- Como no es prospectivo, la capacidad del investigador para distinguir si el factor de riesgo precede realmente al factor respuesta, dependerá de la calidad de la información obtenida retrospectivamente sobre el factor de riesgo.
- Pueden existir tanto sesgos de selección, como sesgos de información.
- No se puede realizar una eliminación efectiva de algunos factores de confusión potenciales en la relación temporal correcta, puesto que los atributos se igualan al final de la evaluación.

Para un estudio cohorte:

- Generalmente es de larga duración y alto costo.
- Se pueden perder sujetos por falta de participación o muerte, por causas ajenas a las estudiadas si el período de seguimiento es largo.
- No es muy útil para estudiar enfermedades poco frecuentes.

Cuando el estudio es transversal se pueden mencionar las siguientes:

- Proporciona solo una "instantánea en el tiempo" de la enfermedad o proceso, que podría conducir a información errónea si la cuestión investigada es en realidad parte del proceso patológico.
- No es útil para estudiar enfermedades raras o de corta duración.

2.5 Datos que se pueden estimar a partir de un estudio de casos y controles.

Los datos que se pueden estimar en un estudio de casos y controles depende mucho del diseño del estudio, así como de la información que se tenga disponible tanto para los casos como para los controles, además de las mediciones que se hayan realizado a cada individuo considerado en el estudio.

Mientras que un número de diferentes mediciones por ejemplo, intensidad, duración, tiempo de exposición a algún factor, etc., puede ser obtenida de un estudio de casos y controles, la colección de datos tiende a concentrarse en la intensidad (nivel de exposición y frecuencia por ejemplo). Otro aspecto de la colección de datos que merece una consideración es la escala de medición, dependiendo de los datos que se quieren obtener, se puede utilizar:

- Una escala ordinal con límites especificados 1 -10 cigarros por día, 11-20 cigarros por día, por ejemplo.
- Escala ordinal con límites no especificados (bajo, medio, alto) si se consideran niveles.
- Escala nominal, preguntas con respuestas definidas sin orden.
- Escala binaria con respuesta (siempre o nunca, si o no) por ejemplo.

Las escalas continuas son generalmente recomendadas para variables continuas, dado que ofrecen los métodos más flexibles de evaluación del comportamiento de riesgo de acuerdo a la exposición. Con mediciones continuas por ejemplo, los efectos lineales y no lineales pueden ser examinados o la medición puede ser reducida a categorías y el riesgo de la enfermedad examinada en base a tales categorías.

En la práctica la intensidad de exposición puede ser medida en ambas escalas, continua o binaria, dependiendo del tipo de exposición. Por ejemplo la intensidad es más comúnmente medida sobre escala continua, por ejemplo, para la dieta el consumo de nutrientes, y en escala binaria la variable podría ser, la exposición a ciertos factores, el uso de anticonceptivos, por ejemplo.

Todas las investigaciones comienzan con la epidemiología descriptiva para caracterizar tiempo, lugar y personas entre los casos identificados inicialmente, esta información llega a ser la base sobre la cuál se hacen recomendaciones para definir los controles, el diseño y las variables a estudiar. De manera general los datos que se pueden obtener son los que se mencionan a continuación.

- Si se obtiene una muestra aleatoria de la población y se le observa durante un período de tiempo (0, T), se pueden estimar los siguientes parámetros:

- p Proporción de individuos expuestos al inicio del período.
 - P_1 Probabilidad o riesgo de que un sujeto expuesto desarrolle la enfermedad durante el período de observación.
 - P_0 Riesgo de que un sujeto no expuesto desarrolle la enfermedad durante el mismo período.
 - Δ_1 Tasa de incidencia de la enfermedad, acumulada para el período de duración τ para el grupo de expuestos.
 - Δ_0 Tasa de incidencia acumulada, para el período de duración τ , para el grupo que no está expuesto.
 - P_1/P_0 Probabilidad de que un sujeto expuesto desarrolle la enfermedad entre la probabilidad de que uno no expuesto desarrolle la enfermedad durante el mismo período.
 - Δ_1/Δ_0 Tasa de incidencia acumulada para el grupo expuesto entre la tasa de incidencia acumulada del grupo no expuesto.
 - P_1' Probabilidad de exposición dado que el individuo está enfermo $P(\text{Expuesto} | \text{Enfermo})$.
 - P_0' Probabilidad de exposición dado que el individuo no está enfermo $P(\text{Expuesto} | \text{No enfermo})$.
 - Ψ razón de momios.
- Si se toman muestras aleatorias de expuestos y no expuestos al principio del estudio, entonces se pueden estimar:
 - P_1 Probabilidad o riesgo de que un sujeto expuesto desarrolle la enfermedad durante el período de observación.
 - P_0 Riesgo de que un sujeto no expuesto desarrolle la enfermedad durante el mismo período.
 - Δ_1 Tasa de incidencia de la enfermedad, acumulada para el período de duración τ para el grupo de expuestos.
 - Δ_0 Tasa de incidencia acumulada, para el período de duración τ , para el grupo que no está expuesto.
 - P_1/P_0 Probabilidad de que un sujeto expuesto desarrolle la enfermedad entre la probabilidad de que uno no expuesto desarrolle la enfermedad durante el mismo período.
 - Δ_1/Δ_0 Tasa de incidencia acumulada para el grupo expuesto entre la tasa de incidencia acumulada del grupo no expuesto.
 - Si se realiza un estudio de casos y controles, y se investiga su exposición a ciertos factores en forma retrospectiva, se pueden estimar:

- $P(\text{Expuesto}|\text{caso})$ Probabilidad de exposición a factores de riesgo dado que es un caso.
- $P(\text{Expuesto}|\text{control})$ Probabilidad de exposición a factores de riesgo dado que es un control.
- Ψ Razón de momios de exposición.

sin embargo en algunas circunstancias es posible estimar algunos otros parámetros a partir de un diseño de casos y controles, por ejemplo:

- ◊ Cuando los casos son todos los que existen en una población finita en un determinado período o son una muestra probabilística de los mismos, y los controles son una muestra probabilística de la misma población pero de los sujetos que no padecen la enfermedad en estudio, es decir cuando las probabilidades de selección de casos y controles son conocidas, entonces es posible estimar los riesgos de enfermedad de manera separada para los grupos de exposición, el riesgo relativo y riesgo atribuible (diferencia en riesgos).
- ◊ Cuando no se conocen las probabilidades de selección de casos y controles, pero se sabe que los controles se obtuvieron de la misma población que originó a los casos, no se puede estimar el riesgo de enfermedad por grupo de exposición. Sin embargo si no hay sesgos de selección y P_1 y P_0 son pequeños (enfermedad poco frecuente), se puede estimar la razón de tasas de incidencia acumulada o el riesgo relativo para el período, aproximados mediante la razón de momios de enfermedad, mediante el estimador de la razón de momios de exposición.

2.5.1 Medición de asociación entre enfermedad y factores de riesgo.

La probabilidad de ocurrencia de una enfermedad durante un período de tiempo dado, conocido también como el riesgo de ocurrencia de la enfermedad, es el número de casos nuevos de la enfermedad que ocurren en el período de tiempo, expresados como una proporción de la población en riesgo.

El principal propósito de la estimación de riesgo es el poder hacer comparaciones entre el riesgo de enfermedad con diferentes niveles de un factor de exposición, y así estimar la asociación entre la enfermedad y éstos factores de riesgo.

2.5.1.1 Riesgo relativo

Suponga que un factor de riesgo particular tiene dos valores o niveles distintos, llamados expuesto y no expuesto. Si denotamos el riesgo de ocurrencia de una enfermedad durante un período de tiempo por p_e para la persona expuesta y p_0 para la no expuesta. El riesgo de ocurrencia de la enfermedad en una persona expuesta relativo a una persona no expuesta es

$$p = \frac{P_e}{P_o}$$

conocido como riesgo relativo de la enfermedad y es una medida de la magnitud para determinar si un individuo que está expuesto a un factor particular tiene mayor o menor probabilidad de desarrollar la enfermedad respecto a alguno que no está expuesto. En particular una persona expuesta es p veces más probable de contraer la enfermedad en un periodo de tiempo dado, que una persona no expuesta. Cuando el riesgo de la enfermedad es similar para ambos grupos, el riesgo relativo será uno. Los valores de p menores que uno indican que una persona no expuesta tiene más riesgo que una persona expuesta, mientras que un valor de p más grande que uno indica que una persona expuesta tiene un riesgo mayor.

2.5.1.2 Riesgo atribuible

El riesgo atribuible, llamado a veces fracción etiológica, fracción atribuible, o porcentaje de riesgo atribuible a la población, genera un estimador de la proporción de casos que están relacionados con un factor de exposición dado. Este riesgo es usualmente interpretado como la fracción de enfermedad en una población que puede ser eludida por reducción o eliminación de exposición a un agente etiológico, sabiendo que es causa de dicha enfermedad.

El riesgo atribuible es definido como:

$$RA = [P(D) - P(D|E)] / P(D)$$

donde $P(D)$ es la probabilidad de enfermedad en la población y $P(D|E)$ es la probabilidad de enfermedad hipotética en la misma población pero sin el factor de exposición. Esta estimación asume que el factor de exposición es causal. Cuando no se hace esta suposición de ajuste para las variables de confusión, la fórmula anterior puede ser escrita como:

$$RA = P(E) (RR - 1) / 1 + P(E) (RR - 1)$$

donde $P(E)$ es la proporción de la población expuesta a el factor y RR es la razón de riesgo.

En la estimación del riesgo atribuible para datos de casos y controles, el RR puede ser reemplazado por la razón de momios y $P(E)$ puede ser reemplazado por $P(ED)$, donde $P(ED)$ es la proporción de casos expuestos a el factor.

2.5.1.3 Razón de momios

Una medida útil de asociación entre la ocurrencia de una enfermedad y un factor de riesgo es el razón de momios. Suponiendo que el factor de riesgo es dicotómico con dos niveles

(expuesto y no expuesto), donde p_e y p_u son los riesgos para el grupo expuesto y no expuesto respectivamente. El momio de una persona con la enfermedad en el grupo expuesto es $p_e / (1-p_e)$, y similarmente, el momio de enfermedad para alguna persona del grupo no expuesto es $p_u / (1-p_u)$. La razón de momios, es la razón de enfermedad para cualquier persona expuesta dividida entre el de una persona no expuesta y esta dado por:

$$\Psi = \frac{p_e / (1-p_e)}{p_u / (1-p_u)},$$

cuando una enfermedad es rara, p_e y p_u son pequeñas y la razón de momios Ψ , es aproximadamente igual al riesgo relativo, p . Para mostrar esto, considere los momios de enfermedad para un individuo en el grupo expuesto, $p_e / (1-p_e)$. Esto puede ser escrito como:

$$p_e (1-p_e)^{-1} = p_e (1 + p_e + p_e^2 + \dots),$$

si p_e es pequeño, p_e^2 y el orden de la potencia mas alta de p_e serán insignificantes, por lo tanto $p_e / (1-p_e)$ será aproximadamente igual a p_e . De la misma manera, $p_u / (1-p_u)$ es aproximadamente igual a p_u , cuando p_u es pequeño. Consecuentemente, el riesgo relativo, p , puede ser aproximado por la razón de momios, Ψ . Esta aproximación es bastante aceptable en la mayoría de los estudios epidemiológicos de asociación entre la enfermedad y los factores de riesgo. Dado que la razón de momios es aproximadamente igual a el riesgo relativo cuando la enfermedad bajo estudio es rara, la razón de momios se usa como una medida de asociación entre la enfermedad y los factores de riesgo especialmente en los estudios cohorte y de casos y controles.

La razón de momios se puede calcular fácilmente cuando el factor de riesgo es dicotómico, supongamos que se tienen datos de un estudio donde los individuos han sido observados por un período de tiempo, y donde el número de individuos en las cuatro combinaciones de estado de la enfermedad y nivel de exposición son a, b, c y d (Tabla 2).

	Enfermo	No enfermo
Expuesto	a	b
No expuesto	c	d

Tabla 2

La probabilidad de éxito estimada para los grupos expuesto y no expuesto esta dada por las siguientes expresiones:

$$\hat{p}_e = \frac{a}{(a+b)} \quad \text{y} \quad \hat{p}_u = \frac{c}{(c+d)},$$

entonces el riesgo relativo esta dado por:

$$\hat{p} = \frac{\hat{p}_c}{\hat{p}_n} = \frac{a(c+d)}{c(a+b)},$$

ya la razón de momios estimada $\hat{\Psi}$ es:

$$\hat{\Psi} = \frac{\hat{p}_c / (1 - \hat{p}_c)}{\hat{p}_n / (1 - \hat{p}_n)} = \frac{ad}{bc}.$$

El estimador es la razón de los productos de 2 pares de elementos diagonales en la tabla de 2×2 , algunos autores la definen como razón de productos cruzados.

• **Inferencia acerca de la razón de momios**

El logaritmo de la razón de momios estimada se aproxima mejor a una distribución normal que la razón de momios misma. Especialmente cuando el número total de observaciones no es muy grande.

El error estándar aproximado del logaritmo de la razón de momios estimada, puede escribirse como:

$$e.s.(\log \hat{\Psi}) = \sqrt{\left(\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d} \right)}.$$

Un intervalo de confianza aproximado $100(1 - \alpha)\%$ para el logaritmo de la razón de momios estimada ($\log \hat{\Psi}$) es:

$$\left[\log \hat{\Psi} - Z_{\alpha/2} e.s.(\log \hat{\Psi}) \right], \left[\log \hat{\Psi} + Z_{\alpha/2} e.s.(\log \hat{\Psi}) \right],$$

donde $z_{\alpha/2}$ es el punto superior de la distribución normal estándar ($100\alpha/2\%$).

El utilizar logaritmos en la construcción de intervalos de confianza para Ψ asegura que ambos límites siempre serán no negativos, siendo un requisito natural ya que por definición, la razón de momios no puede ser negativa.

Si se quiere obtener el error estándar aproximado de la razón de momios estimada $\hat{\Psi}$, puede calcularse usando el resultado de que $e.s.(\hat{\Psi}) \approx \hat{\Psi} e.s.(\log \hat{\Psi})$, por lo tanto:

$$e. s. (\hat{\Psi}) = \hat{\Psi} \sqrt{\left(\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}\right)} \quad \text{ec. (1)}$$

sin embargo, puesto que la razón de momios no puede ser negativa, la ecuación 1 no se debería usar como base para el intervalo de confianza de Ψ , debido a que no se obtendría una buena aproximación a una distribución normal.

La hipótesis nula para el que la razón de momios es igual a 1 es: $H_0: \Psi = 1$, esta hipótesis puede probarse usando los siguientes estadísticos:

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left[\frac{1}{n_1} + \frac{1}{n_2}\right]}} \quad \text{ec. (2)}$$

donde $\hat{p}_1 = y_1/n_1$ y $\hat{p}_2 = y_2/n_2$ son las probabilidades de éxito estimadas para los grupos expuesto y no expuesto respectivamente, (y_1 e y_2 son el número de éxitos para cada muestra) y n_1, n_2 son los tamaños para el grupo expuesto y no expuesto.

Hay un número de alternativas equivalentes para la prueba estadística de la ecuación 2 que pueden ser usadas cuando se tiene una hipótesis alternativa de dos lados. Estas alternativas se derivan de un resultado general el cuál establece que si la variable aleatoria X tiene una distribución normal estándar, entonces el cuadrado de ésta variable aleatoria Z^2 , tiene una distribución ji-cuadrada con un grado de libertad. La fórmula general cuando se quiere comparar más de dos proporciones se escribe como:

$$z^2 = \sum_{i=1}^2 \frac{n_i (\hat{p}_i - \hat{p})^2}{\hat{p}(1-\hat{p})} \quad \text{ec. (3)}$$

$$\text{donde } \hat{p} = (y_1 + y_2) / (n_1 + n_2) = (n_1 \hat{p}_1 + n_2 \hat{p}_2) / (n_1 + n_2)$$

Cuando se tiene una tabla de contingencia⁹ de 2x2, se pueden denotar los totales observados por O_1, O_2, O_3 y O_4 , y los correspondientes valores esperados como E_1, E_2, E_3 y E_4 , entonces el estadístico de la ecuación 3 se puede escribir como:

$$\sum_{i=1}^4 \frac{(O_i - E_i)^2}{E_i}$$

este es el estadístico usado en pruebas de asociación entre dos factores para una tabla de contingencia de 2x2, el estadístico se denota generalmente por χ^2 y, tiene una distribución

⁹ Una tabla de contingencia es un ordenamiento de datos categóricos mediante renglones y columnas, donde intervienen dos variables cualitativas. Cada casilla en la tabla representa una combinación de categorías de las dos variables.

ji-cuadrada con un grado de libertad. Este estadístico fue primero propuesto por Karl Pearson y se conoce generalmente como el estadístico χ^2 de Pearson.

2.6 Consideraciones para el análisis

El objetivo del análisis es determinar el efecto que el factor de riesgo tiene en la respuesta, es también importante estudiar las posibles interacciones que el factor de riesgo o tratamiento y la respuesta pueda tener con otros factores, para poder describir claramente como es que se comporta dicho efecto ante la presencia de esos otros factores.

En primer lugar es importante describir el perfil tanto del grupo de casos como el de los controles, en términos de la distribución y de las interrelaciones que en cada uno de dichos grupos tienen las diferentes variables incluidas en el estudio, así como ciertas características generales que pueden influir en los resultados tales como: edad, raza, sexo, método de diagnóstico, etc.. El conocer estas características permite saber si los procedimientos de apareamiento se realizaron adecuadamente y explorar fuentes posibles de error.

En un análisis inicial se debe explorar el efecto que tiene cada uno de los factores de riesgo o tratamiento de interés en el factor respuesta de manera separada. Esto puede realizarse de varias maneras dependiendo de la escala de medición del factor de riesgo.

- ⇒ Cuando el factor de riesgo es dicotómico, se construye una tabla de 2×2 y a partir de ella se pueden hacer inferencias acerca de la razón de momios.
- ⇒ Cuando el factor de riesgo es politómico, se elige a un nivel como nivel base de comparación contra el cual se comparan los demás. Se realizan inferencias para la razón de momios que relaciona a cada par de niveles y se comparan éstas.

En análisis posteriores se investiga la acción combinada de varios factores de riesgo o tratamiento:

- ⇒ Se puede examinar el efecto de varios factores de riesgo o tratamiento de manera simultánea.

Cuando se trabaja con muchos factores de riesgo y/o confusión, es mejor utilizar modelos de regresión para investigar sus efectos en el factor respuesta.

2.6.1 Confusión e Interacción

En un estudio etiológico diseñado a investigar si una enfermedad está asociada con factores de riesgo, la información será obtenida usualmente sobre un número de factores que se espera estén asociados con la ocurrencia de la enfermedad. Aunque algunos de éstos serán de interés como causas posibles de la enfermedad, otros simplemente podrían disfrazar o exagerar el verdadero efecto de un factor de riesgo de interés. Esto significa que una asociación aparente entre el factor de riesgo y la ocurrencia de la enfermedad es debido a otro factor. Alternativamente, una falta de asociación entre un factor de riesgo y la enfermedad puede ser debido a que el factor exacto (causa) no ha sido tomado en cuenta dentro del análisis de datos.

Confusión

Una variable que explica parcialmente la asociación aparente entre una enfermedad y un factor de exposición es llamada variable confusa. Específicamente una variable de confusión es una variable que es un factor de riesgo para la enfermedad bajo estudio y esta asociada con ésta, pero no es una consecuencia del factor de riesgo. Una variable que inicia como factor de exposición y se vuelve causa de la enfermedad no será tomada en cuenta como una variable de confusión.

Interacción

Cuando el grado de asociación entre una enfermedad y un factor de riesgo es diferente para cada nivel en una variable de confusión, se dice que hay interacción entre el factor de riesgo y la variable de confusión. En otras palabras la variable de confusión modifica el efecto del factor de riesgo en la enfermedad. Una variable de confusión que interactúa con un factor de riesgo es llamado "efecto modificador".

Cuando la razón de momios para individuos expuestos relativo a los no expuestos en diferentes niveles de las variables de confusión son comparables, pueden ser combinados para obtener una razón de momios global. Este valor total es después ajustado para la variable de confusión. Sin embargo cuando la razón de momios difiere entre los niveles de una variable de confusión, esto es; cuando hay interacción, no es apropiado hacer combinaciones, y la razón de momios para cada nivel de la variable de confusión debe ser reportada de manera independiente.

DIAGNÓSTICO DE LA INFORMACIÓN

3. DIAGNÓSTICO DE LA INFORMACIÓN

Antes de definir todos los aspectos considerados para la obtención de los datos y el diagnóstico de la información se hará una descripción breve sobre las características generales del problema en estudio.

La palabra leucemia (sangre blanca) implica una condición de la sangre caracterizada por un gran aumento del número de leucocitos. La alteración fundamental que caracteriza a la leucemia no radica realmente en la sangre, sino más bien en los tejidos, tanto en los que producen las células sanguíneas, que están proliferando de forma indebidamente rápida o, cuando menos anómala, como en aquellos otros en que las células proliferantes exfoliadas pueden depositarse y a menudo acumularse. La palabra "anómala" significa la aparición de una nueva y distinta raza de leucocitos¹¹.

En resumen la *leucemia* puede definirse como una proliferación (lenta o rápida), generalizada, anormal y neoplásica de uno de los tejidos leucopoyéticos, asociada a menudo con cifras anómalas de leucocitos en sangre y que eventualmente conduce a la anemia, hemorragias y muerte del sujeto.

Tipos de leucemia fundamentales:

1. Granulocítica
2. Linfóide
3. Monocítica
4. Plasmocitaria

La edad crítica para la aparición de la leucemia se sitúa entre los 3-4 años, otro momento crítico se da en la adolescencia. Aunque no son bien conocidas las causas de estas variaciones en la edad de inicio de estos procesos, todo parece indicar la existencia de factores de susceptibilidad genética en la niñez y de cambios hormonales en la adolescencia.

Aunque no se sabe exactamente el mecanismo por el cuál se produce, ni los factores que la determinan se mencionarán algunos que según la literatura médica están asociados con el desarrollo y que servirán de base para comparación con los resultados que se obtengan.

• Factores Ambientales:

Radiación.

Durante los últimos años se han recogido pruebas suficientes para afirmar que las radiaciones ionizantes son capaces de causar leucemia en los seres humanos, se sabe que la radiación en dosis excesivamente elevadas es una aseveración fuerte de leucemia mielóide crónica y de leucemia aguda. Algunos otros estudios señalan que las radiaciones diagnósticas preconcepcionales incrementan el riesgo de leucemia en los niños engendrados

¹¹ Este concepto y todos los demás de esta tabla que aparezcan con un asterisco, aunque no son importantes para el estudio pueden consultarse en el apéndice

después de la exposición. Algunos estudios retrospectivos y prospectivos indican que la irradiación diagnóstica durante el embarazo puede incluso doblar el riesgo de leucemia infantil.

Químicos.

El estudio de casos bien documentados indican que la exposición a ciertos químicos como el benzol y cloranfenicol poseen la capacidad de deprimir la médula ósea* y dañar los cromosomas*, por lo tanto no se descarta la posibilidad de que otros agentes químicos estén presentes en los procesos leucemogénicos.

Viricos.

A pesar del descubrimiento de que muchas leucemias de los animales de laboratorio son provocadas por virus, no se ha logrado todavía una clara confirmación epidemiológica que avale la hipótesis de que la leucemia humana es transmitida por agentes infecciosos. Sin embargo estudios retrospectivos efectuados en grupos de niños con leucemia (Stewart 1958), han sugerido la mayor frecuencia de ciertas condiciones entre las personas afectadas de leucemia que entre el resto de la población; por ejemplo: infecciones pulmonares graves, manifestaciones alérgicas, infecciones agudas del tracto respiratorio etc.. La naturaleza de asociación entre tales condiciones y leucemia es hasta cierto punto especulativa, por lo que son necesarios muchos más estudios para confirmar estos hallazgos.

• El huésped

Enfermedades hereditarias.

El gran número de casos de leucemia observados en las personas con síndrome de Down, en el que el defecto cromosómico característico es el resultado de una falta de división meiótica anterior a la concepción, sugiere que un cromosoma extra puede estar implicado en la leucemogénesis. Menos conocida es la relación de la leucemia con otros defectos cromosómicos congénitos; posiblemente sea debido a que estas anomalías son raras, muy letales y/o poco conocidas. El riesgo de leucemia en el mongolismo* es muy grande en todas las edades, y en especial en el período neonatal.

Mientras que los defectos congénitos citogenéticos predisponen a diversos tipos de leucemia, la revisión de algunas investigaciones sugieren que el linfoma se desarrolla principalmente en los síndromes genéticos asociados con deficiencia inmunitaria, tales como el síndrome de Kott-Aldrich, sin embargo queda por determinar si otros trastornos congénitos no asociados con defectos cromosómicos o inmunológicos están asociados con el aumento de riesgo de leucemia.

• Genéticos

Los factores genéticos son más evidentes en los casos de leucemia familiar, algunos de los cuales implican la influencia de los genes recesivos.

3.1 Definición de variables

Al obtener información acerca de los tipos de personas afectadas por una enfermedad, y de las circunstancias en que esa enfermedad se presenta, comúnmente se aprovechan aquellas variables descriptivas respecto a las cuales existen grandes cantidades de datos en relación con la población. entre las mas generales tenemos las siguientes:

EDAD.

La edad es una variable que debe ser siempre tomada en cuenta en los estudios epidemiológicos, el conocimiento de las asociaciones con la edad es importante porque el estudio de esta variación puede ayudar a comprender los factores responsables del desarrollo de la enfermedad. Conceptualmente se define como el tiempo transcurrido en años, entre la fecha de nacimiento de la persona y la realización de la entrevista.

SEXO.

Como para otras variables descriptivas, la asociación entre el sexo y una enfermedad se pone en evidencia, por la diferencia entre las tasas de la enfermedad, es decir, por medio de la comparación de casos encontrados en los hombres con respecto a las mujeres. El sexo se define como la condición orgánica que distingue a las personas en hombres o mujeres.

OCUPACIÓN.

La ocupación se usa como variable descriptiva con un número de propósitos bastante diferentes. Se pueden mencionar los siguientes:

1. Medición del estrato socioeconómico. La ocupación es uno de los índices básicos que sirven para determinar la posición social y económica del individuo.
2. Identificación de riesgos específicos asociados con la exposición a ciertos agentes nocivos peculiares de algunas ocupaciones.
3. Indicación de las condiciones generales bajo las cuáles trabaja un grupo ocupacional, por ejemplo, la cantidad de ejercicio físico, tensión mental o alteración de la rutina.

Conceptualmente la ocupación se define como la actividad laboral que realiza el sujeto y a cambio recibe o no una remuneración económica.

NIVEL SOCIOECONÓMICO.

El nivel socioeconómico es un fenómeno determinado por un amplio espectro de variables que son conceptualizadas como una combinación de características económicas, ocupacionales y educativas.

ESCOLARIDAD.

Se define como el número de años cursados en forma oficial, dentro de cualquier institución educativa. Se ha comprobado que cuando el individuo tiene un alto nivel educativo, existe una mejor respuesta ante los programas de salud comunitaria diseñados para incrementar los conocimientos sobre los factores de riesgo.

Como el objetivo del estudio es determinar posibles factores de riesgo asociados al desarrollo de leucemia en niños, se tomaron en cuenta las variables que proporcionan información sobre exposición a dichos factores así como aquellas que definen los

antecedentes heredo familiares, antecedentes personales no patológicos y patológicos de los padres, antecedentes ginecoobstétricos de la madre y algunos otros.

Puesto que el número de variables es grande, la definición particular de cada una de ellas consideradas en el análisis se definen en la tabla A del apéndice.

El estudio se realizó tomando en consideración como "casos" aquellos pacientes que fueron diagnosticados con leucemia en la Unidad de Epidemiología del hospital de Pediatría del Centro Médico Nacional o del Hospital La Raza. Para ser incluido como control, fue necesario no estar enfermo y tener una edad similar al caso.

La información de ambas muestras está registrada en la base de datos IMSS.WKS que se trabajará mediante el paquete estadístico SPSS-X versión 5.0.

3.2 Tablas de Frecuencias simples

Considerando que la estadística descriptiva permite obtener una visión rápida respecto a un conjunto de datos, y puesto que en la mayoría de las investigaciones epidemiológicas los datos históricos o antecedentes del paciente se consideran importantes, generalmente se hace uso de esta herramienta para explorar los datos y poder describir cuál es el comportamiento típico de la población.

En este trabajo como análisis preliminar de los datos se obtuvieron las frecuencias simples de todas las variables de interés, y se agruparon en tablas de acuerdo las categorías dadas en la definición de variables.

Tabla 1. Distribución de frecuencia por sexo

Valor	Freq	Pct	Pct Cum
1	251	64.0	64.0
2	211	53.0	117.0
Total	374	100.0	100.0

Casos válidos: 374

Val = valor o categoría de la variable (definido en la tabla A del apéndice)

Freq = frecuencia

Pct = porcentaje (donde porcentaje = (frecuencia / total) * 100)

Pct Cum = Suma acumulada de porcentaje

Así como se presenta esta tabla, se obtuvieron las demás para cada una de las variables en estudio, sin embargo se seleccionaron principalmente aquellas que nos proporcionan información sobre los factores de riesgo y los antecedentes del menor, no obstante en la parte B del apéndice se encuentran de manera ordenada y detallada cada una de las tablas de frecuencias utilizadas para obtener las características generales de la población y que se resumen a continuación.

De la tabla 2 del apéndice se puede apreciar que la edad varía mucho, pero se tienen frecuencias bajas para cada valor, siendo la moda 26 meses, es decir hubo una frecuencia mayor para los niños con edad de 2.2 años aproximadamente, sin embargo la media se ubica entre los 6 y 7 años.

Un aspecto importante que se observó es el domicilio del paciente, pues de puede ver de la tabla 4 que hay una frecuencia mayor para algunos municipios específicos respecto a los demás y será interesante analizar el comportamiento de éstos pacientes respecto a la enfermedad.

En cuanto a los antecedentes personales no patológicos del paciente se encontró que si hay exposición a los diferentes factores de riesgo, por lo que dichas variables serán consideradas en el análisis bivariado pero solo en términos de exposición.

De los datos sobre la ocupación de los padres, se encontró que hay frecuencias bajas a los diferentes factores de riesgo, sin embargo se tomarán en cuenta para el análisis bivariado pues nos interesa saber si hay asociación con el desarrollo de leucemia en el niño.

Respecto a los antecedentes del niño, se puede apreciar en las tablas correspondientes que la frecuencia de exposición a los diferentes factores de riesgo no es alta, pero el interés es precisamente ver que grado de asociación guarda dicha exposición con la enfermedad, por lo que serán consideradas en la segunda parte del análisis.

Así pues, antes de realizar el análisis bivariado es bueno comentar que se decidió utilizar tablas de frecuencias simples para el análisis preliminar de los datos debido a la cantidad de información y de las características de medición en las variables, pues el utilizar alguna otra herramienta, haría un poco más difícil el agrupamiento de los datos ya que hay muchas variables que no están acotadas, como es el caso de las distancias respecto a los factores de riesgo tales como cables de alta tensión, transformadores de luz, subestaciones eléctricas, etc., donde se observa que hay frecuencias bajas y diferencias grandes en cuanto a las medidas proporcionadas por las personas entrevistadas, variables como éstas no serán consideradas en el análisis bivariado ya que tienen poca precisión y no serían de mucha utilidad.

3.3 ANÁLISIS BIVARIADO

Lo que interesa ahora es conocer si existe asociación entre los factores de exposición y el desarrollo de la leucemia, en base a los datos obtenidos de la encuesta realizada para los niños en estudio, así como poder establecer si existe relación de la enfermedad con algunas variables específicas tales como el sexo, la edad, domicilio, etc..

En general en el análisis bivariado se utilizará la prueba chi-cuadrada, para probar si hay asociación entre las variables analizadas, además se tomarán en cuenta también las medidas de asociación (riesgo relativo y razón de momios) para describir mejor el comportamiento de las variables en cada caso.

• **Análisis para probar si el desarrollo de leucemia está asociado con el sexo.**

De la sección 2.2 se puede observar en la tabla 1 que del total de niños en estudio, hay una proporción mayor para el sexo masculino tanto en la muestra de casos como la de controles, lo que interesa es ver si el desarrollo de leucemia es independiente del sexo, por lo que la hipótesis que se desea probar es la siguiente:

H_0 = El desarrollo de leucemia es independiente del sexo.

H_1 = El desarrollo de leucemia no es independiente del sexo.

para probar esta hipótesis se agruparon los datos en una tabla de 2x2 donde masculino y femenino representan las categorías de la variable sexo y las columnas caso y control corresponden a las muestra de pacientes con y sin leucemia respectivamente.

Tabla A1 Datos del estudio sobre asociación entre leucemia y sexo

SEXO by PACIENTE	caso	control	total renglón
Masculino	105	108	213
Femenino	82	79	161
total columna	187	187	374

Tabla A2 Cálculo de la estadística chi cuadrada para los datos

Sexo	Grupo	Valor observado (o)	Valor esperado (e)	Valor observado menos esperado (o - e)	Contribución a χ^2 [(o-e) ² /e]
Masculino	Casos	105	106,5	-1,5	.0211
	Controles	108	106,5	1,5	.0211
Femenino	Casos	82	80,5	1,5	.0279
	Controles	79	80,5	-1,5	.0279
Total		374			$\chi^2 = .098$

Tabla A3 Resultado de la prueba chi-cuadrada

Chi - cuadrada	valor	Grados de Libertad	Significancia
Pearson	.09815	1	.75406

De los resultados obtenidos en la tabla A3 se puede ver que el valor del estadístico obtenido es .09815 siendo no significativo a un nivel del 95 %, es decir; si se busca en la tabla de la distribución chi-cuadrada con 1 gl, se puede observar que para $\alpha=.05$ el valor

correspondiente es 3.84, como el resultado obtenido es menor que 3.84 se concluye que el desarrollo de leucemia es independiente del sexo, por lo que se acepta la hipótesis nula.

Si consideramos otras medidas de asociación entre la enfermedad y el factor de riesgo como la razón de momios y el riesgo relativo se tiene lo siguiente:

Estadístico	Valor	Intervalo de confianza a un 95%	
Razón de momios	.93665	.62197	1.41055
Riesgo relativo para el grupo de casos	.96774	.78784	1.18872
Riesgo relativo para el grupo de controles	1.03319	.84269	1.26675

en este caso la razón de momios nos indica que el riesgo de la enfermedad para una persona expuesta respecto a una no expuesta es menor que uno, cuando el valor es menor que uno no se considera un riesgo asociado con la enfermedad.

• **Análisis de asociación de leucemia respecto a la exposición del niño con transformadores de luz.**

Como se menciona al inicio del capítulo 3, la exposición del paciente a campos electromagnéticos puede ser un factor que este asociado con el desarrollo de leucemia, por lo que se considera importante analizar la información obtenida sobre algunos factores de exposición, en este caso se considera que la exposición del niño respecto a transformadores de luz representa un factor de riesgo que puede influir en el desarrollo de leucemia. Para esta variable se tomaron como expuestas a las personas que afirmaron tener instalados transformadores de luz cerca de su hogar.

H₀ = La exposición del menor a transformadores de luz no es factor de riesgo para el desarrollo de leucemia.
 H₁ = La exposición del menor a transformadores de luz es factor de riesgo para el desarrollo de leucemia.

Tabla B1 Datos del paja examinar la asociación entre la ocurrencia de leucemia y la exposición a transformadores de luz.

TRANSFORMADORES DE LUZ	casos	controles	total renglón
Expuestos	139	136	275
No expuestos	48	51	99
total columna	187	187	374

Tabla B2 Cálculo de la estadística chi cuadrada para los datos

Factor de riesgo	Grupo	Valor observado (o)	Valor esperado (e)	Valor observado menos esperado (o - e)	Contribución a χ^2 $((o - e)^2 / e)$
Expuestos	Casos	139	137.5	1.5	.0163
	Controles	136	137.5	-1.5	.0163
No expuestos	Casos	48	49.5	-1.5	.0454
	Controles	51	49.5	1.5	.0454
Total		374			$\chi^2 = .123$

Tabla B3 Resultado de la prueba chi cuadrada

Chi - cuadrada	valor	Grados de Libertad	Significancia
Pearson	.12364	1	.72512

Según el valor del estadístico obtenido 0.12364, se puede decir que no es significativo a un nivel del 95 %, por lo tanto la prueba de bondad de ajuste realizada nos conduce a aceptar la hipótesis nula, es decir el desarrollo de leucemia no esta asociado con la exposición del paciente a transformadores de luz.

Si tomamos en cuenta las medidas de asociación como el riesgo relativo y la razón de momios tenemos lo siguiente:

Estadístico	Valor	Intervalo de confianza a un 95%	
Razón de momios	1.08594	.68581	1.71952
Riesgo relativo para el grupo de casos	1.04167	.83147	1.30500
Riesgo relativo para el grupo de controles	.95923	.75887	1.21249

de estos datos podemos comentar que aunque el riesgo relativo de ocurrencia de la enfermedad para el grupo de los casos es ligeramente mayor respecto al de los controles, sin embargo ambos son muy cercanos a uno, así mismo la razón de momios muestra no asociación entre la enfermedad y el factor de riesgo.

• **Prueba de asociación entre el desarrollo de leucemia y exposición del niño a cables de alta tensión.**

Suponiendo que no hay asociación entre la leucemia y la exposición a cables de alta tensión se tiene lo siguiente:

- H_0 = La exposición a cables de alta tensión no es factor de riesgo para el desarrollo de leucemia.
- H_1 = La exposición a cables de alta tensión es factor de riesgo para el desarrollo de leucemia.

Tabla C1. Datos del estudio para examinar la asociación entre la ocurrencia de leucemia y la exposición del menor a cables de alta tensión.

CABLES DE ALTA TENSION	casos	controles	total renglón
Expuestos	179	155	334
No expuestos	8	32	40
total columna	187	187	374

Tabla C2. Cálculo de la estadística chi cuadrada para los datos

Factor de riesgo Cables de alta tensión	Grupo	Valor observado (o)	Valor esperado (e)	Valor observado menos esperado (o - e)	Contribución a χ^2 [(o - e) ² / e]
Expuestos	Casos	179	167.0	12.0	.8622
	Controles	155	167.0	-12.0	.8622
No expuestos	Casos	8	20.0	-12.0	7.2
	Controles	32	20.0	12.0	7.2
Total		374			$\chi^2 = 16.124$

Tabla C3. Resultado de la prueba chi-cuadrada

Chi - cuadrada	valor	Grados de Libertad	Significancia
Pearson	16.124	1	.00006

El valor del estadístico obtenido es muy grande y el nivel de significancia .00006 es bastante considerable, es decir (p<.001), por lo tanto se rechaza la hipótesis nula y se

concluye que la exposición a cables de alta tensión es un factor de riesgo para el desarrollo de leucemia. Los siguientes estadísticos indican el grado de asociación entre las variables.

Estadístico	Valor	Ase	Significancia
Cramer's V	.20760		.00006
Coefficiente de Contingencia	.20330		.00006
Lambda variable dependiente PACPREL	.12834	.03158	
Coefficiente de Incertidumbre variable dependiente PACPREL	.03307	.01509	.00003
Somers' D variable dependiente PACPREL	-.33593	.06888	

la V de Cramers y el coeficiente de contingencia indican asociación no muy fuerte, la significancia se considera relevante por el hecho de que la frecuencia de expuestos es muy representativa respecto a los no expuestos. Si consideramos el error estándar asimétrico para el estadístico lambda y el coeficiente de incertidumbre vemos son ambos pequeños.

Estadístico	Valor	Intervalo de confianza a un 95%
Razón de momios	4.619	2.066 10.3194
Riesgo relativo para el grupo de casos	2.6796	1.199 5.9864
Riesgo relativo para el grupo de controles	0.5800	0.2595 1.294

Después de obtener la razón de momios vemos que el riesgo es de 4.61, lo que resulta ser un valor bastante significativo para los expuestos, así mismo se ve que el riesgo relativo de enfermedad es mayor para el grupo de casos.

• Prueba de asociación entre leucemia y exposición a subestaciones eléctricas.

En este caso se supondrá que no hay asociación entre el desarrollo de leucemia y la exposición del niño en estudio con subestaciones eléctricas, cerca de su hogar.

H_0 = La exposición a subestaciones eléctricas no es un factor de riesgo para el desarrollo de leucemia.

H_1 = Le exposición a subestaciones eléctricas es factor de riesgo para el desarrollo de leucemia.

Tabla D1. Datos del estudio para examinar la asociación entre la ocurrencia de leucemia y la exposición a subestaciones eléctricas

SUBEST. by PACPREL	casos	controles	total renglon
Expuestos	25	25	50
No expuestos	162	162	324
total columna	187	187	374

Tabla D2. Cálculo de la estadística chi cuadrada para los datos

Factor de riesgo SUBESTACIONES ELÉCTRICAS	Grupo	Valor observado (e)	Valor esperado (e)	Valor observado menos esperado (o - e)	Contribución a χ^2 [(o - e) ² / e]
Expuestos	Casos	25	25.0	0	0
	Controles	25	25.0	0	0
No expuestos	Casos	162	162.0	0	0
	Controles	162	162.0	0	0
Total		374			$\chi^2 = 0$

Tabla D3 Resultado de la prueba chi-cuadrada

Chi - cuadrada	valor	Grados de Libertad	Significancia
Pearson	.00000	1	1.00000

Se puede observar a simple vista que no existe relación alguna entre el factor de exposición y la enfermedad, puesto que el valor del estadístico es cero, en otras palabras el desarrollo de leucemia es independiente de la exposición que el individuo pueda tener con subestaciones eléctricas.

Estadístico	Valor	Intervalo de confianza a un 95%	
Razón de momios	1.00000	.55123	1.81412
Riesgo relativo para el grupo de casos	1.00000	.74245	1.34689
Riesgo relativo para el grupo de controles	1.00000	.74245	1.34689

Considerando el riesgo relativo y la razón de momios para el grupo de casos así como de los controles, se ve que el riesgo de desarrollar leucemia es igual puesto que ambos tienen un valor de 1

• **Prueba de asociación entre el desarrollo de leucemia y exposición a torres de alta tensión.**

Si suponemos que no hay asociación entre desarrollar leucemia y estar expuesto a cables de alta tensión, tenemos lo siguiente:

- H_0 = El desarrollo de leucemia es independiente de la exposición que el niño tenga con torres de alta tensión.
- H_1 = El desarrollo de leucemia no es independiente de la exposición que el niño tenga con torres de alta tensión.

Tabla E1. Datos del estudio para examinar la asociación entre la ocurrencia de leucemia y la exposición del menor a torres de alta tensión

TORRES DE ALTA TENSION	CASOS	CONTROLES	TOTAL
Expuestos	52	51	103
No expuestos	135	136	271
TOTAL COLUMNA	187	187	374

Tabla E2. Cálculo de la estadística chi cuadrada para los datos

Factor de riesgo	Grupo	Valor observado (o)	Valor esperado (e)	Valor observado menos esperado (o - e)	Contribución a χ^2 [(o - e) ² / e]
Expuestos	Casos	52	51.5	0.5	.0048
	Controles	51	51.5	-0.5	.0048
No expuestos	Casos	135	135.5	-0.5	.0018
	Controles	136	135.5	0.5	.0018
Total		374			$\chi^2 = .013$

Tabla E3 Resultado de la prueba chi-cuadrada

Chi - cuadrada	valor	Grados de Libertad	Significancia
Pearson	.01340	1	.90785

El valor del estadístico obtenido es de .01340 siendo no significativo a un nivel del 95 %, por lo tanto se acepta la hipótesis nula , y se puede decir que la exposición a torres de alta tensión no es un factor de riesgo para el desarrollo de leucemia. Se observa además que las frecuencias observadas son casi igual a las esperadas para cada una de las celdas.

Al medir la asociación en términos del riesgo relativo y la razón de momios tenemos:

Estadístico	Valor	Intervalo de confianza a un 95%	
Razón de momios	1.02716	.65249	1.61698
Riesgo relativo para el grupo de casos	1.01353	.80669	1.27340
Riesgo relativo para el grupo de controles	.98673	.78751	1.23635

El valor de la razón de momios indica que no existe riesgo significativo de las personas expuestas respecto a las no expuestas.

El estudio también considera la exposición de los factores de riesgo para las personas que no han vivido siempre en el mismo lugar. Al realizar esta prueba se tomaron en cuenta dos casos, en el primero se realizó la prueba de asociación para:

- Exposición a transformadores de luz.
- exposición a cables de alta tensión,
- exposición a subestaciones eléctricas y
- exposición a torres de alta tensión,

donde se agregaron las personas que afirmaron haber vivido antes en otro lugar y pudieron o no estar expuestas, mientras que en el segundo caso también se realizó la prueba de la chi-cuadrada pero sólo se consideraron aquellas personas que tuvieron exposición siempre (domicilio actual mas domicilio anterior), así como aquellas que en ninguno de los lugares estuvieron expuestas para corroborar si de alguna manera influye el tiempo de exposición.

• **Prueba de asociación entre el desarrollo de leucemia y exposición a transformadores de luz considerando el primer caso.**

Si se supone que no existe asociación entre el desarrollo de leucemia y el factor de riesgo se tiene lo siguiente:

H_0 = El estar expuesto a transformadores de luz no es un riesgo para el desarrollo de leucemia.

H_1 = El estar expuesto a transformadores de luz es un riesgo para el desarrollo de leucemia.

Tabla F1. Datos del estudio para examinar la asociación entre la ocurrencia de leucemia y la exposición del menor a transformadores de luz

TRANSFORMADORES DE LUZ	casos	controles	total renglón
Expuestos	187	182	369
No expuestos	88	89	177
total columna	275	271	546

Tabla F2. Cálculo de la estadística chi cuadrada para los datos.

Factor de riesgo transformadores de luz	Grupo	Valor observado (o)	Valor esperado (e)	Valor observado menos esperado (o - e)	Contribución a χ^2 [(o - e) ² / e]
Expuestos	Casos	187	185.8	1.2	.0077
	Controles	182	183.1	-1.1	.0066
No expuestos	Casos	88	89.1	-1.1	.0135
	Controles	89	87.8	1.2	.0164
Total		546			$\chi^2 = .0443$

Tabla F3. Resultados de la prueba chi-cuadrada

Chi - cuadrada	valor	Grados de Libertad	Significancia
Pearson	0.443	1	.8065

El valor del estadístico nos indica que no es significativo, por lo que se acepta la hipótesis nula, y se puede comentar que aunque el paciente ha vivido en otro lugar, de igual manera la exposición a transformadores no es factor de riesgo para la enfermedad. Si se hace inferencia respecto a la razón de momios, se ve que el riesgo de que una persona desarrolle leucemia dado que ha estado expuesta en algún momento a transformadores de luz no es relevante respecto a una persona que no ha estado expuesta.

