

**UNIVERSIDAD DE CHILE
FACULTAD DE MEDICINA
ESCUELA DE SALUD PÚBLICA**



**USO DE METODOS MULTINIVEL PARA IDENTIFICAR
GRUPOS SUSCEPTIBLES A CONTAMINACION
ATMOSFERICA: APLICACIÓN EN UN ESTUDIO DE
SALUD RESPIRATORIA DE NIÑOS Y NIÑAS EN
SANTIAGO, CHILE EN LOS AÑOS 2010-2011.**

MARCELO MORA AHUMADA

**TESIS PARA OPTAR AL GRADO DE
MAGISTER EN BIOESTADISTICA**

Director de Tesis: Dr. Pablo Ruiz R.

**UNIVERSIDAD DE CHILE
FACULTAD DE MEDICINA
ESCUELA DE SALUD PÚBLICA**



**USO DE METODOS MULTINIVEL PARA IDENTIFICAR
GRUPOS SUSCEPTIBLES A CONTAMINACION
ATMOSFERICA: APLICACIÓN EN UN ESTUDIO DE
SALUD RESPIRATORIA DE NIÑOS Y NIÑAS EN
SANTIAGO, CHILE EN LOS AÑOS 2010-2011.**

MARCELO MORA AHUMADA

**TESIS PARA OPTAR AL GRADO DE
MAGISTER EN BIOESTADISTICA**

Director de Tesis: Dr. Pablo Ruiz R.

Santiago, Junio 2017

**UNIVERSIDAD DE CHILE
FACULTAD DE MEDICINA
ESCUELA DE SALUD PÚBLICA**

**INFORME DE APROBACION TESIS DE
MAGISTER EN BIOESTADISTICA**

Se informa a la Comisión de Grados Académicos de la Facultad de Medicina, que la Tesis de Magíster en Bioestadística presentada por el candidato

MARCELO MORA AHUMADA

ha sido aprobada por la Comisión Informante de Tesis como requisito para optar al Grado de **MAGISTER EN BIOESTADISTICA** en Examen de Defensa de Tesis rendido el día **29 de Junio del 2017**

Dr. Pablo Ruiz R.
Director de Tesis
Escuela de Salud Pública, Facultad de Medicina, Universidad de Chile

COMISION INFORMANTE DE TESIS

Sr. Gabriel Cavada Ch.

Dr. Dante Cáceres L.

Dr. Sergio Alvarado O.
Presidente Comisión de Examen

Agradecimientos

A mis profesores, que además de compartir sus conocimientos me entregaron motivación, comprensión y ánimo.

Al equipo administrativo y de apoyo del Programa de Bioestadística, especialmente a Mabel y Patricia.

A mis compañeros y amigos del Programa (generaciones 2012 y 2013) ya que además de su apoyo compartimos muy buenos momentos y conversaciones.

A todos mis amigos ya que recibí su apoyo y cariño.

A mi prima Marcela, cuasi madrina.

Especialmente a mi bella familia que me acompañó en esta travesía que ha concluido felizmente.

Índice

Resumen	4
Abstract	5
1. Introducción.....	6
2. Marco teórico.....	8
2.1 Salud respiratoria infantil y contaminación atmosférica.....	8
2.2 Modelos Mixtos.....	10
2.2.1 Introducción.....	10
2.2.2. Modelo Mixto de Dos Niveles.....	11
2.2.3 Coeficiente de Correlación Intra Clase.....	13
2.2.4 Estimación de parámetros.....	13
2.2.5 Efectos Fijos y Aleatorios.....	14
2.2.6 Modelo mixto básico o modelo nulo.....	15
2.3 Contraste de hipótesis para los coeficientes de modelos anidados.....	16
2.4 Comparación de modelos no anidados.....	18
2.5 Modelos Auto-regresivos.....	19
2.5.1 Respuesta auto-regresiva y efectos aleatorios.....	19
3. Objetivos de la investigación.....	21
3.1 Objetivo General:.....	21
3.2 Objetivos Específicos:	21
4. Hipótesis.....	22
5. Metodología.....	23
5.1 Diseño de estudio.....	23
5.2 Universo y muestra.....	23
5.3 Métodos o procedimientos de recolección de información e instrumentos usados.....	23
5.4 Variables.....	24
5.4.1 Variable Dependiente.....	24
5.4.2 Variables de Formación de Subgrupos:	24
5.4.3 Variables Independientes:	26
5.5 Técnicas de procesamiento de la información.....	28
5.6 Organización de las Variables.....	28
5.6.1 Definición de los subgrupos	28

5.6.2 Jerarquización de las variables en la etapa de modelación.....	29
5.7 Proceso de modelamiento	30
5.8 Aspectos éticos	32
6. Análisis y resultados.	33
6.1 Etapa de selección y caracterización de subgrupos.	33
6.1.1 Caracterización de los subgrupos en relación a su CCI.	34
6.1.2 Características generales de los subgrupos.....	35
6.2 Exposición a contaminantes por subgrupo.....	39
6.3 Exposición a variables meteorológicas por subgrupo.....