	Estadístico	Valor	Intervalo de confianza a un 95%
Razón de momios	1.039	0.7260	1.4867
Riesgo relativo para el grupo de casos	1.019	0.7120	1.4581
Riesgo relativo para el grupo de controles	.9810	0.6855	1.4039

• **Prueba de asociación entre el desarrollo de leucemia y exposición de las personas a cables de alta tensión considerando el primer caso.**

Si se supone que no existe asociación entre el desarrollo de leucemia y el factor de riesgo se tiene lo siguiente:

H_0 = El estar expuesto a cables de alta tensión no es un riesgo para el desarrollo de leucemia.

H_1 = El estar expuesto a cables de alta tensión es un riesgo para el desarrollo de leucemia.

Tabla G1. Datos del estudio para examinar la asociación entre la ocurrencia de leucemia y la exposición del menor a cables de alta tensión

CABLES2 by PACPRIL	casos	controles	total renglón
Expuestos	255	224	479
No expuestos	16	51	67
total columna	271	275	546

Tabla G2. Cálculo de la estadística chi cuadrada para los datos

Factor de riesgo Cables de alta tensión	Grupo	Valor observado (o)	Valor esperado (e)	Valor observado menos esperado (o - e)	Contribución a χ^2 [(o - e) ² / e]
Expuestos	Casos	255	237.7	17.3	1.259
	Controles	224	241.2	-17.2	1.226
No expuestos	Casos	16	33.2	-17.2	8.910
	Controles	51	33.7	17.3	8.881
Total		546			$\chi^2 = 20.27$

Tabla G3. Resultado de la prueba chi-cuadrada

Chi - cuadrada	valor	Grados de Libertad	Significancia
Pearson	20.27	1	.0000

En este caso se ve que el valor del estadístico es muy grande para un nivel del 95%, por lo tanto se rechaza la hipótesis nula, en la tabla G1 se puede observar que el número de expuestos es mucho mayor respecto a los no expuestos. Al calcular la razón de momios y el riesgo relativo se obtuvieron los siguientes resultados:

Estadístico	Valor	Intervalo de confianza a un 95%
Razón de momios	3.628	2.0105 6.53786
Riesgo relativo para el grupo de casos	2.229	1.2361 4.01935
Riesgo relativo para el grupo de controles	.6143	.3046 1.1076

El valor de la razón de momios indica que el riesgo de desarrollar leucemia para una persona expuesta respecto a una que no estuvo expuesta es 3.1 veces mayor. Si observamos el riesgo relativo vemos que también el riesgo es mayor para los expuestos en ambos grupos.

• **Prueba de asociación entre el desarrollo de leucemia y exposición de las personas a subestaciones eléctricas considerando el primer caso**

Si se supone que no existe asociación entre el desarrollo de leucemia y el factor de riesgo se tiene lo siguiente:

- H_0 = El estar expuesto a subestaciones eléctricas no es un riesgo para el desarrollo de leucemia.
- H_1 = El estar expuesto a subestaciones eléctricas es un riesgo para el desarrollo de leucemia.

Tabla H1. Datos del estudio para examinar la asociación entre la ocurrencia de leucemia y la exposición del menor a subestaciones eléctricas.

SUBEST. by PAF PBL1	casos	controles	total renglón
Expuestos	33	37	70
No expuestos	242	234	476
total columna	275	271	546

Tabla H2. Cálculo de la estadística chi cuadrada para los datos

Factor de riesgo subestaciones eléctricas	Grupo	Valor observado (o)	Valor esperado (e)	Valor observado menos esperado (o - e)	Contribución a χ^2 [(o - e) ² /e]
Expuestos	Casos	33	35.2	-2.2	.1375
	Controles	37	34.7	2.3	.1524
No expuestos	Casos	242	239.7	2.3	.0230
	Controles	234	236.2	-2.2	.0204
Total		246			$\chi^2 = .3325$

Tabla H3. Resultado de la prueba chi-cuadrada

Chi - cuadrada	valor	Grados de Libertad	Significancia
Pearson	.3325	1	.5122

De acuerdo a la tabla H3 se puede comentar que no hay asociación entre las variables consideradas dado que el valor del estadístico es muy pequeño, por lo tanto se considera que la exposición a transformadores no es factor de riesgo para la enfermedad. Si se hace inferencia respecto a la razón de momios, se ve que el riesgo de que una persona desarrolle leucemia dado que en algún momento ha estado expuesta a transformadores de luz no es relevante respecto a una persona que no ha estado expuesta.

	Estadístico	Valor	Intervalo de confianza a un 95%	
Razón de momios		.8624	0.5217	1.4254
Riesgo relativo para el grupo de casos		.9272	0.5609	1.5326
Riesgo relativo para el grupo de controles		1.078	0.6522	1.7817

• **Prueba de asociación entre el desarrollo de leucemia y exposición de las personas a torres de alta tensión considerando el primer caso.**

Si se supone que no existe asociación entre el desarrollo de leucemia y el factor de riesgo se tiene lo siguiente:

H_0 = El estar expuesto a torres de alta tensión no es un riesgo para el desarrollo de leucemia.

H_1 = El estar expuesto a torres de alta tensión es un riesgo para el desarrollo de leucemia.

Tabla 11. Datos del estudio de casos y controles para examinar la asociación entre la ocurrencia de leucemia y la exposición a torres de alta tensión

TORRES DE ALTA TENSIÓN	casos	controles	total renglón
Expuestos	70	68	138
No expuestos	205	203	408
total columna	275	271	546

Tabla 12. Cálculo de la estadística chi cuadrada para los datos

Factor de riesgo Torres de alta tensión	Grupo	Valor observado (o)	Valor esperado (e)	Valor observado menos esperado (o - e)	Contribución a χ^2 {(o - e) ² / e}
Expuestos	Casos	70	69.5	0.5	.0035
	Controles	68	68.4	-0.4	.0023
No expuestos	Casos	205	205.4	-0.4	.0007
	Controles	203	202.5	0.5	.0012
Total		546			$\chi^2 = .0079$

Tabla 13. Resultado de la prueba chi-cuadrada

Chi - cuadrada	valor	Grados de Libertad	Significancia
Pearson	.0079	1	.9688

El valor del estadístico nos indica que no es significativo, por lo que se acepta la hipótesis nula. Si se hace inferencia respecto a la razón de momios, se ve que el riesgo de ocurrencia de la enfermedad es igual para ambos grupos.

	Estadístico	Valor	Intervalo de confianza a un 95%	
Razón de momios		1.019	0.6926	1.4722
Riesgo relativo para el grupo de casos		1.009	0.6858	1.4578
Riesgo relativo para el grupo de controles		0.990	0.6729	1.4303

• **Prueba de asociación entre el desarrollo de leucemia y exposición de las personas a transformadores de luz considerando el segundo caso.**

Si se supone que no existe asociación entre el desarrollo de leucemia y el factor de riesgo se tiene lo siguiente:

H_0 = El estar expuesto a transformadores de luz no es un riesgo para el desarrollo de leucemia.
 H_1 = El estar expuesto a transformadores de luz es un riesgo para el desarrollo de leucemia.

Tabla 11. Datos del estudio para examinar la asociación entre la ocurrencia de leucemia y la exposición a transformadores de luz.

TRANSF. PA/PRII	casos	controles	total renglón
Expuestos	38	35	73
No expuestos	11	12	23
total columna	49	47	96

Tabla 12. Cálculo de la estadística chi cuadrada para los datos

Factor de riesgo Transformadores de luz	Grupo	Valor observado (o)	Valor esperado (e)	Valor observado menos esperado (o - e)	Contribución a χ^2 [(o-e) ² /e]
Expuestos	Casos	38	37.2	0.8	.0172
	Controles	35	35.7	-0.7	.0137
No expuestos	Casos	11	11.2	-0.2	.0035
	Controles	12	11.2	0.8	.0035
Total		96			$\chi^2 = .03804$

Tabla 13. Resultado de la prueba chi cuadrada

Chi - cuadrada	valor	Grados de Libertad	Significancia
Pearson	.03804	1	.8496

El valor del estadístico nos indica que no es significativo, por lo que se acepta la hipótesis nula, sin embargo si comparamos la razón de momios respecto al primer caso vemos que el riesgo de ocurrencia de la enfermedad se incrementa ligeramente de un valor de 1.08 a 1.18, es decir, el riesgo de enfermedad para una persona que ha estado expuesta siempre, es 0.1 mayor respecto al de una persona que se expuso sólo en alguna circunstancia.

	Estadístico	Valor	Intervalo de confianza a un 95%
Razón de momios		1.18	0.51410 2.7070
Riesgo relativo para el grupo de casos		1.08	0.47063 2.4782
Riesgo relativo para el grupo de controles		0.9259	0.40374 2.1249

• **Prueba de asociación entre el desarrollo de leucemia y exposición de las personas a cables de alta tensión considerando el segundo caso.**

Para probar si las variables están asociadas se tiene la siguiente hipótesis:

H_0 = La exposición a cables de alta tensión no es un riesgo para el desarrollo de leucemia.
 H_1 = La exposición a cables de alta tensión es un riesgo para el desarrollo de leucemia.

Tabla K1. Datos del estudio para examinar la asociación entre la ocurrencia de leucemia y la exposición del menor a cables de alta tensión

CABLES2 by PACPREL	casos	controles	total renglón
Expuestos	73	62	135
No expuestos	1	6	7
total columna	74	68	142

Tabla K2. Cálculo de la estadística chi cuadrada para los datos

Factor de riesgo Cables de alta tensión	Grupo	Valor		Valor observado menos esperado (o - e)	Contribución a χ^2 [(o - e) ² / e]
		observado (o)	esperado (e)		
Expuestos	Casos	73	70.3	2.7	.1036
	Controles	62	64.6	-2.6	.1046
No expuestos	Casos	1	3.6	-2.6	1.877
	Controles	6	3.3	2.7	2.209
Total					$\chi^2 = 4.291$

Tabla K3. Resultados de la prueba chi cuadrada

Chi - cuadrada	valor	Grados de Libertad	Significancia
Pearson	4.291	1	.02830

El valor del estadístico obtenido es 4.214, y en este caso tomando a $\alpha = .05$ vemos que si es significativo, esto nos conduce a rechazar la hipótesis nula.

Estadístico	Valor	Intervalo de confianza a un 95%	
Razón de momios	7.0645	0.8292	60.1813
Riesgo relativo para el grupo de casos	3.7851	0.4443	32.2446
Riesgo relativo para el grupo de controles	.5358	.06289	4.5388

Al igual que en el caso I el riesgo es mayor para los expuestos y vemos en este caso que para las personas que han estado expuestas siempre, el riesgo se incrementa en una proporción considerable..

- **Prueba de asociación entre el desarrollo de leucemia y exposición de las personas a subestaciones eléctricas considerando el segundo caso.**
Si suponemos que no hay asociación entre las variables la hipótesis es la siguiente:

H_0 = La exposición a subestaciones eléctricas no es factor de riesgo para el desarrollo de leucemia.
 H_1 = La exposición a subestaciones eléctricas es un factor de riesgo para el desarrollo de leucemia.

Tabla L1. Datos del estudio para examinar la asociación entre la ocurrencia de leucemia y la exposición a subestaciones eléctricas

SUBEST2 by PACPREL	casos	controles	total renglón
Expuestos	3	1	4
No expuestos	68	64	132
total columna	71	65	136

Tabla 1.2. Cálculo de la estadística chi cuadrada para los datos

Factor de riesgo	Grupo	Valor observado (o)	Valor esperado (e)	Valor observado menos esperado (o - e)	Contribución a χ^2 [(o - e) ² / e]
Subestación eléctrica	Casos	3	2.08	0.92	0.4069
	Control	1	1.91	-0.91	0.4335
No expuestos	Casos	68	68.91	-0.91	0.0120
	Controles	64	63.0	-1	0.0158
Total		136			$\chi^2 = 0.8682$

Tabla 1.3. Resultado de la prueba chi-cuadrada

Chi - cuadrada	Grados de Libertad	Significancia
Pearson	1	.2639

Podemos ver que para $\alpha=0.05$ el valor del estadístico es muy pequeño, esto nos permite aceptar la hipótesis nula, se puede comentar sin embargo que al comparar la razón de momios respecto al primer caso si hay una diferencia en riesgo puesto que el valor cambia de .8624 a 2.82, por lo tanto se puede decir que el riesgo se incrementa cuando una persona ha estado expuesta por un tiempo mayor.

Estadístico	Valor	Intervalo de confianza a un 95%
Razón de momios	2.82	0.28573 2.7878
Riesgo relativo para el grupo de casos	1.455	0.14754 14.347
Riesgo relativo para el grupo de controles	0.686	0.06956 9.4972

- **Prueba de asociación entre el desarrollo de leucemia y exposición de las personas a torres de alta tensión considerando el segundo caso.**
La hipótesis nula suponiendo que no hay asociación entre el desarrollo de leucemia y el factor de riesgo se tiene lo siguiente:

H_0 = La exposición del menor a torres de alta tensión no es factor de riesgo para la ocurrencia de leucemia
 H_1 = La exposición del menor a torres de alta tensión es factor de riesgo para el desarrollo de leucemia.

Tabla M1. Datos del estudio para examinar la asociación entre la ocurrencia de leucemia y la exposición del menor a torres de alta tensión

TORRES BY PACTREL	casos	controles	total renglón
Expuestos	9	10	19
No expuestos	57	48	105
total columna	66	58	124

Tabla M2. Cálculo de la estadística chi cuadrada para los datos

Factor de riesgo	Grupo	Valor observado (o)	Valor esperado (e)	Valor observado menos esperado (o - e)	Contribución a χ^2 [(o - e) ² / e]
Torres de alta tensión	Casos	9	10.11	-1.1	.1195
	Control	10	8.8	1.2	.1636
No expuestos	Casos	57	55.8	1.2	.0258
	Controles	48	49.1	-1.1	.0246
Total		124			$\chi^2 = .3336$

Tabla M3 Resultado de la prueba chi-cuadrada

Chi - cuadrada	valor	Grados de Libertad	Significancia
Pearson	.3336	1	.5842

Después de realizar la prueba de bondad de ajuste se puede ver que no hay asociación entre las variables, por lo que se acepta la hipótesis nula.

Estadístico	Valor	Intervalo de confianza a un 95%	
Razón de momios	0.7578	0.28475	2.01677
Riesgo relativo para el grupo de casos	0.8725	0.32785	2.32216
Riesgo relativo para el grupo de controles	1.146	0.43062	3.04978

Se puede ver que la razón de momios es igual a 0.7578, esto indica que el riesgo de desarrollar leucemia cuando una persona esta expuesta por un tiempo mayor no es significativo.

Dentro del análisis bivariado también se consideró la exposición de los padres con ciertos factores de riesgo, antes o durante el embarazo de la madre del niño en estudio.

En la literatura médica se menciona que los factores asociados con leucemia pueden ser prenatales, natales o postnatales¹², por ello se considera importante también incluir en el análisis la ocupación del padre cuando la madre se embarazó del niño así como su exposición a ciertos factores que pudiesen influir en el proceso prenatal.

- **Análisis para ver si hay asociación entre la exposición a solventes derivados del petróleo que tuvo el padre cuando la madre se embarazó del menor respecto al desarrollo de leucemia.**

Suponiendo que la exposición a solventes que pudo tener el padre antes o durante el embarazo de la madre no representa un riesgo para el menor, se tiene lo siguiente:

H_0 = La exposición que el padre del niño tuvo a solventes antes o durante el embarazo de la madre no es un factor de riesgo para el desarrollo de leucemia.

H_1 = La exposición que el padre del niño tuvo a solventes antes o durante el embarazo de la madre es un factor de riesgo para el desarrollo de leucemia.

Tabla N1. Datos del estudio para examinar la asociación entre la ocurrencia de leucemia en el niño y la exposición del padre a solventes derivados del petróleo durante el embarazo de la madre

SOLVENTE y PÉTRÓLEO	casos	controles	total renglón
Expuestos	22	21	43
No expuestos	165	166	331
total columna	187	187	374

¹² Shaw G., Lavey. Association of child-hood leukemia with maternal age, birth order and parenteral occupation. 1984, p.p. 788-795

Tabla N2. Cálculo de la estadística chi cuadrada para los datos.

Factor de riesgo solventes	Grupo	Valor observado (o)	Valor esperado (e)	Valor observado menos esperado (o - e)	Contribución a χ^2 [(o - e) ² / e]
Expuestos	Casos	22	21.5	0.5	.0116
	Control	21	21.5	-0.5	.0116
No expuestos	Casos	165	165.5	-0.5	.0015
	Controles	166	165.5	0.5	.0015
Total		374			$\chi^2 = .0262$

Tabla N3. Resultado de la prueba chi-cuadrada

Chi - cuadrada	valor	Grados de Libertad	Significancia
Pearson	.02628	1	.87123

El valor del estadístico chi-cuadrada es muy pequeño por lo que no es significativo a un nivel del 95 %, por lo tanto podemos decir que la exposición del padre a solventes derivados del petróleo no representa un factor de riesgo para el producto durante el embarazo de la madre.

Estadístico	Valor	Intervalo de confianza a un 95%	
Razón de momios	1.05397	.55821	1.99001
Riesgo relativo para el grupo de casos	1.02690	.74254	1.42017
Riesgo relativo para el grupo de controles	.97432	.71363	1.33023

Por otro lado la razón de momios es de 1.05, como el valor es muy cercano a uno se dice que la probabilidad de que una persona expuesta desarrolle leucemia es la misma que para una persona no expuesta.

- **Análisis para ver si hay asociación entre la exposición a insecticidas que tuvo el padre cuando la madre se embarazó del niño en estudio respecto al desarrollo de leucemia.**

Suponiendo que la exposición a insecticidas del padre antes o durante el embarazo de la madre no representa un riesgo para el menor, se tiene lo siguiente:

H_0 = La exposición que el padre del niño tuvo con insecticidas antes o durante el embarazo de la madre no es un factor de riesgo para el desarrollo de leucemia.

H_1 = La exposición que el padre del niño tuvo con insecticidas antes o durante el embarazo de la madre es un factor de riesgo para el desarrollo de leucemia.

Tabla O1. Datos del estudio para examinar la asociación entre la ocurrencia de leucemia en el niño y la exposición del padre a insecticidas antes o durante el embarazo de la madre

INSECTICIDA	CASOS	CONTROLES	total renglon
Expuestos	13	4	17
No expuestos	174	183	357
total columna	187	187	374

Tabla O2. Cálculo de la estadística chi cuadrada para los datos

Factor de riesgo	Grupo	Valor observado (o)	Valor esperado (e)	Valor observado menos esperado (o - e)	Contribución a χ^2 [(o-e) ² /e]
Expuestos	Casos	13	8.5	-4.5	2.3823
	Control	4	8.5	-4.5	2.3823
No expuestos	Casos	174	178.5	-4.5	.1134
	Controles	183	178.5	4.5	.1134
Total		374			$\chi^2 = 4.991$

Tabla O3. Resultado de la prueba chi cuadrada

Chi - cuadrada	valor	Grados de Libertad	Significancia
Pearson	4.99160	1	.02547

Según el valor obtenido en la tabla O3, se puede ver que para un nivel del 95% es significativo, por lo tanto se rechaza la hipótesis nula, esto indica que la exposición que tuvo el padre con insecticidas antes del embarazo puede ser un factor de riesgo para el niño.

Estadístico	Valor	ASE*	Significancia
Cramer's V	.11553		.02547
Coefficiente de Contingencia	.11476		.02547
Lambda con variable dependiente PACPREL	.04813	.09858	
Coefficiente de Incertidumbre con variable dependiente PACPREL	.01011	.00848	.02203
Somers' D con variable dependiente PACPREL	.27731	.10623	

Al obtener los estadísticos que miden el grado de asociación, podemos ver que el valor de la V de Cramer y el coeficiente de contingencia son pequeños, es decir la asociación entre el factor de riesgo y la enfermedad es pequeña.

Estadístico	Valor	Intervalo de confianza a un 95%	
Razón de momios	3.4181	1.0934	10.6846
Riesgo relativo para el grupo de casos	1.5689	.50190	4.90423
Riesgo relativo para el grupo de controles	.6378	.20403	1.13479

Considerando el análisis en términos del riesgo, según el valor de la razón de momios el riesgo de enfermedad es 3.4 veces mayor para una persona expuesta.

- **Análisis para ver si hay asociación entre la exposición a thiner que tuvo el padre cuando la madre se embarazó del niño en estudio respecto al desarrollo de leucemia.**

Considerando que no existe alguna relación entre la exposición que el padre del niño pudo tener antes o durante el embarazo de la madre y el posible desarrollo de leucemia en el niño, se tiene lo siguiente:

H_0 = La posible exposición del padre a thiner cuando la madre se embarazó del niño en estudio no es un riesgo para desarrollar leucemia.

H_1 = La posible exposición del padre a thiner cuando la madre se embarazó del niño en estudio es un riesgo para desarrollar leucemia.

Tabla P1. Datos del estudio para examinar la asociación entre la ocurrencia de leucemia en el niño y la exposición del padre a thiner antes o durante el embarazo de la madre

THINER y PAPAPEL	casos	controles	total renglón
Expuestos	36	33	69
No expuestos	151	154	305
total columna	187	187	374

Tabla P2. Cálculo de la estadística chi cuadrada para los datos

Factor de riesgo Thiner	Grupo	Valor observado (o)	Valor esperado (e)	Valor observado menos esperado (o - e)	Contribución a χ^2 [(o - e) ² / e]
Expuestos	Casos	36	34.5	1.5	.0652
	Control	33	34.5	-1.5	.0652
No expuestos	Casos	151	152.5	-1.5	.0147
	Controles	154	152.5	1.5	.0147
Total		374			$\chi^2 = .159$

Tabla P3. Resultado de la prueba chi-cuadrada

Chi - cuadrada	valor	Grados de Libertad	Significancia
Pearson	.15944	1	.68921

De acuerdo al valor del estadístico se ve que no es significativo, esto implica que no hay relación entre la exposición del padre y el proceso prenatal del menor para el desarrollo posible de leucemia.

Para la medida de asociación en términos del riesgo se tiene lo siguiente:

Estadístico	Valor	Intervalo de confianza a un 95%	
Razón de momios	1.11258	.65949	1.87697
Riesgo relativo para el grupo de casos	1.05574	.80565	1.38345
Riesgo relativo para el grupo de controles	.94891	.73699	1.22177

se puede ver que no existe asociación entre las variables dado que el valor de la razón de momios es muy cercano a uno, es decir el riesgo de ocurrencia de la enfermedad para una persona expuesta es de apenas 1.11 respecto a una que no esta expuesta.

- **Análisis para ver si hay asociación entre la exposición a pinturas que tuvo el padre cuando la madre se embarazó del niño en estudio respecto al desarrollo de leucemia.**

Si suponemos que la exposición del padre con pinturas antes o durante el embarazo de la madre respecto al niño en estudio, no es un factor de riesgo; se puede definir la hipótesis nula como sigue:

H_0 = La posible exposición del padre a pinturas durante el embarazo de la madre para el niño en estudio no es un riesgo para desarrollar leucemia.

H_1 = La posible exposición del padre a pinturas durante el embarazo de la madre para el niño en estudio es un riesgo para desarrollar leucemia.

Tabla Q1. Datos del estudio para examinar la asociación entre la ocurrencia de leucemia en el niño y la exposición del padre a pinturas antes o durante el embarazo de la madre.

Pinturas y padre(a)	casos	controles	total renglón
Expuestos	31	32	63
No expuestos	156	155	311
total columna	187	187	374

Tabla Q2. Cálculo de la estadística chi cuadrada para los datos

Factor de riesgo	Grupo	Valor observado (o)	Valor esperado (e)	Valor observado menos esperado (o - e)	Contribución a χ^2 $((o - e)^2 / e)$
Expuestos	Casos	31	31.5	-.5	.0079
	Control	32	31.5	.5	.0079
No expuestos	Casos	156	155.5	.5	.0016
	Controles	155	155.5	-.5	.0016
Total					$\chi^2 = .019$

Tabla Q3. Resultado de la prueba chi cuadrada

Chi - cuadrada	valor	Grados de Libertad	Significancia
Pearson	.01909	1	.89011

El valor del estadístico obtenido en la tabla Q3 es muy pequeño, por lo tanto se puede decir que la exposición del padre con pinturas no influye en el posible desarrollo de leucemia del niño. En los siguientes datos se puede ver que la razón de momios se aproxima a la unidad, esto quiere decir que el riesgo de la madre respecto a que el niño pueda desarrollar leucemia dado que el padre estuvo expuesto a pinturas durante el embarazo es el mismo para una madre cuyo esposo no estuvo expuesto durante el embarazo.

Estadístico	Valor	Intervalo de confianza a un 95%
Razón de momios	.96254	.56000 1.65443
Riesgo relativo para el grupo de casos	.98121	.75099 1.28201
Riesgo relativo para el grupo de controles	1.01940	.77488 1.34107

- **Análisis para ver si hay asociación entre la exposición a fertilizantes que tuvo el padre cuando la madre se embarazó del niño en estudio respecto al desarrollo de leucemia.**

Suponiendo que la exposición del padre con fertilizantes cuando la madre estuvo embarazada del niño no es un riesgo, se tiene lo siguiente:

H_0 = La posible exposición del padre a fertilizantes durante el embarazo de la madre para el niño en estudio no es un riesgo para desarrollar leucemia.

H_1 = La posible exposición del padre a fertilizantes durante el embarazo de la madre para el niño en estudio es un riesgo para desarrollar leucemia.

Tabla R1. Datos del estudio para examinar la asociación entre la ocurrencia de leucemia en el niño y la exposición del padre a fertilizantes antes o durante el embarazo de la madre.

FEUILI by PACPRILL	CASOS	controles	total renglón
Expuestos	3	4	7
No expuestos	184	183	367
total columna	187	187	374

Tabla R2. Cálculo de la estadística chi cuadrada para los datos.

Factor de riesgo Fertilizantes	Grupo	Valor observado (o)	Valor esperado (e)	Valor observado menos esperado (o - e)	Contribución a χ^2 [(o - e) ² / e]
Expuestos	Casos	3	3.5	-.5	.0714
	Control	4	3.5	.5	.0714
No expuestos	Casos	184	183.5	.5	.0013
	Controles	183	183.5	-.5	.0013
Total		374			$\chi^2 = .145$

Tabla R3. Resultado de la prueba chi-cuadrada.

Chi - cuadrada	valor	Grados de Libertad	Significancia
Pearson	.14558	1	.70279

Después de realizar la prueba de bondad de ajuste chi-cuadrada, se obtiene un valor de 0.14558 siendo no significativo a un nivel del 95 %, esto nos permite aceptar la hipótesis nula.

Estadístico	Valor	Intervalo de confianza a un 95%
Razón de momios	.74592	.16464 3.37953
Riesgo relativo para el grupo de casos	.87262	.45568 1.67103
Riesgo relativo para el grupo de controles	1.16985	.49431 2.76858

El valor obtenido de la razón de momios es igual a .74592, esto quiere decir el riesgo de la enfermedad es mayor para los no expuestos, si observamos las frecuencias para los no expuestos de la tabla R1 vemos que la mayoría de los niños que desarrollaron leucemia caen dentro del grupo de padres que dijeron no estar expuestos con fertilizantes durante el embarazo de la madre, así mismo para la muestra de controles se tiene una frecuencia mayor para los no expuestos, esto realmente confirma que el desarrollo de leucemia se debe a otros factores.

- **Análisis para ver si hay asociación entre la exposición a gasolina que tuvo el padre durante el embarazo de la madre del niño en estudio respecto al desarrollo de leucemia.**

Considerando que no existe alguna relación entre la exposición a gasolina que el padre del niño pudo tener antes o durante el embarazo de la madre y el posible desarrollo de leucemia en el niño, se tiene lo siguiente:

H_0 = La posible exposición del padre a gasolina durante el embarazo de la madre del niño en estudio no es un riesgo para desarrollar leucemia.

H_1 = La posible exposición del padre a gasolina durante el embarazo de la madre del niño en estudio es un riesgo para desarrollar leucemia.

Tabla S1. Datos del estudio para examinar la asociación entre la ocurrencia de leucemia en el niño y la exposición del padre a gasolina antes o durante el embarazo de la madre.

GALÓN de GASOLINA	casos	controles	total renglón
Expuestos	28	21	49
No expuestos	159	166	325
total columna	187	187	374

Tabla S2. Cálculo de la estadística chi cuadrada para los datos.

Factor de riesgo	Grupo	Valor observado (o)	Valor esperado (e)	Valor observado menos esperado (o - e)	Contribución a χ^2 [(o - e) ² / e]
Expuestos	Casos	28	24.5	3.5	0.5
	Control	21	24.5	-3.5	0.5
No expuestos	Casos	159	162.5	-3.5	.0753
	Controles	166	162.5	3.5	.0753
Total		374			$\chi^2 = 1.150$

Tabla S3. Resultados de la prueba chi-cuadrada.

Chi - cuadrada	valor	Grados de Libertad	Significancia
Pearson	1.15077	1	28339

De la tabla S3 se ve que el valor del estadístico es pequeño, por lo que no se considera significativo a un nivel del 95 %.

Estadístico	Valor	Intervalo de confianza a un 95%
Razón de momios	1.39203	.75928 2.55211
Riesgo relativo para el grupo de casos	1.19179	.84797 1.67503
Riesgo relativo para el grupo de controles	.85615	.65572 1.11786

Es importante comentar que aunque la prueba de bondad de ajuste no es significativa, el valor de la razón de momios nos indica que el riesgo de que un niño desarrolle leucemia dado que el padre estuvo expuesto con gasolina durante el proceso prenatal es 1.3 veces mayor que el riesgo para un niño cuyo padre no se expuso durante el embarazo de la madre.

Como se mencionó al inicio del capítulo tres, la exposición a radiaciones ionizantes del niño puede ser un factor de riesgo que este asociado con el desarrollo de leucemia, para ver este aspecto se consideró en el diseño de la investigación, obtener información acerca de la probable exposición del niño a radiografías o ultrasonidos.

- **Análisis para ver si el desarrollo de leucemia esta relacionado con la toma de ultrasonidos al niño en estudio.**

Considerando que no hay relación entre la ocurrencia de leucemia en el niño y el hecho de que se le hubiese tomado algún ultrasonido (s) se tiene lo siguiente:

- H₀ = La toma de ultrasonidos al niño no es factor de riesgo para el desarrollo de leucemia.
- H₁ = La toma de ultrasonidos al niño es factor de riesgo para el desarrollo de leucemia.

Tabla T1. Datos del estudio para examinar la asociación entre la ocurrencia de leucemia y la toma de ultrasonidos al niño en estudio

LEUCEMI y PACIFRELL	casos	controles	total renglón
Expuestos	12	2	14
No expuestos	175	185	360
total columna	187	187	374

Tabla T2. Cálculo de la estadística chi cuadrada para test de datos

Factor de riesgo Ultrasonido	Grupo	Valor observado (o)	Valor esperado (e)	Valor observado menos esperado (o - e)	Contribución a χ^2 [(o - e) ² /e]
Expuestos	Casos	12	7.0	5.0	3.5714
	Control	2	7.0	-5.0	3.5714
No expuestos	Casos	175	180.0	-5.0	.1388
	Controles	185	180.0	5.0	.1388
Total					$\chi^2 = 7.420$

Tabla T3. Resultado de la prueba chi-cuadrada

Chi - cuadrada	valor	Grados de Libertad	Significancia
Pearson	7.42063	1	.00645

En este caso vemos que el valor del estadístico para un grado de libertad y un nivel del 95% es 7.42063 siendo muy significativo, es decir en este caso p es menor que .01, esto nos conduce a rechazar la hipótesis nula.

Para medir la asociación en términos del riesgo se tiene lo siguiente:

Estadístico	Valor	Intervalo de confianza a un 95%	
Razón de momios	6.34286	1.39959	28.74537
Riesgo relativo para el grupo de casos	1.76326	.99316	13.02912
Riesgo relativo para el grupo de controles	.56713	.44667	.72008

la razón de momios nos indica que el riesgo de que una persona desarrolle leucemia dado que está expuesta al factor de riesgo es 6.342 veces más probable respecto a una persona que no está expuesta. En otras palabras si consideramos el total marginal para el grupo de expuestos vemos que de los 14 solamente dos corresponden al grupo control (14.28 %), mientras que 12 de los 14 expuestos desarrollaron leucemia (85.72 %), esto muestra que el riesgo para una persona expuesta es muy alto. Sin embargo es importante comentar que esta variable no se tomará en cuenta para el análisis multivariado ya que no es importante para el estudio, pues si bien es cierto que la relación entre los casos y controles expuestos es alta, en la tabla 1.25 de la sección 2.2 se observa que la frecuencia de ultrasonidos cae entre 1 y 2 para la mayoría de los casos.

- **Análisis para ver si el desarrollo de leucemia esta asociado con el defecto fisiológico que el niño pueda tener al nacimiento.**

Considerando que no hay relación alguna entre el desarrollo de leucemia y algún defecto físico del niño al nacimiento se tiene lo siguiente:

H_0 = El defecto físico del niño al nacimiento no es un factor de riesgo para posible desarrollo de leucemia.
 H_1 = El defecto físico del niño al nacimiento es un factor de riesgo para posible desarrollo de leucemia.

Tabla U1. Datos del estudio para examinar la asociación entre la ocurrencia de leucemia y el defecto físico del niño al nacimiento.

LEUCEMIA por PAPEL	casos	controles	total renglón
Expuestos	11	5	16
No expuestos	176	182	358
total columna	187	187	374

Tabla U2. Cálculo de la estadística chi cuadrado para los datos

Factor de riesgo Defecto físico	Grupo	Valor observado (o)	Valor esperado (e)	Valor observado menos esperado (o - e)	Contribución a χ^2 [(o - e) ² / e]
Expuestos	Casos	11	8.0	3.0	1.125
	Control	5	8.0	-3.0	1.125
No expuestos	Casos	176	179.0	-3.0	.0502
	Controles	182	179.0	3.0	.0502
Total		374			$\chi^2 = 2.350$

Tabla U3. Resultado de la prueba chi-cuadrada

Chi - cuadrada Pearson	valor	Grados de Libertad	Significancia
	2.35056	1	.12524

De la tabla U3 podemos ver que el valor (2.35056), es menor que el correspondiente al de las tablas de la distribución chi-cuadrada con un grado de libertad (3.84), por lo tanto se acepta la hipótesis nula.

En términos de riesgo se tiene lo siguiente:

Estadístico	Valor	Intervalo de confianza a un 95%	
Razón de momios	2.27500	.77471	6.68069
Riesgo relativo para el grupo de casos	1.39843	.78094	3.38889
Riesgo relativo para el grupo de controles	.71508	.50556	1.01145

En este caso vemos que aunque la prueba de bondad de ajuste no es significativa, la razón de momios (2.27500) indica que el riesgo de que una persona desarrolle leucemia dado que está expuesta es 2.27 mayor respecto a una persona no expuesta. Considerando solo los niños que si tuvieron defectos físicos al nacimiento, se ve que de los 16, 11 desarrollaron leucemia, esto apoya la literatura médica en el hecho de que una persona cuyo sistema inmunológico es mas débil tiene un riesgo mayor de desarrollar leucemia.

- **Análisis para ver si el desarrollo de leucemia esta asociado con la exposición del menor a solventes derivados del petróleo.**

Considerando que no existe relación alguna entre el desarrollo de leucemia y la exposición del menor a solventes derivados del petróleo se tiene lo siguiente:

H_0 = La exposición del menor a solventes no es factor de riesgo para el desarrollo de leucemia.
 H_1 = La exposición del menor a solventes es factor de riesgo para el desarrollo de leucemia.

Tabla V1. Datos del estudio para examinar la asociación entre la ocurrencia de leucemia y la exposición del menor a solventes derivadas del petróleo

Exposición de Paciente	casos	controles	total renglón
Expuestos	5	1	6
No expuestos	182	186	368
total columna	187	187	374

Tabla V2. Cálculo de la estadística chi cuadrada para los datos

Factor de riesgo solventes	Grupo	Valor observado (o)	Valor esperado (e)	Valor observado menos esperado (o - e)	Contribución a χ^2 [(o-e) ² /e]
Expuestos	Casos	5	3.0	2.0	1.33
	Control	1	3.0	-2.0	1.33
No expuestos	Casos	182	184.0	-2.0	.0217
	Controles	186	184.0	2.0	.0217
Total					$\chi^2 = 2.710$

Tabla V3. Resultado de la prueba chi cuadrada

Chi - cuadrada	valor	Grados de Libertad	Significancia
Pearson	2.71014	1	.09971

Según los datos obtenidos en la tabla V3, se ve que el valor del estadístico no es significativo a un nivel del 95 %, esto se puede corroborar con la tabla V1, donde se ve que de los 374 niños incluidos en el estudio solamente 6 tuvieron exposición con solventes (1.6 %). Sin embargo si se analiza la frecuencia de esos seis, el 83.33 % desarrollo la enfermedad, si consideramos la razón de momios vemos que la probabilidad de que una persona expuesta desarrolle leucemia es 5.1 veces mayor que una persona que no está expuesta. Es importante observar que si la prueba de bondad de ajuste se realizara con un intervalo de confianza del 90%, es decir $\alpha=0.1$, entonces la prueba si resultaría significativa.

Estadístico	Valor	Intervalo de confianza a un 95%	
Razón de momios	5.10989	.59123	44.16418
Riesgo relativo para el grupo de casos	1.68498	.50529	18.20087
Riesgo relativo para el grupo de controles	.59348	.40893	.86130

• **Análisis para ver si el desarrollo de leucemia esta asociado con la exposición del menor a insecticidas.**

Suponiendo que no hay asociación entre el desarrollo de leucemia y la exposición del menor a insecticidas se tiene lo siguiente:

- H₀ = El desarrollo de leucemia es independiente de la exposición del niño con insecticidas.
- H₁ = El desarrollo de leucemia no es independiente de la exposición del niño con insecticidas.

Tabla W1. Datos del estudio para examinar la asociación entre la ocurrencia de leucemia y la exposición del menor a insecticidas.

Exposición de Paciente	casos	controles	total renglón
Expuestos	54	52	106
No expuestos	133	135	268
total columna	187	187	374

Tabla W2. Cálculo de la estadística chi cuadrada para los datos

Factor de riesgo	Grupo	Valor observado (o)	Valor esperado (e)	Valor observado menos esperado (o - e)	Contribución a χ^2 [(o - e) ² / e]
Expuestos	Casos	54	53.0	1.0	.0188
	Control	52	53.0	-1.0	.0188
No expuestos	Casos	133	134.0	-1.0	.0074
	Controles	135	134.0	1.0	.0074
Total		374			$\chi^2 = .052$

Tabla W3. Resultado de la prueba chi cuadrada

Chi cuadrada	valor	Grados de Libertad	Significancia
Pearson	.05266	1	.81850

De la tabla W3 podemos ver que el valor del estadístico es muy pequeño, por lo tanto se puede decir que no hay asociación entre el factor de riesgo y la enfermedad, la razón de momios (1.05408) es muy cercano a uno, por lo tanto el riesgo de enfermedad es similar para ambos grupos.

Estadístico	Valor	Intervalo de confianza a un 95%	
Razón de momios	1.05408	.67222	1.65285
Riesgo relativo para el grupo de casos	1.02684	.81790	1.28915
Riesgo relativo para el grupo de controles	.97416	.77993	1.21675

• **Análisis para ver si el desarrollo de leucemia esta asociado con la exposición del menor a thiner.**

Aquí lo que se supone es que la exposición del menor con thiner, no es un riesgo para el desarrollo de leucemia, la hipótesis a probar es la siguiente:

- H₀ = El desarrollo de leucemia es independiente de la exposición del niño con thiner.
 H₁ = El desarrollo de leucemia no es independiente de la exposición del niño con thiner.

Tabla X1. Datos del estudio para examinar la asociación entre la ocurrencia de leucemia y la exposición del menor a thiner.

RESULTADO PACTRIL	casos	controles	total renglón
Expuestos	17	17	34
No expuestos	170	170	340
total columna	187	187	374

Tabla X2. Cálculo de la estadística chi cuadrada para los datos

Factor de riesgo	Grupo	Valor observado (o)	Valor esperado (e)	Valor observado menos esperado (o - e)	Contribución a χ^2 [(o - e) ² / e]
Expuestos	Casos	17	17.0	.0	.0
	Control	17	17.0	.0	.0
No expuestos	Casos	170	170.0	.0	.0
	Controles	170	170.0	.0	.0
Total		374			$\chi^2 = .000$

Tabla X1. Resultado de la prueba chi-cuadrada

Chi - cuadrada	valor	Grados de Libertad	Significancia
Pearson	.00000	1	1.00000

En este caso si observamos la tabla X1, podríamos decir sin necesidad de realizar la prueba de bondad de ajuste que no hay asociación puesto que las frecuencias son las mismas para cada grupo, al realizar la prueba chi-cuadrada nos da un valor de cero siendo no significativo para $\alpha=.05$, por lo tanto se acepta la hipótesis nula. Así mismo la razón de momios es igual a uno lo cual indica que la probabilidad de desarrollar leucemia para una persona expuesta es la misma que para una persona no expuesta.

Estadístico	Valor	Intervalo de confianza a un 95%	
Razón de momios	1.00000	.49407	2.02400
Riesgo relativo para el grupo de casos	1.00000	.49407	2.02400
Riesgo relativo para el grupo de controles	1.00000	.49407	2.02400

• **Análisis para ver si el desarrollo de leucemia esta asociado con la exposición del menor a pinturas.**

Para ver si el desarrollo de leucemia es independiente de la exposición del niño con pinturas se tiene la siguiente hipótesis:

H_0 = El desarrollo de leucemia es independiente de la exposición del niño con pinturas.

H_1 = El desarrollo de leucemia no es independiente de la exposición del niño con pinturas.

Tabla Y1. Datos del estudio s para examinar la asociación entre la exposición de leucemia y la exposición del menor a pinturas.

LEUCEMIA por PAPEREL	casos	controles	total renglón
Expuestos	31	46	77
No expuestos	156	141	297
total columna	187	187	374

Tabla Y2. Cálculo de la estadística chi cuadrada para los datos

Factor de riesgo	Grupo	Valor observado (o)	Valor esperado (e)	Valor observado menos esperado (o - e)	Contribución a χ^2 [(o-e) ² / e]
Expuestos	Casos	31	38.5	-7.5	1.461
	Control	46	38.5	7.5	1.461
No expuestos	Casos	156	148.5	7.5	.3787
	Controles	141	148.5	-7.5	.3787
Total					3.679

Tabla Y3. Resultado de la prueba chi-cuadrada

Chi - cuadrada	valor	Grados de Libertad	Significancia
Pearson	3.67965	1	.05508

El valor del estadístico es de 3.6796, lo cual indica que no hay asociación entre las variables, está consideración debe quedar muy al margen ya que si bien la prueba nos permite aceptar la hipótesis nula, es importante tomar en cuenta que si α se incrementara a 0.1, la prueba si resultaría significativa.

Si se mide la asociación mediante la razón de momios se ve que no hay riesgo de desarrollar leucemia para un niño expuesto con pinturas respecto a uno no expuesto.

Estadístico	Valor	Intervalo de confianza a un 95%	
Razón de momios	.6911	.3666	1.01355
Riesgo relativo para el grupo de casos	.76648	.63843	.98918
Riesgo relativo para el grupo de controles	1.30466	.97352	1.74843

- Análisis para ver si el desarrollo de leucemia esta asociado con la exposición del menor a gasolina.

Para ver si el desarrollo de leucemia es independiente de la exposición del niño con gasolina se tiene la siguiente hipótesis:

H_0 = El desarrollo de leucemia es independiente de la exposición del niño con gasolina.

H_1 = El desarrollo de leucemia no es independiente de la exposición del niño con gasolina.

Tabla 21. Datos del estudio para examinar la asociación entre la ocurrencia de leucemia y la exposición del menor a gasolina.

GRUPO o PACIENTE	casos	controles	total renglón
Expuestos	10	8	18
No expuestos	177	179	356
total columna	187	187	374

Tabla 22. Cálculo de la estadística chi cuadrado para los datos

Factor de riesgo	Grupo	Valor observado (o)	Valor esperado (e)	Valor observado menos esperado (o - e)	Contribución a χ^2 [(o - e) ² / e]
Expuestos	Casos	10	9.0	1.0	.1111
	Control	8	9.0	-1.0	.1111
No expuestos	Casos	177	178.0	-1.0	.0056
	Controles	179	178.0	1.0	.0056
Total					$\chi^2 = .2334$

Tabla 23. Resultados de la prueba chi-cuadrada

Chi - cuadrada	valor	Grados de Libertad	Significancia
Pearson	.23346	1	.62897

El resultado de la prueba chi-cuadrada para un nivel del 95% y un grado de libertad es no significativo para las muestras, por lo tanto la exposición del niño con gasolina no está relacionado con el desarrollo de leucemia, esto nos permite aceptar la hipótesis nula. Si medimos el riesgo en términos de la razón de momios podemos ver que el riesgo de desarrollar leucemia para una persona expuesta es de 1.26 siendo algo significativo respecto a la persona no expuesta.

Estadístico	Valor	Intervalo de confianza a un 95%	
Razón de momios	1.26412	.48761	3.27725
Riesgo relativo para el grupo de casos	1.13132	.66809	1.91575
Riesgo relativo para el grupo de controles	.89494	.58439	1.37054

- **Análisis para ver si existe alguna relación entre el desarrollo de leucemia y las enfermedades que el niño puede tener en el primer año de vida.**

Para ver si el desarrollo de leucemia es independiente de las enfermedades que el niño pueda tener en su primer año de vida se tiene lo siguiente:

H_0 = Las enfermedades del niño en el primer año de vida no son riesgo para el desarrollo de leucemia

H_1 = Las enfermedades del niño en el primer año de vida no son riesgo para el desarrollo de leucemia

Tabla AA1. Datos del estudio para examinar la asociación entre la ocurrencia de leucemia y las enfermedades del niño en el primer año de vida

TIPO DE ENFERMEDAD	casos	controles	total renglón
Expuestos	59	114	174
No expuestos	128	132	260
total columna	187	187	374

Tabla AA2. Cálculo de la estadística chi cuadrada para los datos

Factor de riesgo (enfermedad)	Grupo	Valor observado (o)	Valor esperado (e)	Valor observado menos esperado (o - e)	Contribución a χ^2 $[(o - e)^2 / e]$
Expuestos	Casos	59	57.0	-2.0	.0701
	Control	55	57.0	-2.0	.0701
No expuestos	Casos	128	130.0	-2.0	.0307
	Controles	132	130.0	2.0	.0307
Total		374			$\chi^2 = .201$

Tabla AA3. Resultado de la prueba chi cuadrada

Chi - cuadrada	valor	Grados de Libertad	Significancia
Pearson	.20189	1	.65320

Después de realizar la prueba de bondad de ajuste se ve que no existe asociación entre las variables, por lo tanto se acepta la hipótesis nula. Al considerar la razón de momios se ve que el riesgo de desarrollar leucemia cuando la persona esta expuesta es de 1.106 por lo tanto no se considera relevante.

Estadístico	Valor	Intervalo de confianza a un 95%	
Razón de momios	1.10625	.71209	1.71859
Riesgo relativo para el grupo de casos	1.05231	.84057	1.31739
Riesgo relativo para el grupo de controles	.95124	.76646	1.18057

- **Análisis para ver si el desarrollo de leucemia tiene alguna relación con los niños que por algún motivo se hospitalizaron el primer año de vida.**

Si suponemos que hay alguna relación entre el desarrollo de leucemia y los niños que se hospitalizaron por algún motivo en su primer año de vida, la hipótesis a probar es:

H_0 = La posible hospitalización del niño en su primer año de vida no es un riesgo para el desarrollo de leucemia.

H_1 = La posible hospitalización del niño en su primer año de vida es riesgo para el desarrollo de leucemia.

Tabla AB1. Datos del estudio para examinar la asociación entre la ocurrencia de leucemia y la hospitalización del niño en su primer año de vida

HOSPITALIA by PACPREL	casos	controles	total renglon
Expuestos	24	39	63
No expuestos	163	148	311
total columna	187	187	374

Tabla AB2. Cálculo de la estadística chi cuadrada para los datos

Factor de riesgo Hospitalización	Grupo	Valor observado (o)		Valor observado menos esperado (o - e)	Contribución a χ^2 [(o - e) ² / e]
		valor observado (o)	Valor esperado (e)		
Expuestos	Casos	24	31.5	-7.5	1.7857
	Control	39	31.5	7.5	1.7857
No expuestos	Casos	163	155.5	7.5	.3617
	Controles	148	155.5	-7.5	.3617
Total		374			$\chi^2 = 4.294$

Tabla AB1. Resultados de la prueba chi-cuadrada

Chi - cuadrada	valor	Grados de Libertad	Significancia
Pearson	4.29490	1	.03823

De acuerdo al valor del estadístico y considerando un nivel del 95%, se ve que si es significativo (.03823), por lo tanto se rechaza la hipótesis nula.

Estadístico	Valor	ASE ¹	Significancia
Cramer's V	.10716		.03823
Coefficiente de Contingencia	.10655		.03823
Lambda			
con variable dependiente PACPREL	.08021	.09044	
Coefficiente de Incertidumbre			
con variable dependiente PACPREL	.00835	.00797	.03745
Somers' D			
con variable dependiente PACPREL	-.14316	.06742	

De los estadísticos obtenidos se puede ver que la asociación es pequeña ya que la V de Cramer y el Coeficiente de contingencia es de apenas 0.107.

Estadístico	Valor	Intervalo de confianza a un 95%
Razón de momios	.55875	.32075 .97337
Riesgo relativo para el grupo de casos	.72687	.61317 .96378
Riesgo relativo para el grupo de controles	1.37580	.98701 1.91775

Respecto a la razón de momios vemos que aparentemente es un riesgo protector, sin embargo es un dato interesante ya que el haber encontrado que el número de niños hospitalizados fue menor en el grupo de los casos, apoya la literatura¹³ en el hecho de que los pacientes al no exponerse a factores infecciosos, no hacen que su sistema inmunológico madure y pudiera ser el inicio para el desarrollo de leucemia, o puede ser también que los

¹³Van Steensel-Moll HA, Van Zanen GE. Childhood leukemia and infectious diseases in the first year of life: a register-based case-control study. 1986, p. 590-594

controles que se estudiaron son niños que han resistido mejor a cualquier reto del medio ambiente que pudiese alterar su aparato inmunológico.

- **Análisis para ver si el desarrollo de leucemia en el niño esta asociado con la exposición de la madre durante su embarazo a rayos x.**
Aquí lo que se supone es que la exposición de la madre a rayos x durante su embarazo del niño en estudio puede ser un factor de riesgo para el posible desarrollo de leucemia..

H_0 = La exposición de la madre a rayos x durante su embarazo del niño en estudio no es un factor de riesgo para el posible desarrollo de leucemia.
 H_1 = La exposición de la madre a rayos x durante su embarazo del niño en estudio es un factor de riesgo para el posible desarrollo de leucemia.

Tabla AC1. Datos del estudio para examinar la asociación entre la ocurrencia de leucemia en el niño y la exposición de la madre

KAMADSE. BY PACTRIL	casos	controles	total renglón
Expuestos	27	18	45
No expuestos	160	169	329
total columna	187	187	374

Tabla AC2. Cálculo de la estadística chi cuadrada para los datos

Factor de riesgo Rayos x	Grupo	Valor observado (o)	Valor esperado (e)	Valor observado menos esperado (o - e)	Contribución a χ^2 [(o - e) ² / e]
Expuestos	Casos	27	22.8	4.5	.9
	Control	18	22.5	-4.5	.9
No expuestos	Casos	160	164.5	-4.5	.1231
	Controles	169	164.5	4.5	.1231
Total					$\chi^2 = 2.046$

Tabla AC3. Resultado de la prueba chi cuadrada

Chi - cuadrada	valor	Grados de Libertad	Significancia
Pearson	2.04620	1	.15259

El valor del estadístico es 2.04620, considerando un nivel del 95 %, no es significativo para las muestras, por lo tanto se rechaza la hipótesis nula.

De la tabla AC2 podemos ver que aunque el número total de expuestos sólo representa el 12% del total del total de niños, de los 45 expuestos el 60 % desarrollo leucemia.

Estadístico	Valor	Intervalo de confianza a un 95%
Razón de momios	1.58438	.84014 2.98788
Riesgo relativo para el grupo de casos	1.28419	.88441 1.86469
Riesgo relativo para el grupo de controles	.81054	.62301 1.05452

Si medimos el riesgo en términos de la razón de momios podemos ver que el valor es 1.5843, es decir que una persona expuesta a rayos x durante su embarazo tiene un riesgo mayor de que el niño pueda desarrollar leucemia.

- **Análisis para ver si el desarrollo de leucemia esta asociado con la exposición de la madre a ultrasonidos durante su embarazo.**

Lo que se supone es que la exposición de la madre a ultrasonidos durante su embarazo del niño en estudio puede ser un factor de riesgo para el posible desarrollo de leucemia.

H_0 = La exposición de la madre a ultrasonidos durante su embarazo del niño en estudio no es un factor de riesgo para el posible desarrollo de leucemia.

H_1 = La exposición de la madre a ultrasonidos durante su embarazo del niño en estudio es un factor de riesgo para el posible desarrollo de leucemia.

Tabla AD1. Datos del estudio para examinar la asociación entre la ocurrencia de leucemia en el niño y la exposición de la madre a ultrasonidos.

ULMADNE by PACPRIL	casos	controles	total renglón
Expuestos	39	44	83
No expuestos	148	143	329
total columna	187	187	374

Tabla AD2. Cálculo de la estadística chi cuadrada para los datos

Factor de riesgo Ultrasonidos	Grupo	Valor observado (o)	Valor esperado (e)	Valor observado menos esperado (o - e)	Contribución a χ^2 [(o - e) ² / e]
Expuestos	Casos	39	41.5	-2.5	.1506
	Control	44	41.5	2.5	.1506
No expuestos	Casos	148	145.5	2.5	.0429
	Controles	143	145.5	-2.5	.0429
Total		374			$\chi^2 = .387$

Tabla AD3. Resultado de la prueba chi cuadrada.

Chi - cuadrada	Valor	Grados de Libertad	Significancia
Pearson	.38712	1	.53382

El valor de la prueba chi-cuadrada es .38712 siendo no significativo a un nivel del 95%, por lo tanto se rechaza la hipótesis y se dice que no hay asociación entre las variables.

En términos del riesgo se tiene lo siguiente:

Estadístico	Valor	Intervalo de confianza a un 95%	
Razón de momios	.85642	.52546	1.39584
Riesgo relativo para el grupo de casos	.92698	.73368	1.17119
Riesgo relativo para el grupo de controles	1.08239	.83885	1.39663

de estos valores se ve que no hay asociación entre la exposición de la madre a ultrasonidos durante el embarazo y el posible desarrollo de leucemia.

En la sección 2.2 se encontró que la frecuencia de niños que asisten a la escuela es de 221, los cuales representan el 60 % del total de pacientes incluidos en el estudio. El interés ahora es realizar un análisis mas detallado considerando dos casos, en el primero se incluyen los 374 datos sobre la exposición a ciertos factores del niño en su domicilio mas los datos de exposición para los niños que asisten a la escuela, el objetivo del primer caso es ver cual es el comportamiento de las variables. En el segundo caso el análisis solo incluye aquellas personas que han estado siempre expuestas (expuesta en su hogar y

expuesta en su escuela), así como aquellas que nunca han estado expuestas para corroborar si de alguna manera influye el tiempo de exposición.

- **Análisis para ver si hay relación entre el desarrollo de leucemia y al exposición del menor a transformadores de luz considerando el primer caso.**
Si suponemos que no existe asociación entre la ocurrencia de leucemia y la exposición del niño a transformadores de luz se tiene lo siguiente:

H_0 = La exposición del niño a transformadores de luz no es un factor de riesgo para el desarrollo de leucemia.

H_1 = La exposición del niño a transformadores de luz es un factor de riesgo para el desarrollo de leucemia.

Tabla AE1. Datos del estudio para examinar la asociación entre la ocurrencia de leucemia y la exposición del menor a transformadores de luz

TRANSFORMADORES DE LUZ	CASOS	CONTROLES	TOTAL RENGLÓN
Expuestos	183	184	367
No expuestos	115	113	228
TOTAL COLUMNA	298	297	595

Tabla AE2. Cálculo de la estadística chi cuadrada para los datos

Factor de riesgo	Grupo	Valor observado (o)	Valor esperado (e)	Valor observado menos esperado (o - e)	Contribución a χ^2 [(o - e) ² / e]
Expuestos	Casos	183	183.8	-0.8	.0034
	Control	184	183.1	0.9	.0043
No expuestos	Casos	115	114.1	0.9	.0070
	Controles	113	113.8	-0.8	.0056
Total		595			$\chi^2 = 0.2006$

Tabla AE3. Resultado de la prueba chi cuadrada

Chi - cuadrada	valor	Grados de Libertad	Significancia
Pearson	.02006	1	.8714

En la tabla AE3 se puede ver que el nivel de significancia para el valor del estadístico obtenido no es relevante cuando $\alpha=0.05$, por lo tanto se dice que la exposición a transformadores de luz no representa un riesgo para el desarrollo de leucemia en el niño.

De acuerdo al valor de la razón de momios (0.9772) obtenido, indica que no hay riesgo para el grupo de los expuestos respecto a los no expuestos ya que es muy cercano a uno.

Estadístico	Valor	Intervalo de confianza a un 95%	
Razón de momios	.9772	0.6662	1.2904
Riesgo relativo para el grupo de casos	.9886	0.7103	1.3757
Riesgo relativo para el grupo de controles	1.011	0.7264	1.9682

- **Análisis para ver si hay relación entre el desarrollo de leucemia y la exposición del menor a cables de alta tensión considerando el primer caso.**

Si suponemos que no existe asociación entre la ocurrencia de leucemia y la exposición que del niño con cables de alta tensión se tiene lo siguiente:

H_0 = La exposición del niño a cables de alta tensión no es un factor de riesgo para el desarrollo de leucemia.

H_1 = La exposición del niño a cables de alta tensión es un factor de riesgo para el desarrollo de leucemia.

Tabla AF1. Datos del estudio para examinar la asociación entre la ocurrencia de leucemia y la exposición del menor a cables de alta tensión

CABLES DE ALTA TENSIÓN	casos	controles	total renglón
Expuestos	283	245	528
No expuestos	14	53	67
total columna	297	298	595

AF2. Cálculo de la estadística chi cuadrada para los datos

Factor de riesgo Cables de alta tensión	Grupo	Valor observado (o)	Valor esperado (e)	Valor observado menos esperado (o - e)	Contribución a χ^2 [(o - e) ² / e]
Expuestos	Casos	283	263.5	19.5	1.443
	Control	245	264.4	-19.4	1.423
No expuestos	Casos	14	33.4	-19.4	11.268
	Controles	53	33.5	19.5	11.350
Total					$\chi^2 = 25.484$

Tabla AF3. Resultado de la prueba chi-cuadrada

Chi - cuadrada	valor	Grados de Libertad	Significancia
Pearson	25.484	1	(.000)

La prueba chi-cuadrada resultó ser altamente significativa para un nivel del 95%, con un grado de libertad, por lo tanto se rechaza la hipótesis nula, esto quiere decir que la exposición del niño a cables de alta si es un factor de riesgo.

Al calcular las medidas de riesgo se tiene lo siguiente:

Estadístico	Valor	Intervalo de confianza a un 95%	
Razón de momios	4.3728	2.36817	8.0743
Riesgo relativo para el grupo de casos	2.5650	1.38912	4.73623
Riesgo relativo para el grupo de controles	.5865	.31770	1.0826

De acuerdo a los valores obtenidos se puede comentar que en el riesgo de que una persona desarrolle leucemia dado que ha estado expuesta en algún momento a cables de luz es aproximadamente 4 veces mayor que para una persona no expuesta.

- **Análisis para ver si hay relación entre el desarrollo de leucemia y la exposición del menor a subestaciones eléctricas considerando el primer caso.**

Si suponemos que no existe asociación entre la ocurrencia de leucemia y la exposición que del niño a subestaciones eléctricas se tiene lo siguiente:

H_0 = La exposición del niño a subestaciones eléctricas no es un factor de riesgo para el desarrollo de leucemia.
 H_1 = La exposición del niño a subestaciones eléctricas es un factor de riesgo para el desarrollo de leucemia.

Tabla AG1. Datos del estudio de para examinar la asociación entre la ocurrencia de leucemia y la exposición del menor a subestaciones eléctricas

SUBJETOS by PACAPREL	casos	controles	total reneión
Expuestos	29	33	62
No expuestos	269	264	533
total columna	298	297	595

AG2. Cálculo de la estadística χ^2 cuadrada para los datos

Factor de riesgo	Grupo	Valor observado (o)	Valor esperado (e)	Valor observado menos esperado (o - e)	Contribución a χ^2 [(o - e) ² / e]
Expuestos	Casos	29	31.0	-2.0	.1290
	Control	33	30.9	2.1	.1427
No expuestos	Casos	269	266.9	2.1	.0165
	Controles	264	266.0	-2.0	.0150
Total		595			$\chi^2 = .3033$

Tabla AG3. Resultado de la prueba χ^2 cuadrada

χ^2 - cuadrada	valor	Grados de Libertad	Significancia
Pearson	.3033	1	.5947

Después de haber realizado la prueba de bondad de ajuste vemos que no hay asociación entre las variables por lo que se acepta la hipótesis nula.

El valor de la razón de momios (.8624), indica que no hay riesgo significativo para el grupo de los expuestos respecto a los no expuestos.

Estadístico	Valor	Intervalo de confianza a un 95%	
Razón de momios	.8624	0.5097	1.4606
Riesgo relativo para el grupo de casos	.9269	0.5741	1.5695
Riesgo relativo para el grupo de controles	1.0789	0.6144	1.8273

• **Análisis para ver si hay relación entre el desarrollo de leucemia y al exposición del menor a torres de alta tensión considerando el primer caso.**

Si suponemos que no existe asociación entre la ocurrencia de leucemia y la exposición que del niño con torres de alta tensión se tiene lo siguiente:

H_0 = La exposición del niño con torres de alta tensión no es un factor de riesgo para el desarrollo de leucemia.
 H_1 = La exposición del niño con torres de alta tensión es un factor de riesgo para el desarrollo de leucemia.

Tabla AH1. Datos del estudio para examinar la asociación entre la ocurrencia de leucemia y la exposición del menor con torres de alta tensión

TORRES BY PACPRILL	casos	controles	total renglón
Expuestos	73	62	135
No expuestos	225	235	460
total columna	298	297	595

AH2. Cálculo de la estadística chi cuadrada para los datos

Factor de riesgo	Grupo	Valor observado (o)	Valor esperado (e)	Valor observado menos esperado (o - e)	Contribución a χ^2 [(o - e) ² / e]
Torres de alta tensión	Expuestos	73	67.6	5.4	.4313
	Control	62	67.3	-5.3	.4173
No expuestos	Casos	225	230.3	-5.3	.1219
	Controles	235	229.6	5.4	.1270
Total					$\chi^2 = 1.097$

Tabla AH3. Resultado de la prueba chi-cuadrada

Chi - cuadrada	valor	Grados de Libertad	Significancia
Pearson	1.097	1	.2963

De acuerdo a la tabla AH3, el valor no es significativo para un nivel del 95%, con un grados de libertad, por lo tanto se acepta la hipótesis nula.

Si medimos el riesgo en términos de la razón de momios el valor obtenido es 1.229, esto quiere decir que aunque el número de expuestos no es relevante respecto a cada muestra, el riesgo que una persona tiene de desarrollar leucemia cuando esta expuesta es 1.2 veces mayor que el que tendría una persona no expuesta.

Estadístico	Valor	Intervalo de confianza a un 95%
Razón de momios	1.2297	0.8369 1.8066
Riesgo relativo para el grupo de casos	1.1055	0.7524 1.6241
Riesgo relativo para el grupo de controles	0.9045	0.6156 1.3288

- **Análisis para ver si hay relación entre el desarrollo de leucemia y al exposición del menor a transformadores de luz para el segundo caso.**

Si suponemos que no existe asociación entre la ocurrencia de leucemia y la exposición del niño con transformadores de luz se tiene lo siguiente:

H_0 = La exposición del niño a transformadores de luz no es un factor de riesgo para el desarrollo de leucemia.

H_1 = La exposición del niño a transformadores de luz es un factor de riesgo para el desarrollo de leucemia.

Tabla AH1. Datos del estudio para examinar la asociación entre la ocurrencia de leucemia y la exposición del menor a transformadores de luz

TRANSFORMADORES BY PACPRILL	casos	controles	total renglón
Expuestos	42	42	84
No expuestos	28	21	49
total columna	70	63	133

Tabla A12. Cálculo de la estadística chi cuadrada para los datos.

Factor de riesgo Transformaciones de las	Grupo	Valor observado (o)	Valor esperado (e)	Valor observado menos esperado (o - e)	Contribución a χ^2 [(o - e) ² /e]
Expuestos	Casos	42	44.2	-2.2	.1095
	Control	42	39.7	2.3	.1332
No expuestos	Casos	28	25.7	2.3	.2058
	Controles	21	23.21	-2.2	.2085
Total		133			$\chi^2 = .657$

Tabla A13. Resultado de la prueba chi-cuadrada

Chi - cuadrada	valor	Grados de Libertad	Significancia
Pearson	.657	1	.4287

La tabla A13 muestra que el valor del estadístico es muy pequeño, considerando que se tiene un grado de libertad y un valor de $\alpha = .05$, el resultado de la prueba chi-cuadrada nos indica que no hay asociación entre las variables y se acepta la hipótesis nula.

Si medimos el riesgo en términos de la razón de momios el valor que se obtiene es de .750, como es menor que uno nos indica que no hay riesgo significativo para las personas expuestas, la importancia de esto es comentar que aunque las personas estuvieron siempre expuestas el riesgo sigue siendo bajo.

Estadístico	Valor	Intervalo de confianza a un 95%
Razón de momios	0.750	0.36902 2.70959
Riesgo relativo para el grupo de casos	0.875	0.43053 1.77837
Riesgo relativo para el grupo de controles	1.142	0.56190 2.32095

- **Análisis para ver si hay relación entre el desarrollo de leucemia y la exposición del menor a cables de alta tensión para el segundo caso**

Si suponemos que no existe asociación entre la ocurrencia de leucemia y la exposición que el niño con cables de alta tensión instalados cerca de la escuela a la que asiste se tiene lo siguiente:

H_0 = La exposición del niño a cables de alta tensión no es factor de riesgo para el desarrollo de leucemia.

H_1 = La exposición del niño a cables de alta tensión es un riesgo para el desarrollo de leucemia.

Tabla A11. Datos del estudio para examinar la asociación entre la ocurrencia de leucemia y la exposición del menor a cables de alta tensión

CABLES BY PACIFIC	controles	casos	total renglón
Expuestos	99	86	185
No expuestos	1	10	11
total columna	100	96	196

AJ2. Cálculo de la estadística chi cuadrada para los datos

Factor de riesgo Categoría de alta tensión	Grupo	Valor observado (o)	Valor esperado (e)	Valor observado menos esperado (o - e)	Contribución a χ^2 [(o - e) ² / e]
Expuestos	Casos	99	94.3	4.7	.2342
	Control	86	90.6	-4.6	.2335
No expuestos	Casos	1	5.6	-4.6	3.778
	Controles	10	5.3	4.7	4.167
Total		196			$\chi^2 = 8.4132$

Tabla AJ3. Resultado de la prueba chi-cuadrada

Chi - cuadrada	valor	Grados de Libertad	Significancia
Pearson	8.4132	1	.00356

Después de realizar la prueba de bondad de ajuste vemos que el valor del estadístico es 8.4132, como es mayor que el que se tiene en la tabla de la distribución chi-cuadrada con un grado de libertad a un nivel del 95%, se dice que el resultado obtenido es significativo para las muestras, esto nos conduce a rechazar la hipótesis nula.

Si calculamos el valor de la razón de momios se obtiene un riesgo bastante significativo para las personas expuestas.

Estadístico	Valor	Intervalo de confianza a un 95%
Razón de momios	11.51162	10.5112 91.7615
Riesgo relativo para el grupo de casos	5.8864	.73845 46.921
Riesgo relativo para el grupo de controles	.51135	.06414 1.03532

- **Análisis para ver si hay relación entre el desarrollo de leucemia y al exposición del menor a subestaciones eléctricas para el segundo caso.**

Si suponemos que no existe asociación entre la ocurrencia de leucemia y la exposición que el niño con subestaciones eléctricas instalados cerca de la escuela a la que existe se tiene lo siguiente:

H_0 = La exposición del niño a subestaciones eléctrica no es un factor de riesgo para el desarrollo de leucemia.

H_1 = La exposición del niño a subestaciones eléctricas es un factor de riesgo para el desarrollo de leucemia.

Tabla AK1. Datos del estudio para examinar la asociación entre la ocurrencia de leucemia y la exposición del menor a subestaciones eléctricas

SUBESTO DE TENSIÓN	CASOS	CONTROLES	TOTAL REGLÓN
Expuestos	3	5	8
No expuestos	97	90	187
total columna	100	95	195

Tabla AK2. Cálculo de la estadística chi cuadrada para los datos

Factor de riesgo	Grupo	Valor observado (o)	Valor esperado (e)	Valor observado menos esperado (o - e)	Contribución a χ^2 [(o - e) ² / e]	
Expuestos a subestaciones eléctricas	Casos	3	4.1	-1.1	.2951	
	Control	5	3.8	1.2	.3789	
No expuestos	Casos	97	95.8	1.2	.0150	
	Controles	91	91.1	-1.1	.0132	
Total					195	$\chi^2 = .7022$

Tabla AK3. Resultado de la prueba chi-cuadrada

Chi - cuadrada	valor	Grados de Libertad	Significancia
Pearson	.7022	1	.22878

En la tabla AK3 se puede ver que el nivel de significancia para el valor del estadístico obtenido no es relevante para $\alpha=.05$, por lo tanto se dice que la exposición a subestaciones eléctricas no representa un riesgo para el desarrollo de leucemia en el niño.

Según el valor de la razón de momios (0.556), podríamos decir que aparentemente el riesgo es mayor para el grupo de los no expuestos, de la tabla AK1 se ve que la frecuencia es mayor en el grupo de los no expuestos (187), esto indica que la enfermedad pudo darse por otros factores y no por la exposición a subestaciones eléctricas.

Estadístico	Valor	Intervalo de confianza a un 95%	
Razón de momios	0.556	0.12927	2.39599
Riesgo relativo para el grupo de casos	0.7229	0.16791	3.11055
Riesgo relativo para el grupo de controles	1.383	0.32143	5.95005

- **Análisis para ver si hay relación entre el desarrollo de leucemia y al exposición del menor a torres de alta tensión para el segundo caso.**

Si suponemos que no existe asociación entre la ocurrencia de leucemia y la exposición que el niño con torres de alta tensión instalados cerca de la escuela a la que asiste se tiene lo siguiente:

H_0 = La exposición del niño con torres de alta tensión no es un factor de riesgo para el desarrollo de leucemia.

H_1 = La exposición del niño con torres de alta tensión es un factor de riesgo para el desarrollo de leucemia.

Tabla AL1. Datos del estudio para examinar la asociación entre la ocurrencia de leucemia y la exposición del menor con torres de alta tensión.

TORRES DE ALTA TENSIÓN	CASOS	CONTROLES	total renglón
Expuestos	19	10	29
No expuestos	73	74	147
total columna	92	84	176

AL2. Cálculo de la estadística chi cuadrada para los datos

Factor de riesgo Tiempo de exposición	Grupo	Valor observado (o)	Valor esperado (e)	Valor observado menos esperado (o - e)	Contribución a χ^2 [(o - e) ² / e]
Expuestos	Casos	19	15.1	3.9	1.007
	Control	10	13.8	-3.8	1.046
No expuestos	Casos	73	76.8	-3.8	.1880
	Controles	74	70.1	3.9	.2169
Total		221			$\chi^2 = 2.457$

Tabla AL3. Resultado de la prueba chi-cuadrada

Chi - cuadrada Pearson	valor	Grados de Libertad	Significancia
	2.457	1	.1106

De acuerdo a la tabla AL3, el valor no es significativo para un nivel del 95%, con un grado de libertad, por lo tanto se acepta la hipótesis nula. Al calcular la razón de momios el valor obtenido es 1.926, esto quiere decir que aunque el número de expuestos no es relevante respecto a cada muestra, el riesgo que una persona tiene de desarrollar leucemia tiende a incrementarse con el tiempo, en este caso vemos que el riesgo de desarrollar leucemia es 1.9 veces mayor para una persona que ha estado expuesta siempre respecto a una persona que nunca ha estado expuesta.

Estadístico	Valor	Intervalo de confianza a un 95%
Razón de momios	1.926	0.83853 - 4.41982
Riesgo relativo para el grupo de casos	1.3193	0.57464 - 3.02890
Riesgo relativo para el grupo de controles	0.7579	0.33011 - 1.74002

• **Análisis para ver si el tipo de leucemia es independiente del sexo.**

Tomando en cuenta únicamente la muestra de los casos, lo que nos interesa es ver si existe alguna relación entre el tipo de leucemia desarrollada y el sexo del paciente. Por lo tanto si suponemos que no existe asociación entre el tipo de leucemia y el sexo se tiene lo siguiente:

- H₀ = El tipo de leucemia es independiente del sexo.
- H₁ = El tipo de leucemia no es independiente del sexo.

Tabla AM1. Datos de la muestra de casos para examinar la asociación entre el tipo de leucemia respecto al sexo del paciente.

SEXO DEL PACIENTE	Tipo 1	Tipo 2	total renglón
Masculino	94	11	105
Femenino	73	9	82
total columna	167	20	187

Tabla AM2. Cálculo de la estadística chi cuadrada para los datos

Factor de riesgo Sexo	Grupo	Valor observado (o)	Valor esperado (e)	Valor observado menos esperado (o - e)	Contribución a χ^2 [(o - e) ² / e]
Masculino	1	94	93.8	.2	.0004
	0	11	11.2	-.2	.0035
Femenino	1	73	73.2	-.2	.0005
	0	9	8.8	.2	.0045
Total		187			$\chi^2 = .012$

Tabla AM3. Resultado de la prueba chi-cuadrada

Chi - cuadrada	valor	Grados de Libertad	Significancia
Pearson	.01202	1	.91269

Después de realizar la prueba chi-cuadrada se ve que no hay asociación entre las variables, por lo tanto se acepta la hipótesis nula.

Considerando el valor de la razón de momios vemos que es muy cercano a uno, esto nos indica que la probabilidad de desarrollar algún tipo de leucemia es la misma para ambos sexos, sin embargo es importante comentar que si se analiza la tabla AM1 se puede ver que hay una tendencia mayor para el tipo 1 en ambos sexos, se tiene que del total de casos (187), el 89.3% desarrolló leucemia linfoblástica aguda mientras que solo el 10.7 desarrolló leucemia mieloblástica, el comportamiento es similar inclusive para el sexo, es decir; para el sexo masculino el 89.52% desarrolló leucemia tipo 1, mientras que el 10.48% desarrolló leucemia tipo 2, para el sexo femenino el 89.02% desarrolló leucemia linfoblástica aguda mientras que sólo el 10.98% desarrolló leucemia mieloblástica.

Estadístico	Valor	Intervalo de confianza a un 95%
Razón de momios	1.05355	.41465 2.67689
Riesgo relativo para el grupo de casos	1.00560	.45582 2.40803
Riesgo relativo para el grupo de controles	.99442	.89953 1.09932

Al analizar el tipo de leucemia para la muestra de casos, se vio que la incidencia es mayor para el tipo1, lo que ahora se desea conocer es si hay alguna relación entre el tipo de leucemia y el municipio al que pertenece el niño, estos datos nos darían una pista acerca de que factores pueden influir en el niño.

Tabla AN1. Datos de la muestra de casos para examinar si hay relación entre el tipo de leucemia y el municipio al que pertenece el niño leucémico.

TIPLEU by DUMIC	01	02	03	04	05	06	07	08
Tipo1	9	11	5	10	5	12	2	16
Tipo2	1	0	0	1	1	4	0	3
total columna	10	11	5	11	6	16	2	19

(continuación tabla AN1)

TIPLEU by DUMIC	09	10	11	12	13	14	15	17
Tipo1	2	5	6	1	2	1	2	18
Tipo2	0	0	1	0	0	1	0	1
total columna	2	5	7	1	2	2	2	19

(continuación tabla AN1)

TIPLEU by DUMIC	18	19	20	21	22	23	24	31
Tipo1	12	8	7	12	4	1	2	0
Tipo2	1	0	1	2	0	0	0	1
total columna	13	8	8	14	4	1	2	1

(continuación tabla AN1)

TIPLEU by DUMIC	33	34	35	36	37	39	46	47	48	total renglón
Tipo1	1	4	2	1	1	1	1	0	3	168
Tipo2	0	0	0	0	1	0	0	1	0	20
total columna	1	4	2	1	2	1	1	1	3	187

Si analizamos la tabla anterior en términos de frecuencias, podemos ver que hay algunos municipios en donde la frecuencia es alta respecto a los demás, por ejemplo para el municipio de Alvaro Obregón se tienen 10 casos (5.3%), de los cuáles 9 desarrollaron leucemia tipo1 y sólo uno tipo2, en el municipio de Azcapotzalco encontramos que se registraron 11 casos (5.9%), de estos los 11 corresponden al tipo1, para el municipio de Benito Juárez se tienen 5 casos (2.6%) todos ellos para el tipo1, el municipio de Coyacoacán registró 11 casos (5.9%), en el municipio de Gustavo A. Madero se tienen 16 casos (8.6%), en el caso de Iztapalapa también se encontró una frecuencia considerable (10.16%), para el municipio de Tlalpan es importante comentar que todos los casos observados (5), corresponde al tipo1, en el municipio de Ecatepec también se encuentra una frecuencia alta (10.16%), en el municipio de Naucalpan se tienen 13 casos (6.7%), siendo 12 para el tipo1 y sólo uno para el tipo 2, otros de los municipios con frecuencias más o menos considerables son Cuahuitlán Izcalli y Netzahualcoyotl con 8 casos cada uno, en el municipio de Tlalnepantla se registraron 14 casos (7.5%), en los municipios de Valle de Aragón y Atizapan de Zaragoza se tienen 4 casos en cada uno.

Los municipios que se mencionaron son los que se consideran más importantes por el número de casos registrados, sin embargo es importante comentar que aunque los demás municipios representan una frecuencia más baja, en la mayoría el desarrollo de leucemia linfoblástica aguda es más común. Es importante comentar que no se realizó la prueba de la chi-cuadrada dado que se tienen muchos casos donde la frecuencia es muy baja, y en estos casos el dato representa el 100 por ciento para el municipio respectivo.

Para finalizar con el análisis bivariado se analizará la variable "CIALUZ", considerada como la más importante para el estudio ya que nos permitirá conocer el nivel de exposición a campos electromagnéticos.

La importancia de ésta variable radica en el hecho de que fue medida por personal experto de la Comisión Federal de Electricidad, cegado a conocimiento de que la medición realizada correspondía a un "caso" o "control", en otras palabras se ignoraba el estado de salud del paciente. Otro aspecto considerado es que se validó el esquema de medición para determinar el nivel de exposición dependiendo del tipo de cable y de la distancia de la casa a las líneas eléctricas, dado que la exposición depende inclusive del tipo de vivienda y nivel de construcción respecto a los cables.

Por estudios ya realizados¹⁴ se sabe que para una casa cuyo techo es de loza los campos tienden a ser rechazados más fácilmente que para una casa cuyo techo es de lámina o cartón, se considera además que una casa cuyo nivel de construcción está bajo suelo o raíz de suelo tiene un nivel de exposición diferente respecto a una cuyo nivel de construcción está sobre suelo. Los tres niveles de exposición considerados son los siguientes:

- A = Exposición alta
- M = Exposición media
- B = Exposición baja

¹⁴Neglia ME. Examine role of electromagnetic fields in childhood leukemia. 1990, p.p. 175-176

En primera instancia se tomaron en cuenta los tres niveles y después se dicotomizó la variable.

• **Análisis para ver si el nivel de exposición a campos electromagnéticos esta asociado con la leucemia en el niño.**

Considerando que no hay asociación entre la exposición a campos electromagnéticos y la ocurrencia de leucemia se tiene lo siguiente:

H_0 = El nivel de exposición a campos electromagnéticos no es un factor de riesgo para el desarrollo de leucemia

H_1 = El nivel de exposición a campos electromagnéticos es un factor de riesgo para el desarrollo de leucemia

Tabla AO1. Datos del estudio para examinar la asociación entre el nivel de exposición a campos electromagnéticos y la ocurrencia de leucemia en el niño

Categoría de Pacientes	Casos	controles	total renglón
Alta	82	65	147
Media	90	102	192
Baja	15	20	35
total columna:	187	187	374

Tabla AO2. Cálculo de la estadística chi cuadrada para los datos

Exposición a Campos electromagnéticos	Grupo	Valor observado (o)	Valor esperado (e)	Valor observado menos esperado (o - e)	Contribución a χ^2 $[(o - e)^2 / e]$
Alta	Casos	82	73.5	8.5	.9829
	Controles	65	73.5	-8.5	.9829
Media	Casos	90	96.0	-6.0	.3750
	Controles	102	96.0	6.0	.3750
Baja	Casos	15	17.5	-2.5	.3571
	Controles	20	17.5	2.5	.3571
Total		374			$\chi^2 = 3.4302$

Tabla AO3. Resultado de la prueba chi cuadrada

Chi - cuadrada	valor	Grados de Libertad	Significancia
Pearson	3.4302	2	.1799

La prueba de bondad de ajuste nos indica que para un nivel del 95% con dos grados de libertad no es significativa, por lo tanto se acepta la hipótesis nula. Si comparamos los riesgos de ocurrencia de la enfermedad en los diferentes niveles de exposición tenemos lo siguiente:

- Riesgo de ocurrencia de la enfermedad para las personas que tuvieron exposición alta respecto a los de baja exposición 1.682.
- Riesgo de ocurrencia de la enfermedad para las personas que tuvieron exposición media respecto a los de baja exposición 1.176.

Después de comparar los niveles vemos que en el primer caso se tiene un riesgo mayor para los de alta exposición, mientras que en el segundo caso el riesgo para los de exposición media respecto a los de baja no es muy significativo, esto quiere decir que el riesgo de ocurrencia de leucemia tiende a ser mayor cuando la exposición a campos

electromagnéticos es alta, para probar esto se dicotomizó la variable de exposición considerando a todos los de exposición media y baja, como MB (exposición media-baja)

- **Análisis para ver si el nivel de exposición a campos electromagnéticos esta asociado con la leucemia en el niño después de dicotomizar la variable Cialuz.**
Considerando que no hay asociación entre la exposición a campos electromagnéticos y la ocurrencia de leucemia se tiene lo siguiente:

H_0 = El nivel de exposición a campos electromagnéticos no es un factor de riesgo para el desarrollo de leucemia
 H_1 = El nivel de exposición a campos electromagnéticos es un factor de riesgo para el desarrollo de leucemia

Tabla AP1. Datos del estudio para examinar la asociación entre el nivel de exposición a campos electromagnéticos y la ocurrencia de leucemia en el niño.

	Casos	Controles	total renglón
Alta	82	65	147
Media-baja	105	122	227
total columna	187	187	374

Tabla AP2. Cálculo de la estadística chi cuadrada para los datos

Exposición a Campos electromagnéticos	Grupo	Valor observado (o)	Valor esperado (e)	Valor observado menos esperado (o - e)	Contribución a χ^2 [(o - e) ² / e]
Alta	Casos	82	73.5	8.5	.9829
	Controles	65	73.5	-8.5	.9829
Media-baja	Casos	105	113.5	-8.5	.6365
	Controles	122	113.5	8.5	.6365
Total					$\chi^2 = 3.2391$

Tabla AP3. Resultados de la prueba chi-cuadrada

Chi - cuadrada	valor	Grados de Libertad	Significancia
Pearson	3.2391	1	.07190

El valor del estadístico para un intervalo de confianza del 95% no resultó ser significativo, sin embargo si $\alpha=0.1$, entonces el resultado obtenido si sería significativo. Por otro lado al obtener la razón de momios de la enfermedad vemos que después de dicotomizar la variable se conserva la relación de riesgo para los de alta exposición, siendo un riesgo de casi 1.5 veces mayor respecto a los de exposición media-baja.

Estadístico	Valor	Intervalo de confianza a un 95%
Razón de momios	1.465	0.9653 2.223
Riesgo relativo para el grupo de exposición alta	1.205	0.7940 1.828
Riesgo relativo para el grupo de exposición media	.8227	0.5421 1.248

Como punto final de este trabajo se realizará un análisis multivariado utilizando regresión logística, es por ello que los primeros incisos del siguiente capítulo están enfocados a definir y describir las características del modelo logístico, ya que esto nos permitirá interpretar los resultados obtenidos del análisis multivariado.

REGRESIÓN LOGÍSTICA

4. REGRESIÓN LOGÍSTICA

4.1 INTRODUCCIÓN AL ANÁLISIS MULTIVARIADO

Se define al análisis multivariado como el conjunto de técnicas estadísticas que analizan simultáneamente más de dos variables en una muestra de observaciones. Para algunos autores, el análisis multivariado es la rama de la estadística y del análisis de datos que estudia, interpreta y elabora el material estadístico sobre la base de un conjunto de $n > 1$ variables, que pueden ser de tipo cuantitativo, cualitativo o una mezcla de ambos.

Uno de los objetivos fundamentales del análisis multivariado, consiste en resumir y sintetizar grandes conjuntos de datos y variables, en función de ciertos objetivos, de cara a obtener informaciones válidas que permitan una mejor comprensión del fenómeno objeto de estudio.

4.1.1 Orígen

El origen del análisis multivariado se remonta a las primeras generalizaciones de la correlación y regresión. Los pioneros de la Estadística moderna fueron ingleses (Galton, Pearson, Fisher, Student)¹⁴. Posteriormente el centro de gravedad se desplazó a los Estados Unidos (Hotelling, Wilks, Bartlett). Pero también surgen nuevas escuelas como la India (Mahalanobis, Roy, Krishnaiah, etc.).

En los años 60 empiezan a utilizarse las computadoras en el de análisis de datos. El proceso mecánico de cálculo deja de tener importancia y empieza a aplicarse el análisis multivariado a la Psicología, Biología, Medicina etc.. En nuestro país la utilización de las computadoras en la investigación empírica se inicia a principios de los 70 y no es sino hasta los años 80 cuando la informática pasa a ser de dominio público.

4.1.2 Recomendaciones previas para el estudio del análisis multivariado

Para que el análisis multivariado tenga sentido se recomienda hacer uso de lo siguiente:

1. Análisis exploratorio de los datos: Su objetivo es la comprensión de cada una de las variables por separado, y puede realizarse mediante la estadística descriptiva univariada, como medidas de tendencia central, tablas de frecuencias, asimetría, representaciones gráficas etc..
2. Estadística bivariada: Su objetivo consiste en buscar la relación entre pares de variables.
3. Análisis multivariado: Su objetivo consiste en analizar simultáneamente muchas variables.

¹⁴Bizquerria Alzina. Introducción conceptual al análisis multivariable. P.p. 3-5

4.1.3 Clasificación de los métodos

Diversos autores han propuesto una clasificación de los métodos multivariados. Sin embargo dada su diversidad y los enfoques clasificatorios, la labor se hace difícil. Tomando las aportaciones de algunos autores como Shet, Kinnear y Taylor, Dillon y Goldstein¹⁵, se presenta en la figura 1 la clasificación general de estos métodos.

Los criterios clasificatorios son:

1. Métodos descriptivos o explicativos.
2. Número de variables.
3. Escalas de medición.

En esta clasificación se incluyen los principales métodos de análisis estadístico, y puede considerarse como un intento de algoritmo conducente a la determinación del método más adecuado para las características del problema objeto de estudio.

En la figura 2 se presenta otro tipo de clasificación, tomando como primer criterio el propósito del análisis.

¹⁵ Bizquerria Alzina. Introducción conceptual al análisis multivariable, p.p. 5-15

Clasificación de los Métodos Multivariados

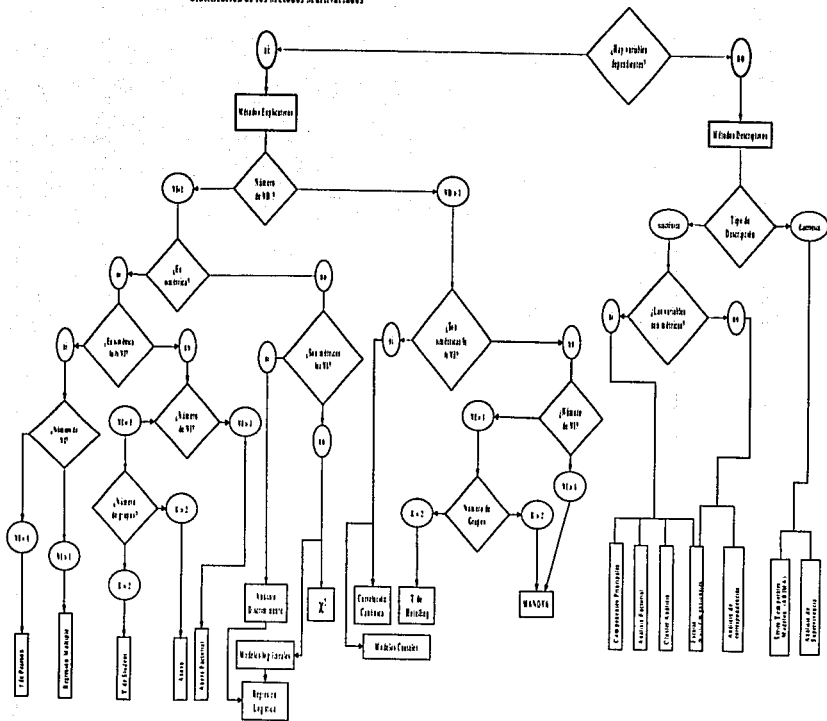


Figura 1

ESTA TESIS NO DEBE
SALIR DE LA BIBLIOTECA

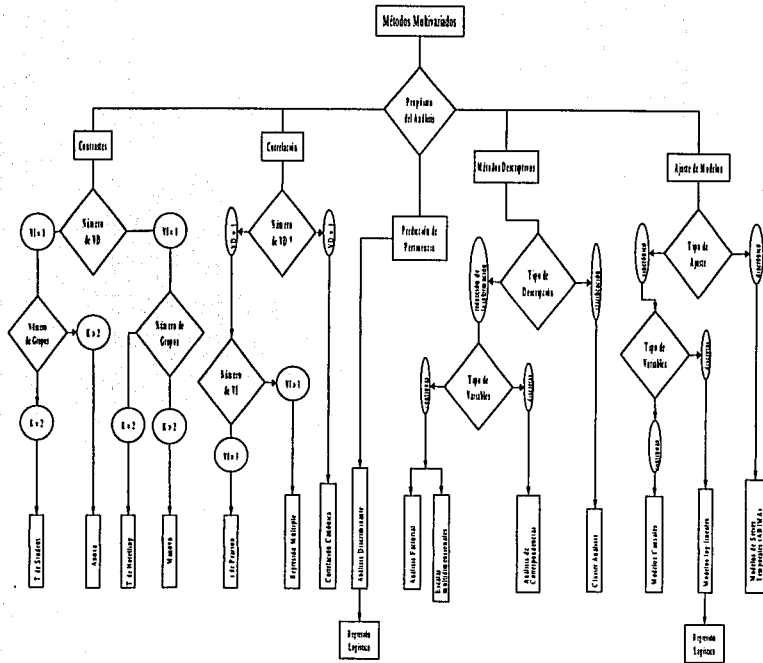


Figura 2. Clasificación de los métodos multivariados según el propósito

Dentro de la estadística se pueden distinguir dos grandes vertientes: Estadística Matemática y Estadística Aplicada. Se entiende por estadística aplicada aquella parte que se ocupa de las aplicaciones de la estadística a las investigaciones de carácter empírico. Deriva de la estadística matemática, pero no suele ser objeto de estudio de los matemáticos. Son los investigadores empíricos los que se han preocupado de la estadística aplicada. Por ejemplo, las investigaciones de carácter empírico en Psicología, Economía, Medicina, Biología, etc., suelen recurrir a métodos estadísticos aplicados en la fase de análisis de datos.

La estadística aplicada a la investigación empírica ha recibido un impulso revolucionario con la generalización del uso de computadoras. La informática realiza el proceso mecánico de cálculo. Así, análisis estadísticos que antes eran difíciles, se pueden realizar con cierta facilidad y rapidez mediante el uso de paquetes estadísticos.

La estadística aplicada se vale de los análisis más adecuados para cada investigación, de la forma de realizarlos mediante paquetes de programas estadísticos y de la interpretación de resultados, sin preocuparse demasiado del proceso mecánico de cálculo ni de la demostración de las fórmulas matemáticas. Dentro de la estadística aplicada se pueden distinguir dos características importantes:

- Utilización y manejo de paquetes de programas estadísticos.
- Interpretación de resultados.

La estadística aplicada cobra sentido cuando se considera en la fase de análisis de datos dentro del proceso de investigación empírica. Los cálculos estadísticos hoy en día se realizan por computadora mediante el uso de paquetes estadísticos. Los más conocidos son el SPSS-X, BMDP, BOX-JENKINS, LISREL, SAS, SPLUS, STATA, EGRET, pero existen muchos más como el ALICE, LINPACK, MATH-PACK etc.

4.2 Modelación estadística

El objetivo básico del modelado es obtener una representación matemática de las relaciones entre una variable respuesta observada y un número de variables explicativas, así como la medición de la incertidumbre inherente de algunas relaciones. Los modelos estadísticos contruidos para la variable respuesta son una aproximación representativa de la manera en la cual algunas variables observables dependen de otras variables. Algunos modelos serán más apropiados que otros, pero para algún grupo de datos, habrá un número de modelos que posiblemente satisfacen de manera similar el objetivo del estudio y las bases para seleccionar algún modelo del resto, no descansa sólo en la estadística.

Los modelos estadísticos son esencialmente descriptivos, pero si los modelos están basados en datos experimentales u observacionales, se describen mejor como modelos empíricos, por consiguiente la estructura de un modelo empírico dependerá de los datos.

Un modelo estadístico ampliamente usado es aquel donde la variable respuesta es expresada como la suma de dos componentes:

Variable respuesta = componente sistemático + componente aleatorio

El componente sistemático resume la variabilidad en la variable respuesta de una unidad bajo estudio debido a los valores de ciertas variables o los niveles de ciertos factores medidos para esta unidad. La variación residual es definida como el componente aleatorio, este componente contabiliza la magnitud de desviación de la respuesta observada respecto a la media o valor esperado, descrita por la parte sistemática del modelo.

4.2.1 Modelos lineales

Cuando una variable respuesta es continua, los modelos para n observaciones y_1, y_2, \dots, y_n son a menudo de la forma:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ki} + \varepsilon_i, \quad \text{ec. (4.1)}$$

esta ecuación muestra en realidad que la respuesta para la i -ésima observación depende linealmente de los valores de las k variables explicativas llamadas x_1, x_2, \dots, x_k , y de los parámetros desconocidos $\beta_0, \beta_1, \dots, \beta_k$. El término ε_i es una variable aleatoria no observable que representa la variación residual, asumiendo que tiene media cero, implica que toda la variación sistemática en la variable respuesta ha sido explicada. Suponiendo que la media y la varianza de ε_i están dadas por:

$$E(\varepsilon_i) = 0 \text{ y } \text{Var}(\varepsilon_i) = \sigma^2$$

respectivamente, para $y=1, 2, \dots, n$, el valor esperado de la i -ésima observación y_i , es:

$$E(y_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ki} \quad \text{ec. (4.2)}$$

y la varianza de esta observación es $\text{Var}(y_i) = \sigma^2$. Estrictamente hablando, esta es una variable aleatoria Y_i asociada con la i -ésima observación y_i . Si se asume que el componente aleatorio está distribuido normalmente con media 0 y varianza σ^2 , escrita como $N(0, \sigma^2)$, entonces como consecuencia, y_i tendrá una distribución $N(E(y_i), \sigma^2)$, donde cada una de las n observaciones tiene una media diferente que depende de los valores de las k variables explicativas x_{ij} , $j=1, 2, \dots, k$, para esta observación.

El lado derecho de la expresión para $E(y_i)$, es el componente lineal sistemático del modelo, denotado generalmente por η_i . En general, $\eta_i = \sum_{j=0}^k \beta_j x_{ij}$, donde $x_{i0} = 1$, y el modelo puede ser escrito como $y_i = \eta_i + \varepsilon_i$. Un ejemplo particular es el método de regresión lineal simple $y_i = \beta_0 + \beta_1 x_{i1} + \varepsilon_i$, donde el componente lineal sistemático del modelo es $\eta_i = \beta_0 + \beta_1 x_{i1}$.

Un modelo lineal puede también incluir términos correspondientes a variables cualitativas conocidas como factores, los cuales pueden tomar un número determinado de valores conocidos como los niveles del factor. Un modelo lineal puede incluir un término α_j , $j=1, 2, \dots, m$, representando el efecto del j -ésimo nivel del factor A sobre la respuesta

esperada. Los valores α_i se conocen como los efectos principales del factor. Los modelos que contienen factores pueden ser expresados como una combinación lineal de las variables explicativas o indicadoras.

• **Uso de variables indicadoras**

A menudo es necesario incluir términos mezclados en un modelo para resolver situaciones donde el coeficiente de una variable explicativa varía de acuerdo a los niveles del mismo factor. Los modelos que contienen términos que son combinaciones de factores y variables se encuentran generalmente, cuando en un modelo de regresión, hay comparación de relación entre grupos de individuos.

Suponga que hay seis pares de observaciones $(x_i, y_i) = 1, 2, \dots, 6$, de los cuales los primeros tres son observados en el primer nivel del factor A, y los siguientes tres en el segundo nivel de A. Si la pendiente de una línea recta entre "x" y "y" depende del nivel de A, la variable explicativa x es incluida en el modelo con un coeficiente diferente para cada nivel de A, esto se logra mediante la definición de dos variables indicadoras u_1 y u_2 , como se muestra en la siguiente tabla:

observación	y	x	nivel de A	u_1	u_2	u_1x	u_2x
1	y_1	x_1	1	1	0	x_1	0
2	y_2	x_2	1	1	0	x_2	0
3	y_3	x_3	1	1	0	x_3	0
4	y_4	x_4	2	0	1	0	x_4
5	y_5	x_5	2	0	1	0	x_5
6	y_6	x_6	2	0	1	0	x_6

Tabla 1

los productos u_1x y u_2x en las dos columnas finales se deben incluir como términos en el modelo. Los coeficientes de estos dos términos son los coeficientes requeridos de x en cada nivel del factor A.

La elección de variables indicadoras significa que cuando se incluye un término constante en el modelo, junto con un término tal como α_j ; o cuando dos o más factores están presentes, no es posible obtener un único estimador de todos los parámetros en el modelo. Por ejemplo considere el modelo que incluye una constante y el término $\alpha_j = 1, 2, \dots, m$, representando el efecto del j-ésimo nivel de un factor con m niveles. Este modelo contiene mas términos para el factor α_j , que el número disponible de grados de libertad, se dice entonces que el modelo esta sobrep parametrizado. Esta característica puede resolverse definiendo justamente m-1 variables indicadoras, x_2, x_3, \dots, x_m como se ve en la tabla 2.

nivel del factor	x_2	x_3	x_4	...	x_m
1	0	0	0	...	0
2	1	0	0	...	0
3	0	1	0	...	0
4	0	0	1	...	0
m	0	0	0	...	1

tabla 2

El término α_j es entonces equivalente a la inclusión de x_2, x_3, \dots, x_m en el modelo con coeficientes $\alpha_1, \alpha_2, \dots, \alpha_m$.

4.2.2 Ajuste de modelos lineales para datos binomiales

En el caso especial de datos binomiales, se sabe que una distribución apropiada para la i -ésima observación y_i , es la distribución binomial con parámetros n_i, p_i , denotada por $B(n_i, p_i)$ donde, p_i representa la probabilidad de éxito y n_i se define como el denominador binomial. Bajo esta consideración el valor esperado de la i -ésima observación es $E(y_i) = n_i p_i$ y la varianza se define como $\text{var}(y_i) = n_i p_i (1 - p_i)$. En vez de modelar directamente la dependencia de $E(y_i)$ sobre las variables explicativas, es más importante explorar como la probabilidad de éxito $p_i = E(y_i/n_i)$ puede ser descrita por las variables explicatorias observadas.

Una forma de modelar datos binomiales es el análisis de regresión lineal, apoyandose en software estadístico, en este caso se utiliza el modelo $p_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ik}$ y se aplica el método de los mínimos cuadrados para obtener los valores de los parámetros desconocidos $\beta_0, \beta_1, \dots, \beta_k$.

$$\sum_i \left(\frac{y_i}{n_i} - p_i \right)^2 = \sum_i (\hat{p}_i - \beta_0 - \beta_1 x_{i1} - \dots - \beta_k x_{ik})^2 \quad (\text{ec. 4.3})$$

la ecuación 4.3 minimiza la suma de las desviaciones cuadradas de los valores esperados de las observaciones. Existen sin embargo algunos problemas que se tienen que considerar para el caso de datos binomiales, una dificultad es que no se puede afirmar que la variable respuesta está distribuida como una normal, sin embargo esta restricción no es severa, en vista de que la distribución binomial tiende a una normal cuando el tamaño de las muestras es grande. Otra dificultad de importancia se refiere a los valores ajustados \hat{p}_i bajo el modelo, en este caso no se tiene ninguna restricción para el valor de los parámetros β_j , y pueden tomar algún valor positivo o negativo, grande o pequeño, así que la combinación lineal de ellos puede en un principio caer en el rango $(-\infty, \infty)$, lo que implica que al obtener las probabilidades ajustadas mediante $\hat{p}_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ik}$, puede no garantizarnos

que los valores estén en el intervalo (0,1). Un procedimiento que se utiliza para transformar la escala de probabilidad de (0,1) a $(-\infty, \infty)$ es la transformación logística, ésta asegura que las probabilidades ajustadas estarán entre cero y uno.

4.2.3 La transformación logística

La transformación logística de una probabilidad de éxito p es $\log\{p/(1-p)\}$, y puede ser escrito como $\text{logit}(p)$. Recordando que $p/(1-p)$ es la razón de momios de un éxito, entonces la transformación logística de p es el logaritmo de la razón de momios de un éxito. Por lo tanto un valor de p en el rango (0,1) corresponde a un valor de $\text{logit}(p)$ en $(-\infty, \infty)$. Cuando $p \rightarrow 0$, $\text{logit}(p) \rightarrow -\infty$; cuando $p \rightarrow 1$, $\text{logit}(p) \rightarrow \infty$, y para $p=0.5$, $\text{logit}(p) = 0$. La función $\text{logit}(p)$ es una curva sigmoideal que es simétrica alrededor de $p=0.5$ (ver figura 3), es esencialmente lineal entre $p=0.2$ y $p=0.8$ pero fuera de este rango es marcadamente no lineal.

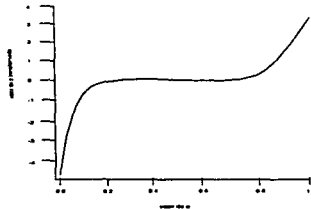


figura 3

Hay dos razones por las que la transformación logística es más utilizada que otras transformaciones. La primera es que tiene una interpretación directa en términos del logaritmo de la razón en favor de un éxito. La segunda es que los modelos basados en la transformación logística son particularmente apropiados para el análisis de datos que han sido colectados retrospectivamente tal como un estudio de casos y controles, pero lo más importante es que los parámetros son función de la estadística suficiente.

4.3 Definición del modelo logístico

Se quiere estudiar un fenómeno cuyos posibles resultados son sólo dos (éxito o fracaso, enfermedad o no enfermedad), es decir la variable a estudiar es una variable binaria. Se dispone además de ciertas variables que se supone están relacionadas con la variable de interés y que por lo tanto ayudan a explicar su comportamiento.

Supóngase que se tienen Y_1, \dots, Y_n variables aleatorias independientes que se distribuyen como Bernoulli con parámetros n_i, p_i . El objetivo es encontrar métodos que permitan establecer la dependencia de p_i con ciertas variables explicativas.

Se podría intentar modelar dichas probabilidades p_i considerando un modelo de regresión usual:

$$p_i = \sum_{k=1}^K \beta_k x_{ik}$$

donde x_{ik} es el vector de variables explicativas para la i -ésima observación.

este procedimiento tiene algunas limitaciones:

- Como la $\text{var}(y_i) = n_i p_i (1 - p_i)$, la condición de igualdad de varianzas que se requiere para aplicar mínimos cuadrados no se cumple, salvo en el caso poco interesante de que $p_i = p$ para toda y . Sin embargo como se sabe que cambios moderados en la $\text{var}(y_i)$ producen una pérdida modesta de eficiencia, si $0.2 \leq p_i \leq 0.8$, se podría aplicar el procedimiento.
- Como la distribución de las Y_i 's no es normal, ningún método de estimación que sea lineal en las Y_i 's será completamente eficiente.
- La restricción más fuerte del uso de éste método surge del requerimiento de que $0 \leq p_i \leq 1$. Si se aplican mínimos cuadrados ponderados o no, directamente a las observaciones binarias, puede ocurrir que se obtenga un vector de valores ajustados con algún componente que no satisfaga esta condición.

Una manera más simple de representar la dependencia de una probabilidad p_i , en variables explicativas, de tal forma que se cumpla la condición $0 \leq p_i \leq 1$, es mediante el siguiente modelo.

Suponiendo que se tienen n observaciones binomiales de la forma $y_i/n_i, y_i = 1, 2, \dots, n$, donde $E(y_i) = n_i p_i$ y p_i es la probabilidad de éxito correspondiente a la i -ésima observación. El modelo logístico lineal para la dependencia de p_i sobre los valores de las k variables explicativas, $x_{i1}, x_{i2}, \dots, x_{ik}$, asociados con esta observación, es:

$$\text{logit}(p_i) = \log(p_i/(1 - p_i)) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik}$$

que es equivalente a:

$$p_i = \frac{\exp(\beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ik})}{1 + \exp(\beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ik})} \quad (\text{ec 4.4})$$

si escribimos $\eta_i = \sum_j \beta_j x_{ij}$, entonces:

$$p_i = e^{n_i} / (1 + e^{n_i}) .$$

Si y_i es una observación de una distribución binomial con media $n_i p_i$, un modelo correspondiente para el valor esperado y_i es:

$$E(y_i) = n_i e^{n_i} / (1 + e^{n_i}) .$$

El nombre del modelo se debe al hecho de que a la transformación logística de p_i , se le expresa como función lineal de ciertos parámetros asociados con variables explicativas. Las variables pueden ser de cualquier tipo (continuas o discretas).

El modelo logístico lineal es miembro de la clase de modelos conocidos como modelos lineales generalizados, introducido por Nelder y Wedderburn en 1972. Esta clase de modelos también incluye el modelo lineal para datos distribuidos normalmente y el modelo log-lineal para datos poisson. La función que relaciona p con el componente lineal del modelo es conocida como la función liga, así que en el modelo logístico lineal se usa esta función.

Los modelos logísticos lineales son muy útiles para el análisis de estudios de casos y controles, ya que permiten explorar el efecto individual y conjunto de muchos factores de riesgo, controlar simultáneamente la presencia de factores de confusión y analizar interacciones. Además, los parámetros del modelo tienen una interpretación directa en términos de la razón de momios, que es la única medida del efecto que el factor de riesgo tiene en el factor respuesta estrictamente estimable a partir de estudios como estos.

Una aplicación particular del modelo logístico lineal es modelar la dependencia de una proporción respecto a una variable explicativa simple. Cuando un modelo logístico lineal es ajustado para explorar la relación entre la variable respuesta binaria y una o mas variables explicativas, el modelo se conoce como "modelo de regresión logística".

4.3.1 Ajuste del modelo logístico lineal para datos binomiales

Suponga que se tienen disponibles datos binomiales con éxito de la forma y_i en n_i pruebas, $y=1, 2, \dots, n$, donde la transformación logística de la probabilidad de éxito correspondiente p_i , se modela como una combinación lineal de k variables explicativas, $x_{1i}, x_{2i}, \dots, x_{ki}$, por lo tanto:

$$\text{logit}(p_i) = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki} .$$

las observación binomial y_i tienen media $n_i p_i$ y puede ser expresada como $y_i = n_i p_i + \epsilon_i$, el componente aleatorio $\epsilon_i = y_i - n_i p_i$ tienen media cero, pero no tiene una distribución binomial.

Para ajustar un modelo logístico a un grupo de datos, tienen que estimarse primero los $k+1$ parámetros desconocidos $\beta_0, \beta_1, \dots, \beta_k$. Estos parámetros son estimados usando el método de máxima verosimilitud. La función de verosimilitud está dado por:

$$L(\beta) = \prod_{i=1}^n \binom{n_i}{x_i} p_i^{x_i} (1 - p_i)^{n - x_i}.$$

esta función depende de las probabilidades de éxito desconocidas p_i , la cual a su vez depende de los valores β , así la función de verosimilitud puede ser considerada como una función de β . El problema es obtener los valores $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k$ los cuales maximizan $L(\beta)$ o el equivalente $\log L(\beta)$. A veces es mas apropiado usar un algoritmo conocido como el método de notación de Fisher para obtener los estimadores de máxima verosimilitud $\hat{\beta}$, este procedimiento es equivalente a usar el método iterativo de los mínimos cuadrados donde los valores para una variable dependiente ajustada $z_i = n_i + (y_i - n_i p_i) / \{n_i p_i (1 - p_i)\}$ son regresados sobre las k variables independientes, $x_{i1}, x_{i2}, \dots, x_{ik}$, usando ponderaciones $w_i = n_i p_i (1 - p_i)$.

Cuando se obtiene $\hat{\beta}$, el valor estimado del componente sistemático lineal del modelo es:

$$\hat{\eta}_i = \hat{\beta}_0 + \beta_1 x_{i1} + \hat{\beta}_2 x_{i2} + \dots + \hat{\beta}_k x_{ik}$$

llamado predictor lineal. Las probabilidades ajustadas \hat{p}_i pueden ser encontradas usando

$$\hat{p}_i = \exp(\hat{\eta}_i) / [1 + \exp(\hat{\eta}_i)]$$

• Bondad de ajuste de un modelo logístico lineal

Después de ajustar un modelo a un grupo de datos, es común hablar de la bondad del modelo para la variable respuesta. Si el ajuste entre las observaciones y los correspondientes valores es bueno, el modelo puede ser adecuado, de otra manera el modelo propuesto tendrá que ser revisado; este proceso de adecuación de un modelo es conocido como bondad de ajuste.

Hay algunos procesos estadísticos que miden la discrepancia entre las proporciones binomiales observadas y_i/n_i , y las proporciones ajustadas \hat{p}_i , bajo algún modelo considerado para cierta probabilidad de éxito p_i . De estos el mas ampliamente usado es el que se basa en la función de probabilidad del modelo.

El valor de la probabilidad puede usarse para resumir la magnitud de ajuste del modelo para la muestra de datos cuando los parámetros desconocidos son igual a la máxima verosimilitud estimada, esta máxima verosimilitud se denota por \hat{L}_n . Este estadístico no puede ser usado propiamente para estimar el ajuste del modelo, puesto que no depende del número de observaciones en la muestra. Es por tanto necesario comparar el modelo con

otros modelos alternativos basados en los mismos datos. Un segundo modelo que se utiliza para comparación, es aquel donde valores ajustados coinciden con las observaciones actuales, es decir el modelo ajusta los datos perfectamente, este modelo es llamado modelo saturado y la máxima verosimilitud se denota por \hat{L}_j .

Para comparar los dos modelos \hat{L}_c y \hat{L}_j es necesario utilizar dos veces el logaritmo de la razón de máxima verosimilitud, denotado por:

$$D = -2 \log \left(\frac{\hat{L}_c}{\hat{L}_j} \right) = -2 [\log \hat{L}_c - \log \hat{L}_j]$$

Los valores de D son grandes cuando \hat{L}_c es pequeño en relación con \hat{L}_j , indicando que el modelo actual no es bueno, por otro lado si se obtienen valores pequeños para D cuando \hat{L}_c es casi igual a \hat{L}_j , el modelo es bueno. El estadístico D mide además la magnitud de variación del modelo actual respecto al modelo saturado, esta magnitud se llama devianza.

Dentro de la modelación de n observaciones binomiales, donde p_i es la probabilidad de éxito correspondiente a la i-ésima observación y_i/n_i , $y_i=1, 2, \dots, n$, la función de probabilidad es:

$$\prod_{i=1}^n \binom{n_i}{y_i} p_i^{y_i} (1-p_i)^{n_i-y_i}$$

en el ajuste de un modelo lineal logístico con k+1 parámetros desconocidos $\beta_0, \beta_1, \dots, \beta_k$, los valores ajustados \hat{p}_i son obtenidos, de la siguiente ecuación:

$$\log \text{it}(\hat{p}_i) = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \dots + \hat{\beta}_k x_{ik} \quad (\text{ec. 4.5})$$

el logaritmo de la función de verosimilitud maximizada bajo el modelo esta dada por:

$$\log \hat{L}_c = \sum_i \left\{ \log \binom{n_i}{y_i} + y_i \log \hat{p}_i + (n_i - y_i) \log(1 - \hat{p}_i) \right\}$$

Bajo el modelo saturado, la probabilidad ajustada será la misma que la proporción observada $\hat{p}_i = y_i / n_i$, $y_i=1, 2, \dots, n$ y el logaritmo de la función de verosimilitud maximizada para este modelo es:

$$\log \hat{L}_j = \sum_i \left\{ \log \binom{n_i}{y_i} + y_i \log \hat{p}_i + (n_i - y_i) \log(1 - \hat{p}_i) \right\}$$

la devianza esta dada por:

$$D = -2 \left[\log \hat{L}_c - \log \hat{L}_f \right]$$

$$= 2 \sum_i \left\{ y_i \log \left(\frac{\hat{p}_i}{\hat{\beta}_i} \right) + (n_i - y_i) \log \left(\frac{1 - \hat{p}_i}{1 - \hat{\beta}_i} \right) \right\}$$

si el número de éxitos ajustados bajo el modelo actual es $\hat{y}_i = n_i \hat{p}_i$, la devianza puede ser escrita como:

$$D = 2 \sum_i \left\{ y_i \log \left(\frac{y_i}{\hat{y}_i} \right) + (n_i - y_i) \log \left(\frac{n_i - y_i}{n_i - \hat{y}_i} \right) \right\}$$

Este es un estadístico que compara las observaciones y_i con sus correspondientes valores ajustados \hat{y}_i .

• **La desviación para datos binarios**

En el caso especial de datos binarios, donde $n_i=1$ para $y=1, 2, \dots, n$, la desviación depende solamente de las probabilidades calculadas p_i , y no da información suficiente acerca del ajuste del modelo. La probabilidad para n observaciones binarias, vista como una función de los parámetros β , es

$$L(\beta) = \prod_i p_i^{y_i} (1 - p_i)^{1 - y_i}$$

y el logaritmo de la función maximizada bajo el modelo actual es

$$\log \hat{L}_c = \{ y_i \log \hat{p}_i + (1 - y_i) \log(1 - \hat{p}_i) \}.$$

Para el modelo saturado, $\hat{p}_i = y_i$ y puesto que $y_i \log y_i$ y $(1 - y_i) \log(1 - y_i)$ son ambos cero para los dos posibles valores de y_i , 0 y 1. $\log \hat{L}_c = 0$. La devianza de los datos es:

$$D = -2 \sum_i \{ \hat{p}_i \log it(\hat{p}_i) + \log(1 - \hat{p}_i) \}.$$

esta devianza de las observaciones binarias y_i , depende solo de las probabilidades ajustadas, consecuentemente la desviación de un modelo para datos con respuesta binaria no puede ser usada como una medida de bondad de ajuste del modelo. Para evaluar la magnitud en el cual un modelo adoptado adecua un grupo de datos binomiales, se requiere la distribución de la devianza, bajo la consideración de que el modelo es correcto.

4.3.2 Comparación de modelos logísticos lineales

Se pueden comparar modelos alternativos lineales logísticos sobre la base de una bondad de ajuste estadística tal como la devianza. Cuando un modelo contiene términos que son

adicionales en comparación con otro, se dice que están anidados. La diferencia en las devianzas de los dos modelos anidados mide la magnitud para la cual los términos adicionales mejoran el ajuste del modelo para la variable respuesta observada.

Examinando el efecto de inclusión o exclusión de términos en un modelo, es importante reconocer que el cambio en la devianza dependerá de que términos estén en el modelo al inicio. Por ejemplo considere la siguiente secuencia de modelos:

$$\text{Modelo (1) : } \text{logit}(p) = \beta_0$$

$$\text{Modelo (2) : } \text{logit}(p) = \beta_0 + \beta_1 x_1$$

$$\text{Modelo (3) : } \text{logit}(p) = \beta_0 + \beta_2 x_2$$

$$\text{Modelo (4) : } \text{logit}(p) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

La diferencia en la devianza entre el modelo 4 y el modelo 2 y entre el modelo 3 y el modelo 1, es que ambos miden el efecto de inclusión x_2 en el modelo. En la primera comparación entre el modelo 4 y el modelo 2, el efecto x_2 es ajustado para x_1 , puesto que x_2 está siendo agregado a un modelo que realmente incluye x_1 . En la segunda comparación entre el modelo 3 y el modelo 1, el efecto de x_2 es desajustado para x_1 , la magnitud para el de la devianza, sobre la inclusión de x_2 en el modelo con y sin x_1 , depende de la magnitud en la que x_1 y x_2 están asociados. Si x_1 y x_2 están actuando independientemente, la diferencia en las devianzas serán bastante similares, por otro lado si x_1 y x_2 están altamente correlacionados, estas serán muy diferentes.

De manera general suponiendo que dos modelos logísticos lineales modelo 1 y modelo 2, van a ser comparados, se pueden escribir de manera general como:

$$\text{Modelo (1) : } \text{logit}(p) = \beta_0 + \beta_1 x_1 + \dots + \beta_h x_h$$

$$\text{Modelo (2) : } \text{logit}(p) = \beta_0 + \beta_1 x_1 + \dots + \beta_h x_h + \beta_{h+1} x_{h+1} + \dots + \beta_k x_k$$

La devianza para cada modelo se denotará como D_1 y D_2 , por lo tanto D_1 y D_2 tienen $(n - h - 1)$ y $(n - k - 1)$ grados de libertad respectivamente, donde n es el número de observaciones binomiales. El modelo 1 está anidado con el modelo 2 y puesto que el modelo 2 contiene más términos que el modelo 1, este se ajustará inevitablemente mejor a los datos, y D_2 será más pequeño que D_1 . La diferencia en las devianzas $D_1 - D_2$ reflejará el efecto combinado de las variables $x_{h+1}, x_{h+2}, \dots, x_k$, después de que x_1, x_2, \dots, x_h han sido realmente incluidos en el modelo. La diferencia en la devianza es definida como la desviación ajustada para $x_{h+1}, x_{h+2}, \dots, x_k$ o eliminación de x_1, x_2, \dots, x_h .

Para comparar dos modelos anidados para datos binomiales, no hay una distribución teórica exacta disponible. Sin embargo, puesto que la devianza para cada modelo tiene una distribución aproximada a la distribución χ^2 , la diferencia entre las dos devianzas estará

distribuida aproximadamente como una χ^2 , por ejemplo suponga que se comparan dos modelos y que el modelo 1 está anidado con el modelo 2; el modelo 1 tendrá por lo tanto menos términos que el modelo 2. Si la devianza para el modelo 1 es D_1 con γ_1 grados de libertad y para el modelo 2 es D_2 con γ_2 grados de libertad, y si denotamos la máxima verosimilitud bajo el modelo 1 y modelo 2 por $\hat{L}_{1,1}$ y $\hat{L}_{1,2}$ respectivamente, las dos devianzas son:

$$D_1 = -2[\log \hat{L}_{1,1} - \log \hat{L}_I]$$

$$D_2 = -2[\log \hat{L}_{1,2} - \log \hat{L}_I]$$

donde \hat{L}_I es la probabilidad maximizada bajo el modelo saturado. Puesto que D_1 tiene una distribución χ^2 aproximada con γ_1 grados de libertad y D_2 tiene una distribución χ^2 aproximada con γ_2 grados de libertad, la diferencia en las devianzas $D_1 - D_2$ tendrán una distribución χ^2 aproximada con $\gamma_1 - \gamma_2$ grados de libertad.

Cuando se comparan dos devianzas, el término involucrado \hat{L}_I desaparece y $D_1 - D_2 = -2[\log \hat{L}_{1,1} - \log \hat{L}_{1,2}]$, la aproximación chi-cuadrada para la diferencia entre dos devianzas puede entonces usarse para comparar modelos anidados para datos binarios.

Cuando hay un gran número de grupos de datos y una sola variable explicativa x , se pueden definir diferentes modelos para ver la dependencia lineal de la variable respuesta. El modelo más general (modelo 1) especifica que hay diferentes líneas de regresión logística para cada grupo de datos, el modelo puede ser escrito como:

$$\text{Modelo (1) : } \text{logit}(p) = \alpha_j + \beta_j x,$$

en este caso es más conveniente denotar la intercepción y pendiente de la línea de regresión logística para el j -ésimo grupo por α_j, β_j , en vez de β_{0j}, β_{1j} para evitar el tener que usar doble descripción.

Un modelo algo más simple que el modelo 1 podría tener una pendiente común para los grupos, pero diferentes intercepciones sobre la escala logística, así el modelo (2) es:

$$\text{Modelo (2) : } \text{logit}(p) = \alpha_j + \beta x.$$

Este modelo conduce a una serie de líneas de regresión paralelas con pendiente β . Un modelo con el mismo número de parámetros que el modelo 2, podría tener intercepciones comunes, pero diferentes pendientes, este tercer modelo puede expresarse como:

$$\text{Modelo(3) : } \text{logit}(p) = \alpha + \beta_j x.$$

Bajo este modelo, las líneas son concurrentes en la respuesta probable de α cuando $x=0$. En un problema particular, uno podría no estar interesado en ambos modelos (2) y (3). Para la mayoría de las aplicaciones, esto es importante para investigar si las líneas de regresión logística son paralelas o si son concurrentes.

El modelo más simple es el siguiente:

$$\text{Modelo (4) : } \text{logit}(p) = \alpha + \beta x.$$

En este modelo, la relación entre la probabilidad de la respuesta y "x" es la misma para cada uno de los diferentes grupos de datos.

Los modelos ajustados pueden ser comparados en base a sus devianzas. Si D(1) a D(4) denotan las devianzas de los cuatro modelos ajustados, y suponiendo que sus respectivos grados de libertad son γ_1 a γ_4 . Para probar si las líneas de regresión logística pueden ser paralelas, los modelos 1 y 2 podrían ser comparados usando el resultado de que D(2)-D(1) tienen una distribución χ^2 con $\gamma_2 - \gamma_1$ grados de libertad, cuando las β_j son iguales. Si el valor observado del estadístico chi-cuadrada es muy significativo, los dos modelos no son comparables. Si estos son comparables, se podría entonces probar si son coincidentes usando el resultado de que D(4)-D(2) tienen una distribución χ^2 con $\gamma_4 - \gamma_2$ grados de libertad, una y otra vez cuando la hipótesis de intercepción constante es cierta. En este procedimiento secuencial, el modelo (3) podría ser substituído por el modelo (2) si este fue diseñado para probar la concurrencia seguida por la coincidencia.

4.3.3 No convergencia y sobreajuste

Cuando usamos un paquete estadístico para ajustar modelos lineales para datos con respuesta binaria, se pueden encontrar mensajes acerca de la no convergencia del proceso iterativo usado para producir el parámetro estimado.

Cuando se tiene un caso así, implica que el modelo contiene más parámetros desconocidos de los que pueden ser justificados por los datos. Cuando el sobreajuste ocurre, significa que los parámetros están estimados de manera incorrecta, y por lo tanto tendrán un error estándar muy grande.

Los problemas de no convergencia y sobreajuste generalmente ocurren cuando los datos observados están esparcidos, es decir cuando hay algunas combinaciones de variables explicativas con muy pocas observaciones binarias. La aproximación a la distribución chi-cuadrada de la desviación no será entonces muy confiable, podría no necesariamente indicar que el modelo es malo. En estas circunstancias, alguna tentativa para incluir términos adicionales en el modelo para ocasionar que disminuya la desviación a un nivel más aceptable podría generar un modelo sobreajustado.

4.3.4 Consideraciones para seleccionar el modelo

Antes de seleccionar un modelo es importante estudiar los factores de riesgo particulares sobre la ocurrencia de la enfermedad así como las variables de confusión potenciales. Hay dos procedimientos estadísticos posibles que se pueden usar para determinar cuando, una variable de confusión potencial puede ser incluida en el modelo logístico lineal. En primera instancia, los efectos de posibles variables de confusión sobre la probabilidad de ocurrencia de la enfermedad es calculada agregando estas variables a un modelo logístico lineal que contiene un término constante. La decisión de incluir o no una variable de confusión particular podría estar basado en la diferencia de la devianza resultante con la inclusión de dicha variable en el modelo. Puesto que es importante incluir todas las variables de confusión, estas pueden ser adicionadas si el cambio resultante en la devianza no es relevante para los niveles de significancia. Cuando hay un número considerable de variables de confusión potenciales, cada una de ellas podría ser agregada al modelo que incluye un término constante para determinar cual de ellas es la más importante. Combinaciones de estas variables así como las interacciones entre ellas, pueden entonces incluirse para identificar un grupo de variables de confusión que serán ajustadas para los factores de exposición.

Un procedimiento mas directo para ver si se valora o no una variable de confusión potencial, es examinar el efecto de la confusión sobre el estimador de la razón de momios de interés. Por ejemplo si al ajustar la variable de confusión en el modelo tiene un efecto insignificante sobre la razón de momios estimada, esta no necesita tomarse en cuenta en el modelo. Por otro lado, si el efecto de inclusión de la confusión en el modelo, varía respecto a la razón de momios estimada, será importante incluir la variable de confusión en el modelo.

Después de que se ha identificado una variable de confusión relevante, el siguiente paso es ver cuáles factores de riesgo son importantes. El interés es ahora ver el tamaño de los efectos debidos a cada uno de los factores de exposición, la importancia de un factor de riesgo puede ser calculado en forma usual por comparación del cambio en la desviación sobre la inclusión de un término con puntos porcentuales de la distribución χ^2 .

Para determinar si se requieren o no los términos que corresponden a las interacciones entre los factores de exposición y las variables de confusión en el modelo, se examina el efecto de adición de tales términos. En particular, si la reducción en la devianza debido a la inclusión de una interacción particular es muy significativa, se podría concluir que la razón de momios de enfermedad, para las personas expuestas relativas a las no expuestas, no es consistente sobre los diferentes niveles de las variables de confusión.

4.4 Interpretación de los parámetros en un modelo logístico lineal

Una de las razones para usar el modelo logístico lineal en el análisis de datos de un estudio etiológico, es que los coeficientes de las variables explicativas en el modelo pueden ser interpretados como logaritmos de la razón de momios. Esto significa que la estimación del

riesgo relativo de una enfermedad y su correspondiente error estándar puede fácilmente ser obtenido del modelo ajustado.

4.4.1 Factor de exposición simple dicotómico

En este modelo la probabilidad de ocurrencia de la enfermedad está relacionada con un factor de exposición simple dicotómico. Se considerará que los dos niveles de este factor son llamados expuesto y no expuesto respectivamente. Suponga que las respuestas binarias que han sido registradas para cada individuo están agrupadas, así que los datos básicos consisten de proporciones de individuos quienes desarrollaron la enfermedad en los grupos expuestos y no expuestos respectivamente. Un modelo logístico lineal para p_j , es la probabilidad de ocurrencia de la enfermedad en un individuo en el j -ésimo grupo expuesto, $j=1,2$ y puede ser escrito como:

$$\text{logit}(p_j) = \beta_0 + \gamma_j$$

donde γ_j es el efecto debido al j -ésimo nivel del factor de exposición. Este modelo está sobreparametrizado y por lo tanto es necesario imponer una restricción de estimabilidad, que puede ser por ejemplo $\gamma_1 = 0$, entonces el modelo puede escribirse en la siguiente forma:

$$\text{logit}(p_j) = \beta_0 + \beta_1 x_j \quad (\text{ec. 4.6})$$

donde x es una variable indicadora que toma el valor cero para individuos dentro del grupo no expuesto (nivel 1 del factor de exposición) y 1 para individuos dentro del grupo expuesto (nivel 2 del factor de exposición). En otras palabras, $x_1=0$ y $x_2=1$, y el parámetro β_1 es equivalente a γ_2 , p_1 es la probabilidad de ocurrencia de enfermedad en un individuo quien no está expuesto, y la razón de momios de este evento es $p_1/(1-p_1)$. De la ecuación anterior, $\text{logit}(p_1)=\beta_0$, puesto que $x_1=0$, y así la razón de momios de la enfermedad para una persona no expuesta está dada por:

$$\frac{p_1}{1-p_1} = e^{\beta_0}$$

Similarmente, la razón de momios para una persona en el grupo expuesto es:

$$\frac{p_2}{1-p_2} = e^{\beta_0 + \beta_1}$$

La razón de momios de enfermedad para una persona expuesta relativo a una persona no expuesta es entonces:

$$\psi = \frac{p_2 / (1-p_2)}{p_1 / (1-p_1)} = e^{\beta_1}$$

y consecuentemente, $\beta_1 = \log \psi$ es el logaritmo de la razón de momios.

Quando el modelo descrito en la ecuación (4.6) es ajustado a un grupo de datos observados, se obtiene la máxima verosimilitud estimada de los parámetros desconocidos β_0 y β_1 , denotando estos estimadores por $\hat{\beta}_0$ y $\hat{\beta}_1$ respectivamente, la máxima verosimilitud estimada de la razón de momios es $\hat{\psi} = \exp(\hat{\beta}_1)$. Para el modelo descrito por la ecuación 4.6 el error estándar de $\hat{\beta}_1$, será el error estándar del logaritmo de la razón de momios estimada, es decir *e.s.*($\hat{\beta}_1$) = *e.s.*{ $\log \hat{\psi}$ }. Usando este resultado, un intervalo de confianza aproximado 100(1 - α)% para el logaritmo de la razón de momios, es el intervalo con límites:

$$\log \hat{\psi} \pm z_{\alpha/2} \text{ e.s. } \{ \log \hat{\psi} \}$$

donde $z_{\alpha/2}$ es el punto superior $\alpha/2$ de la distribución normal estándar.

4.4.2 Factor de exposición politémica

Ahora consideremos la situación donde la probabilidad de ocurrencia de la enfermedad está relacionada con un factor de exposición simple que tiene m niveles. Suponiendo que los datos están agrupados, el modelo logístico lineal para la probabilidad de ocurrencia de enfermedad en un individuo dentro del j -ésimo grupo expuesto, $j=1, 2, \dots, m$, puede ser escrito de la siguiente manera:

$$\text{logit}(p_j) = \beta_0 + \gamma_j$$

donde γ_j es el efecto debido a el j -ésimo nivel del factor de exposición. Nuevamente el modelo está sobreparametrizado y se necesita imponer una restricción de estimabilidad dentro del modelo, la restricción $\gamma_1=0$ es conveniente. Para ajustar el modelo con esta restricción, se definen las variables indicadoras x_1, x_2, \dots, x_m con sus valores respectivos como se muestran en la tabla 3

nivel del factor de exposición	x_1	x_2	x_3	...	x_m
1	0	0	0	...	0
2	1	0	0	...	0
3	0	1	0	...	0
...
m	0	0	0	...	1

Tabla 3

El modelo puede ser expresado como:

$$\text{logit}(p_j) = \beta_0 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_m x_m \quad (\text{ec. 4.7})$$

Los parámetros $\beta_2, \beta_3, \dots, \beta_m$ en la ecuación 4.7 pueden ser interpretados como el logaritmo de la razón de momios para los individuos expuestos al nivel j del factor de exposición, $j=2, 3, \dots, m$, relativos a los expuestos en el primer nivel. Para ver esto primero considere la razón de momios de enfermedad en un individuo expuesto al primer nivel del factor de exposición, para quien la probabilidad de ocurrencia de enfermedad es p_1 .

Los valores de cada una de las $m-1$ variables indicadoras que corresponden al primer nivel del factor son cero y así $\text{logit}(p_1) = \beta_0$. En el j -ésimo nivel del factor de exposición, $j=2, 3, \dots, m$, solamente la variable que no tiene un valor cero es x_j , por lo tanto la razón de momios de ocurrencia para un individuo en el j -ésimo grupo de exposición será $\text{logit}(p_j) = \beta_0 + \beta_j$. La razón de momios de ocurrencia de enfermedad para individuos en el j -ésimo grupo de exposición relativo al primero, estará dado por:

$$\psi_{j1} = \frac{p_j / (1 - p_j)}{p_1 / (1 - p_1)} = e^{\beta_j}.$$

Cuando el modelo descrito por la ecuación 4.7 es ajustado a los datos observados, se obtienen los parámetros estimados $\hat{\beta}_2, \hat{\beta}_3, \dots, \hat{\beta}_m$, de los cuáles se pueden obtener los valores estimados de la razón de momios $\hat{\psi}_{j1}$.

4.4.3 Variable de exposición continua

Suponga que el valor de una variable de exposición simple continua es registrada para cada uno de los n individuos en un estudio cohorte. Un modelo logístico lineal para p_i , es la probabilidad de que el i -ésimo individuo, $y=1, 2, \dots, n$, desarrolle la enfermedad, y se puede representar como:

$$\text{logit}(p_i) = \beta_0 + \beta_1 x_i$$

donde x_i es el valor de la variable de exposición para el i -ésimo individuo. Nuevamente, el coeficiente de x_i en este modelo puede ser interpretado como el logaritmo de la razón de momios. Considere la razón de momios de enfermedad de un individuo para cuyo valor $x+1$ es relativo al valor x obtenido, la expresión está dado por:

$$\psi = \frac{\exp\{\beta_0 + \beta_1(x+1)\}}{\exp\{\beta_0 + \beta_1 x\}} = e^{\beta_1}$$

y así $\hat{\beta}_1$ en el modelo ajustado es el cambio estimado en el logaritmo de la razón de momios cuando x es incrementada por una unidad. Usando un argumento similar, el cambio estimado en el logaritmo de la razón de momios cuando la variable de exposición x es incrementada por r unidades es $r\hat{\beta}_1$, y el estimador correspondiente de la razón de momios es $\exp(r\hat{\beta}_1)$.

Recordando que la razón de momios es una aproximación a el riesgo relativo, este podría ser interpretado como el cambio aproximado en el riesgo de la enfermedad por cada incremento de r unidades en el valor de la variable x . El error estándar de la razón de momios logarítmica estimada será $r * e.s.(\beta_1)$ del cual se puede derivar los intervalos de confianza para cierta razón de momios.

El argumento anterior muestra que cuando una variable continua x , se incluye en un modelo logístico lineal, el riesgo relativo aproximado de la enfermedad cuando x es cambiada por r unidades no depende del valor actual de x . Por ejemplo, si x fue la edad de una persona, el riesgo relativo de enfermedad para un individuo de 65 años, relativo a alguno de 60 años, debería ser el mismo que para una persona de 25 relativo a una de 20.

Esta característica es un resultado directo de la suposición de que hay una relación lineal entre la transformación logística de la probabilidad de enfermedad y x .

4.4.4 Modelos con combinación de términos de diferentes tipos

La mayoría de los datos provenientes de un estudio etiológico proporcionan información sobre más de una variable, por lo tanto el modelo logístico lineal ajustado a los datos de tales estudios incluye generalmente un número de términos diferentes. En particular, el modelo debería incluir mas de un factor de exposición, términos que confunden las relaciones entre la probabilidad de ocurrencia de la enfermedad y el factor de exposición, y términos que representan interacciones entre los factores de exposición.

4.5 Regresión logística aplicado a un estudio de casos y controles

Después de haber definido las características principales del análisis multivariado así como del modelo de regresión logística, se hará la aplicación real para el estudio de casos y controles de leucemia.

Para realizar el análisis multivariado de los datos mediante regresión logística se utilizó el paquete estadístico SPSS-X, el cual ofrece tres opciones para estimar los modelos y son los siguientes:

- Método Estándar: este método introduce todas las variables seleccionadas en la primera iteración.
- Regresión hacia delante (Forward Logistic Regression): en este método se definen las variables al inicio, sin embargo se va agregando una a una en cada iteración de mayor a menor significancia.
- Regresión hacia atrás (Backward Logistic Regression): en este método se seleccionan las variables a incluir en el análisis, y a diferencia del método anterior, en cada iteración se elimina una variable tomando en cuenta la menos significativa.

Para estos tres métodos se tiene la opción de realizar la regresión de manera normal o condicionada, sin embargo en la estimación de los modelos se utilizó el método de regresión logística no condicionada puesto que es el más adecuado cuando las variables no están apareadas.

De acuerdo al objetivo del estudio se tomó como variable dependiente a la variable PACPREL cuyas categorías son 1=leucemia, 0=no leucemia y como variables independientes se consideraron en primer lugar las variables que resultaron ser significativas en el análisis bivariado, de las cuáles tenemos que fueron las siguientes:

- * Exposición del padre a cables de alta tensión, (cables1).
- * Exposición del padre a insecticidas, (Insect1).
- * Exposición del niño a solventes, (Lcosolv1).
- * Exposición del niño a pinturas, (Lpintur1)
- * Hospitalización del niño en el primer año de vida, (Hospit1a).
- * Exposición del niño a campos electromagnéticos, (Cialuz).

Además de las variables anteriores también se incluirá en la estimación de los modelos aquellas variables que no fueron significativas en el análisis bivariado pero si tienen un riesgo alto de desarrollar leucemia dado el factor de riesgo, y son las siguientes:

- * Exposición del padre a thinner, (Thiner1).
- * Exposición del padre a gasolinas, (Gasol1).
- * Defecto físico del niño al nacimiento, (Lenacdef).
- * Exposición del niño con gasolinas, (Gasol1).

Es importante comentar que las variables de interés para el modelo son principalmente aquellas para las cuáles se encontró alguna asociación con la enfermedad en el análisis bivariado, sin embargo después de obtener el modelo para las variables significativas, se estimará algún otro donde se incluyan aquellas que no están asociadas con la enfermedad, esto con el fin de verificar el comportamiento del modelo, pues en primer lugar, se podría pensar que la inclusión de ellas no afectaría los resultados del modelo obtenido.

El criterio a seguir para el análisis será el siguiente:

En primer lugar se usará el método estándar tomando como variable independiente principal a la variable "Cialuz", por las razones descritas al final del capítulo 3, esta variable se analizará con cada una de las variables que se mencionaron en la lista anterior de manera separada para ver si los parámetros estimados en cada modelo varían respecto a los estadísticos obtenidos en el análisis bivariado. Si resulta alguna diferencia importante se hará un análisis más detallado para ver si hay interacción o confusión entre las variables consideradas. Después de realizar el análisis con cada par de variables, se obtendrá otro modelo que incluya todas las variables que fueron significativas en el análisis bivariado, además de éste, se obtendrá otro que considere tanto las variables significativas como las de

riesgo alto y un último modelo que incluya las significativas, las de alto riesgo y algunas otras para las que no se encontró asociación en el análisis bivariado.

En segundo lugar se usará el método hacia delante donde se obtendrán dos modelos, en el primero se definirán al inicio todas las variables significativas y el segundo además de las significativas se agregarán las de riesgo alto.

Por último se usará el método hacia atrás para obtener dos modelos, uno para incluir todas las variables significativas y un segundo con las significativas más las de alto riesgo.

El hecho de utilizar los tres métodos es para comparar los resultados obtenidos y poder seleccionar el que mejor se ajuste a los datos. La comparación de los resultados se hará en el capítulo cinco, por lo que a continuación solo se muestran los valores obtenidos para cada modelo estimado.

Estimación de los modelos usando el método estándar.

Modelo 1

Variable Dependiente: **Paeprel**
 Variables independientes: **Cialuz**

Variables incluidas en la primera iteración: Cialuz
 Logaritmo inicial de probabilidad: 518.47409
 Número de iteraciones: 2 (Log. de probabilidad decrece a un valor menor que .01%)
 Logaritmo final de probabilidad: 515.22
 Bondad de ajuste: 374.00

Modelo chi-cuadrada	Chi-cuadrada	Grados de libertad	Significancia
	3.245	1	.0717
Mejoramiento	3.245	1	.0717

Variables en la ecuación.

	B	E.S.	Wald	df	Sig.	R	Exp(B)
Cialuz	.3824	.2128	3.2279	1	.0724	.0487	1.465
Constante	-.2323	.1661	1.9572	1	.1618		

B = Parámetro estimado, E.S. = Error estándar, Wald = Estadístico, df = Grados de libertad, R = Resultado, Exp(B) = Momento de probabilidad.

Modelo 2

Variable dependiente: **Paeprel**
 Variables independientes: **Cialuz, Cables1**

Variables incluidas en la primera iteración: Cialuz, Cables1
 Logaritmo inicial de probabilidad: 518.47409
 Número de iteraciones: 3 (Log. de probabilidad decrece a un valor menor que .01%)
 Logaritmo final de probabilidad: 496.863
 Bondad de ajuste: 372.988

	Chi-cuadrada	Grados de libertad	Significancia
Modelo chi-cuadrada	21.611	2	.0000
Mejoramiento	21.611	2	.0000

Variables en la ecuación.

	B	E.S.	Wald	df	Sig.	R	Exp(B)
Cialuz	-.4953	.2182	4.4308	1	.0353	.0685	1.5829
Cables1	1.5927	.4132	14.8571	1	.0001	.1575	4.9171
Constante	-1.7209	.4296	16.0507	1	.0001		

Modelo 3

Variable dependiente: Paeprel

Variables independientes: Cialuz, Insect1

Variables incluidas en la primera iteración:	Cialuz, Insect1
Logaritmo inicial del probabilidad:	518.47409
Número de iteraciones:	3 (Log. de probabilidad decrece a un valor menor que .01%)
Logaritmo final de probabilidad:	510.087
Bondad de ajuste:	374.368

	Chi-cuadrada	Grados de libertad	Significancia
Modelo chi-cuadrada	8.387	2	.0151
Mejoramiento	8.387	2	.0151

Variables en la ecuación.

	B	E.S.	Wald	df	Sig.	R	Exp(B)
Cialuz	-.3789	.2143	3.1269	1	.0770	.0466	1.4606
Insect1	1.2224	.5833	4.3918	1	.0361	.0679	3.3955
Constante	-.2803	.1683	2.7727	1	.0959		

Modelo 4

Variable dependiente: Paeprel

Variables independientes: Cialuz, Hospitia

Variables incluidas en la primera iteración:	Cialuz, Hospitia
Logaritmo inicial del probabilidad:	518.47409
Número de iteraciones:	2 (Log. de probabilidad decrece a un valor menor que .01%)
Logaritmo final de probabilidad:	510.946
Bondad de ajuste:	373.894

	Chi-cuadrada	Grados de libertad	Significancia
Modelo chi-cuadrada	7.528	2	.0232
Mejoramiento	7.528	2	.0232

Variables en la ecuación.

	B	E.S.	Wald	df	Sig.	R	Exp(B)
Cialuz	.3819	.2141	3.1818	1	.0745	.0477	1.4650
Thiner1	-.5814	.2844	4.1799	1	.0409	-.0648	.5591
Constante	.2530	.2904	.7590	1	.3837		

Modelo 5

Variable dependiente: Pacprel

Variables independientes: Cialuz, Lcosolv1

Variables incluidas en la primera iteración:	Cialuz, Lcosolv1
Logaritmo inicial de probabilidad:	518.47409
Número de iteraciones:	3 (Log. de probabilidad decrece a un valor menor que .01%)
Logaritmo final de probabilidad:	512.414
Bondad de ajuste:	374.710

	Chi-cuadrada	Grados de libertad	Significancia
Modelo chi-cuadrada	6.060	2	.0483
Mejoramiento	6.060	2	.0483

Variables en la ecuación.

	B	E.S.	Wald	df	Sig.	R	Exp(B)
Cialuz	.3576	.2136	3.0910	1	.0787	.0459	1.4558
Lcosolv1	1.6019	1.1031	2.1089	1	.1464	.0145	4.9623
Constante	-1.8088	1.1046	2.6815	1	.1015		

Modelo 6

Variable dependiente: Pacprel

Variables independientes: Cialuz, thiner1

Variables incluidas en la primera iteración:	Cialuz, Thiner1
Logaritmo inicial de probabilidad:	518.47409
Número de iteraciones:	2 (Log. de probabilidad decrece a un valor menor que .01%)
Logaritmo final de probabilidad:	514.997
Bondad de ajuste:	373.992

	Chi-cuadrada	Grados de libertad	Significancia
Modelo chi-cuadrada	3.477	2	.1758
Mejoramiento	3.477	2	.1758

Variables en la ecuación.

	B	E.S.	Wald	df	Sig.	R	Exp(B)
Cialuz	.3871	.2132	3.2988	1	.0693	.0501	1.478
Thiner1	-.1292	.2683	.2318	1	.6302	.0000	1.1379
Constante	-.3406	.2798	1.4817	1	.2235		

Modelo 7

Variable dependiente: **Pacprel**

Variable independiente: **Cialuz, Gasol1**

Variables incluidas en la primera iteración: **Cialuz, Gasol1**
 Logaritmo inicial del probabilidad: **518.47409**
 Número de iteraciones: **2 (Log. de probabilidad decrece a un valor menor que .01%)**
 Logaritmo final de probabilidad: **514.049**
 Bondad de ajuste: **373.997**

	Chi-cuadrada	Grados de libertad	Significancia
Modelo chi-cuadrada	4.425	2	.1094
Mejoramiento	4.425	2	.1094

Variables en la ecuación.

	B	E.S.	Wald	df	Sig.	R	Exp(B)
Cialuz	.3846	.2132	3.2539	1	.0713	.0492	1.4690
Thiner1	.3360	.3106	1.1701	1	.2794	.0000	1.3993
Constante	-.5239	.3192	2.7134	1	.0995		

Modelo 8

Variable dependiente: **Pacprel**

Variables independientes: **Cialuz, Lenacdef**

Variables incluidas en la primera iteración: **Cialuz, Lenacdef**
 Logaritmo inicial del probabilidad: **518.47409**
 Número de iteraciones: **2 (Log. de probabilidad decrece a un valor menor que .01%)**
 Logaritmo final de probabilidad: **513.049**
 Bondad de ajuste: **374.200**

	Chi-cuadrada	Grados de libertad	Significancia
Modelo chi-cuadrada	5.425	2	.0664
Mejoramiento	5.425	2	.0664

Variables en la ecuación.

	B	E.S.	Wald	df	Sig.	R	Exp(B)
Cialuz	.3702	.2136	3.0049	1	.0830	.0440	1.4481
Lenacdef	.7867	.5519	2.0324	1	.1540	.0079	2.1962
Constante	-.9796	.5532	3.1357	1	.0766		

Modelo 9

Variable dependiente: **Pacprel**

Variables independientes: **Cialuz, Gasol1**

Variables incluidas en la primera iteración: **Cialuz, Gasol1**
 Logaritmo inicial del probabilidad: **518.47409**
 Número de iteraciones: **2 (Log. de probabilidad decrece a un valor menor que .01%)**
 Logaritmo final de probabilidad: **514.943**
 Bondad de ajuste: **373.991**

Modelo chi-cuadrada	Chi-cuadrada	Grados de libertad	Significancia
Mejoramiento	3.531	2	.1711

Variables en la ecuación.

	B	E.S.	Wald	df	Sig.	R	Exp(B)
Cialuz	.3858	.2130	3.2799	1	.0701	.0497	1.4708
Gasoll	-.2606	.4883	2.8409	1	.5935	.0000	1.2977
Constante	-.4825	.4977	.9399	1	.3323		

Modelo 10

Variable dependiente: **Pacprel**

Variables independientes: **Cialuz, Lpintur1**

Variables incluidas en la primera iteración: Cialuz, Lpintur1
 Logaritmo inicial del probabilidad: 518.47409
 Número de iteraciones: 2 (Log. de probabilidad decrece a un valor menor que .01%)
 Logaritmo final de probabilidad: 511.617
 Bondad de ajuste: 374.053

Modelo chi-cuadrada	Chi-cuadrada	Grados de libertad	Significancia
Mejoramiento	6.857	2	.0324
	6.857	2	.0324

Variables en la ecuación.

	B	E.S.	Wald	df	Sig.	R	Exp(B)
Cialuz	-.3791	.2139	3.1246	1	.0763	.0469	1.4610
Lpintur1	-.4921	.2269	3.5573	1	.0593	-.0548	.6114
Constante	.1614	.2671	.3650	1	.5158		

Modelo 11

Variable dependiente: **Pacprel**

Variables independientes: **Cialuz, Rxxmadne**

Variables incluidas en la primera iteración: Cialuz, Rxxmadne
 Logaritmo inicial del probabilidad: 518.47409
 Número de iteraciones: 2 (Log. de probabilidad decrece a un valor menor que .01%)
 Logaritmo final de probabilidad: 513.265
 Bondad de ajuste: 374.103

Modelo chi-cuadrada	Chi-cuadrada	Grados de libertad	Significancia
Mejoramiento	5.209	2	.0740
	5.209	2	.0740

Variables en la ecuación.

	B	E.S.	Wald	df	Sig.	R	Exp(B)
Cialuz	-.3778	.2134	3.1346	1	.0766	.0468	1.4591
Rxxmadne	-.3515	.3520	1.9299	1	.1648	.0000	1.5707
Constante	-.6274	.3311	3.5910	1	.0581		

Modelo 12Variable dependiente: **Pacprel**Variables independientes: **Cialuz, Cables1, Insect1, Lcosolv1, Lpintur1, Hospitia**

Variables incluidas en la primera iteración: **Cialuz, Cables1, Insect1, Lcosolv1, Lpintur1, Hospitia**
 Logaritmo inicial del probabilidad: **518.47409**
 Número de iteraciones: **3** (Log. de probabilidad decrece a un valor menor que .01%)
 Logaritmo final de probabilidad: **483.191**
 Bondad de ajuste: **372.054**

Modelo chi-cuadrada	Chi-cuadrada	Grados de libertad	Significancia
	35.283	6	.0000
Mejoramiento	35.283	6	.0000

Variables en la ecuación.

	B	E.S.	Wald	df	Sig.	R	Exp(B)
Cialuz	-44000	22240	3.9164	1	.0478	.0608	1.528
Cables1	1.5220	.41900	13.1947	1	.0003	.1469	4.5812
Insect1	1.238	.59870	4.2812	1	.0385	.0663	3.4516
Lcosolv1	1.2570	1.1350	1.2265	1	.2681	.0000	3.5148
Lpintur1	-.4955	.2716	3.3277	1	.0681	-.0506	.6093
hospitia	-.5723	.2954	3.7525	1	.0537	-.0581	.5462
Constante	-2.0646	1.2286	2.8242	1	.0929		

Modelo 13Variable dependiente: **Pacprel**Variables independientes: **Cialuz, Cables1, Insect1, Lcosolv1, Lpintur1, Hospitia, Thiner1, Gasol1, Lenacdef, Gasol11,**

Variables incluidas en la primera iteración: **Cialuz, Cables1, Insect1, Lcosolv1, Lpintur1, Hospitia, Thiner1, Gasol1, Lenacdef, Gasol11**
 Logaritmo inicial del probabilidad: **518.47409**
 Número de iteraciones: **3** (Log. de probabilidad decrece a un valor menor que .01%)
 Logaritmo final de probabilidad: **478.498**
 Bondad de ajuste: **373.011**

Modelo chi-cuadrada	Chi-cuadrada	Grados de libertad	Significancia
	40.049	10	.0000
Mejoramiento	40.049	10	.0000

Variables en la ecuación.

	B	E.S.	Wald	df	Sig.	R	Exp(B)
Cialuz	-4172	.2241	3.4661	1	.0626	.0532	1.5178
Cables1	1.5206	.4214	13.0223	1	.0003	.1458	4.5748
Insect1	1.3573	.6092	4.9642	1	.0259	.0756	3.8856
Lcosolv1	1.2512	1.1601	1.1631	1	.2808	.0000	3.4994
Lpintur1	-.5489	.2991	3.3664	1	.0665	-.0648	.5776
Hospitia	-.6115	.3012	4.1775	1	.0410	-.0648	.5404
Thiner1	-.0284	.4382	.0823	1	.9774	.0000	.8832
Gasol1	-.3576	.3367	1.1280	1	.2882	.0000	1.4299
Lenacdef	.9770	.3891	2.7503	1	.0972	.0367	2.6564
Gasol11	-.3921	.5564	.4835	1	.4868	.0000	1.4802
Constante	-3.4706	1.4177	6.0238	1	.0141		

Modelo 14

Variable dependiente: **Pacprel**

Variables independientes: **Cialuz, Cables1, Insect1, Lcosolv1, Lpintur1, Hospitia, Thiner1, Gasol1, Lenacdef, Gasol1, Sexo, Transfl, Subest1, Solven1, Pintur1, Rxmadne**

Variables incluidas en la primera iteración: **Cialuz, Cables1, Insect1, Lcosolv1, Lpintur1, Hospitia, Thiner1, Gasol1, Lenacdef, Gasol1, Sexo, Transfl, Subest1, Solven1, Pintur1, Rxmadne.**
518.47409

Logaritmo inicial de probabilidad:

Número de iteraciones:

Logaritmo final de probabilidad:

Bondad de ajuste:

3 (Log. de probabilidad decrece a un valor menor que .01%)
475.657
372.733

Modelo	Chi-cuadrada	Grados de libertad	Significancia
chi-cuadrada	42.816	16	.0003
Mejoramiento	42.816	16	.0003

Variables en la ecuación.

	B	E.S.	Wald	df	Sig.	R	Exp(B)
Cialuz	-.4346	.2286	3.6151	1	.0573	.0558	1.5444
Cables1	1.4951	.4237	12.4553	1	.0004	1.420	4.4600
Insect1	1.3446	.6179	4.7385	1	.0295	.0726	3.8367
Lcosolv1	1.2075	1.1681	1.0685	1	.3013	.0000	3.3450
Lpintur1	-.5684	.3009	3.5677	1	.0589	-.0550	.5665
Hospitia	-.6424	.3047	4.4437	1	.0350	-.0687	.5260
Thiner1	-.1094	.4357	.0631	1	.8017	.0000	.8964
Gasol1	.3846	.3698	1.0819	1	.2983	.0000	1.4691
Lenacdef	1.0449	.5922	3.1130	1	.0777	.0463	2.8431
Gasol1	.3883	.5674	.4685	1	.4937	.0000	1.4745
Sexo	.1054	.2260	.2192	1	.6396	.0000	1.1116
Transfl	-.0765	.2519	.0924	1	.7612	.0000	.9263
Subest1	-.0612	.3300	.0344	1	.8529	.0000	.9466
Solven1	.0842	.3639	.0175	1	.8947	.0000	1.0493
Pintur1	-.0225	.3239	.0049	1	.9444	.0000	.9777
Rxmadne	.2615	.1726	.2955	1	.1298	.0239	1.2989
Constante	-3.7065	1.4727	6.3345	1	.0118		

Obtención de los modelos usando el método hacia delante.

Estimación de los modelos utilizando el método de regresión logística hacia delante (Forward Stepwise LR). En este método se incluye en cada iteración una variable nueva de mayor a menor significancia, el método termina cuando se han considerado todas las variables significativas y ya no es posible agregar variables.

Modelo 1

Variable dependiente: **Pacprel**

Variables independientes: **Cialuz, Cables1, Insect1, Lcosolv1, Lpintur1, Hospitia.**

Variables incluidas en la primera iteración: **Cables1**

Logaritmo inicial de probabilidad:

Número de iteraciones:

518.47409
3 (Log. de probabilidad decrece a un valor menor que .01%)

Logaritmo final de probabilidad: 501.328
 Bondad de ajuste: 373.999

	Chi-cuadrada	Grados de libertad	Significancia
Modelo chi-cuadrada	17.146	1	.0000
Mejoramiento	17.146	1	.0000

Variables en la ecuación

	B	E.S.	Wald	df	Sig.	R	Exp(B)
Cables1	1.5302	.4102	13.9144	1	.0002	.1516	4.6192
Constante	-1.3863	.3953	12.2993	1	.0005		

Variables no consideradas en la primera iteración

Variable	Valor	df	Significancia	R
Cialuz	4.4530	1	.0348	.0688
Insect1	4.5259	1	.0334	.0698
Leosolv1	1.8514	1	.1736	.0000
Lpintur1	3.0945	1	.0786	.0459
Hospitia	4.1831	1	.0408	.0649

Variables incluidas en la segunda iteración:

Insect1
 Número de iteraciones: 3 (Log. de probabilidad decrece a un valor menor que .01%)
 Logaritmo de probabilidad: 496.530
 Bondad de ajuste: 373.843

	Chi-cuadrada	Grados de libertad	Significancia
Modelo chi-cuadrada	21.944	2	.0000
Mejoramiento	4.799	1	.0285

Variables en la ecuación

	B	E.S.	Wald	df	Sig.	R	Exp(B)
Cables1	1.5195	.4123	13.5853	1	.0002	.1495	4.5690
Insect1	1.1999	.5923	4.1046	1	.0428	.0637	3.3199
Constante	-1.4528	.3978	12.8435	1	.0003		

Variables no consideradas en la segunda iteración

Variable	Valor	df	Significancia	R
Cialuz	4.3307	1	.0374	.0670
Leosolv1	1.7108	1	.1909	.0000
Lpintur1	3.5124	1	.0609	.0540
Hospitia	4.2051	1	.0403	.0652

Variables incluidas en la tercera iteración:

Cialuz
 Número de iteraciones: 3 (Log. de probabilidad decrece a un valor menor que .01%)
 Logaritmo de probabilidad: 492.187
 Bondad de ajuste: 373.677

Modelo chi-cuadrada	Chi-cuadrada	Grados de libertad	Significancia
Mejoramiento	26.287 4.343	3 1	.0000 .0372

Variables en la ecuación

	B	E.S.	Wald	df	Sig.	R	Exp(B)
Cialuz	-.4558	.2196	4.3091	1	.0379	.0667	1.5775
Cables1	1.5833	.4157	14.5076	1	.0001	-.1553	4.8710
Insect1	1.1884	.5929	4.0170	1	.0450	.0624	3.2818
Constante	-1.7594	.4300	16.5066	1	.0000		

Variables no consideradas en la tercera iteración

Variable	Valor	df	Significancia	R
Lesosol1	1.4423	1	.2298	.0000
Lpintur1	3.3306	1	.0680	.0507
Hospitia	4.2378	1	.0395	.0657

Variables incluidas en la cuarta iteración:

Hospitia

Número de iteraciones: 3 (Log. de probabilidad decrece a un valor menor que .01%)

Logaritmo de probabilidad: -487.907

Bondad de ajuste: 375.256

Modelo chi-cuadrada	Chi-cuadrada	Grados de libertad	Significancia
Mejoramiento	30.567 4.280	4 1	.0000 .0386

Variables en la ecuación

	B	E.S.	Wald	df	Sig.	R	Exp(B)
Cialuz	-.4604	.2210	4.3383	1	.0373	.0672	1.5847
Cables1	1.5911	.4180	14.4889	1	.0001	-.1552	4.9092
Insect1	1.2095	.5978	4.0930	1	.0431	.0635	3.3518
Hospitia	-.5997	.2937	4.1695	1	.0412	-.0647	.5490
Constante	-1.2697	.4933	6.6238	1	.0101		

Variables no consideradas en la cuarta iteración

Variable	Valor	df	Significancia	R
Lesosol1	1.1992	1	.2735	.0000
Lpintur1	3.1787	1	.0746	.0477

Fin de la estimación dado que no se pueden agregar más variables al modelo.

Modelo 2

Variable dependiente: Paeprel

Variables independientes: Cialuz, Cables1, Insect1, Lcosolv1, Lpintur1, Hospitia, Thiner1, Gasol1, Lenacedf, Gasol1

Variables incluidas en la primera iteración: Cables1
 Logaritmo inicial del probabilidad: 518.47409
 Número de iteraciones: 3 (Log. de probabilidad decrece a un valor menor que .01%)
 Logaritmo final de probabilidad: 501.328
 Bondad de ajuste: 373.999

	Chi-cuadrada	Grados de libertad	Significancia
Modelo chi-cuadrada	17,146	1	.0000
Mejoramiento	17,146	1	.0000

Variables en la ecuación

	B	E.S.	Wald	df	Sig.	R	Exp(B)
Cables1	1,5302	.4102	13,9144	1	.0002	.1516	4,6192
Constante	-1,3863	.3953	12,2993	1	.0005		

Variables no consideradas en la primera iteración

Variable	Valor	df	Significancia	R
Cialuz	4,4530	1	.0348	-.0688
Insect1	4,5259	1	.0334	-.0698
Lcosolv1	1,8514	1	.1736	.0000
Lpintur1	3,0945	1	.0786	-.0459
Hospitia	4,1831	1	.0408	-.0649
Thiner1	.0648	1	.7937	.0000
Gasol1	.6640	1	.4151	.0000
Lenacedf	2,3150	1	.1281	-.0246
Gasol1	.2321	1	.6300	.0000

Variables incluidas en la segunda iteración: Insect1
 Número de iteraciones: 3 (Log. de probabilidad decrece a un valor menor que .01%)
 Logaritmo de probabilidad: 496.530
 Bondad de ajuste: 373.843

	Chi-cuadrada	Grados de libertad	Significancia
Modelo chi-cuadrada	21,944	2	.0000
Mejoramiento	4,799	1	.0285

Variables en la ecuación

	B	E.S.	Wald	df	Sig.	R	Exp(B)
Cables1	1,5195	.4123	13,5853	1	.0002	.1495	4,5699
Insect1	1,1999	.5923	4,1046	1	.0428	.0637	3,3194
Constante	-1,4528	.3978	12,8435	1	.0003		

Variables no consideradas en la segunda iteración

Variable	Valor	df	Significancia	R
Cialuz	4.3307	1	.0374	.0670
Leosolv1	1.7108	1	.1909	.0300
Lpintur1	3.5124	1	.0609	.0540
Hospita	4.2051	1	.0403	.0652
Thiner1	.1623	1	.6870	.0000
Gasol1	1.3353	1	.2479	.0000
Lenacdef	2.4841	1	.1150	.0306
Gasol1	.2604	1	.6099	.0000

Variables incluidas en la tercera iteración:

Cialuz

Número de iteraciones: 3 (Log. de probabilidad decrece a un valor menor que .01%)

Logaritmo de probabilidad:

492.187

Bondad de ajuste:

373.677

Modelo	Chi-cuadrada	Grados de libertad	Significancia
Modelo chi-cuadrada	26.287	3	.0000
Mejoramiento	4.343	1	.0372

Variables en la ecuación

	B	E.S.	Wald	df	Sig.	R	Exp(B)
Cialuz	.4558	.2196	4.3091	1	.0379	.0667	1.5775
Cables1	1.5833	.4157	14.5076	1	.0001	.1553	4.8710
Insect1	1.1884	.5929	4.0170	1	.0450	.0624	3.2818
Constante	-1.7594	.4330	16.5066	1	.0000		

Variables no consideradas en la tercera iteración

Variable	Valor	df	Significancia	R
Leosolv1	1.4423	1	.2298	.0000
Lpintur1	3.3306	1	.0680	.0507
Hospita	4.2378	1	.0395	.0657
Thiner1	.2283	1	.6328	.0000
Gasol1	1.2670	1	.2603	.0000
Lenacdef	2.1721	1	.1452	.0153
Gasol1	.3049	1	.5808	.0000

Variables incluidas en la cuarta iteración:

Hospita

Número de iteraciones: 3 (Log. de probabilidad decrece a un valor menor que .01%)

Logaritmo de probabilidad:

487.907

Bondad de ajuste:

375.256

Modelo	Chi-cuadrada	Grados de libertad	Significancia
Modelo chi-cuadrada	30.567	4	.0000
Mejoramiento	4.280	1	.0386

Variables en la ecuación

	B	E.S.	Wald	df	Sig.	R	Exp(B)
Cialuz	.4604	.2210	4.3383	1	.0373	.0672	1.5847
Cables1	1.5911	.4180	14.4889	1	.0001	.1552	4.9092
Insect1	1.2095	.5978	4.0930	1	.0431	.0635	3.2518
Hospita	-.5997	.2937	4.1695	1	.0412	-.0647	.5490
Constante	-1.2697	.4933	6.6238	1	.0101		

Variables no consideradas en la cuarta iteración

Variable	Valor	df	Significancia	R
Lcosolv1	1.1992	1	.2735	.0000
Lpintur1	3.1787	1	.0746	.0477
Talser1	.0580	1	.8096	.0000
Gasol1	1.1141	1	.2912	.0000
Lenacedef	2.8002	1	.0943	.0393
Gasol1	.2422	1	.6226	.0000

Fin de la estimación dado que no se pueden agregar más variables al modelo.

Obtención de los modelos usando el método hacia atrás.

Estimación de los modelos utilizando el método hacia atrás (Backward Stepwise LR), en este caso se toman en cuenta todas las variables en la primera iteración y después de obtener los parámetros iniciales se va eliminando una variable no significativa en cada iteración, el método finaliza cuando ya no es posible eliminar más variables.

Modelo 1

Variable dependiente: Paeprel

Variables independientes: Cialuz, Cables1, Insect1, Lcosolv1, Lpintur1, Hospita1.

Variables incluidas en la primera iteración:	Cialuz, Cables1, Insect1, Lcosolv1, Lpintur1, Hospita1
Logaritmo inicial del probabilidad:	518.470409
Número de iteraciones:	3 (Log. de probabilidad decrece a un valor menor que .01%)
Logaritmo final de probabilidad:	483.191
Bondad de ajuste:	372.054

Modelo chi-cuadrada	Chi-cuadrada	Grados de libertad	Significancia
	35.283	6	.0000
Mejoramiento	35.283	6	.0000

Variables en la ecuación

	B	E.S.	Wald	df	Sig.	R	Exp(B)
Cialuz	.4400	.2224	3.9164	1	.0478	.0608	1.528
Cables1	1.5220	.4190	13.1947	1	.0003	.1469	4.5812
Insect1	1.2388	.5987	4.2812	1	.0385	.0653	3.4516
Lcosolv1	1.2570	1.1350	1.2265	1	.2681	.0000	3.5148
Lpintur1	-.4955	.2716	3.3277	1	.0681	-.0506	.6093
Hospita1	-.5723	.2954	3.7525	1	.0527	-.0581	.5642
Constante	-2.0646	1.2286	2.8242	1	.0929		

Variables eliminada en la segunda iteración: Lcosolv1

iteración:

Número de iteraciones:	3 (Log. de probabilidad decrece a un valor menor que .01%)
Logaritmo de probabilidad:	484.709
Bondad de ajuste:	372.315

Modelo chi-cuadrada	Chi-cuadrada	Grados de libertad	Significancia
Mejoramiento	33,765	5	.0000
	-1,518	1	.2179

Variables en la ecuación

	B	E.S.	Wald	df	Sig.	R	Exp(B)
Cialuz	.4521	.2219	4.1506	1	.0416	.0644	1,5716
Cables1	1.5611	.4178	13.9616	1	.0002	.1519	4,7640
Insect1	1.2550	.5993	4.3849	1	.0363	.0678	3,5080
Lpintur1	-.4795	.2702	3.1500	1	.0759	-.0471	.6191
Hospit1a	-.5922	.2953	4.0227	1	.0449	-.0625	.5531
Constante	-.8635	.5430	2.5291	1	.1118		

Variables no consideradas en la ecuación

Variable	Valor	df	Significancia	R
Lcosolv1	1,3630	1	.2430	.0000

Fin de la estimación dado que no se pueden eliminar mas variables del modelo.

Modelo 2

Variable dependiente: **Pacpre1**

Variables independientes: **Cialuz, Cables1, Insect1, Lcosolv1, Lpintur1, Hospit1a, Thiner1, Gasol1, Lenacdef, Gasol1**

Variables incluidas en la primera iteración:	Cialuz, Cables1, Insect1, Lcosolv1, Lpintur1, Hospit1a, Thiner1, Gasol1, Lenacdef, Gasol1
Logaritmo inicial de probabilidad:	518.470409
Número de iteraciones:	3 (Log. de probabilidad decrece a un valor menor que .01%)
Logaritmo final de probabilidad:	478.498
Bondad de ajuste:	373.011

Modelo chi-cuadrada	Chi-cuadrada	Grados de libertad	Significancia
Mejoramiento	40,049	10	.0000
	40,049	10	.0000

Variables en la ecuación

	B	E.S.	Wald	df	Sig.	R	Exp(B)
Cialuz	.4172	.2241	3.4661	1	.0626	.0532	1,5178
Cables1	1.5206	.4214	13.022	1	.0003	.1458	4,5748
Insect1	1.3573	.6092	4.9462	1	.0259	.0756	3,8856
Lcosolv1	1.2512	1,1601	1,1631	1	.2808	.0000	3,4944
Lpintur1	-.5489	.2991	3,3664	1	.0665	-.0513	.5776
Hospit1a	-.6155	.3012	4,1774	1	.0410	-.0648	.5404
Thiner1	-.1242	.4348	.0823	1	.7742	.0000	.8832
Gasol1	.3576	.3367	1,1280	1	.2882	.0000	1,4299
Lenacdef	.9770	.5891	2,7503	1	.0972	.0380	2,6564
Gasol11	.3921	.5640	.4385	1	.4868	.0000	1,4802
Constante	-3.4796	1,4177	6,0238	1	.0141		

Variables eliminada en la segunda iteración: Thiner I
 Número de iteraciones: 3 (Log. de probabilidad decrece a un valor menor que .01%)
 Logaritmo de probabilidad: 478.507
 Bondad de ajuste: 372.973

Modelo chi-cuadrada	Chi-cuadrada	Grados de libertad	Significancia
Mejoramiento	39.967	9	.0000
	-.082	1	.7740

Variables en la ecuación

	B	E.S.	Wald	df	Sig.	R	Exp(B)
Cialuz	.4169	.2240	3.4627	1	.0628	.0531	1.5173
Cables1	1.5124	.4200	12.9700	1	.0003	.1455	4.5377
Insect1	1.3642	.6101	4.5997	1	.0254	.0761	3.9126
Lcosolv1	1.1946	1.1423	1.0937	1	.2957	.0000	3.3023
Lpintur1	-.5721	.2881	3.9428	1	.0471	-.0612	.5644
Hospit1a	-.6146	.3011	4.1671	1	.0412	-.0647	.5409
Gasol1	.3520	.3358	1.0988	1	.2945	.0000	1.4219
Lenacdef	.9639	.5876	2.6908	1	.1009	.0365	2.6218
Constante	-.3673	.5580	.4333	1	.5104	.0000	1.4438
Gasol1	-3.4708	1.4177	5.9940	1	.0144		

Variables eliminada en la tercera iteración: Gasol1
 Número de iteraciones: 3 (Log. de probabilidad decrece a un valor menor que .01%)
 Logaritmo de probabilidad: 478.940
 Bondad de ajuste: 372.927

Modelo chi-cuadrada	Chi-cuadrada	Grados de libertad	Significancia
Mejoramiento	39.534	8	.0000
	-.433	1	.5105

Variables en la ecuación

	B	E.S.	Wald	df	Sig.	R	Exp(B)
Cialuz	.4138	.2238	3.4183	1	.0643	.0523	1.5126
Cables1	1.5071	.4199	12.8842	1	.0003	.1449	4.5135
Insect1	1.3649	.6114	4.9834	1	.0256	.0759	3.9154
Lcosolv1	1.2801	1.1462	1.2473	1	.2641	.0000	3.5972
Lpintur1	-.5166	.2744	3.5462	1	.0597	-.0546	.5965
Hospit1a	-.6192	.3006	4.2425	1	.0394	-.0658	.5384
Gasol1	.3692	.3338	1.2332	1	.2687	.0000	1.4466
Lenacdef	.9969	.5837	2.9176	1	.0876	.0421	2.7100
Constante	-3.2855	1.3920	5.570	1	.0183		

Variables eliminada en la cuarta iteración: Gasol1
 Número de iteraciones: 3 (Log. de probabilidad decrece a un valor menor que .01%)
 Logaritmo de probabilidad: 480.177
 Bondad de ajuste: 372.231

Modelo chi-cuadrada	Chi-cuadrada	Grados de libertad	Significancia
Mejoramiento	38.298	7	.0000
	-1.236	1	.2662

Variables en la ecuación

	B	E.S.	Wald	df	Sig.	R	Exp(B)
Cialuz	.4179	.2234	3.4985	1	.0614	.0538	1.5188
Cables1	1.5368	.4203	13.3707	1	.0003	.1481	4.6497
Insect1	1.2760	.6025	4.4862	1	.0342	.0692	3.5823
Lcosolv1	1.2415	1.1442	1.1774	1	.2779	.0000	3.4608
Lpintur1	-.5168	.2735	3.5702	1	.0588	-.0550	.5964
Hospit1a	-.6200	.3004	4.3975	1	.0360	-.0660	.5326
Lenacdef	.9787	.5824	2.8264	1	.0928	.0399	2.6611
Constante	-2.9249	1.3478	4.7093	1	.0300		

Variables eliminada en la quinta iteración: Lcosolv1

Número de iteraciones: 3 (Log. de probabilidad decrece a un valor menor que .01%)

Logaritmo de probabilidad: 481.615

Bondad de ajuste: 372.518

	Chi-cuadrada	Grados de libertad	Significancia
Modelo chi-cuadrada	36.860	6	.0000
Mejoramiento	-1.438	1	.2305

Variables en la ecuación

	B	E.S.	Wald	df	Sig.	R	Exp(B)
Cialuz	-.4309	.2229	3.7371	1	.0532	.0579	1.5387
Cables1	1.5653	.4193	13.9342	1	.0002	.1517	4.7841
Insect1	1.2916	.6029	4.5988	1	.0322	.0707	3.6384
Lpintur1	-.5001	.2720	3.3807	1	.0660	-.0516	.6065
Hospit1a	-.6494	.3002	4.6793	1	.0305	-.0719	.5223
Lenacdef	.9842	.5795	2.8847	1	.0894	.0413	2.6756
Constante	-1.7355	.7525	5.3189	1	.0211		

Variables no consideradas en el modelo

Variable	Valor	df	Significancia	R
Lcosolv1	1.3044	1	.2534	.0000
Thiner1	1.1219	1	.7270	.0000
Gasoll1	1.1424	1	.2851	.0000
Gasoll1	.7790	1	.3774	.0000

Fin de la estimación dado que no se pueden eliminar mas variables del modelo.

ANÁLISIS DE RESULTADOS

5. ANÁLISIS DE RESULTADOS

El objetivo del presente capítulo es interpretar los resultados obtenidos en el análisis multivariado y comparar los diferentes modelos para poder seleccionar el que mejor se ajuste a los datos.

5.1 Interpretación de resultados

En la interpretación de los datos para cada modelo se tomará como punto de comparación los resultados obtenidos en el análisis bivariado ya que esto nos permitirá ver si las variables ajustadas en el respectivo modelo siguen el mismo patrón.

Los primeros modelos que se obtuvieron por el método de regresión logística estándar se hicieron con el fin de verificar si el comportamiento de la variable independiente principal (exposición a campos electromagnéticos) sufría alguna modificación importante cuando se agregaba una variable más.

El modelo 1 es el que servirá de base para comparación respecto a los demás modelos, los valores a comparar serán los siguientes:

- Estadístico ajustado (Wald) = 3.2274
- Significancia (Sig.) = .0724
- Razón de momios (Exp(B)) = 1.465

para propósitos de comparación a estos valores se les llamó "valores crudo" y la variación que pueda resultar al agregar variables en el análisis se considerará relevante si en el modelo respectivo, el parámetro ajustado $\exp(B)$ menos el crudo, entre el ajustado es mayor a 0.20, es decir: $[(\text{Exp}(B) \text{ ajustado} - \text{Exp}(B) \text{ crudo}) / \text{Exp}(B) \text{ ajustado}] > 0.20$, si el valor resultante es mayor implica que la variable agregada modifica el efecto de la variable principal, en este caso se dice que puede haber confusión o interacción entre las variables.

Para comparación de resultados se agruparon los datos en tablas de manera que permitan visualizar los principales valores obtenidos en el análisis multivariado respecto al análisis bivariado.

Las primero cuatro tablas corresponden a los modelos obtenidos mediante el método estándar, la siguiente es para el método hacia adelante y las dos últimas son para comparar los modelos obtenidos con el método hacia atrás.

Tabla 1. Comparación de resultados usando el método estándar

Modelo No.	Análisis multivariado Valores obtenidos					Análisis bivariado Valores obtenidos		
	Variable(s)	Estadístico	Sig.	Exp(B)	Diferencia	Estadístico	Sig.	Razón de mórmos
1	Cialuz	3.2274	.0724	3.465		3.2391	.0719	1.465
2	Cialuz Cables 1	4.4308 14.8571	.0353 .0001	1.5829 4.9171	.075	16.1245	.00006	4.619
3	Cialuz Insect 1	3.1269 4.3918	.0770 .0361	1.4606 3.3955	-.003	4.9916	.0254	3.4181
4	Cialuz Hospit 1a	3.1818 4.1799	.0745 .0409	1.465 5.591	0	4.29490	.03823	.55875
5	Cialuz Lecosol 1	3.0910 2.1089	.0787 .1464	1.4558 4.9623	-0.12	2.7101	.09971	5.1098
6	Cialuz Thiner 1	3.2988 2.318	.0893 .6302	1.478 1.1379	.008	15994	68921	1.1125
7	Cialuz Gasol 1	3.2539 1.1701	.0712 .2794	1.4690 1.3993	.002	1.1507	28339	1.3920
8	Cialuz Lenacdef	3.0049 2.0324	.0830 .1540	1.4481 2.1962	-.011	2.3505	.12524	2.2750
9	Cialuz Gasol 1	3.2799 2849	.0701 .5935	1.4708 1.2977	.003	23346	62897	1.26412
10	Cialuz Lpintur 1	3.1246 3.5573	.0763 .0593	1.4610 6.114	-.002	3.67965	.05508	.60911
11	Cialuz Rxadne	3.1346 1.9299	.0766 .1648	1.4591 1.5707	-0.004	2.04620	.15259	1.58438
12	Cialuz Cables 1 Insect 1 Lecolol 1 Lpintur 1 Hospit 1a	3.9164 13.1947 4.2812 1.2265 3.3277 3.7525	.0478 .0003 .0385 .2681 .0681 .0527	1.5528 4.5812 3.4516 3.5148 6.093 5.462	.056	16.124 4.9916 0.254 2.7101 3.6796 4.2949	.00006 .0254 .0997 .0550 .0550 .0382	4.619 3.4181 5.1098 6.0911 6.0911 5.5875
13	Cialuz Cables 1 Insect 1 Lecolol 1 Lpintur 1 Hospit 1a Thiner 1 Gasol 1 Lenacdef Gasol 1	3.4661 13.022 4.9642 1.1631 3.3666 4.1775 .0823 1.1280 2.7503 .4835	.0625 .0003 .0259 .2808 .0685 .0410 .7742 .2882 .0972 .4868	1.5178 4.5748 3.8856 3.4994 5.7766 5.404 8.832 1.4299 2.6564 1.4802	.034	16.124 4.9916 0.254 2.7101 3.6796 4.2949 1.599 1.1507 2.3505 2.334	.00006 .0254 .0997 .0550 .0382 .0382 .6892 2833 1.252 .6289	4.619 3.4181 5.1098 6.0911 6.0911 5.5875 1.1125 1.3920 2.2750 1.2641
14	Cialuz Cables 1 Insect 1 Lecolol 1 Lpintur 1 Hospit 1a Thiner 1 Gasol 1 Lenacdef Gasol 1	3.6151 12.6412 4.7385 1.0685 3.5677 4.4437 7.0631 1.0819 3.1130 .4685	.0573 .0004 .0295 .3013 .0589 .0350 .8017 .2983 .0777 .4937	1.5444 4.5212 3.8367 3.3450 5.665 5.260 8.064 1.4691 2.8431 1.4745	.035	16.124 4.9916 2.7101 3.6796 4.2949 1.599 1.1507 2.350 2.334	.00006 .0254 .0997 .0550 .0382 .6892 2833 1.252 .6289	4.619 3.4181 5.1098 6.0911 5.5875 1.1125 1.3920 2.2750 1.2641

Sexo	.2192	.6396	1.1116		.0981	.7540	.93660
Transf1	.0924	.7612	.9263		.1236	.7251	1.0859
Subest1	.0344	.8529	.9406		.0000	1.000	1.0000
Solven1	.0175	.8947	1.03493		.0262	.8712	1.0539
Pintur1	.0049	.9444	.9777		.0520	.8185	1.0540
Rxmadne	2.2955	.1298	1.2989		2.046	.1525	1.5843

Considerando el primer método se puede verificar en la tabla 1 que para los primeros 11 modelos no hay efecto relevante sobre la variable base cuando se agraga una variable mas al modelo y los datos obtenidos son similares a los del análisis bivariado. Para el modelo 12 donde se incluyen todas las variables que fueron significativas en el análisis bivariado, el incremento en la variable base es de solo .056, sin embargo todas las variables siguen siendo significativas.

Para el modelo 13 cuando se agregan las de riesgo alto la diferencia es de .034, y se puede ver que tampoco resultan significativas sin embargo la relación en riesgo se mantiene respecto al análisis bivariado. En el modelo 14 al agregar las variables que no están asociadas con la enfermedad, no se encontró ningún efecto relevante respecto a la variable base pues la diferencia es de sólo .035, esto quiere decir que no hay confusión o interacción entre las variables.

Tabla 2. Comparación de resultados usando el método de regresión hacia adelante

Análisis multivariado						Análisis bivariado		
Valores obtenidos						Valores obtenidos		
Modelo No.	Variable(s)	Estadístico	Sig.	Exp(B)	Diferencia	Estadístico	Sig.	Razón de momios
1	Cialuz	4.3383	.0373	1.5847	.075	3.2391	.0719	1.465
	Cables1	14.4889	.0001	4.9092		16.124	.00006	4.619
	Insect1	4.0930	.0431	3.3518		4.9916	.0254	3.4181
	Hospita	4.1695	.0412	5.490		4.2949	.0382	.5587

En el método de regresión hacia adelante se puede ver que de todas las variables que se definieron en el modelo 1, se fueron agregando las variables de una en una considerando el nivel de significancia de mayor a menor por lo que en el modelo resultante no aparecen ya las variables Lcosolv1 y Lpintur1, la primera no es significativa y la segunda al 95% tampoco, aunque muy marginalmente. De este modelo se puede comentar que aunque todas son significativas el riesgo de ocurrencia de la enfermedad es muy alto para las variables relacionadas con la exposición a campos electromagnéticos y cables de alta tensión.

Ahora bien si agregamos las variables que resultaron tener un OR (razón de momios) alto en el análisis bivariado aunque no fueron significativas, los resultados para el modelo ajustado son los mismos que en el primer modelo, es decir el modelo finalmente considera solo las significativas.

Tabla 3. Comparación de resultados usando el método de regresión hacia atrás

Modelo No.	Variable(s)	Análisis multivariado Valores obtenidos				Análisis bivariado Valores obtenidos		
		Estadístico	Sig.	Exp(B)	Diferencia	Estadístico	Sig.	Razón de momios
1	Cialuz	4.1506	.0416	1.5716	.067	3.2391	.0719	1.465
	CablesI	13.9616	.0002	4.7640		16.1245	.00006	4.619
	InsectI	4.3849	.0363	3.5080		4.9916	.02547	3.4181
	LpinturI	3.1500	.0759	.6191		3.6796	.05508	.6091
	Hospita	4.0227	.0449	.5531		4.2949	.03823	.5587
2	Cialuz	3.7371	.0532	1.5387	.047	3.2391	.0719	1.465
	CablesI	13.9342	.0002	4.7841		16.1245	.00006	4.619
	InsectI	4.5898	.0322	3.6384		4.9916	.02547	3.4184
	LpinturI	3.3807	.0660	.6065		3.6796	.05508	.6091
	Hospita	4.6793	.0305	.5223		4.2949	.03823	.5587
	Lenacdef	2.8847	.0894	2.6756		2.3505	.12524	2.2750

Cuando se utiliza el método de regresión hacia atrás vemos que a diferencia del modelo 1 en el método de regresión hacia delante, se incluye la variable LpinturI en el modelo ajustado cuya significancia es .0759, y solo se elimina la variable de exposición a solventes derivados del petróleo. Así mismo vemos que el comportamiento de las variables es similar que el obtenido en el análisis bivariado y el efecto sobre la variable base es de sólo .067.

En el segundo modelo se agregaron las variables que tienen un OR (razón de momios) considerable y vemos que los resultados obtenidos varían respecto al modelo 1 por la inclusión de la variable Lenacdef (defecto fisiológico en el niño) pues aunque tiene una significancia de .0894 también es considerada en el modelo. El valor modificado sobre la variable base es de .047 por lo tanto el haber agregado estas variables no repercute en el ajuste de los parámetros. Es importante comentar que para estos dos métodos no se estimaron modelos con las variables que no están asociadas a leucemia pues los resultados serían los mismos, dado que el método de regresión hacia delante toma en cuenta en cada iteración sólo las variables significativas, mientras que el método de regresión hacia atrás elimina en cada iteración las no significativas.

5.2 Selección del modelo

Antes de determinar cual modelo es el mas apropiado para el estudio, se puede decir que se obtuvo un buen ajuste en los datos ya que los resultados obtenidos son consistentes con los del análisis bivariado.

Ahora bien, considerando los tres métodos de estimación para los modelos, podemos ver que aunque siguen un criterio diferente, se obtuvieron resultados similares, en este sentido podemos decir que para el primer caso, como el método considera incluir todas las variables en la primera iteración, el modelo que se seleccione dependerá del objetivo del

estudio, bajo esta consideración se propone el modelo 13 como el que mejor se ajusta a los datos ya que considera tanto las variables significativas como aquellas que aunque no fueron significativas si tienen un riesgo alto. El modelo 14 se elimina dado que contiene variables que no aportan nada al modelo puesto que no se encontró asociación alguna, en este sentido se tendrían parámetros de más.

Considerando el segundo y tercer método de estimación, se considera al modelo 2 del tercer método como el que mejor se ajusta a los datos, debido a que incluye la variable *Lenacdef*, que como se había mencionado al inicio del capítulo 3, el defecto físico del menor tal como el Síndrome de Down puede ser un factor de riesgo importante en el posible desarrollo de leucemia.

CONCLUSIONES

CONCLUSIONES

Para mostrar la importancia del presente trabajo se tomarán en cuenta los aspectos más interesantes que se pudieron observar en las diferentes etapas del proceso de análisis de los datos. En primer lugar al analizar las características generales de los 374 niños incluidos en el estudio se pudo observar que la edad promedio se sitúa entre los 6 y 7 años, el 28.34% de los niños han tenido familiares con cáncer, y la característica interesante es que la mayoría de éstos familiares desarrolló cáncer cervicouterino, leucemia y cáncer de estomago. En general se encontró frecuencia de exposición alta para transformadores de luz, cables de alta tensión y torres de alta tensión. Respecto a la escolaridad de los padres se encontró que la mayoría sólo tienen primaria o secundaria. En general aunque para algunas variables sobre exposición a factores de riesgo se obtuvieron frecuencias bajas, si se encontró asociación con la leucemia.

En el análisis bivariado no se encontró asociación estadísticamente significativa en lo referente al sexo del paciente, exposición a transformadores de luz, subestaciones eléctricas y torres de alta tensión, sin embargo se pudo observar en el análisis más detallado sobre éstos factores, que el riesgo de desarrollar la enfermedad puede incrementarse con el tiempo de exposición. En lo referente a la exposición a productos químicos no se encontró asociación con la exposición del padre a solventes derivados del petróleo y pinturas, pero es importante comentar que si se encontró un riesgo alto para los expuestos a thinner y gasolinas. En lo que se refiere a los factores de riesgo prenatales no se encontró asociación entre la exposición de la madre a ultrasonidos pero si se encontró un riesgo mayor para las que estuvieron expuestas a rayos x. Cuando se realizó el análisis respecto a los antecedentes natales y durante los primeros años de vida se encontró que el nacer con alguna malformación es un factor de riesgo para el desarrollo de leucemia pues aunque de los 374 niños sólo el 4.27% nació con algún defecto, el riesgo fue de 2.72. Así mismo se encontró asociación entre los niños que se hospitalizaron el primer año y desarrollaron leucemia.

Al aplicar el método de regresión logística no condicionada, para las variables con una significancia menor a .05 obtenida en el análisis bivariado, permitió confirmar que las variables más significativas que están asociadas con el desarrollo de leucemia son la exposición a campos electromagnéticos, exposición a cables de alta tensión, el nacer con alguna malformación, el haber estado hospitalizado en el primer año de vida y la exposición del menor a insecticidas. Se puede decir que hubo consistencia en los datos ya que independientemente del método de regresión utilizado, se obtuvieron resultados similares.

Así mismo, mediante el análisis de regresión se pudo observar en los resultados obtenidos que no hubo confusión o interacción entre las variables ya que el efecto obtenido al agregar las diferentes variables en la estimación de los modelos, sobre la variable base fue muy pequeño en la mayoría de los casos y, en aquellos donde se obtuvo un riesgo protector como fue el caso de los niños que estuvieron hospitalizados el primer año de vida, se pudo aclarar con la literatura médica que es normal, aunque en el caso de la exposición a pinturas o algún otro que aparentemente resultó ser un riesgo protector, sería necesario realizar mayores investigaciones.

Se puede comentar que la pequeña aportación contenida en esta investigación es útil no solo porque es un material que puede servir como información para la población, sino también porque puede servir como base de comparación para otras investigaciones.

En general podemos señalar que se cumplió con el objetivo planteado ya que se pudieron encontrar factores de riesgo importantes que se encuentran presentes en la población pediátrica y que están asociados al desarrollo de leucemia.

Finalmente, es apropiado comentar que fue una satisfacción personal ya que tuve la oportunidad de aplicar los conocimientos adquiridos y de darme cuenta que tan importante es el perfil de nuestra carrera.

APÉNDICE

APÉNDICE

TABLA A. CODIFICACIÓN DE VARIABLES

VARIABLES GENERALES			
No.	NOMBRE	DESCRIPCIÓN	ESCALA DE MEDICIÓN
1	SEXO	Sexo	1 = Masculino 0 = Femenino
2	EDAD	Edad del niño	Se registrará en número de meses
3	EDADNAC	Mes de nacimiento	Se tomará el número de mes en que nació el niño en estudio
4	DOMIC	Domicilio	Se clasificará por municipio según la siguiente tipificación: 01 = Alvaro Obregón 02 = Arcatazaco 03 = Benito Juárez 04 = Coyoacán 05 = Cuauhtémoc 06 = Gustavo A. Madero 07 = Izacalco 08 = Iztapalapa 09 = Miguel Hidalgo 10 = Tlalpan 11 = Venustiano Carranza 12 = Xochimilco 13 = Magdalena Contreras 14 = Cuajimalpa 15 = Tlalhuac 16 = Milpa Alta 17 = Ecatepec 18 = Naucalpan 19 = Cuahuhtlán Izcalli 20 = Nezahualcoyotl 21 = Tlalapanilla 22 = Coacalco 23 = Tultitlán 24 = Chalco 25 = Chimalhuacán 26 = Colima 27 = San Martín Mahinalco 28 = Tulpetlac 29 = Acólman 30 = Zumpango 31 = Cd. Azteca 32 = Valle de Aragón 33 = Xalostoc 34 = Atzacapan de Zaragoza 35 = Texcoco 36 = Tepic 37 = Villa Nicolás Romero 38 = Cuatzingo 39 = Cuau. de Romero Rubio 40 = Huixquilucan 41 = Tecamachalco 42 = Coyoacán 43 = Chiconcoac 44 = Chicoloapan 45 = Chihuahua 46 = Los Reyes 47 = Ixtapalaca 48 = Tultitlán 49 = Tepic 50 = Ayotlán 51 = Toluca 52 = Cuahuhtepic 53 = Tequisquiac 54 = Tulpetlac 55 = Teoloyucan 56 = Tequesquahuac
5	PARENTLS	Parentesco del entrevistado	Se codificará de acuerdo a lo siguiente: 1 = Abuelo 2 = Padre 3 = Madre 4 = Hermano 5 = Tío 6 = Otro

ANTECEDENTES HEREDO FAMILIARES			
No.	NOMBRE	DESCRIPCIÓN	ESCALA DE MEDICIÓN
6	TDANCLR	Variable para conocer si alguno de los familiares del niño en estudio ha padecido cáncer	1 = sí 0 = no
7	CUANTOS1	Número de familiares que han padecido cáncer	Se tomará en relación al número que proporcione el entrevistado
8	TCAN1	Parentesco del familiar con cáncer respecto al niño en estudio	Se especificará de acuerdo a la siguiente clasificación: 1 = Abuelo 2 = Padre 3 = Madre 4 = Hermano
9	TCAN1	Tipo de cáncer	Se codificará según la siguiente clasificación: 10 = Leucemia 11 = Esófago 15 = Ovario 16 = Testicular

			<p>Polución:</p> <p>00 = Desconocido 01 = Leucemia 02 = Cervicouterino 03 = Pulmón 04 = Ovario 05 = Mama 06 = De piel 07 = Neuroblastoma 08 = S. N. C. 09 = Laringe</p>	<p>12 = Faringe 13 = L. de Hoshaki 14 = Páncreas 15 = Oso 16 = Testicular 17 = Tiroides 18 = Intestinal 19 = Leucemia 20 = Hígado 21 = Ojo 22 = Encef. 23 = Lengua 24 = Cerebro 25 = Sarcoma 26 = Oído</p>	<p>17 = Tiroides 18 = Intestinal 19 = Hígado 20 = Hepático 21 = Ojo 22 = Encef. 23 = Lengua 24 = Cerebro 25 = Sarcoma 26 = Oído</p>	Alveolar
10	DEFECEN	Variable para ver si el niño tiene familiares con algún defecto físico	1 = si 0 = no			
11	CUANTOS2	Número de familiares	Se nomara según el número que proporcione el entrevistado			
12	IDEP1	Parentesco del familiar que tiene el defecto físico respecto al niño	Se hará de acuerdo a la siguiente tipificación	1 = Abuelo 2 = Padre 3 = Madre 4 = Hermano	5 = Tio 6 = Otros	
13	IDEP1	Tipo de defecto	00 = Desconocido 01 = Pie plano 02 = Cuerdas bucales 03 = Cardiopatía congénita 04 = Hidrocefalia 05 = Labio y paladar hendido 06 = Anencefalia 07 = Ceguera 08 = Enanismo 09 = Oligotemia 10 = Parálisis Cerebral 11 = Hombro pegado al cuello 12 = Acortamiento del brazo 13 = Sordomudo 14 = Acortamiento de cuello 15 = Polidactilia	16 = Soplo en el corazón 17 = Atrofia muscular 18 = Convulsiones 19 = Ano imperforado 20 = Reflujo 21 = Testículo 22 = Malformación intestinal 23 = Extravismo 24 = Tumor en el cerebro 25 = Malformación en oído 26 = Meningitis 27 = Brazo incompleto 28 = Obstrucción pilórica		
14	DOWNOMU	Variable para identificar si algún familiar tuvo Síndrome de Down o Mongolismo	1 = si 0 = no			
15	DOWNPAREN	Parentesco respecto al niño	1 = Abuelo 2 = Padre 3 = Madre 4 = hermano 5 = Tio			

ANTECEDENTES PERSONALES NO PATOLOGICOS

16	CASA	Característica de la casa	1 = Propia 2 = Rentada 3 = Prestada 4 = Desconoce
17	RENTAS	Precio de la renta	Se registrará en miles de pesos
18	CASILEN	Servicios que tiene la casa	1 = Agua 2 = Luz 3 = Luz y Agua 4 = Drenaje 5 = Agua y Drenaje 6 = Luz y Drenaje 7 = Agua, Luz y Drenaje
19	CUARTOS*	Número de cuartos para dormir	Se tomara en cuenta el total de cuartos exclusivos para dormir
20	DUERMIEN*	Número de personas que duermen en cada cuarto	Se tomara de acuerdo al dato que proporcione
21	NIOPERS	Total de personas que habitan en la casa	Se tomara el total de individuos que viven en la casa
22	INGRESO	Ingreso familiar mensual	Se codificará en miles de pesos
23	ESTRATFI	Estrato socioeconómico	1 = Alto y Muy Alto 2 = Medio Normal y Medio Bajo 3 = Medio Normal y Medio Bajo 4 = Popular
24	PARRU2	Precio mensual por pago de luz	Se tomara en miles de pesos
25	COCALEAL	Equipo para calentar los alimentos	01 = Gas 02 = Pétroleo 03 = Carbon 04 = Parral 05 = Electrica 16 = Horno electrico 32 = Horno de microondas ed = Leña
26	TVINLAC	Tiempo que tiene la persona de residir en su domicilio	Calculado en meses
27	TRANN1	Transformador de luz	1 = si 0 = no
28	DIRANS1	Distancia respecto a su casa	El dato se registrará en metros
29	CABLET	Cables de alta tensión	1 = si 0 = no

30	DCABLES1	Distancia	El dato se registrara en metros
31	SUBEST1	Subestacion electrica	1 = si 0 = no
32	DISUBEST1	Distancia	Se medira en metros
33	TORRES1	Torres de alta tension	1 = si 0 = no
34	DTORRES1	Distancia	Se medira en metros
35	VAENIS1	Ya existian cuando llego a vivir a ese lugar	1 = si 0 = no
36	TIEMPO1	Cuanto tiempo despues fueron colgarlos	El dato se registrara en meses
37	DESEHAI1	Si ya existian se preguntara el tiempo	El dato se registrara en meses
38	SUBVIAQ	Variable para saber si ha vivido siempre en el mismo lugar	1 = si 0 = no
39	TRANS1	Si vivia en otro lugar se investigara la existencia de Transformadores de luz	1 = si 0 = no
40	DTRANS1	Distancia	Se medira en metros
41	CCABLES2	Cables de alta tension	1 = si 0 = no
42	DCABLES2	Distancia	Se medira en metros
43	SUBEST2	Subestacion electrica	1 = si 0 = no
44	DISUBEST2	Distancia	Se medira en metros
45	TORRES2	Torres de alta tension	1 = si 0 = no
46	DTORRES2	Distancia	Se medira en metros
47	VAENIS2	Variable para saber si cuando llego a vivir alli ya existian	1 = si 0 = no
48	TIEMPO2	Cuanto tiempo despues los colgaron	Se medira en meses
49	TIEMPO2	Tiempo de residir en ese lugar domcilio	Se medira en meses

ESCOLARIDAD DEL NIÑO

50	GUARDER	Variable para saber si el niño asisto a la guarderia	1 = si 0 = no
51	EDADMILAK	Ejald de ingreso a la guarderia	Se medira en meses
52	ESCEVA	Variable para conocer si el niño va a la escuela	1 = si 0 = no
53	TRANS1	Expuesto a transformador de luz	1 = si 0 = no
54	DTTRANS1	Distancia	El registro se hara en metros
55	CABLES1	Expuesto a cables de alta tension	1 = si 0 = no
56	DCABLES1	Distancia	El registro se hara en metros
57	SUBEST1	Expuesto a subestacion electrica	1 = si 0 = no
58	DISUBEST1	Distancia	El registro se hara en metros
59	TORRES1	Expuesto a torres de alta tension	1 = si 0 = no
60	DTORRES1	Distancia	El registro se hara en metros
61	ANOSESC	Número de años en esa escuela	Se medira en meses
62	HORASESC	Tiempo que permanencia en la escuela	Horas
63	TORNOS1	Turno	1 = Matutino 2 = Vespertino 3 = Nocturno 4 = Mixto
64	OTROESC1	Variable para conocer si estuvo en alguna otra escuela	1 = si
65	TRANS2	Expuesto a transformador de luz	1 = si
66	DTTRANS2	Distancia	Se medira en metros
67	CABLES2	Expuesto a cables de alta tension	1 = si
68	DCABLES2	Distancia	Se medira en metros
69	SUBEST2	Expuesto a Subestacion electrica	1 = si
70	DISUBEST2	Distancia	Se medira en metros
71	TORRES2	Expuesto a torres de alta tension	1 = si
72	DTORRES2	Distancia	Se medira en metros
73	TIEMPO3	Tiempo que estuvo en esa escuela	Meses
74	HORAS2	Tiempo que permanencia en esa escuela	Horas
75	TORNOS2	Turno	1 = Matutino 2 = Vespertino 3 = Nocturno 4 = Nocturno 5 = Mixto

76	ESCOLAR1	Escolaridad del padre	1 = Analfabeta 2 = Primaria 3 = Secundaria 4 = Bachillerato	5 = Técnico 6 = Lic. o Posgrado
77	ANOCUR1	Número de años cursados	Se registrará de acuerdo al dato proporcionado por el entrevistado	
78	ESCOLARM	Escolaridad de la madre	Se clasificara de acuerdo a	1 = Analfabeta 2 = Primaria 3 = Secundaria 4 = Bachillerato
79	ANOCUR2	Número de años cursados	Se registrará de acuerdo al dato proporcionado por el entrevistado	

OCUPACIÓN

80	OCUPALMI	Ocupación del padre cuando se embarazo la madre del niño en estudio	1 = Administrativo 3 = Servicio de limpieza o intendencia 5 = Hogar 7 = Obrero 9 = Repartidor y/o mensajería 11 = Químico	2 = Agrícola 4 = Construcción 6 = Comercio 8 = Técnico o Ing 10 = Profesional 12 = Ninguna Activ
81	ACTIV1	Actividad realizada	Se tomara de acuerdo a la respuesta que proporcione el individuo	
82	FABR1P1	Cargos de la empresa	Se tomara de acuerdo a la respuesta que proporcione el individuo	
83	TAR1AP1	Tarea mas importante que desempeñaba	Se tomara de acuerdo a la respuesta que proporcione el individuo	
84	SOLVEN1	Exposición a Solventes	1 = si 2 = mucho	0 = no 3 = regular
85	FRENO1	Frecuencia	1 = mucho 2 = poco	0 = no 3 = regular
86	HORSO1	No. de horas	Se tomara de acuerdo a la respuesta que proporcione el individuo	
87	INACT1	Exposición a insecticidas	1 = si 2 = mucho	0 = no 3 = regular
88	FREINS1	Frecuencia	1 = mucho 2 = poco	0 = no 3 = regular
89	HORINS1	No. de horas	Se tomara de acuerdo a la respuesta que proporcione el individuo	
90	THINR1	Exposición a thimer	1 = si 2 = mucho	0 = no 3 = regular
91	FREINS1	Frecuencia	1 = mucho 2 = poco	0 = no 3 = regular
92	HORINS1	No. de horas	Se tomara de acuerdo a la respuesta que proporcione el individuo	
93	PINT1R1	Exposición a pinturas	1 = si 2 = mucho	0 = no 3 = regular
94	FREPS1	Frecuencia	1 = mucho 2 = poco	0 = no 3 = regular
95	HORINS1	No. de horas	Se tomara de acuerdo a la respuesta que proporcione el individuo	
96	FERT1L1	Exposición a fertilizantes	1 = si 2 = mucho	0 = no 3 = regular
97	FREPS1	Frecuencia	1 = mucho 2 = poco	0 = no 3 = regular
98	HOR1R1	No. de horas	Se tomara de acuerdo a la respuesta que proporcione el individuo	
99	GAOL1	Exposición a gasolina	1 = si 2 = mucho	0 = no 3 = regular
100	FRE1GAS1	Frecuencia	1 = mucho 2 = poco	0 = no 3 = regular
101	HOR1GAS1	No. de horas	Se tomara de acuerdo a la respuesta que proporcione el individuo	
102	QU1M1S1	Exposición a otros químicos	1 = si 2 = mucho	0 = no 3 = regular
103	FRE1QU1	Frecuencia	1 = mucho 2 = poco	0 = no 3 = regular
104	HOR1QU1	No. de horas	Se tomara de acuerdo a la respuesta que proporcione el individuo	
105	HS1MPS	Tiempo laborando en ese lugar	Meses	
106	EQ1PROT	Utilizaba equipo de protección para laborar	1 = si	0 = no
107	EQ1QUIP	Que equipo utilizaba	01 = Mascartilla 02 = Lentes 03 = Guantes 04 = Careta 05 = Botas 06 = Tapones 07 = Uniforme de algodón 08 = Gorro 09 = Calzabeca 10 = Zapatos con cascudillo	11 = Arma de fuego 12 = Obrero 13 = Chaleco 14 = Mascartilla y guantes 15 = Cinturón, guantes y casco 16 = Equipo para bombero 17 = Bata, guantes y lentes 18 = Lentes para sol
108	EQ1MAQ1	Equipo con el que laboraba	00 = Ninguno 01 = Papelería 02 = Camión, o camioneta 03 = Herramientas 04 = Maquina para cover zapatos 05 = Equipo medico 06 = Planta para soldar 07 = Equipo automotriz 08 = Yunque	35 = Artículos de limpieza 36 = Maquina neumatica manejada por aire 37 = Amasadoras 38 = Soldadura 39 = Folletos y propaganda 40 = Gs, borrar y pizarrón 41 = Maquina pesada 42 = Tierra s. baseado

			10 = Máquina trapeadora 11 = Máquina fresadora (torno) 12 = Máquina de construcción 13 = Máquina de montacarga 14 = Impresora y fotocopiadora 15 = Equipo para pintar 16 = Equipo para carpintería 17 = Hornos 18 = Extintores 19 = Máquina registradora y de escribir 20 = Lámparas de gasolina 21 = Proyector de películas 22 = Máquina cortadora 23 = Computadora 24 = Máquina que taladra 25 = Palas y escobas 26 = Pistolas de aire 27 = Soplete 28 = Válvulas 29 = Máquina de costura 30 = Fumigadora 31 = Máquina tejedora 32 = Compresoras 33 = Gruas y palas 34 = Máquina para procesar refresco	43 = Planchadora 44 = Diable 45 = Escritorio 46 = Herra, mano 47 = Equipo de alta tensión 48 = Equipo para zapatero 49 = Ultrasonógrafo 50 = Arado, asadores 51 = Microscopio 52 = Cuchillona 53 = Instrumentos químicos 54 = Equipo de carpintería 55 = Motor de alta velocidad 56 = Órgano musical 57 = Cinta métrica, tripe, 58 = Bascula 59 = Instrumentos médicos 60 = Hotiquin 61 = Aparatos eléctricos 62 = Cámara fotográfica
109	OTRAAC1	Otras actividades.		
110	ACTIVP2	Ocupación del padre cuando el niño enfermó de Leucemia	Se clasificara según la respuesta obtenida por el individuo 1 = Administrativo 2 = Agrícola 3 = Servicio de limpieza y/o mantenimiento 4 = Construcción 5 = Hogar 6 = Comercio 7 = Obrero 8 = Técnico e Ingeniero 9 = Reparadores y/o mantenimiento 10 = Profesionistas 11 = Químico 12 = Ninguna actividad	
111	QACTIV2	Actividad realizada con mayor frecuencia	Se clasificara según la respuesta proporcionada por el individuo	
112	CONQUIM1	Estuvo en contacto con químicos	1 = si 0 = no	
113	CUALEN1	Tipo de químicos.	01 = Azufre 02 = Saca 03 = Clinko y pino 04 = Soldadura 05 = Resinas 06 = Alcohol 07 = Etileno, óxido y cloruro 08 = Pegamento 500X 09 = Yeso 10 = Insecticidas 11 = Fertilizantes 12 = Ácidos 13 = Cal 14 = Tintes para el cabello 15 = Fijadores para el cabello 16 = Solventes derivados del petróleo 17 = Sulfato de aluminio 18 = Líquido para revelar rollos fot. 19 = Selladores, lacas 20 = Productos para hacer hule 21 = DDT 22 = Enzimas 23 = Aluminio 24 = Substancias químicas 25 = Aceites para transformadores 26 = Fenólicos 27 = Silicon 28 = Aromatizantes	
114	ACTIVM1	Ocupación de la madre antes del embarazo del niño	Se tomara la misma tipificación que para la ocupación del padre	
115	HORASM1	Horas.	Se tomara en base al número de horas laboradas.	
116	ACTIVM2	Ocupación durante el embarazo	Se tomara la misma tipificación que para la ocupación del padre	
117	HORASM2	No. de horas.	Se tomara de acuerdo a la respuesta que proporcione el individuo	
118	CONQUIM2	Durante el embarazo hubo contacto con químicos.	1 = si 0 = no	
119	CUALEN2	Tipo de químicos.	Se utilizara la misma tipificación que para el de los padres.	
120	ACTIVM3	Actividad de la madre cuando el niño se enfermó de Leucemia	Se tomara como patrón el mismo que se describió para la ocupación del padre.	
121	HORASM3	Horas.	Se tomara en cuenta el número de horas laboradas.	

ANTECEDENTES PATOLÓGICOS
ALCOHOLISMO

122	ALCOLPI	El esposo consume bebidas alcohólicas	1 = si	0 = no
123	EDADALPI	Edad en la que se inició el consumo de alcohol	Años	
124	EMBALEPI	Tomaba durante el embarazo	1 = si	0 = no
125	FREALEPI	Frecuencia	01 = 1 vez al año 02 = Ocasionalmente 03 = Diariamente 04 = Cada 8 días 05 = Cada 20 días 06 = Cada mes	07 = Dos veces al mes 08 = Cada 2 meses 09 = Cada tres meses 10 = Cada 4 meses 11 = Cada 6 meses
126	QMBIDAP	Bebida de mayor frecuencia	1 = Vino 2 = cerveza 3 = Brandy 4 = Ron 5 = Tequila	6 = Pulque 7 = Otros
127	ALCOLMI	La madre consume bebidas alcohólicas	1 = si	0 = no
128	EDADALMI	Edad en que se inició el consumo de alcohol	Años	
129	QMBIDAM	Bebida de mayor consumo	1 = Vino 2 = Cerveza 3 = Brandy 4 = Ron 5 = Tequila	6 = Pulque 7 = Otros
130	EMBORMI	Se emborrachaba	1 = si	0 = no
131	FREALEMI	Cuántas veces al año	01 = 1 vez al año 04 = Cada 8 días 07 = Dos veces al mes 10 = Cada 4 meses	02 = Ocasionalmente 05 = Cada 20 días 08 = Cada 2 meses 11 = Cada 5 meses
132	ELLEBAMI	Variable para ver si consumo bebidas alcohólicas durante el embarazo	1 = si	0 = no
133	QMBIDMAI	Mes del embarazo	1 = Primero 4 = Cuarto 7 = Séptimo 10 = Todo el embarazo	2 = Segundo 5 = Quinto 8 = Octavo 9 = Noveno
134	FRCPASMI	No de copas que ingería cada vez	Se considera el total aproximado de copas	

TABAQUISMO

135	TUMAPI	Variable para saber si el padre del menor fuma actualmente	1 = si	0 = no
136	TUMMEMI	El esposo fumo durante el embarazo del niño	1 = si	0 = no
137	ELRCADMI	Variable para saber si durante el embarazo de la madre, el padre fumaba cerca	1 = si	0 = no
138	QMBIDMA2	Mes del embarazo en que fumo mas	1 = Primero 4 = Cuarto 7 = Séptimo 10 = Todo el embarazo	2 = Segundo 5 = Quinto 8 = Octavo 9 = Noveno
139	FRUGIDA	No de cigarrillos al día	Se registrara segun el total aproximado de cigarrillos	
140	TUMALMI	Variable para saber si la madre fumaba antes de su embarazo	1 = si	0 = no
141	TUMODMI2	Variable para ver si fumo durante su embarazo	1 = si	0 = no
142	MILUNIAS	Mes del embarazo en que fumo mas	1 = Primero 4 = Cuarto 7 = Séptimo 10 = Todo el embarazo	2 = Segundo 5 = Quinto 8 = Octavo 9 = Noveno
143	FRUGIDA	No de cigarrillos al día	Se registrara segun el total aproximado de cigarrillos	
144	TUMACEPI	Variable para saber si la madre fuma actualmente	1 = si	0 = no

OTROS ANTECEDENTES DEL MENOR				
145	LEDRX1	Radiografías tomadas antes del diagnóstico de leucemia	1 = si 0 = no	
146	FILELRX1	No. de radiografías	Se tomará según la respuesta proporcionada	
147	LEDAFRX1	Edad del menor cuando ve leonno alguna radiografía	Meses	
148	LEULU1	Ultrasonidos tomados antes del diagnóstico de leucemia	1 = si 0 = no	
149	FILEUL1	No. de ultrasonidos	Se registrará según la respuesta proporcionada	
150	NIHTUL1	Metastasis del ultrasonido	01 = Tumor en boca 02 = Para diagnóstico de leucemia 03 = Dolor de estómago 04 = Presencia de ganglios* en el cuello 05 = Detección de soplo en corazón 06 = For pancreatitis 07 = Para conocer estado de salud de órganos internos 08 = Dependientes	09 = Amígdalas 10 = Hepatitis* 11 = Atresia esofágica 12 = Cefalaea 13 = Probable tumor de riñon 14 = Insuficiencia respiratoria
151	LENADEL1	El niño nació con algún defecto	1 = si 0 = no	
152	TIPODEF1	Tipo de defecto	01 = Pie plano 02 = Enfermedad de Olfier 03 = Sopllo en el corazón 04 = Síndrome de Down 05 = Hipertemia 06 = Atresia esofágica 07 = Ganglios cervicales 08 = Extravisión 09 = Hernia inguinal 10 = Disfonia 11 = Oligofrenia 12 = Hemiplegia	13 = Labio y paladar hendido 14 = Cretinismo 15 = Displasia ósea 16 = Raquitismo 17 = Reflujo 18 = Síndrome de Núnson 19 = Inmadurez de cartilago 20 = Dientes desviados 21 = Anmia
153	LECONOVI	Exposición a solventes	1 = si 0 = no	
154	FRISOEL1	Frecuencia	1 = mucho 2 = poco 3 = regular	
155	FRISOEL1	No. de horas	Se registrará de acuerdo al dato proporcionado por el entrevistado	
156	FRINSUL1	Exposición a insecticidas	1 = si 0 = no	
157	FRINSUL1	Frecuencia	1 = mucho 2 = poco 3 = regular	
158	FRINSUL1	No. de horas	Se registrará de acuerdo al dato proporcionado por el entrevistado	
159	TIPODEF1	Exposición a thiner	1 = si 0 = no	
160	FRINSUL1	Frecuencia	1 = mucho 2 = poco 3 = regular	
161	FRINSUL1	No. de horas	Se registrará de acuerdo al dato proporcionado por el entrevistado	
162	FRINSUL1	Exposición a pinturas	1 = si 0 = no	
163	FRINSUL1	Frecuencia	1 = mucho 2 = poco 3 = regular	
164	FRINSUL1	No. de horas	Se registrará de acuerdo al dato proporcionado por el entrevistado	
165	FRINSUL1	Exposición a fertilizantes	1 = si 0 = no	
166	FRINSUL1	Frecuencia	1 = mucho 2 = poco 3 = regular	
167	FRINSUL1	No. de horas	Se registrará de acuerdo al dato proporcionado por el entrevistado	
168	GASOEL1	Exposición a gasolina	1 = si 0 = no	
169	FRINSUL1	Frecuencia	1 = mucho 2 = poco 3 = regular	
170	FRINSUL1	No. de horas	Se registrará de acuerdo al dato proporcionado por el entrevistado	
171	COLOREL1	Exposición a colorantes para el cabello	1 = si 0 = no	
172	FRINSUL1	Frecuencia	1 = mucho 2 = poco 3 = regular	
173	FRINSUL1	No. de horas	Se registrará de acuerdo al dato proporcionado por el entrevistado	
174	FRINSUL1	Exposición a tintadores para el cabello	1 = si 0 = no	
175	FRINSUL1	Frecuencia	1 = mucho 2 = poco 3 = regular	
176	FRINSUL1	No. de horas	Se registrará de acuerdo al dato proporcionado por el entrevistado	
177	FRINSUL1	Otros	1 = si 0 = no	
178	FRINSUL1	Frecuencia	1 = mucho 2 = poco 3 = regular	
179	FRINSUL1	No. de horas	Se registrará de acuerdo al dato proporcionado por el entrevistado	
180	ENANLI1	Variable para ver si algún niño en la muestra de casos padeció alguna enfermedad	1 = Varicela 2 = Rubéola 3 = Paperas 4 = Sarampión 00 = Ninguna	

181	ENANCI	Variable para ver si algun niño en la muestra de controles padeció alguna enfermedad	1 = Varicela 8 = Paperas	2 = Rubéola 00 = Ninguna	4 = Sarampión
182	FULENFI	El niño fue muy enfermizo en el primer año de vida	1 = si	0 = no	
183	FUEMQI	Fue más enfermizo que sus otros hermanos	1 = si	0 = no	
184	HOSPITIA	Durante el primer año de vida se enfermó en algun hospital por algun padecimiento	1 = si	0 = no	
185	PADICINI	Tipo de padecimiento	00 = Se ignora 01 = Enfermedades respiratorias 02 = Circuncisión 03 = Diarreas 04 = Transfusiones sanguíneas 05 = Bajo peso al nacer 06 = Fiebre tifoidea 07 = Hematúria 08 = Intoxicación medicamentosa 09 = Convulsiones 10 = Sople en el corazón 11 = Traumatismos 12 = Malformaciones 13 = Cardiopatías congénita 14 = Problemas testiculares 15 = Sordomudo 16 = Parálisis cerebral 17 = Atrofia Muscular 18 = Artritis Cardíaca 19 = Quemaduras 20 = Anomaliada 21 = Meningoencefalitis 22 = Intolerancia a la lactosa 23 = Reflujo gastroesofágico 24 = Anencefalia 25 = Hidrocefalia 26 = Hipotermia 27 = Ictericia 28 = Hipertermia 29 = Infección de cordón umbilical 30 = Infección de ojos y garganta 31 = Alergia a medicamentos 32 = Infección en la sangre 33 = Cianodermia 34 = Arteria esofágica 35 = Sarampión 36 = Síndrome de Kinsbore 37 = Tensión del lagrimal tapado 38 = Labios y paladar hendido 39 = Edema facial		

ANTECEDENTES GINECOOBSTETRICOS Y PATOLÓGICOS DE LA MADRE

186	NOEMBAR	No de embarazos que ha tenido	Se tomara de acuerdo al numero total de embarazos		
187	HDOLGI	No. de hijos que ha tenido	Se tomara el total de embarazos menos los abortos		
188	NOABORT	No. de abortos	Se tomara según el dato proporcionado		
189	CAUSASBOR	Hubo algun motivo	1 = si	0 = no	
190	CAUSAT	Causa	01 = Se ignora 02 = Embarazo extrauterino 03 = Caidas 04 = Levantar objeto pesado 05 = Golpes 06 = Debilidad de la matriz 07 = Alteraciones circulatorias 08 = Provocado por fuerbas 09 = Hemorragias 10 = Consumo de pastillas 11 = Sarampión 12 = DIU (dispositivo intra uterino) 13 = Infección de vías urinarias 14 = Infección 15 = Accidentes 16 = Evacuante 17 = Llorera en el cuello de la matriz 18 = DPP (desprendimiento de placenta previa)		
191	ABORANNE	Hubo abortos antes del niño enfermo	1 = si	0 = no	
192	NABORANE	Numero	Se tomara de acuerdo al dato proporcionado por el entrevistado		
193	NMUEBAL	Algun hijo nacio muerto	1 = si	0 = no	
194	CUANACM	Numero	Se tomara de acuerdo al dato proporcionado por el entrevistado		
195	ALFUEANE	Alguno fue antes del niño enfermo	1 = si	0 = no	
196	RXMADNE	A la madre le tomaron alguna radiografía durante el embarazo	1 = si	0 = no	
197	RXMADNEI	Numero	Se tomara de acuerdo al dato proporcionado por el entrevistado		
198	RXMEMBIA	En que mes del embarazo	Se tomara de acuerdo al dato proporcionado por el entrevistado		
199	ULMADNE	Se le tomo a la madre algun ultrasonido durante el embarazo	1 = si	0 = no	
200	NULMADNE	Numero	Para propósitos de la investigación se tomo el dato tal cual		
201	ULMOTIV	Motivo	00 = Se ignora 01 = Para conocer tiempo de gestac. de nacimiento 02 = Embarazo de alto riesgo 03 = Comprobar embarazo 04 = Amenaza de aborto 09 = Para verificar retardo de nacimiento 10 = Cálculos vesicales 11 = Para saber posición fetal 12 = Por probables quistes		

			05 = Para verificar embarazo genéar 06 = Para detectar malformaciones 07 = Por infección de vías urinarias 08 = Por probable diagnóstico de toxoplasmosis	13 = Para saber sexo de producto 14 = Hipertensión 15 = Por ausencia de movimiento fetal
202	ENFMADNI	Alguna enfermedad padecida por la madre al embarazo	1 = Rubéola 2 = Varicela 3 = Paperas 4 = Refruido	5 = Sarampión 6 = Ninguna
203	QDIAGNOS	Quien le hizo el diagnóstico	1 = Médico 2 = Auto diagnóstico	3 = Otro
204	MESIMHNE	En que mes se embarazó del niño en estudio	Se tomara en cuenta el número del mes	
205	ANODIAGL	Fecha del diagnóstico de leucemia	Se considera solo el año	
206	TIPDLU	Tipo de leucemia	1 = Leucemia linfoblástica aguda	2 = Leucemia mieloblástica
207	EDADIAGL	Edad al diagnóstico de leucemia	Meses	
208	EDMUMNI	Edad de la madre al embarazo del menor	Se consideran los años cumplidos	
209	EPADUMNI	Edad del padre al embarazo de la madre	Se consideran los años cumplidos	
210	NCEMNE	Número de embarazo	Se considera el lugar del menor en la familia	
211	PESOALNA	Peso del niño al nacimiento	Se maneja en gramos	
212	DCMNSAC	De cuantos meses nació	Se toma en cuenta el total de meses de la gestación	
213	PACPREL	El paciente por el que se pregunta tiene leucemia	1 = si	0 = no
214	DIAGINT	Diagnóstico de internamiento o motivo por el que está en el hospital	01 = Consulta 02 = Internado 03 = Aplicación de cuestionario 04 = Infección de vías respiratorias 05 = Quimioterapia 06 = Tratamiento 07 = Por medicamento 08 = Neumonía 09 = Valoración médica 10 = Mialgia 11 = Infección estomacal 12 = Irritación de piel 13 = Desnutrición 14 = Certificado médico	15 = Laboratorio 16 = Cines dental 17 = Ortopedia 18 = Cirugía maxilo-facial 19 = Leucemia linfoblástica aguda 20 = Cobro de incapacidad 21 = Visita domiciliaria 22 = Alta del paciente 23 = Fractura 24 = Vacuna 25 = Solicitud de cita
215	HOSPROC	Hospital del procedencia	01 = Centro Médico Nacional Siglo XXI 03 = Hospital Infantil de México	02 = Centro Médico la Raza
216	CIALUZ	Nivel de exposición a campos electromagnéticos	A = Alta M = Media	B = Baja

Tabla 4. Distribución de frecuencia para domicilio de los pacientes por municipio.

Valor	Frec.	Pct.	Pct. Cum.
01	21	6	16
02	21	6	32
03	9	2	34
04	26	7	41
05	11	3	44
06	42	11	55
07	11	3	58
08	11	3	61
09	6	2	63
10	11	3	66
11	4	1	67
12	1	0	67
13	1	0	67
14	1	0	67
15	1	0	67
16	1	0	67
17	1	0	67
18	21	6	73
19	14	4	77
20	14	4	81
21	26	7	88
22	8	2	90
23	2	0	90
24	2	0	90
25	1	0	90
26	1	0	90
27	1	0	90
28	1	0	90
29	1	0	90
30	1	0	90
31	1	0	90
32	1	0	90
33	1	0	90
34	1	0	90
35	2	0	90
36	1	0	90
37	1	0	90
38	1	0	90
39	1	0	90
40	1	0	90
41	1	0	90
42	1	0	90
43	1	0	90
44	1	0	90
45	1	0	90
46	1	0	90
47	1	0	90
48	1	0	90
49	1	0	90
50	1	0	90

C. de los valores: 174

Tabla 5. Distribución de frecuencia para identificar si los familiares de los pacientes han tenido algún tipo de cáncer.

Valor	Frec.	Pct.	Pct. Cum.
0	10	2	100
1	98	22.5	100

C. de los valores: 174

Tabla 6. Frecuencia para saber cuántos familiares han padecido cáncer.

Valor	Frec.	Pct.	Pct. Cum.
1	1	0	100
2	2	0	100
3	2	0	100
Total	5	0	100

C. de los valores: 174

Tabla 7. Distribución de frecuencia para el parentesco con el niño de aquellos que tienen un hijo familiar con cáncer.

Valor	Frec.	Pct.	Pct. Cum.
1	16	13.8	100
2	1	0.8	100
3	1	0.8	100
4	2	1.7	100
5	1	0.8	100
Total	21	18.1	100

C. de los valores: 174

Fuente: Unidad de Epidemiología, Hospital de Pediatría del Centro Médico Nacional Siglo XXI

Tabla 8. Distribución de frecuencia para el tipo de cáncer desarrollado

Valor	Frec.	Pct.	Pct. Cum.
01	14	13.3	13.3
02	14	13.2	26.5
03	22	20.7	47.2
04	7	6.7	53.9
05	5	4.7	58.6
06	2	1.9	60.5
07	1	1.1	61.6
08	4	3.8	65.4
09	10	10.1	75.5
10	1	1.1	76.6
11	1	1.1	77.7
12	1	1.1	78.8
13	1	1.1	79.9
14	1	1.1	81.0
15	1	1.1	82.1
16	1	1.1	83.2
17	1	1.1	84.3
18	1	1.1	85.4
19	1	1.1	86.5
20	1	1.1	87.6
21	1	1.1	88.7
Total	105	100.0	100.0

Casos válidos: 105

Tabla 9. Frecuencia del parentesco con el niño para el segundo familiar con cáncer.

Valor	Frec.	Pct.	Pct. Cum.
1	2	2.3	2.3
2	10	11.8	14.1
3	19	22.1	36.2
Total	31	100.0	100.0

Casos válidos: 32

Tabla 10. Distribución de frecuencia para el tipo de cáncer desarrollado.

Valor	Frec.	Pct.	Pct. Cum.
01	2	19.0	19.0
02	1	9.1	28.1
03	2	17.7	45.8
04	1	9.1	54.9
05	1	9.1	64.0
06	1	9.1	73.1
07	1	9.1	82.2
08	1	9.1	91.3
09	1	9.1	100.0
Total	11	100.0	100.0

Casos válidos: 11

Tabla 11. Frecuencia del parentesco con el niño para el tercer familiar con cáncer.

Valor	Frec.	Pct.	Pct. Cum.
1	4	50.0	50.0
2	4	50.0	100.0
Total	8	100.0	100.0

Casos válidos: 8

Tabla 12. Distribución de frecuencia para el tipo de cáncer desarrollado.

Valor	Frec.	Pct.	Pct. Cum.
01	2	25.0	25.0
02	1	12.5	37.5
03	1	12.5	50.0
04	1	12.5	62.5
05	2	25.0	87.5
Total	8	100.0	100.0

Casos válidos: 8

Tabla 13. Distribución de frecuencia para ubicar si los pacientes tienen familiares que nacieron con algún defecto.

Valor	Frec.	Pct.	Pct. Cum.
01	140	92.8	92.8
02	8	5.0	97.8
Total	148	100.0	100.0

Casos válidos: 148

Fuente: Unidad de Epidemiología, Hospital de Pediatría del Centro Médico Nacional Siglo XXI

Tabla 14. Frecuencia para ubicar si los niños en estudio tienen familiares que nacieron con algún defecto

Valor	Frec.	Por.	Por. Cum.
1	24	5.5	5.5
2	2	0.5	6.0
Total	26	100.0	

Casos válidos = 26

Tabla 15. Distribución de frecuencia del parentesco que tiene el familiar que nació con un defecto respecto al niño incluido en el estudio

Valor	Frec.	Por.	Por. Cum.
1	1	3.8	3.8
2	2	7.7	11.5
3	1	3.8	15.4
4	1	3.8	19.2
5	1	3.8	23.1
Total	6	100.0	

Casos válidos = 26

Tabla 16. Distribución por tipo de defecto

Valor	Frec.	Por.	Por. Cum.
01	1	3.8	3.8
02	1	3.8	7.7
03	2	7.7	15.4
04	1	3.8	19.2
05	1	3.8	23.1
06	1	3.8	26.9
07	1	3.8	30.8
08	1	3.8	34.6
09	1	3.8	38.5
10	1	3.8	42.3
11	1	3.8	46.2
12	1	3.8	49.9
13	1	3.8	53.8
14	1	3.8	57.7
15	1	3.8	61.5
16	1	3.8	65.4
17	1	3.8	69.2
18	1	3.8	73.1
19	1	3.8	76.9
20	1	3.8	80.8
Total	20	100.0	

Casos válidos = 26

Tabla 17. Distribución de frecuencia del parentesco con el niño del estudio familiar con defecto físico

Valor	Frec.	Por.	Por. Cum.
1	1	25.0	25.0
2	1	25.0	50.0
Total	2	100.0	

Casos válidos = 2

Tabla 18. Distribución de frecuencia por tipo de defecto

Valor	Frec.	Por.	Por. Cum.
01	1	25.0	25.0
02	1	25.0	50.0
Total	2	100.0	

Casos válidos = 2

Tabla 19. Distribución de frecuencia para ver si algún familiar del niño tiene Síndrome de Down o Mucopolisacaridosis

Valor	Frec.	Por.	Por. Cum.
1	151	93.9	93.9
2	2	1.2	95.1
Total	153	100.0	

Casos válidos = 153

FUENTE: Unidad de Epidemiología, Hospital de Pediatría del Centro Médico Nacional Siglo XXI

Tabla 25. Frecuencia de la distancia (metros) de dicha instalación

Val	Frecs.	Pct.	Pct. Cum.	Val	Frecs.	Pct.	Pct. Cum.	Val.	Frecs.	Pct.	Pct. Cum.
1.00	40	12	12	13.00	5	1	16	10000	1	11	91
2.00	60	21	33	14.00	1	11	60	10001	2	2	93
3.00	18	6	39	15.00	12	4	72	10002	3	3	96
4.00	20	6	45	16.00	1	2	74	10003	2	3	99
5.00	16	5	50	20.00	10	5	79	10004	1	11	100
6.00	10	3	53	25.00	2	1	81	10005	2	13	102
7.00	1	1	54	30.00	1	1	82	10006	1	14	103
8.00	5	2	56	35.00	4	2	84	10007	1	15	104
9.00	1	1	57	40.00	1	1	85	10008	1	16	105
10.00	0	0	57	45.00	1	1	86	10009	1	17	106

C. caso válidos: 334 No-Registra: 1

Tabla 26. Distribución de frecuencia de las familias que tienen alguna subestación eléctrica instalada cerca de su hogar

Valor	Frec.	Pct.	Pct. Cum.
0	174	52.0	52.0
1	160	48.0	100.0
Total	334	100.0	100.0

C. caso válidos: 334

Tabla 27. Frecuencia de la distancia (metros) de dicha instalación

Valor	Frec.	Pct.	Pct. Cum.
10.00	1	2	2
20.00	1	2	4
30.00	1	2	6
40.00	1	2	8
50.00	2	4	12
60.00	1	2	14
70.00	1	2	16
80.00	1	2	18
90.00	2	4	22
100.00	2	4	26
110.00	1	2	28
120.00	13	24	52
130.00	1	2	54
140.00	1	2	56
150.00	7	13	69
160.00	5	9	78
170.00	2	4	82
180.00	4	7	89
190.00	1	2	91

C. caso válidos: 51

Tabla 28. Distribución de frecuencia de las familias que tienen torres de alta tensión instaladas cerca de su hogar

Valor	Frec.	Pct.	Pct. Cum.
0	271	72.5	72.5
1	103	27.5	100.0
Total	374	100.0	100.0

C. caso válidos: 374

Tabla 29. Frecuencia de la distancia (metros) de dicha instalación

Val.	Frecs.	Pct.	Pct. Cum.	Val.	Frecs.	Pct.	Pct. Cum.	Val.	Frecs.	Pct.	Pct. Cum.
1.00	1	1	1	10000	1	1	1	10000	1	1	1
10.00	4	4	5	10001	1	1	2	10001	1	2	2
20.00	1	1	6	10002	1	1	3	10002	1	3	3
30.00	1	1	7	10003	1	1	4	10003	1	4	4
40.00	1	1	8	10004	1	1	5	10004	1	5	5
50.00	1	1	9	10005	1	1	6	10005	1	6	6
60.00	1	1	10	10006	1	1	7	10006	1	7	7
70.00	1	1	11	10007	1	1	8	10007	1	8	8
80.00	1	1	12	10008	1	1	9	10008	1	9	9
90.00	1	1	13	10009	1	1	10	10009	1	10	10
100.00	2	2	15	10010	1	1	11	10010	1	11	11

C. caso válidos: 103

Tabla 30. Distribución de frecuencia para saber si la familia del niño en estudio ha vivido siempre en el mismo lugar

Valor	Frec.	Pct.	Pct. Cum.
0	16	4.8	4.8
1	278	85.2	90.0
Total	294	100.0	100.0

C. caso válidos: 274

Fuente: Unidad de Epidemiología, Hospital de Pediatría del Centro Médico Nacional Sique XXI

Tabla 31. Frecuencia de las personas que han vivido en otro lugar y tenían transformadores instalados e en su casa

Valor	Freq.	Pct.	Pct. Cum.
1	76	41.5	41.5
2	99	52.9	94.4
Total	175	100.0	100.0

Casos válidos=172

Tabla 32. Distancia (metros) de donde se encontraba su hogar

Val.	Freq.	Pct.	Pct. Cum.	Val.	Freq.	Pct.	Pct. Cum.	Val.	Freq.	Pct.	Pct. Cum.
1.11	4	2.3	2.3	1.11	1	0.6	0.6	1.11	1	0.6	0.6
2.11	1	0.6	2.9	2.11	2	1.2	1.8	1.11	4	2.3	3.1
3.11	2	1.2	4.1	3.11	2	1.2	3.0	2.11	2	1.2	4.3
4.11	3	1.7	5.8	4.11	2	1.2	4.2	3.11	1	0.6	4.8
5.11	5	2.8	8.6	5.11	1	0.6	4.8	4.11	1	0.6	5.4
6.11	2	1.2	9.8	6.11	3	1.7	6.5	5.11	1	0.6	7.1
7.11	1	0.6	10.4	7.11	10	5.7	12.2	6.11	1	0.6	7.7
8.11	1	0.6	11.0	8.11	2	1.2	13.4	7.11	1	0.6	8.3
11.11	6	3.5	28	11.11	1	0.6	14.0	8.11	1	0.6	8.9
Total	175	100.0	100.0	11.11	1	0.6	14.6	9.11	1	0.6	9.5

Casos válidos=164

Tabla 33. Frecuencia de las personas que han vivido en otro lugar y tenían cables de alta tensión instalados e en su casa

Valor	Freq.	Pct.	Pct. Cum.
1	145	77.2	77.2
2	17	9.2	86.4
Total	162	100.0	100.0

Casos válidos=152

Tabla 34. Distancia (metros) de donde se encontraba su hogar

Val.	Freq.	Pct.	Pct. Cum.	Val.	Freq.	Pct.	Pct. Cum.	Val.	Freq.	Pct.	Pct. Cum.
1.11	6	3.3	3.3	1.11	2	1.1	1.1	1.11	2	1.1	1.1
2.11	12	6.4	9.7	2.11	1	0.5	1.6	1.11	4	2.2	3.3
3.11	12	6.4	16.1	3.11	2	1.1	2.7	2.11	2	1.1	4.4
4.11	17	9.3	25.4	4.11	5	2.7	5.4	3.11	1	0.5	5.9
5.11	11	6.1	31.5	5.11	4	2.2	7.6	4.11	1	0.5	8.1
6.11	12	6.4	37.9	6.11	1	0.5	8.1	5.11	1	0.5	8.6
7.11	12	6.4	44.3	7.11	1	0.5	8.6	6.11	1	0.5	9.1
11.11	11	6.1	50.4	8.11	1	0.5	9.1	7.11	1	0.5	9.6
Total	111	7	60	9.11	1	0.5	9.6	8.11	1	0.5	10.1

Casos válidos=146 Faltantes=1

Tabla 35. Frecuencia de las personas que han vivido en otro lugar y tenían subestaciones eléctricas instaladas e en su casa

Valor	Freq.	Pct.	Pct. Cum.
1	113	59.9	59.9
2	77	41.1	101.0
Total	190	100.0	100.0

Casos válidos=172

Tabla 36. Distancia (metros) de donde se encontraba su hogar

Valor	Freq.	Pct.	Pct. Cum.
1.11	1	0.5	0.5
2.11	1	0.5	1.0
3.11	1	0.5	1.5
4.11	1	0.5	2.0
5.11	1	0.5	2.5
6.11	1	0.5	3.0
7.11	1	0.5	3.5
8.11	1	0.5	4.0
9.11	1	0.5	4.5
10.11	1	0.5	5.0
11.11	1	0.5	5.5
12.11	1	0.5	6.0
13.11	1	0.5	6.5
14.11	1	0.5	7.0
15.11	1	0.5	7.5
16.11	1	0.5	8.0
17.11	1	0.5	8.5
18.11	1	0.5	9.0
19.11	1	0.5	9.5
20.11	1	0.5	10.0

Casos válidos=14 Faltantes=1

ESCOLARIDAD

Tabla 36. Distribución de personas para ver si los niños incluidos en el estudio (estudiantes) continúan

Valor	Freq.	Pct.	Pct. Cum.
1	113	60.4	60.4
2	77	41.6	102.0
Total	190	100.0	100.0

Casos válidos=174

Fuente: Unidad de Epidemiología Hospital de Pediatría del Centro Médico Nacional Siglo XXI

Tabla 37. Distribución de frecuencia para los niños considerados en el estudio que asisten a la escuela.

Valor	Freq.	Pct.	Pct. Cum.
0	253	46,9	46,9
1	322	59,4	106,3
Total	575	100,0	100,0

¿ A qué edad? 174

Tabla 38. Frecuencia para ver si existen transformadores instalados cerca de la escuela.

Valor	Freq.	Pct.	Pct. Cum.
0	278	50,1	50,1
1	322	56,6	106,7
Total	600	100,0	100,0

¿ A qué edad? 221 Faltantes 1

Tabla 39. Distancia (metros) respecto a la escuela.

Val.	Freq.	Pct.	Pct. Cum.	Val.	Freq.	Pct.	Pct. Cum.	Val.	Freq.	Pct.	Pct. Cum.
0,00	1	1	1	1,00	2	2	2	5,00	1	1	1
1,00	2	2	3	15,00	1	1	3	10,00	12	12	22
2,00	6	6	12	20,00	6	6	13	10,00	2	2	24
3,00	1	1	13	25,00	5	5	18	20,00	1	1	25
4,00	1	1	17	30,00	4	4	22	10,00	1	1	26
5,00	3	3	20	35,00	3	3	25	5,00	1	1	27
6,00	1	1	24	50,00	13	13	38	10,00	1	1	28
Total	6	6	30	70,00	70	70	100	20,00	1	1	100

¿ A qué edad? 221 Faltantes 1

Tabla 40. Frecuencia para ver si existen cables de alta tensión instalados cerca de la escuela.

Valor	Freq.	Pct.	Pct. Cum.
0	40	6,7	6,7
1	559	93,3	100,0
Total	599	100,0	100,0

¿ A qué edad? 221 Substancia 1

Tabla 41. Distancia (metros) respecto a la escuela.

Valor	Freq.	Pct.	Pct. Cum.
0,00	12	12	12
2,00	8	8	20
3,00	7	7	27
4,00	12	12	39
5,00	10	10	49
6,00	6	6	55
7,00	8	8	63
10,00	13	13	76
12,00	7	7	83
15,00	1	1	84
20,00	13	13	97
25,00	7	7	104
30,00	9	9	113
100,00	16	16	129
150,00	13	13	142
200,00	12	12	154

¿ A qué edad? 183 Faltantes 1

Tabla 42. Frecuencia para ver si existen subestaciones eléctricas instaladas cerca de la escuela.

Valor	Freq.	Pct.	Pct. Cum.
0	214	37,1	37,1
1	366	62,9	100,0
Total	580	100,0	100,0

¿ A qué edad? 224 Substancia 1

Tabla 43. Distancia (metros) respecto a la escuela.

Valor	Freq.	Pct.	Pct. Cum.
0,00	1	0,1	0,1
10,00	2	0,2	0,3
150,00	1	0,1	0,4
200,00	1	0,1	0,5
300,00	1	0,1	0,6
500,00	2	0,2	0,8
1000,00	1	0,1	0,9
2000,00	1	0,1	1,0

¿ A qué edad? 12

Tabla 44. Frecuencia para ver si existen torres de alta tensión instaladas cerca de la escuela

Valor	Freq.	Pct.	Pct. Cum.
0	110	53,2	53,2
1	33	15,8	69,0
Total	223	100,0	100,0

Casos válidos: 223 (Subgrupos: 1)

Tabla 45. Distancia (metros) respecto a la escuela

Val.	Freq.	Pct.	Pct. Cum.	Val.	Freq.	Pct.	Pct. Cum.	Val.	Freq.	Pct.	Pct. Cum.
1500	1	0,4	0,4	1500	2	0,9	1,3	1500	1	0,4	0,4
2000	2	0,9	1,3	1500	4	1,8	2,2	1500	1	0,4	0,8
3000	1	0,4	1,7	1500	2	0,9	3,1	1500	1	0,4	1,2
3500	3	1,3	3,0	1500	1	0,4	3,5	1500	1	0,4	1,6
4000	1	0,4	3,4	1500	5	2,3	5,8	1500	1	0,4	2,0
4500	2	0,9	4,3	1500	1	0,4	6,2	1500	1	0,4	2,4

Casos válidos: 33

Tabla 46. Distribución de frecuencia del número de meses que ha estado en esa escuela

Valor	Freq.	Pct.	Pct. Cum.
1	43	20,9	20,9
2	40	19,7	40,6
3	18	8,9	49,5
4	22	10,8	60,3
5	18	8,7	69,0
6	15	7,3	76,3
7	1	0,5	76,8
8	1	0,5	77,3
90	0	0,0	77,3

Casos válidos: 202

Tabla 47. Distribución de frecuencia de tiempo (horas) que permanece en la escuela diariamente.

Valor	Freq.	Pct.	Pct. Cum.
2	3	1,4	1,4
3	23	10,5	11,9
4	120	54,9	66,8
5	32	14,6	81,4
6	21	9,5	90,9
7	18	8,2	99,1
8	5	2,3	100,0

Casos válidos: 222

Tabla 48. Frecuencia para el turno en el que existe

Valor	Freq.	Pct.	Pct. Cum.
1	170	76,2	76,2
2	23	10,3	86,5
3	1	0,4	86,9
Total	222	100,0	100,0

Casos válidos: 222

Tabla 49. Distribución de frecuencia para identificar si el niño asiste antes a otra escuela.

Valor	Freq.	Pct.	Pct. Cum.
0	20	9,0	9,0
1	197	88,0	97,0
Total	222	100,0	100,0

Casos válidos: 222

Tabla 50. Distribución de frecuencia para las escuelas que tienen transformadores instalados cerca

Valor	Freq.	Pct.	Pct. Cum.
0	100	45,5	45,5
1	85	38,1	83,6
Total	185	100,0	100,0

Casos válidos: 185 (Subgrupos: 1)

Tabla 56. Frecuencia para las escuelas que tienen torres de alta tensión instaladas cerca.

Valor	Freq.	Por.	Por. Cum.
0	130	83.3	83.3
1	23	14.2	97.5
Total	153	100.0	100.0

Casos válidos: 153. Total Casos: 1

Tabla 57. Distancia (metros) de donde se encuentran respecto a la escuela.

Valor	Freq.	Por.	Por. Cum.
0	1	4.3	4.3
100	1	4.3	8.6
250	1	4.3	12.9
500	1	4.3	17.1
750	1	4.3	21.4
1000	1	4.3	25.7
1250	1	4.3	30.0
1500	1	4.3	34.3
1750	1	4.3	38.6
2000	1	4.3	42.9
2250	1	4.3	47.1
2500	1	4.3	51.4
2750	1	4.3	55.7
3000	1	4.3	60.0
3250	1	4.3	64.3
3500	1	4.3	68.6
3750	1	4.3	72.9
4000	1	4.3	77.1
4250	1	4.3	81.4
4500	1	4.3	85.7
4750	1	4.3	90.0
5000	1	4.3	94.3
5250	1	4.3	98.6
5500	1	4.3	100.0

Casos válidos: 23. Total Casos: 23

Tabla 58. Distribución de frecuencias del número de horas que permanece en la escuela.

Valor	Freq.	Por.	Por. Cum.
0	51	37.2	37.2
1	15	10.7	47.9
2	15	10.7	58.6
3	19	13.8	72.4

Casos válidos: 136. Total Casos: 1

Tabla 59. Frecuencia del turno en que asista a la escuela.

Valor	Freq.	Por.	Por. Cum.
0	174	83.9	83.9
1	1	0.5	84.4
2	1	0.5	84.9
Total	176	100.0	100.0

Casos válidos: 176. Total Casos: 1

Tabla 60. Distribución de frecuencias para la escolaridad del padre.

Valor	Freq.	Por.	Por. Cum.
0	154	77.0	77.0
1	11.5	5.7	82.7
2	11	5.5	88.2
3	12	6.0	94.2
4	11	5.5	99.7
5	11	5.5	100.0

Casos válidos: 174. Total Casos: 1

Tabla 61. Distribución de frecuencias para escolaridad de la madre.

Valor	Freq.	Por.	Por. Cum.
0	141	70.5	70.5
1	11	5.5	76.0
2	11	5.5	81.5
3	11	5.5	87.0
4	11	5.5	92.5
5	11	5.5	98.0
6	11	5.5	100.0

Casos válidos: 174. Total Casos: 1

Tabla 65. Frecuencia de exposición a solventes

Valor	Freci.	Pct.	Pct. Cum.
1	17	22.4	22.4
2	100	131.6	154.0
3	14	18.2	172.2
Total	129	164.2	164.2

Casos válidos de Faltantes 3

Tabla 66. Distribución de frecuencias para conocer si el padre en su trabajo estaba expuesto a insecticidas

Valor	Freci.	Pct.	Pct. Cum.
0	157	24.5	24.5
1	477	74.5	99.0
Total	634	100.0	100.0

Casos válidos 374

Tabla 66. Frecuencia de exposición a insecticidas

Valor	Freci.	Pct.	Pct. Cum.
1	5	11.1	11.1
2	5	11.2	22.3
3	12	27.8	50.1
Total	22	100.0	100.0

Casos válidos 16 Faltantes 1

Tabla 67. Distribución de frecuencias para conocer si el padre en su trabajo estaba expuesto a thiner

Valor	Freci.	Pct.	Pct. Cum.
0	303	41.6	41.6
1	401	54.4	96.0
Total	704	100.0	100.0

Casos válidos 374

Tabla 68. Frecuencia de exposición a thiner

Valor	Freci.	Pct.	Pct. Cum.
0	211	31.2	31.2
1	27	40.2	71.4
2	28	41.6	113.0
Total	266	100.0	100.0

Casos válidos 16 Faltantes 5

Tabla 69. Distribución de frecuencias para conocer si el padre en su trabajo estaba expuesto a pinturas

Valor	Freci.	Pct.	Pct. Cum.
0	41	63.2	63.2
1	23	36.8	100.0
Total	64	100.0	100.0

Casos válidos 374

Tabla 70. Frecuencia de exposición a pinturas

Valor	Freci.	Pct.	Pct. Cum.
1	4	21.4	21.4
2	20	100.0	121.4
3	17	85.7	107.1
Total	41	100.0	100.0

Casos válidos 36 Faltantes 9

Tabla 71. Distribución de frecuencias para conocer si el padre en su trabajo estaba expuesto a fertilizantes

Valor	Freci.	Pct.	Pct. Cum.
0	36	6.4	6.4
1	512	93.6	100.0
Total	548	100.0	100.0

Casos válidos 374

Tabla 72. Frecuencia de exposición a fertilizantes

Valor	Freq.	Pct.	Pct. Cum.
1	2	11.1	
2	2	11.1	22.2
3	2	11.1	33.3
Total	6	100.0	100.0

Casos válidos: 6 Faltantes: 1

Tabla 73. Distribución de frecuencias para conocer si el padre en su trabajo estaba expuesto a pesticidas

Valor	Freq.	Pct.	Pct. Cum.
1	45	13.1	
2	314	86.9	100.0
Total	359	100.0	100.0

Casos válidos: 359

Tabla 74. Frecuencia de exposición a pesticidas

Valor	Freq.	Pct.	Pct. Cum.
1	14	31.1	
2	15	33.3	64.4
3	10	22.2	86.7
Total	45	100.0	100.0

Casos válidos: 45 Faltantes: 4

Tabla 75. Distribución de frecuencia para la ocupación del padre cuando el niño enfermo de leucemia

Valor	Freq.	Pct.	Pct. Cum.
1001	3	4.8	4.8
1011	3	4.8	9.6
1022	5	7.7	17.3
1031	3	4.8	22.1
1041	12	17.6	39.7
1050	6	8.8	48.5
1077	2	3.2	51.7
1086	2	3.2	54.9
1099	12	17.6	72.5
1104	15	22.4	94.9
1112	14	20.6	100.0
1126	14	20.6	100.0

Casos válidos: 100 Sobrantes: 1

Tabla 76. Distribución de frecuencia para los padres que estuvieron expuestos a químicos.

Valor	Freq.	Pct.	Pct. Cum.
10	100	100.0	
Total	100	100.0	100.0

Casos válidos: 100 Sobrantes: 1

Tabla 77. Distribución de frecuencia por Tipo de químico

Val.	Freq.	Pct.	Pct. Cum.	Val.	Freq.	Pct.	Pct. Cum.	Val.	Freq.	Pct.	Pct. Cum.
101	1	2.1		101	2	4.2	8.2	17	2	4.2	8.2
102	2	4.2	8.2	102	1	2.1	10.3	19	3	6.3	16.5
103	2	4.2	12.4	103	2	4.2	14.6	20	1	2.1	16.7
104	1	2.1	14.5	104	1	2.1	16.7	21	1	2.1	18.8
105	1	2.1	16.6	105	1	2.1	18.8	22	1	2.1	20.9
107	1	2.1	18.7	107	1	2.1	20.9	23	1	2.1	23.0

Casos válidos: 47 Sobrantes: 2

Tabla 78. Distribución de frecuencia para la ocupación de la madre antes del embarazo del niño.

Valor	Freq.	Pct.	Pct. Cum.
101	1	4.8	
102	1	4.8	9.6
103	18	80.4	90.0
104	1	4.8	94.8
105	1	4.8	99.6
106	2	9.6	100.0
107	1	4.8	100.0
108	1	4.8	100.0
109	28	128.0	100.0
110	28	128.0	100.0
111	28	128.0	100.0
112	47	205.2	100.0
113	109	481.2	100.0
114	18	81.6	100.0

Casos válidos: 374

Fuente: Unidad de Epidemiología, Hospital de Pediatría del Centro Médico Nacional Siglo XXI

Tabla 79. Distribución de Frecuencias para la actividad de la madre cuando estaba embarazada del niño

Valor	Freq.	Pct.	Pct. Cum.
101	11	3	3
104	11	4	7
105	9	7	14
106	2	1	15
107	9	11	27
109	9	2	30
110	15	4	34
112	22	6	40
114	269	70	100
14	0	0	100

Casos válidos: 374

Tabla 80. Distribución de frecuencias para ver si las madres experimentan espasmos o quimeras

Valor	Freq.	Pct.	Pct. Cum.
11	102	23.5	23.5
12	22	5	28.5
Total	374	100.0	100.0

Casos válidos: 374

Tabla 81. Tipo de quimeras

Valor	Freq.	Pct.	Pct. Cum.
101	4	4	4
104	159	48.5	52.5
105	1	1	53.5
107	1	1	54.5
108	4	1.4	55.9
109	4	1.5	57.4
110	0	0	57.4
112	1	0.3	57.7
114	1	0.3	58
24	2	0.6	58.6
26	2	0.6	59.2

Casos válidos: 212

Tabla 82. Distribución de frecuencias de la actividad de la madre cuando el niño entraba de leu emu

Valor	Freq.	Pct.	Pct. Cum.
101	1	1.5	1.5
104	2	3	4.5
105	2	3	7.5
106	12	17.5	25
107	6	9	34
109	6	9	43
110	11	16.5	59.5
112	11	16.5	76
114	11	16.5	92.5
14	0	0	92.5

Casos válidos: 187

ANTECEDENTES PATOLÓGICOS

Tabla 83. Distribución de frecuencia de consumo de alcohol de la madre durante el embarazo

Valor	Freq.	Pct.	Pct. Cum.
11	347	92.3	92.3
12	20	5.3	97.6
Total	374	100.0	100.0

Casos válidos: 373. 1 atípico ()

Fuente: Unidad de Epidemiología, Hospital de Pediatría del Centro Médico Nacional Sige XXI

Tabla 84. Frecuencia para ver en que mes se hizo más consumo de alcohol

Valor	Freq.	Pct.	Pct. Cum.
0	1	3,0	3,0
1	6	18,2	21,2
2	6	18,2	39,4
3	5	15,2	54,6
4	5	15,2	69,8
5	10	30,4	100,0
Total	34	100,0	

Casos válidos: 28 Subgrupos: 2

Tabla 85. Frecuencia de consumo de tabaco del padre durante el embarazo del niño.

Valor	Freq.	Pct.	Pct. Cum.
0	20,1	58,3	58,3
1	14,3	41,7	100,0
Total	34,4	100,0	

Casos válidos: 37*

Tabla 86. Distribución de frecuencia para ver si el consumo de tabaco se hizo o no de la madre del niño.

Valor	Freq.	Pct.	Pct. Cum.
0	12,9	37,5	37,5
1	12,9	37,5	75,0
Total	33,8	100,0	

Casos válidos: 37* Subgrupos: 3

Tabla 87. Distribución de frecuencia de consumo de tabaco de la madre antes del embarazo.

Valor	Freq.	Pct.	Pct. Cum.
0	22,9	65,9	65,9
1	9,1	26,1	92,0
Total	34,0	100,0	

Casos válidos: 37*

Tabla 88. Distribución de frecuencia de consumo de tabaco de la madre durante el embarazo.

Valor	Freq.	Pct.	Pct. Cum.
0	15,9	46,4	46,4
1	21,1	60,6	107,0
Total	33,0	100,0	

Casos válidos: 37*

Tabla 89. Frecuencia para el mes de mayor consumo de tabaco

Valor	Freq.	Pct.	Pct. Cum.
0	22,9	65,9	65,9
1	6	17,3	83,2
2	6	17,3	100,5
3	11	31,4	131,9
4	11	31,4	163,3
Total	57,5	163,3	

Casos válidos: 22* Subgrupos: 1

OTROS ANTECEDENTES DEL MENOR

Tabla 90. Distribución de frecuencia para ver si se tomaron radiografías al niño en estudio.

Valor	Freq.	Pct.	Pct. Cum.
0	22,9	65,9	65,9
1	14,7	42,1	108,0
Total	34,6	108,0	

Casos válidos: 37*

Tabla 91. Frecuencia para el número de radiografías.

Valor	Freq.	Pct.	Pct. Cum.
1	67	46	46
2	26	18	64
3	14	10	74
4	11	8	82
5	11	8	90
6	9	6	96
7	5	4	100
111	2	1	100
12	1	1	100
13	1	1	100
20	1	1	100
54	1	1	100
Total	144	100.0	

Casos válidos: 144. Tablas: 1.

Tabla 92. Distribución de frecuencia para ver si se le tomó algún ultrasonido al niño en estudio.

Valor	Freq.	Pct.	Pct. Cum.
0	6	4.2	4.2
1	138	95.8	100.0
Total	144	100.0	

Casos válidos: 144.

Tabla 93. Frecuencia del número de ultrasonidos tomados.

Valor	Freq.	Pct.	Pct. Cum.
1	6	4.2	4.2
2	6	4.2	8.4
3	2	1.4	9.8
4	1	0.7	10.5
5	1	0.7	11.2
Total	14	9.7	

Casos válidos: 14.

Tabla 94. Distribución de frecuencia para comer si el niño nació con algún defecto fetal.

Valor	Freq.	Pct.	Pct. Cum.
0	65	45.1	45.1
1	79	54.9	100.0
Total	144	100.0	

Casos válidos: 144.

Tabla 95. Distribución de frecuencia por tipo de defecto.

Valor	Freq.	Pct.	Pct. Cum.
111	1	1.4	1.4
112	1	1.4	2.8
113	1	1.4	4.3
114	2	2.8	7.1
115	1	1.4	8.5
116	1	1.4	9.9
117	1	1.4	11.3
118	1	1.4	12.7
119	1	1.4	14.1
120	1	1.4	15.5
121	1	1.4	16.9
122	1	1.4	18.3
123	1	1.4	19.7
Total	16	11.1	

Casos válidos: 16.

Tabla 96. Distribución de frecuencia de los niños que estuvieron o no con el síndrome de Down del estudio.

Valor	Freq.	Pct.	Pct. Cum.
0	66	45.8	45.8
1	78	54.2	100.0
Total	144	100.0	

Casos válidos: 144.

Tabla 97. Frecuencia de exposición a solventes

Valor	Freq.	Pct.	Pct. Cum.
1	4	66,7	66,7
2	2	33,3	100,0
Total	6	100,0	

Casos válidos = 6

Tabla 98. Distribución de frecuencia de los niños que estuvieron expuestos a insecticidas

Valor	Freq.	Pct.	Pct. Cum.
1	25	71,4	71,4
2	10	28,6	100,0
Total	35	100,0	

Casos válidos = 35

Tabla 99. Frecuencia de exposición a insecticidas

Valor	Freq.	Pct.	Pct. Cum.
1	11	32,4	32,4
2	20	60,6	93,0
3	1	3,0	96,0
Total	32	100,0	

Casos válidos = 32. Faltantes = 4

Tabla 100. Distribución de frecuencia de los niños que estuvieron expuestos a thinner

Valor	Freq.	Pct.	Pct. Cum.
1	10	50,0	50,0
2	12	60,0	100,0
Total	22	100,0	

Casos válidos = 22

Tabla 101. Frecuencia de exposición a thinner

Valor	Freq.	Pct.	Pct. Cum.
1	11	33,3	33,3
2	22	66,7	100,0
Total	33	100,0	

Casos válidos = 33. Faltantes = 4

Tabla 100. Distribución de frecuencia de los niños que estuvieron expuestos a pinturas

Valor	Freq.	Pct.	Pct. Cum.
1	20	50,0	50,0
2	10	25,0	75,0
Total	30	100,0	

Casos válidos = 30

Tabla 102. Frecuencia de exposición a pinturas

Valor	Freq.	Pct.	Pct. Cum.
1	6	18,2	18,2
2	10	30,3	48,5
3	12	36,4	84,9
Total	28	100,0	

Casos válidos = 28. Faltantes = 3

Tabla 103. Distribución de frecuencia de los niños que estuvieron expuestos a fertilizantes

Valor	Freq.	Pct.	Pct. Cum.
1	17	59,3	59,3
2	8	27,4	86,7
Total	25	100,0	

Casos válidos = 25

Tabla 104. Frecuencia de exposición a fertilizantes

Valor	Freq.	Pct.	Pct. Cum.
1	1	25,0	25,0
2	3	75,0	100,0
Total	4	100,0	

Casos válidos = 4

Fuente: Unidad de Epidemiología Hospital de Pediatría del Centro Médico Nacional Siglo XXI

Tabla 105. Distribución de frecuencias de los niños que estuvieron expuestos a casullina.

Valor	Frec.	Pct.	Pct. Cum.
0	50	45.2	45.2
1	18	16.4	61.6
Total	112	100.0	100.0

Casos válidos: 57a

Tabla 106. Frecuencia de exposición a casullina.

Valor	Frec.	Pct.	Pct. Cum.
1	3	61.2%	61.2%
2	3	57.5%	118.7%
Total	16	100.0	100.0

Casos válidos: 10 - Frecuencia: 2

Tabla 107. Exposición de los niños a colorantes para el cabello.

Valor	Frec.	Pct.	Pct. Cum.
0	65	57.2	57.2
1	20	17.8	75.0
Total	112	100.0	100.0

Casos válidos: 57a

Tabla 108. Frecuencia de exposición a colorantes para el cabello.

Valor	Frec.	Pct.	Pct. Cum.
1	1	5.0	5.0
2	1	5.0	10.0
3	2	10.0	20.0
Total	4	100.0	100.0

Casos válidos: 26 - Frecuencia: 1

Tabla 109. Distribución de frecuencias de niños expuestos a tinturas para el cabello.

Valor	Frec.	Pct.	Pct. Cum.
0	21	57.9	57.9
1	10	26.3	84.2
Total	37	100.0	100.0

Casos válidos: 57a

Tabla 110. Frecuencia de exposición a tinturas para el cabello.

Valor	Frec.	Pct.	Pct. Cum.
1	1	10.0	10.0
2	1	9.1	19.1
3	2	18.2	37.3
Total	4	100.0	100.0

Casos válidos: 15a

Tabla 111. Distribución de frecuencias para conocer si el niño "caso" padecía alguna enfermedad.

Valor	Frec.	Pct.	Pct. Cum.
181	1	4.2%	4.2%
111	1	4.2%	8.4%
112	1	4.2%	12.6%
113	1	4.2%	16.8%
114	1	4.2%	21.0%
115	1	4.2%	25.2%
116	1	4.2%	29.4%
117	1	4.2%	33.6%
118	1	4.2%	37.8%
119	1	4.2%	42.0%
120	1	4.2%	46.2%
121	1	4.2%	50.4%
122	1	4.2%	54.6%
123	1	4.2%	58.8%
124	1	4.2%	63.0%
125	1	4.2%	67.2%
126	1	4.2%	71.4%
127	1	4.2%	75.6%
128	1	4.2%	79.8%
129	1	4.2%	84.0%
130	1	4.2%	88.2%
131	1	4.2%	92.4%
132	1	4.2%	96.6%
133	1	4.2%	100.0%

Casos válidos: 18a - Subcategoría: 2

Fuente: Unidad de Epidemiología, Hospital de Pediatría del Centro Médico Nacional Siglo XXI

Tabla 112. Distribución de frecuencias para conocer si el niño "control" padece alguna enfermedad

Valor	Frec.	Pct.	Pct. Cum.
101	1	0,1	0,1
102	9	19,6	20,7
103	4	9,1	29,8
104	9	20,2	50,0
105	12	26,8	76,8
106	1	2,2	79,0
107	4	9,1	88,1
108	1	2,2	90,3
109	6	13,4	103,7
110	1	2,2	105,9
111	1	2,2	108,1
112	1	2,2	110,3
113	1	2,2	112,5
114	1	2,2	114,7
115	1	2,2	116,9
Total	7	15,4	100,0

Casos válidos: 185 Faltantes: 2

Tabla 113. Distribución de frecuencia para ver si el niño fue enfermo en el primer año de vida

Valor	Frec.	Pct.	Pct. Cum.
0	234	66,8	66,8
1	114	32,2	100,0
Total	348	100,0	

Casos válidos: 374

Tabla 114. Distribución de frecuencia para ver si el niño en estudio fue más enfermo que sus otros hermanos

Valor	Frec.	Pct.	Pct. Cum.
0	261	77,8	77,8
1	81	24,2	100,0
Total	342	100,0	

Casos válidos: 374

Tabla 115. Distribución de frecuencia para ver si el niño estuvo hospitalizado en el primer año de vida

Valor	Frec.	Pct.	Pct. Cum.
0	111	32,2	32,2
1	211	62,8	100,0
Total	322	100,0	

Casos válidos: 374

Tabla 116. Frecuencia por tipo de desplazamiento

Valor	Frec.	Pct.	Pct. Cum.
101	18	24,0	24,0
102	29	38,7	62,7
103	4	5,3	68,0
104	4	5,3	73,3
105	4	5,3	78,6
106	2	2,7	81,3
107	2	2,7	84,0
111	1	1,3	85,3
112	1	1,3	86,6
113	1	1,3	87,9
114	1	1,3	89,2
115	1	1,3	90,5
116	1	1,3	91,8
117	1	1,3	93,1
118	1	1,3	94,4
119	1	1,3	95,7
120	1	1,3	97,0
121	1	1,3	98,3
122	1	1,3	99,6
123	1	1,3	100,0
Total	75	100,0	

Casos válidos: 92 Faltantes: 1

ANTECEDENTES GINECOOBSTÉTRICOS Y PATOLÓGICOS DE LA MADRE

Tabla 117. Distribución de frecuencia del número total de embarazos de la madre.

Valor	Frec.	Por.	Por. Cum.
1	61,1	1,1	1,1
2	91,1	2,4	3,5
3	98,1	2,7	6,2
4	75,1	2,1	8,3
5	12,1	0,3	8,6
6	12,1	0,3	8,9
7	9,1	0,2	9,1
8	9,1	0,2	9,3
9	4,1	0,1	9,4
10	2,1	0,0	9,4
11	2,1	0,0	9,4
12	1,1	0,0	9,4
13	1,1	0,0	9,4
14	1,1	0,0	9,4
15	1,1	0,0	9,4
16	1,1	0,0	9,4
17	1,1	0,0	9,4
18	1,1	0,0	9,4
19	1,1	0,0	9,4
20	1,1	0,0	9,4
21	1,1	0,0	9,4
22	1,1	0,0	9,4
23	1,1	0,0	9,4
24	1,1	0,0	9,4
25	1,1	0,0	9,4
26	1,1	0,0	9,4
27	1,1	0,0	9,4
28	1,1	0,0	9,4
29	1,1	0,0	9,4
30	1,1	0,0	9,4
31	1,1	0,0	9,4
32	1,1	0,0	9,4
33	1,1	0,0	9,4
34	1,1	0,0	9,4
35	1,1	0,0	9,4
36	1,1	0,0	9,4
37	1,1	0,0	9,4
38	1,1	0,0	9,4
39	1,1	0,0	9,4
40	1,1	0,0	9,4
41	1,1	0,0	9,4
42	1,1	0,0	9,4
43	1,1	0,0	9,4
44	1,1	0,0	9,4
45	1,1	0,0	9,4
46	1,1	0,0	9,4
47	1,1	0,0	9,4
48	1,1	0,0	9,4
49	1,1	0,0	9,4
50	1,1	0,0	9,4
51	1,1	0,0	9,4
52	1,1	0,0	9,4
53	1,1	0,0	9,4
54	1,1	0,0	9,4
55	1,1	0,0	9,4
56	1,1	0,0	9,4
57	1,1	0,0	9,4
58	1,1	0,0	9,4
59	1,1	0,0	9,4
60	1,1	0,0	9,4
61	1,1	0,0	9,4
62	1,1	0,0	9,4
63	1,1	0,0	9,4
64	1,1	0,0	9,4
65	1,1	0,0	9,4
66	1,1	0,0	9,4
67	1,1	0,0	9,4
68	1,1	0,0	9,4
69	1,1	0,0	9,4
70	1,1	0,0	9,4
71	1,1	0,0	9,4
72	1,1	0,0	9,4
73	1,1	0,0	9,4
74	1,1	0,0	9,4
75	1,1	0,0	9,4
76	1,1	0,0	9,4
77	1,1	0,0	9,4
78	1,1	0,0	9,4
79	1,1	0,0	9,4
80	1,1	0,0	9,4
81	1,1	0,0	9,4
82	1,1	0,0	9,4
83	1,1	0,0	9,4
84	1,1	0,0	9,4
85	1,1	0,0	9,4
86	1,1	0,0	9,4
87	1,1	0,0	9,4
88	1,1	0,0	9,4
89	1,1	0,0	9,4
90	1,1	0,0	9,4
91	1,1	0,0	9,4
92	1,1	0,0	9,4
93	1,1	0,0	9,4
94	1,1	0,0	9,4
95	1,1	0,0	9,4
96	1,1	0,0	9,4
97	1,1	0,0	9,4
98	1,1	0,0	9,4
99	1,1	0,0	9,4
100	1,1	0,0	9,4
TOTAL	574	100,0	100,0

Casos válidos: 574

Tabla 118. Distribución de frecuencia del número de embarazos letrados.

Valor	Frec.	Por.	Por. Cum.
1	65,1	2,4	2,4
2	109,1	2,4	4,8
3	88,1	2,4	7,2
4	69,1	1,6	8,8
5	22,1	0,2	9,0
6	12,1	0,1	9,1
7	9,1	0,1	9,2
8	7,1	0,1	9,3
9	7,1	0,1	9,4
10	7,1	0,1	9,5
11	2,1	0,0	9,5
12	1,1	0,0	9,5
13	1,1	0,0	9,5
14	1,1	0,0	9,5
15	1,1	0,0	9,5
16	1,1	0,0	9,5
17	1,1	0,0	9,5
18	1,1	0,0	9,5
19	1,1	0,0	9,5
20	1,1	0,0	9,5
21	1,1	0,0	9,5
22	1,1	0,0	9,5
23	1,1	0,0	9,5
24	1,1	0,0	9,5
25	1,1	0,0	9,5
26	1,1	0,0	9,5
27	1,1	0,0	9,5
28	1,1	0,0	9,5
29	1,1	0,0	9,5
30	1,1	0,0	9,5
31	1,1	0,0	9,5
32	1,1	0,0	9,5
33	1,1	0,0	9,5
34	1,1	0,0	9,5
35	1,1	0,0	9,5
36	1,1	0,0	9,5
37	1,1	0,0	9,5
38	1,1	0,0	9,5
39	1,1	0,0	9,5
40	1,1	0,0	9,5
41	1,1	0,0	9,5
42	1,1	0,0	9,5
43	1,1	0,0	9,5
44	1,1	0,0	9,5
45	1,1	0,0	9,5
46	1,1	0,0	9,5
47	1,1	0,0	9,5
48	1,1	0,0	9,5
49	1,1	0,0	9,5
50	1,1	0,0	9,5
51	1,1	0,0	9,5
52	1,1	0,0	9,5
53	1,1	0,0	9,5
54	1,1	0,0	9,5
55	1,1	0,0	9,5
56	1,1	0,0	9,5
57	1,1	0,0	9,5
58	1,1	0,0	9,5
59	1,1	0,0	9,5
60	1,1	0,0	9,5
61	1,1	0,0	9,5
62	1,1	0,0	9,5
63	1,1	0,0	9,5
64	1,1	0,0	9,5
65	1,1	0,0	9,5
66	1,1	0,0	9,5
67	1,1	0,0	9,5
68	1,1	0,0	9,5
69	1,1	0,0	9,5
70	1,1	0,0	9,5
71	1,1	0,0	9,5
72	1,1	0,0	9,5
73	1,1	0,0	9,5
74	1,1	0,0	9,5
75	1,1	0,0	9,5
76	1,1	0,0	9,5
77	1,1	0,0	9,5
78	1,1	0,0	9,5
79	1,1	0,0	9,5
80	1,1	0,0	9,5
81	1,1	0,0	9,5
82	1,1	0,0	9,5
83	1,1	0,0	9,5
84	1,1	0,0	9,5
85	1,1	0,0	9,5
86	1,1	0,0	9,5
87	1,1	0,0	9,5
88	1,1	0,0	9,5
89	1,1	0,0	9,5
90	1,1	0,0	9,5
91	1,1	0,0	9,5
92	1,1	0,0	9,5
93	1,1	0,0	9,5
94	1,1	0,0	9,5
95	1,1	0,0	9,5
96	1,1	0,0	9,5
97	1,1	0,0	9,5
98	1,1	0,0	9,5
99	1,1	0,0	9,5
100	1,1	0,0	9,5
TOTAL	574	100,0	100,0

Casos válidos: 574

Tabla 119. Distribución de frecuencia para el número de abortos que tuvo la madre antes del niño en estudio.

Valor	Frec.	Por.	Por. Cum.
1	4,1	0,1	0,1
2	1,1	0,0	0,1
TOTAL	57	100,0	100,0

Casos válidos: 57

Tabla 119. Distribución de frecuencia para conocer si la madre tuvo algún niño muerto o nacimientos.

Valor	Frec.	Por.	Por. Cum.
1	10,1	1,7	1,7
2	1,1	0,2	1,9
TOTAL	574	100,0	100,0

Casos válidos: 574

Tabla 120. Distribución de frecuencia para conocer si algunos fue antes del niño en estudio.

Valor	Frec.	Por.	Por. Cum.
1	10,1	1,7	1,7
2	1,1	0,2	1,9
TOTAL	574	100,0	100,0

Casos válidos: 574

Tabla 121. Distribución de frecuencia que muestra si se le tomó alguna radiografía si la madre durante el embarazo del niño.

Valor	Frec.	Por.	Por. Cum.
1	12,1	2,1	2,1
2	18,1	3,1	5,2
TOTAL	574	100,0	100,0

Casos válidos: 574

Tabla 122 *Números de radiografías tomadas a la madre durante el embarazo*

Valor	Freq.	Pct.	Pct. Cum.
112	1	2.2	2.2
113	1	2.2	4.3
114	2	4.3	8.6
115	2	4.3	12.9
116	1	2.2	15.1
117	1	2.2	17.3
118	15	32.0	49.3
119	15	32.0	81.3

Casos válidos= 45

Tabla 123 *Frecuencia para el mes de embarazo en que se le tomaron radiografías.*

Valor	Freq.	Pct.	Pct. Cum.
112	1	2.3	2.3
113	1	2.3	4.6
114	2	4.6	9.1
115	1	2.3	11.4
116	1	2.3	13.7
117	1	2.3	16.0
118	15	33.3	49.3
119	15	33.3	82.6

Casos válidos= 45

Tabla 124 *Frecuencia para el número de ultrasonidos tomadas a la madre del niño en estudio.*

Valor	Freq.	Pct.	Pct. Cum.
1	104	77.8	77.8
2	29	21.2	99.0
3 ó más	172	124.4	100.0

Casos válidos= 374

Tabla 125 *Frecuencia del número de ultrasonidos*

Valor	Freq.	Pct.	Pct. Cum.
1	40	55.4	55.4
2	20	27.7	83.1
3	14	19.2	92.3
4	6	8.3	100.0
5	1	1.4	100.0

Casos válidos= 72

Tabla 126 *Distribución de frecuencia para saber si la madre padeció alguna enfermedad durante el embarazo del niño en estudio.*

Valor	Freq.	Pct.	Pct. Cum.
104	1	0.9	0.9
105	1	0.9	1.8
106	1	0.9	2.7
107	1	0.9	3.6
108	1	0.9	4.5
109	1	0.9	5.4
110	1	0.9	6.3
111	1	0.9	7.2
112	1	0.9	8.1
113	1	0.9	9.0
114	1	0.9	9.9
115	1	0.9	10.8
116	1	0.9	11.7
117	1	0.9	12.6
118	1	0.9	13.5
119	1	0.9	14.4
120	1	0.9	15.3
121	1	0.9	16.2
122	1	0.9	17.1
123	1	0.9	18.0
124	1	0.9	18.9
125	1	0.9	19.8
126	1	0.9	20.7
127	1	0.9	21.6
128	1	0.9	22.5
129	1	0.9	23.4
130	1	0.9	24.3
131	1	0.9	25.2
132	1	0.9	26.1
133	1	0.9	27.0
134	1	0.9	27.9
135	1	0.9	28.8
136	1	0.9	29.7
137	1	0.9	30.6
138	1	0.9	31.5
139	1	0.9	32.4
140	1	0.9	33.3
141	1	0.9	34.2
142	1	0.9	35.1
143	1	0.9	36.0
144	1	0.9	36.9
145	1	0.9	37.8
146	1	0.9	38.7
147	1	0.9	39.6
148	1	0.9	40.5
149	1	0.9	41.4
150	1	0.9	42.3
151	1	0.9	43.2
152	1	0.9	44.1
153	1	0.9	45.0
154	1	0.9	45.9
155	1	0.9	46.8
156	1	0.9	47.7
157	1	0.9	48.6
158	1	0.9	49.5
159	1	0.9	50.4
160	1	0.9	51.3
161	1	0.9	52.2
162	1	0.9	53.1
163	1	0.9	54.0
164	1	0.9	54.9
165	1	0.9	55.8
166	1	0.9	56.7
167	1	0.9	57.6
168	1	0.9	58.5
169	1	0.9	59.4
170	1	0.9	60.3
171	1	0.9	61.2
172	1	0.9	62.1
173	1	0.9	63.0
174	1	0.9	63.9
175	1	0.9	64.8
176	1	0.9	65.7
177	1	0.9	66.6
178	1	0.9	67.5
179	1	0.9	68.4
180	1	0.9	69.3
181	1	0.9	70.2
182	1	0.9	71.1
183	1	0.9	72.0
184	1	0.9	72.9
185	1	0.9	73.8
186	1	0.9	74.7
187	1	0.9	75.6
188	1	0.9	76.5
189	1	0.9	77.4
190	1	0.9	78.3
191	1	0.9	79.2
192	1	0.9	80.1
193	1	0.9	81.0
194	1	0.9	81.9
195	1	0.9	82.8
196	1	0.9	83.7
197	1	0.9	84.6
198	1	0.9	85.5
199	1	0.9	86.4
200	1	0.9	87.3
201	1	0.9	88.2
202	1	0.9	89.1
203	1	0.9	90.0
204	1	0.9	90.9
205	1	0.9	91.8
206	1	0.9	92.7
207	1	0.9	93.6
208	1	0.9	94.5
209	1	0.9	95.4
210	1	0.9	96.3
211	1	0.9	97.2
212	1	0.9	98.1
213	1	0.9	99.0
214	1	0.9	99.9
215	1	0.9	100.0

Casos válidos= 374

Tabla 127 *Mes de la concepción del niño en estudio*

Valor	Freq.	Pct.	Pct. Cum.
1	12	3.2	3.2
2	20	5.6	8.8
3	11	3.1	11.9
4	23	6.4	18.3
5	27	7.5	25.8
6	36	10.1	35.9
7	28	7.8	43.7
8	27	7.5	51.2
9	26	7.3	58.5
10	41	11.5	70.0
11	28	7.8	77.8
12	11	3.1	80.9

Casos válidos= 374

Tabla 128. Distribución de frecuencia para el año de diagnóstico de la leucemia únicamente para los casos

Valor	Freq.	Pct.	Pct. Cum.
01	24	27.3	27.3
02	23	26.1	53.4
03	22	25.0	78.4
04	14	16.0	94.4
05	2	2.3	100.0

Casos válidos: 87

Tabla 129. Distribución de frecuencia por tipo de leucemia desarrollada

Valor	Freq.	Pct.	Pct. Cum.
1	85	97.7	97.7
2	2	2.3	100.0

Casos válidos: 87

Tabla 130. Distribución de frecuencia por edad (meses) al diagnóstico de la leucemia

Val.	Freq.	Pct.	Pct. Cum.	Val.	Freq.	Pct.	Pct. Cum.	Val.	Freq.	Pct.	Pct. Cum.
1	1	1.1	1.1	5	2	2.3	4.4	11	2	2.3	4.4
2	1	1.1	2.2	6	2	2.3	6.7	12	1	1.1	7.8
3	1	1.1	3.3	7	1	1.1	7.8	13	1	1.1	8.9
4	1	1.1	4.4	8	1	1.1	8.9	14	1	1.1	10.0
5	1	1.1	5.6	9	1	1.1	10.0	15	1	1.1	11.1
6	1	1.1	6.7	10	1	1.1	11.1	16	1	1.1	12.2
7	1	1.1	7.8	11	1	1.1	12.2	17	1	1.1	13.3
8	1	1.1	8.9	12	1	1.1	13.3	18	1	1.1	14.4
9	1	1.1	10.0	13	1	1.1	14.4	19	1	1.1	15.6
10	1	1.1	11.1	14	1	1.1	15.6	20	1	1.1	16.7
11	1	1.1	12.2	15	1	1.1	16.7	21	1	1.1	17.8
12	1	1.1	13.3	16	1	1.1	17.8	22	1	1.1	18.9
13	1	1.1	14.4	17	1	1.1	18.9	23	1	1.1	20.0
14	1	1.1	15.6	18	1	1.1	20.0	24	1	1.1	21.1
15	1	1.1	16.7	19	1	1.1	21.1	25	1	1.1	22.2
16	1	1.1	17.8	20	1	1.1	22.2	26	1	1.1	23.3
17	1	1.1	18.9	21	1	1.1	23.3	27	1	1.1	24.4
18	1	1.1	20.0	22	1	1.1	24.4	28	1	1.1	25.6
19	1	1.1	21.1	23	1	1.1	25.6	29	1	1.1	26.7
20	1	1.1	22.2	24	1	1.1	26.7	30	1	1.1	27.8
21	1	1.1	23.3	25	1	1.1	27.8	31	1	1.1	28.9
22	1	1.1	24.4	26	1	1.1	28.9	32	1	1.1	30.0
23	1	1.1	25.6	27	1	1.1	29.9	33	1	1.1	31.1
24	1	1.1	26.7	28	1	1.1	31.1	34	1	1.1	32.2
25	1	1.1	27.8	29	1	1.1	32.2	35	1	1.1	33.3
26	1	1.1	28.9	30	1	1.1	33.3	36	1	1.1	34.4
27	1	1.1	30.0	31	1	1.1	34.4	37	1	1.1	35.6
28	1	1.1	31.1	32	1	1.1	35.6	38	1	1.1	36.7
29	1	1.1	32.2	33	1	1.1	36.7	39	1	1.1	37.8
30	1	1.1	33.3	34	1	1.1	37.8	40	1	1.1	38.9
31	1	1.1	34.4	35	1	1.1	38.9	41	1	1.1	40.0
32	1	1.1	35.6	36	1	1.1	39.9	42	1	1.1	41.1
33	1	1.1	36.7	37	1	1.1	41.1	43	1	1.1	42.2
34	1	1.1	37.8	38	1	1.1	42.2	44	1	1.1	43.3
35	1	1.1	38.9	39	1	1.1	43.3	45	1	1.1	44.4
36	1	1.1	40.0	40	1	1.1	44.4	46	1	1.1	45.6
37	1	1.1	41.1	41	1	1.1	45.6	47	1	1.1	46.7
38	1	1.1	42.2	42	1	1.1	46.7	48	1	1.1	47.8
39	1	1.1	43.3	43	1	1.1	47.8	49	1	1.1	48.9
40	1	1.1	44.4	44	1	1.1	48.9	50	1	1.1	50.0
41	1	1.1	45.6	45	1	1.1	50.0	51	1	1.1	51.1
42	1	1.1	46.7	46	1	1.1	51.1	52	1	1.1	52.2
43	1	1.1	47.8	47	1	1.1	52.2	53	1	1.1	53.3
44	1	1.1	48.9	48	1	1.1	53.3	54	1	1.1	54.4
45	1	1.1	50.0	49	1	1.1	54.4	55	1	1.1	55.6
46	1	1.1	51.1	50	1	1.1	55.6	56	1	1.1	56.7
47	1	1.1	52.2	51	1	1.1	56.7	57	1	1.1	57.8
48	1	1.1	53.3	52	1	1.1	57.8	58	1	1.1	58.9
49	1	1.1	54.4	53	1	1.1	58.9	59	1	1.1	60.0
50	1	1.1	55.6	54	1	1.1	60.0	60	1	1.1	61.1
51	1	1.1	56.7	55	1	1.1	61.1	61	1	1.1	62.2
52	1	1.1	57.8	56	1	1.1	62.2	62	1	1.1	63.3
53	1	1.1	58.9	57	1	1.1	63.3	63	1	1.1	64.4
54	1	1.1	60.0	58	1	1.1	64.4	64	1	1.1	65.6
55	1	1.1	61.1	59	1	1.1	65.6	65	1	1.1	66.7
56	1	1.1	62.2	60	1	1.1	66.7	66	1	1.1	67.8
57	1	1.1	63.3	61	1	1.1	67.8	67	1	1.1	68.9
58	1	1.1	64.4	62	1	1.1	68.9	68	1	1.1	70.0
59	1	1.1	65.6	63	1	1.1	70.0	69	1	1.1	71.1
60	1	1.1	66.7	64	1	1.1	71.1	70	1	1.1	72.2
61	1	1.1	67.8	65	1	1.1	72.2	71	1	1.1	73.3
62	1	1.1	68.9	66	1	1.1	73.3	72	1	1.1	74.4
63	1	1.1	70.0	67	1	1.1	74.4	73	1	1.1	75.6
64	1	1.1	71.1	68	1	1.1	75.6	74	1	1.1	76.7
65	1	1.1	72.2	69	1	1.1	76.7	75	1	1.1	77.8
66	1	1.1	73.3	70	1	1.1	77.8	76	1	1.1	78.9
67	1	1.1	74.4	71	1	1.1	78.9	77	1	1.1	80.0
68	1	1.1	75.6	72	1	1.1	80.0	78	1	1.1	81.1
69	1	1.1	76.7	73	1	1.1	81.1	79	1	1.1	82.2
70	1	1.1	77.8	74	1	1.1	82.2	80	1	1.1	83.3
71	1	1.1	78.9	75	1	1.1	83.3	81	1	1.1	84.4
72	1	1.1	80.0	76	1	1.1	84.4	82	1	1.1	85.6
73	1	1.1	81.1	77	1	1.1	85.6	83	1	1.1	86.7
74	1	1.1	82.2	78	1	1.1	86.7	84	1	1.1	87.8
75	1	1.1	83.3	79	1	1.1	87.8	85	1	1.1	88.9
76	1	1.1	84.4	80	1	1.1	88.9	86	1	1.1	90.0
77	1	1.1	85.6	81	1	1.1	90.0	87	1	1.1	91.1
78	1	1.1	86.7	82	1	1.1	91.1	88	1	1.1	92.2
79	1	1.1	87.8	83	1	1.1	92.2	89	1	1.1	93.3
80	1	1.1	88.9	84	1	1.1	93.3	90	1	1.1	94.4
81	1	1.1	90.0	85	1	1.1	94.4	91	1	1.1	95.6
82	1	1.1	91.1	86	1	1.1	95.6	92	1	1.1	96.7
83	1	1.1	92.2	87	1	1.1	96.7	93	1	1.1	97.8
84	1	1.1	93.3	88	1	1.1	97.8	94	1	1.1	98.9
85	1	1.1	94.4	89	1	1.1	98.9	95	1	1.1	100.0

Casos válidos: 87

Tabla 131. Distribución de frecuencia para la edad de la madre cuando se embarazó del niño

Valor	Freq.	Pct.	Pct. Cum.
14	1	1.6	1.6
15	1	1.6	3.2
16	1	1.6	4.8
17	1	1.6	6.4
18	2	3.2	9.6
19	1	1.6	11.2
20	1	1.6	12.8
21	1	1.6	14.4
22	1	1.6	16.0
23	1	1.6	17.6
24	1	1.6	19.2
25	1	1.6	20.8
26	1	1.6	22.4
27	1	1.6	24.0
28	1	1.6	25.6
29	1	1.6	27.2
30	1	1.6	28.8

Fuente: Unidad de Epidemiología Hospital de Pediatría del Centro Médico Nacional Siská XXI

Continuación tabla 131

23	201	8.5	699.7
24	171	6.9	551.1
25	147	5.9	464.2
26	121	4.9	381.8
27	99	4.0	315.9
28	78	3.1	252.8
29	61	2.4	195.5
30	47	1.9	150.7
31	36	1.4	111.1
32	27	1.1	84.2
33	20	0.8	62.6
34	15	0.6	46.7
35	11	0.4	34.1
36	8	0.3	24.9
37	6	0.2	46.9
38	4	0.1	31.1
39	3	0.1	23.3
40	2	0.1	15.5
41	1	0.0	7.7
42	1	0.0	7.7
43	1	0.0	7.7
44	1	0.0	7.7
45	1	0.0	7.7
46	1	0.0	7.7
47	1	0.0	7.7
48	1	0.0	7.7
49	1	0.0	7.7
50	1	0.0	7.7
51	1	0.0	7.7
52	1	0.0	7.7
53	1	0.0	7.7
54	1	0.0	7.7
55	1	0.0	7.7
56	1	0.0	7.7
57	1	0.0	7.7
58	1	0.0	7.7
59	1	0.0	7.7
60	1	0.0	7.7
61	1	0.0	7.7
62	1	0.0	7.7
63	1	0.0	7.7
64	1	0.0	7.7
65	1	0.0	7.7
66	1	0.0	7.7
67	1	0.0	7.7
68	1	0.0	7.7
69	1	0.0	7.7
70	1	0.0	7.7
71	1	0.0	7.7
72	1	0.0	7.7
73	1	0.0	7.7
74	1	0.0	7.7
75	1	0.0	7.7
76	1	0.0	7.7
77	1	0.0	7.7
78	1	0.0	7.7
79	1	0.0	7.7
80	1	0.0	7.7
81	1	0.0	7.7
82	1	0.0	7.7
83	1	0.0	7.7
84	1	0.0	7.7
85	1	0.0	7.7
86	1	0.0	7.7
87	1	0.0	7.7
88	1	0.0	7.7
89	1	0.0	7.7
90	1	0.0	7.7
91	1	0.0	7.7
92	1	0.0	7.7
93	1	0.0	7.7
94	1	0.0	7.7
95	1	0.0	7.7
96	1	0.0	7.7
97	1	0.0	7.7
98	1	0.0	7.7
99	1	0.0	7.7
100	1	0.0	7.7
Total	3311	133.0	10611.0

Casos válidos: 3311 Faltantes: 0

Tabla 132 Distribución de frecuencias para la edad del padre al embarazo de la madre.

Valor	Frec.	Por.	Por. Cum.
16	1	0.0	1.7
17	1	0.0	3.4
18	1	0.0	5.1
19	1	0.0	6.8
20	1	0.0	8.5
21	1	0.0	10.2
22	1	0.0	11.9
23	1	0.0	13.6
24	1	0.0	15.3
25	1	0.0	17.0
26	1	0.0	18.7
27	1	0.0	20.4
28	1	0.0	22.1
29	1	0.0	23.8
30	1	0.0	25.5
31	1	0.0	27.2
32	1	0.0	28.9
33	1	0.0	30.6
34	1	0.0	32.3
35	1	0.0	34.0
36	1	0.0	35.7
37	1	0.0	37.4
38	1	0.0	39.1
39	1	0.0	40.8
40	1	0.0	42.5
41	1	0.0	44.2
42	1	0.0	45.9
43	1	0.0	47.6
44	1	0.0	49.3
45	1	0.0	51.0
46	1	0.0	52.7
47	1	0.0	54.4
48	1	0.0	56.1
49	1	0.0	57.8
50	1	0.0	59.5
51	1	0.0	61.2
52	1	0.0	62.9
53	1	0.0	64.6
54	1	0.0	66.3
55	1	0.0	68.0
56	1	0.0	69.7
57	1	0.0	71.4
58	1	0.0	73.1
59	1	0.0	74.8
60	1	0.0	76.5
61	1	0.0	78.2
62	1	0.0	79.9
63	1	0.0	81.6
64	1	0.0	83.3
65	1	0.0	85.0
66	1	0.0	86.7
67	1	0.0	88.4
68	1	0.0	90.1
69	1	0.0	91.8
70	1	0.0	93.5
71	1	0.0	95.2
72	1	0.0	96.9
73	1	0.0	98.6
74	1	0.0	100.3
75	1	0.0	102.0
76	1	0.0	103.7
77	1	0.0	105.4
78	1	0.0	107.1
79	1	0.0	108.8
80	1	0.0	110.5
81	1	0.0	112.2
82	1	0.0	113.9
83	1	0.0	115.6
84	1	0.0	117.3
85	1	0.0	119.0
86	1	0.0	120.7
87	1	0.0	122.4
88	1	0.0	124.1
89	1	0.0	125.8
90	1	0.0	127.5
91	1	0.0	129.2
92	1	0.0	130.9
93	1	0.0	132.6
94	1	0.0	134.3
95	1	0.0	136.0
96	1	0.0	137.7
97	1	0.0	139.4
98	1	0.0	141.1
99	1	0.0	142.8
100	1	0.0	144.5
Total	3311	133.0	10611.0

Casos válidos: 3311 Faltantes: 0

Tabla 133 Distribución de la número de embarazos de la madre respecto al niño en estudio.

Valor	Frec.	Por.	Por. Cum.
1	1	0.0	2.1
2	1	0.0	4.2
3	1	0.0	6.3
4	1	0.0	8.4
5	1	0.0	10.5
6	1	0.0	12.6
7	1	0.0	14.7
8	1	0.0	16.8
9	1	0.0	18.9
10	1	0.0	21.0
11	1	0.0	23.1
12	1	0.0	25.2
13	1	0.0	27.3
14	1	0.0	29.4
15	1	0.0	31.5
16	1	0.0	33.6
17	1	0.0	35.7
18	1	0.0	37.8
19	1	0.0	39.9
20	1	0.0	42.0
21	1	0.0	44.1
22	1	0.0	46.2
23	1	0.0	48.3
24	1	0.0	50.4
25	1	0.0	52.5
26	1	0.0	54.6
27	1	0.0	56.7
28	1	0.0	58.8
29	1	0.0	60.9
30	1	0.0	63.0
31	1	0.0	65.1
32	1	0.0	67.2
33	1	0.0	69.3
34	1	0.0	71.4
35	1	0.0	73.5
36	1	0.0	75.6
37	1	0.0	77.7
38	1	0.0	79.8
39	1	0.0	81.9
40	1	0.0	84.0
41	1	0.0	86.1
42	1	0.0	88.2
43	1	0.0	90.3
44	1	0.0	92.4
45	1	0.0	94.5
46	1	0.0	96.6
47	1	0.0	98.7
48	1	0.0	100.8
49	1	0.0	102.9
50	1	0.0	105.0
51	1	0.0	107.1
52	1	0.0	109.2
53	1	0.0	111.3
54	1	0.0	113.4
55	1	0.0	115.5
56	1	0.0	117.6
57	1	0.0	119.7
58	1	0.0	121.8
59	1	0.0	123.9
60	1	0.0	126.0
61	1	0.0	128.1
62	1	0.0	130.2
63	1	0.0	132.3
64	1	0.0	134.4
65	1	0.0	136.5
66	1	0.0	138.6
67	1	0.0	140.7
68	1	0.0	142.8
69	1	0.0	144.9
70	1	0.0	147.0
71	1	0.0	149.1
72	1	0.0	151.2
73	1	0.0	153.3
74	1	0.0	155.4
75	1	0.0	157.5
76	1	0.0	159.6
77	1	0.0	161.7
78	1	0.0	163.8
79	1	0.0	165.9
80	1	0.0	168.0
81	1	0.0	170.1
82	1	0.0	172.2
83	1	0.0	174.3
84	1	0.0	176.4
85	1	0.0	178.5
86	1	0.0	180.6
87	1	0.0	182.7
88	1	0.0	184.8
89	1	0.0	186.9
90	1	0.0	189.0
91	1	0.0	191.1
92	1	0.0	193.2
93	1	0.0	195.3
94	1	0.0	197.4
95	1	0.0	199.5
96	1	0.0	201.6
97	1	0.0	203.7
98	1	0.0	205.8
99	1	0.0	207.9
100	1	0.0	210.0
Total	3311	133.0	10611.0

Casos válidos: 3311 Faltantes: 0

Fuente: Unidad de Epidemiología, Hospital de Pediatría del Centro Médico Nacional Sotelo XXI

Tabla 134. Frecuencia para diagnósticos de internamiento.

Valor	Freq.	Por.	Por. Cum.
111	244	79.5	79.5
112	15	4.7	84.2
113	2	0.6	84.8
114	2	0.6	85.4
115	2	0.6	86.0
116	1	0.3	86.3
117	4	1.2	87.5
118	1	0.3	87.8
119	4	1.2	89.0
120	6	1.9	90.9
121	11	3.4	94.3
122	2	0.6	94.9
123	2	0.6	95.5
124	2	0.6	96.1
TOTAL	307	100.0	100.0

Casos válidos=301. Datos faltantes= 6

Tabla 135. Distribución de los pacientes por hospital de procedencia.

Valor	Freq.	Por.	Por. Cum.
111	152	47.4	47.4
112	253	77.6	125.0
TOTAL	314	100.0	

Casos válidos=314

Suente: Unidad de Epidemiología, Hospital de Pediatría del Centro Médico Nacional Siglo XXI

C. PORCENTILES DE LA DISTRIBUCIÓN CHI-CUADRA DA

d.f.	$\chi^2_{.99}$	$\chi^2_{.95}$	$\chi^2_{.90}$	$\chi^2_{.85}$	$\chi^2_{.80}$	$\chi^2_{.75}$	$\chi^2_{.70}$	$\chi^2_{.65}$	$\chi^2_{.60}$	$\chi^2_{.55}$	$\chi^2_{.50}$	$\chi^2_{.45}$	$\chi^2_{.40}$	$\chi^2_{.35}$	$\chi^2_{.30}$	$\chi^2_{.25}$	$\chi^2_{.20}$	$\chi^2_{.15}$	$\chi^2_{.10}$	$\chi^2_{.05}$	$\chi^2_{.01}$	$\chi^2_{.001}$	
1	.00016	.00063	.0039	.016	.064	.15	.46	1.07	1.64	2.71	3.84	5.41	6.64	10.83	10.83								
2	.02	.04	.10	.21	.45	.71	1.39	2.41	3.22	4.60	5.99	7.82	9.21	13.82	13.82								
3	.12	.18	.35	.58	1.00	1.42	2.37	3.66	4.64	6.25	7.82	9.84	11.34	16.27	16.27								
4	.30	.43	.71	1.06	1.65	2.20	3.36	4.88	5.99	7.78	9.49	11.67	13.28	18.46	18.46								
5	.55	.75	1.14	1.61	2.34	3.00	4.35	6.06	7.29	9.24	11.07	13.39	15.09	20.52	20.52								
6	.87	1.13	1.64	2.20	3.07	3.83	5.35	7.23	8.56	10.64	12.59	15.03	16.81	22.46	22.46								
7	1.24	1.56	2.17	2.83	3.82	4.67	6.35	8.38	9.80	12.02	14.07	16.62	18.48	24.32	24.32								
8	1.65	2.03	2.73	3.49	4.59	5.53	7.34	9.52	11.03	13.36	15.51	18.17	20.09	26.12	26.12								
9	2.09	2.53	3.32	4.17	5.38	6.39	8.34	10.66	12.24	14.68	16.92	19.68	21.67	27.88	27.88								
10	2.56	3.06	3.94	4.86	6.18	7.27	9.34	11.78	13.44	15.99	18.31	21.16	23.21	29.59	29.59								
11	3.05	3.61	4.58	5.58	6.99	8.15	10.34	12.90	14.63	17.28	19.68	22.62	24.72	31.26	31.26								
12	3.57	4.18	5.23	6.30	7.81	9.04	11.34	14.01	15.81	18.55	21.03	24.05	26.22	32.91	32.91								
13	4.11	4.76	5.89	7.04	8.63	9.93	12.34	15.12	16.98	19.81	22.36	25.47	27.69	34.53	34.53								
14	4.66	5.37	6.57	7.79	9.47	10.82	13.34	16.22	18.15	21.06	23.68	26.87	29.14	36.12	36.12								
15	5.23	5.98	7.26	8.55	10.31	11.72	14.34	17.32	19.31	22.31	25.00	28.26	30.58	37.70	37.70								
16	5.81	6.61	7.96	9.31	11.15	12.62	15.34	18.42	20.46	23.54	26.30	29.43	32.00	39.29	39.29								
17	6.41	7.26	8.67	10.08	12.00	13.53	16.34	19.51	21.62	24.77	27.59	31.00	33.41	40.78	40.78								
18	7.02	7.91	9.39	10.86	12.86	14.44	17.34	20.60	22.76	25.99	28.87	32.35	34.80	42.31	42.31								
19	7.63	8.57	10.12	11.65	13.72	15.35	18.34	21.69	23.90	27.20	30.14	33.69	36.19	43.82	43.82								
20	8.26	9.24	10.85	12.44	14.58	16.27	19.34	22.78	25.04	28.41	31.43	35.02	37.57	45.32	45.32								
21	8.90	9.92	11.59	13.24	15.44	17.18	20.34	23.86	26.17	29.62	32.67	36.34	38.93	46.80	46.80								
22	9.54	10.60	12.34	14.04	16.31	18.10	21.34	24.84	27.30	30.84	33.92	37.66	40.29	48.27	48.27								
23	10.20	11.29	13.09	14.85	17.19	19.02	22.34	26.02	28.43	32.07	35.17	38.97	41.64	49.73	49.73								
24	10.86	11.99	13.85	15.66	18.06	19.94	23.34	27.10	29.55	33.20	36.42	40.27	42.98	51.18	51.18								
25	11.52	12.70	14.61	16.47	18.84	20.87	24.34	28.17	30.68	34.38	37.65	41.57	44.31	52.62	52.62								
26	12.20	13.41	15.38	17.29	19.82	21.79	25.34	29.25	31.80	35.56	38.88	42.86	45.64	54.05	54.05								
27	12.88	14.12	16.15	18.11	20.70	22.72	26.34	30.32	32.91	36.74	40.11	44.14	46.96	55.48	55.48								
28	13.56	14.85	16.93	18.94	21.59	23.65	27.34	31.39	34.03	37.92	41.34	45.42	48.28	56.89	56.89								
29	14.26	15.57	17.71	19.77	22.48	24.58	28.34	32.46	35.14	39.09	42.56	46.69	49.50	58.30	58.30								
30	14.95	16.31	18.49	20.60	23.36	25.51	29.34	33.53	36.25	40.26	43.77	47.86	50.89	60.70	60.70								

Adaptado de la tabla IV de Fisher y Yates: *Statistical Tables for Biological, Agricultural and Medical Research*, publicado por Logman Group Ltd., Londres.

D. DEFINICIÓN DE TÉRMINOS MÉDICOS

Anuria	Ausencia de orina que puede presentarse, entre otros casos, cuando el enfermo está en choque.
Cardiopatía congénita	Enfermedad del corazón perteneciente al nacimiento.
Cefalea	Forma de jaqueca producida por dilatación de las arterias cerebrales.
Cianoderma	Azulado en la piel o epitelio, sucede cuando los niños nadan mucho en agua fría.
Convulsión	Estado inconsciente del paciente caracterizado por una postura rígida del cuerpo y cese en la respiración, aunque en una segunda fase puede haber movimientos rítmicos de los miembros.
Cretinismo	Ausencia total de la función tiroidea en la infancia
Cromosoma	Estructura proteica localizada en el núcleo de las células y que contiene cadenas de DNA (Ácido desoxirribonucleico) que transmite la información genética, haciéndose particularmente visible durante la fase de división celular.
Disfonia	Voz débil.
Displasia	Alteración de una determinada función fisiológica.
Enzima	Grupo de sustancias químicas de carácter proteico que, producidas por el organismo aceleran o catalizan infinidad de reacciones bioquímicas que son imprescindibles para la realización de los múltiples procesos metabólicos que continuamente se están realizando en el organismo
Ganglio linfático	Órgano que se intercala en el trayecto de los vasos linfáticos y funciona como filtro biológico, aparte de poseer función linfocitopoyética.
Hematuria	Emisión por la orina de sangre pura o mezclada a la secreción urinaria, síntoma de enfermedades diversas.
Hemiplejía	Pálisis de un lado del cuerpo.
Hepatitis	Inflamación del hígado.
Hidrocefalia	Aumento de tamaño de la cavidad craneana por exceso de producción o defecto de absorción del líquido cefalorraquídeo que se acumula en el encefalo.
Hipertermia	Temperatura por encima de lo normal.
Hipotermia	Temperatura por debajo de lo normal.
Ictericia	Coloración amarilla de la piel, mucosas y secreciones debida a la presencia de pigmentos biliares en la sangre.
Leucocitos	Glóbulos blancos.
Linfoma de Hodgkin	Aumento de tamaño y consistencia de los ganglios linfáticos.
Médula ósea	Sitio donde se realiza la formación y maduración de los eritrocitos, leucocitos, granulocitos y megacariocitos.
Meningitis	Inflamación de las membranas meníngeas.
Meningoencefalitis	Inflamación simultánea del encefalo y las meninges.
Mialgia	Dolor muscular.

Mongolismo	Estado que se refiere a idiotez.
Neumonía	Enfermedad infecciosa localizada en el pulmón, producida por diversos gérmenes; caracterizada por la congestión aguda de un lóbulo pulmonar con exudado fibrinoso que llena los alvéolos.
Neuroblastoma	Tumor de células embrionarias que originan neuronas.
Oligofrenia	Deficiencia mental
Polidactilia	Enfermedad genética (exceso de dedos)
Quimioterapia	Tratamiento por sustancias químicas.
Quiste	Forma latente de vida de un parásito
Raquitismo	Enfermedad causada por deficiencia de vitamina D y afecta principalmente a los niños e infantes
Reflujo	Retroceso de un líquido.
Soplo en el corazón	Fenómeno producido cuando la corriente sanguínea pasa de una porción estrecha a otra de mayor amplitud, formando torbellinos fluidos que al chocar contra las paredes del corazón o de los vasos grandes las hacen vibrar produciendo un ruido de mayor intensidad.
Tiroides	Glándula de secreción interna situada en la región cervical a ambos lados de la tráquea.

**UNIDAD DE INVESTIGACIÓN EN EPIDEMIOLOGÍA
CLÍNICA SUBJEFATURA DE INVESTIGACIÓN
I.M.S.S.**

La leucemia es el principal cáncer en el niño y aún no se conoce el mecanismo por el cual se produce, ni los factores que la determinan, por lo cual esta unidad está llevando a cabo un estudio para saber cuáles son los factores de la misma, y ampliar el conocimiento existente sobre este problema de salud.

INSTRUCCIONES DE LLENADO:

1. Lea cuidadosamente cada pregunta
2. Si la pregunta es abierta debe contestarse en forma breve
3. No invada los espacios que se encuentran en la columna derecha.
4. En las preguntas que tienen opciones marque con una cruz la respuesta que corresponde
5. En las preguntas que se requiera contestar cifras en años, su respuesta será en años cumplidos, si la respuesta es menos de un año anote cero en los años y escriba en los meses.
6. No deje casillas vacías, anote cero a la izquierda cuando su respuesta sólo sea de un dígito.

Ejemplo:	Fecha	05	03	92
		día	mes	año

Día 5 del mes de Marzo de 1992.

Nota: Esta información es de carácter confidencial.

**PROYECTO
"EXPOSICIÓN A CAMPOS ELECTROMAGNÉTICOS Y SU ASOCIACIÓN CON LEUCEMIA EN NIÑOS
RESIDENTES DEL DISTRITO FEDERAL".**

CUESTIONARIO

Anoté el número correspondiente a cada pregunta en la casilla correspondiente.

• **FICHA DE IDENTIFICACIÓN**

- 1.- Folio _____
- 2.- Nombre del paciente _____
- 3.- Sexo _____ 1) masc. 0) fem.
- 4.- Edad del niño _____ años _____ meses
- 5.- Fecha exacta de nacimiento _____ día _____ mes _____ año
- 6.- Domicilio del paciente _____
- 7.- Teléfono _____
- 8.- Nombre del entrevistado _____
- 9.- Parentesco del entrevistado _____
- 10.- Fecha de la entrevista _____ día _____ mes _____ año

• **ANTECEDENTES HEREDO FAMILIARES.**

- 11.- ¿Algún de sus familiares ha padecido algún tipo de cáncer? _____ 1) si 0) no
- Si la contestación es no pase a la pregunta 15
- 12.- Cuántos _____
- | | |
|--|---------------------|
| 13.- Parentesco que tienen o tenían con el niño: | 14.- Tipo de cáncer |
| 1) _____ | 1) _____ |
| 2) _____ | 2) _____ |
| 3) _____ | 3) _____ |
- 15.- ¿Alguien de su familia nació con algún defecto o enfermedad? _____ 1) si 0) no
- 16.- ¿Cuántos? _____
- | | |
|------------------------------|----------------------|
| 17.- Parentesco con el niño: | 18.- Tipo de defecto |
| 1) _____ | 1) _____ |
| 2) _____ | 2) _____ |
| 3) _____ | 3) _____ |
- 19.- ¿Alguien de su familia tiene o tuvo síndrome de Down o Mongóllamo? _____ 1) si 0) no

Si la contestación es no pase a la pregunta 21

20.- ¿Cuál es el parentesco con el niño? _____

*** ANTECEDENTES PERSONALES NO PATOLÓGICOS**

21.- La casa donde vive es:

1) Propia _____ 2) Rentada _____ 3) Prestada _____ 4) Desconoce _____

22.- Si es rentada cuánto pagan mensualmente (miles) _____

23.- La casa donde vive tiene:

1) Agua _____ 2) Luz _____ 4) Drenaje _____

24.- ¿Cuántos cuartos tiene para dormir (sin contar cocina, sala y comedor)? _____

25.- ¿Cuántas personas duermen en cada cuarto?

26.- Total de personas que habitan la casa _____

27.- Ingreso familiar mensual (en miles) _____

28.- Estrato socioeconómico:

1) Alto _____ 2) Medio _____ 3) Bajo _____

29.- ¿Cuanto paga de luz mensualmente aproximadamente (en miles) _____

30.- ¿Con que calienta sus alimentos?

01) Gas _____	16) Hornos eléctricos _____
02) Parríleo _____	32) Hornos de microondas _____
04) Carbón _____	64) Lefla _____
08) Parrilla eléctrica _____	

31.- ¿Cuánto tiempo tiene de vivir en el lugar actual?

_____ años _____ meses

32.- ¿Existe cerca de donde usted vive alguna de las siguientes instalaciones y a cuántos metros se encuentran de su casa?

1) Transformador de luz _____	1) si	0) no	Distancia _____
2) Cables de alta tensión _____	1) si	0) no	Distancia _____
3) Subestación eléctrica _____	1) si	0) no	Distancia _____
4) Torres y cables de alta tensión _____	1) si	0) no	Distancia _____

Si la contestación fue NO para todos los incisos anteriores pase a la pregunta 36

33.- ¿Ya existían cuando usted llegó a vivir ahí? _____ 1) si 0) no

34.- ¿Cuanto tiempo después lo pusieron?

_____ años _____ meses _____ semanas

35.- Desde hace cuánto tiempo existen _____ años _____ meses

36.- ¿Siempre ha vivido aquí? _____ 1) si 0) no

Si la contestación es SI pase a la pregunta 41

37.- Donde vivía antes existía alguna de las siguientes instalaciones, y en caso de existir a cuántos metros se encontraban de su casa?

- | | | | | | |
|------------------------------------|-------|-------|-------|-----------|-------|
| 1) Transformador de luz | _____ | 1) sí | 0) no | Distancia | _____ |
| 2) Cables de alta tensión | _____ | 1) sí | 0) no | Distancia | _____ |
| 3) Subestación eléctrica | _____ | 1) sí | 0) no | Distancia | _____ |
| 4) Torres y cables de alta tensión | _____ | 1) sí | 0) no | Distancia | _____ |

38.- ¿Cuándo llegó a vivir allí ya existían? _____ 1) sí 0) no

Si la respuesta es SÍ pase a la pregunta 41

39.- ¿Cuanto tiempo después los pusieron?

_____ años _____ meses _____ semanas

40.- ¿Cuánto tiempo vivió allí? _____ años _____ meses

*** ESCOLARIDAD**

41.- ¿Su hijo fue o va a la guardería? _____ 1) sí 0) no

Si la respuesta es NO pase a la pregunta 43

42.- ¿Desde que edad? (meses) _____

43.- ¿Su hijo va a la escuela? _____ 1) sí 0) no

Si la respuesta es NO pase a la pregunta 53

44.- Existe cerca de la escuela del niño alguna de las siguientes instalaciones, y en caso de existir alguna ¿a cuántos metros se encuentra de la misma?

- | | | | | | |
|------------------------------------|-------|-------|-------|-----------|-------|
| 1) Transformador de luz | _____ | 1) sí | 0) no | Distancia | _____ |
| 2) Cables de alta tensión | _____ | 1) sí | 0) no | Distancia | _____ |
| 3) Subestación eléctrica | _____ | 1) sí | 0) no | Distancia | _____ |
| 4) Torres y cables de alta tensión | _____ | 1) sí | 0) no | Distancia | _____ |

Si la respuesta ha sido NO para todas las opciones anteriores pase a la pregunta 48

45.- ¿Cuántos años ha estado en esa escuela? _____

46.- ¿Cuántas horas permanece al día en la escuela? _____

47.- ¿En que turno? _____

48.- Antes de estar en esa escuela ¿estuvo en otra? _____ 1) sí 0) no

Si la contestación es NO pase a la pregunta 53

49.- ¿Existían cerca de la escuela anterior alguna de las siguientes instalaciones siguientes, y en caso de existir ¿a que distancia (en metros) se encontraba de la misma?

- | | | | | | |
|------------------------------------|-------|-------|-------|-----------|-------|
| 1) Transformador de luz | _____ | 1) sí | 0) no | Distancia | _____ |
| 2) Cables de alta tensión | _____ | 1) sí | 0) no | Distancia | _____ |
| 3) Subestación eléctrica | _____ | 1) sí | 0) no | Distancia | _____ |
| 4) Torres y cables de alta tensión | _____ | 1) sí | 0) no | Distancia | _____ |

Si la respuesta es NO para todas las opciones anteriores pase a la pregunta 53

50.- ¿Cuánto tiempo estuvo en esa escuela? _____
 años meses

51.- ¿Cuántas horas permanecía en la escuela? _____

52.- ¿En que turno? _____

*** ESCOLARIDAD DE LOS PADRES**

53.- Escolaridad del padre: _____

54.- Número de años cursados _____

55.- Escolaridad de la madre _____

56.- Número de años cursados _____

*** OCUPACIÓN**

57.- Cuando la madre se embarazó ¿cuál era la ocupación del padre? _____

58.- ¿Que actividad realizaba? _____

59.- A que se dedicaba la empresa, negocio, institución o fábrica donde trabajaba? _____

60.- Tarea mas importante que desempeñaba _____

61.- Señale si en su trabajo se exponía a alguno de los siguientes productos, y especifique a cuáles, así como la frecuencia de exposición a cada uno de ellos, indicando además el número de horas en promedio que se exponía a cada uno de ellos:

	1) si	0) no	Frecuencia (Mucho Poco Regular)	Horas
1) Solventes (derivados del petróleo)	_____	_____	_____	_____
2) Insecticidas	_____	_____	_____	_____
3) Thiner	_____	_____	_____	_____
4) Pinturas	_____	_____	_____	_____
5) Fertilizantes	_____	_____	_____	_____
6) Gasolina	_____	_____	_____	_____
7) Otros	_____	_____	_____	_____

62.- Tiempo laborando en ese lugar _____
 años meses semanas

63.- Utilizaba algún tipo de protección para desempeñar sus actividades? _____ 1) si 0) no

64.- ¿Qué equipo? _____

65.- Equipo o maquinaria con la que laboraba _____

66.- Otras actividades _____

LAS SIGUIENTES 4 PREGUNTAS UNICAMENTE SE REALIZARÁN A LOS PAPAS DE LOS NIÑOS CON LEUCEMIA, DE NO TENERLA PASAR A LA PREGUNTA 71

67.- ¿En qué trabajaba el padre cuando el niño enfermó de leucemia? _____

68.- ¿Que actividad realizaba con mayor frecuencia el padre del niño cuando se enfermo de leucemia? _____

69.- ¿Estuvo en contacto con químicos? _____ 1) si 0) no

Si la contestación es NO pase a la pregunta 71

70.- ¿Cuáles? _____

71.- ¿En que trabajaba la madre antes del embarazo de su hijo? _____

72.- ¿Durante cuántas horas al día? _____

73.- ¿A que se dedicaba la madre del niño cuando estaba embarazada? _____

74.- ¿Cuántas horas al día? _____

75.- ¿Durante el embarazo del niño estuvo en contacto con productos químicos? _____ 1) si 0) no

Si la pregunta es NO pase a la pregunta 79

76.- ¿Cuáles? _____

NOTA: Las siguientes dos preguntas son únicamente para los papas de los niños con leucemia.

77.- ¿A que se dedicaba la madre del niño cuando su hijo enfermó de leucemia? _____

78.- ¿Durante cuántas horas? _____

***ANTECEDENTES PATOLÓGICOS**

DEL PADRE:

79.- Su esposa ¿consume bebidas alcohólicas? _____ 1) si 0) no

Si la contestación es NO pase a la pregunta 81

80.- Desde que edad (años) _____ (si no lo hace poner 00)

81.- ¿Cuándo estuvo embarazada su esposa se emborrachaba? _____ 1) si 0) no

82.- ¿Con que frecuencia? _____

83.- ¿Que bebida es la que con mayor frecuencia toma?

1) Vino _____ 4) Ron _____ 7) Otros _____
 2) Cerveza _____ 5) Tequila _____
 3) Brandy _____ 6) Pulque _____

84.- En caso de haber respondido otros especifique _____

DE LA MADRE:

85.- ¿Consumo o toma bebidas alcohólicas? _____ 1) si 0) no

86.- ¿Desde que edad? _____ años

87.- ¿Qué bebida es la que con mayor frecuencia toma?

- 1) Vino _____ 4) Ron _____ 7) Otros _____
 2) Cerveza _____ 5) Tequila _____
 3) Brandy _____ 6) Pulque _____

88.- En caso de haber respondido otros especifique _____

89.- ¿Se emborrachaba? _____ 1) si 0) no

Si la contestación es NO pase a la pregunta 91

90.- ¿Cuántas veces al año? _____

91.- ¿Tomó o consumió bebidas alcohólicas durante su embarazo? _____ 1) si 0) no

Si la contestación es NO pase a la pregunta 94

92.- ¿En que mes del embarazo?

- 1) Primero _____ 4) Cuarto _____ 7) Séptimo _____
 2) Segundo _____ 5) Quinto _____ 8) Octavo _____
 3) Tercero _____ 6) Sexto _____ 9) Noveno _____
 10) Todo el embarazo _____

93.- ¿Cuántas copas ingería en promedio cada vez que tomaba? _____

*** TABAQUISMO**

DEL PADRE:

94.- Su esposo ¿fuma actualmente? _____ 1) si 0) no

95.- Su esposo ¿fumó durante el embarazo del niño? _____ 1) si 0) no

Si la respuesta es NO pase a la pregunta 99

96.- ¿Fumaba cerca de usted? _____ 1) si 0) no

97.- ¿En que mes de su embarazo fumó mas?

- 1) Primero _____ 4) Cuarto _____ 7) Séptimo _____
 2) Segundo _____ 5) Quinto _____ 8) Octavo _____
 3) Tercero _____ 6) Sexto _____ 9) Noveno _____
 10) Todo el embarazo _____

98.- ¿Cuántos cigarrillos fumaba al día? _____

DE LA MADRE:

99.- ¿Fumaba antes del embarazo? _____ 1) si 0) no

100.- ¿Fumó durante su embarazo? _____ 1) si 0) no

Si la contestación es NO pase a la pregunta 103

101.- ¿En que mes de su embarazo fumó mas?

- 1) Primero _____ 4) Cuarto _____ 7) Séptimo _____
 2) Segundo _____ 5) Quinto _____ 8) Octavo _____
 3) Tercero _____ 6) Sexto _____ 9) Noveno _____
 10) Todo el embarazo _____

102.- ¿Cuántos cigarrillos fumaba? _____

103.- ¿Fuma actualmente? _____ 1) si 0) no

LAS SIGUIENTES PREGUNTAS SE HARAN A CASOS Y CONTROLES, ÚNICAMENTE SE DEJARA DE HACER MENSIÓN A LA PALABRA, LEUCEMIA EN EL NIÑO CONTROL.

OTROS ANTECEDENTES DEL MENOR:

104.- Desde que nació su hijo hasta antes de que le diagnosticaran leucemia ¿le tomaron alguna radiografía (rayos x)? _____ 1) si 0) no

105.- ¿Cuántas? _____

106.- ¿A que edad? _____ años _____ meses _____ años _____ meses

107.- Desde que nació su hijo hasta antes de que le diagnosticaran leucemia ¿le tomaron algún ultrasonido? _____ 1) si 0) no

108.- ¿Cuántos? _____

109.- ¿Cuál fue el motivo del ultrasonido? _____

110.- Su hijo ¿nació con algún defecto? _____ 1) si 0) no

11.- ¿Cuál defecto? _____

112.- Antes de que su hijo (a) le diagnosticaran leucemia estuvo el niño (a) en contacto con alguno de los siguientes productos, en caso de que haya estado, especifique cuáles, así como la frecuencia de exposición, indicando además el número promedio de horas que estaba en contacto:

	1) si	0) no	Frecuencia (Mucho Poco Regular)	Horas
1) Solventes (derivados del petróleo)	_____	_____	_____	_____
2) Insecticidas	_____	_____	_____	_____
3) Thiner	_____	_____	_____	_____
4) Pinturas	_____	_____	_____	_____
5) Fertilizantes	_____	_____	_____	_____
6) Gasolina	_____	_____	_____	_____
7) Colorantes para el cabello	_____	_____	_____	_____
8) Fijadores para el cabello	_____	_____	_____	_____
9) Otros	_____	_____	_____	_____

NOTA: La siguiente pregunta sólo se realizara a los niños con leucemia.

113.- Su hijo ¿padeció algunas de las siguientes enfermedades antes de que le diera leucemia?

1) Varicela _____ 2) Rubéola _____
 4) Sarampión _____ 8) Paperas _____
 00) Ninguna _____

NOTA: La siguiente pregunta se realizara únicamente a los niños controles

114.- Su hijo ¿ha padecido alguna de las siguientes enfermedades?
 1) Varicela _____ 2) Rubéola _____
 4) Sarampión _____ 8) Paperas _____
 00) Ninguna _____

115.- Su hijo ¿fue muy enfermizo durante su primer año de vida? _____ 1) si 0) no

116.- ¿Fue mas enfermizo que sus otros hermanos? _____ 1) si 0) no

117.- Durante el primer año de vida, ¿estuvo el niño internado en algún hospital por algún padecimiento? _____ 1) si 0) no

Si la respuesta es NO pase a la pregunta 119

118.- ¿Cuál fue el padecimiento? _____

***ANTECEDENTES GINECOOBSTETRICOS Y PATOLÓGICOS DE LA MADRE**

119.- Cuantas veces se ha embarazado? _____

120.- ¿Cuantos hijos ha tenido? _____

121.- ¿Cuantos abortos ha tenido? _____

Si no ha tenido ninguno pase a la pregunta 126

122.- ¿Hubo algo que los provocara? _____ 1) si 0) no

123.- ¿Cuál fue la causa? _____

124.- ¿Tuvo abortos antes del niño enfermo? _____ 1) si 0) no

125.- ¿Cuántos? _____

126.- ¿Alguno de sus hijos nació muerto? _____ 1) si 0) no

Si la contestación es NO pase a la pregunta 129

127.- ¿Cuantos? _____

128.- ¿Alguno fue antes del niño enfermo? _____ 1) si 0) no

129.- A la madre del niño durante su embarazo le tomaron alguna radiografía (rayos x)? _____ 1) si 0) no

Si la contestación es NO pase a la pregunta 132

130.- ¿Cuantas? _____

131.- ¿En que mes del embarazo? _____

132.- A la madre del niño durante su embarazo ¿le tomaron algún ultrasonido? _____ 1) si 0) no

Si la respuesta es NO pase a la pregunta 135

133.- ¿Cuantos? _____

134.- ¿Cual fue el motivo? _____

135.- Cuando estuvo embarazada del niño que estamos entrevistando ¿padeció alguna de las siguientes enfermedades?

- | | |
|---------------------------------|-------------------|
| 1) Varicela _____ | 2) Rubéola _____ |
| 4) Sarampión _____ | 8) Paperas _____ |
| 16) Resfriados o catarros _____ | 00) Ninguna _____ |

Si no padeció ninguna de las anteriores pase a la pregunta 137

136.- ¿Quién le hizo el diagnóstico? _____

137.- ¿En que mes se embarazó del niño enfermo? (mes de la concepción) _____

NOTA: Las siguientes dos preguntas son únicamente para los padres de los niños con leucemia, si no pase a la pregunta 141

138.- Fecha del diagnóstico de la enfermedad (leucemia) en el niño

____ año ____ mes ____ día

139.- ¿Tipo de leucemia? _____

140.- ¿Edad del niño cuando se le diagnosticó la enfermedad? _____

____ años ____ meses

141.- ¿Edad de la madre cuando se embarazó del niño (años cumplidos)? _____

142.- ¿Edad del padre cuando se embarazó la madre del niño? _____

143.- ¿Número de embarazos? _____

144.- ¿Cuanto fue el peso al nacer? _____

145.- ¿Cuanto midió al nacer? _____

146.- ¿De cuántos meses nació el niño? _____

147.- El paciente por el que se pregunta ¿tiene leucemia? _____ 1) si 0) no

Si la respuesta es No pase a la pregunta siguiente

148.- Diagnóstico de internamiento o motivo por el cual se encuentra en el hospital? _____

149.- ¿Hospital de procedencia? _____

BIBLIOGRAFÍA

- 1) Bisquerra Alzina;
"Introducción conceptual al análisis multivariable, un enfoque informático con los paquetes estadísticos SPSS-X, BMDP, LISREL y SPAD";
Vol. 1, Ed. PPU.
- 2) Bothman Kenneth J.;
"Epidemiología. Métodos Estadísticos";
Ed. Díaz de Santos, Madrid 1987.
- 3) David W. Hosmer, Stanley Lemeshow;
"Applied Logistic Regression";
Ed. Wiley Interscience, 1989.
- 4) D. Collet
"Modelling Binary Data";
Ed. Chapman & Hall.
- 5) Holland, James F. & Miescher, Peter A.;
"Leucemia Linfomas";
Ed. Científico Médica.
- 6) Jencer Milos;
"Epidemiología. Principios, técnicas y aplicaciones";
Ed. Salvat, Barcelona.
- 7) Joseph F. hair,
Ralph E. Anderson;
"Multivariate Data Analysis with Readings";
Ed. Macmillan Publishing Company.
- 8) Mazzafero, Vicente Enrique;
"Epidemiología Fundamental y Aplicada a la Evaluación de Servicios Hospitalarios y Acciones de Salud Pública";
Saubert, Leslie, Buenos Aires.

- 9) Pochedly Carl;
"Leucemia en el niño";
Ed. Barcelona , 1974.
- 10) Robert C. Elston, William D. Johnson;
"Principios de Bioestadística";
Ed. Manual Moderno.
- 11) Robert G. Trapp, Beth Dawson-Sanders;
"Bioestadística Médica";
Ed. Manual Moderno.
- 12) Ruiz Torres Francisco;
"Diccionario de terminos médicos";
Ed. Científico Médica.
- 13) Ruthman Kenneth J.
"Epidemiología, Métodos Estadísticos";
Ed. Diaz de Santos.
- 14) Shelly Hernández, Ricardo;
"Estadística Aplicada a las Ciencias Biológicas";
Ed. Grafo.
- 15) Steve Selvin, D. ;
"Statistical Analysis of Epidemiologic Data";
Ed. Monographs in Epidemiology and Biostatistics.
- 16) William Dameshek;
"La leucemia";
Ed. Científico Médica.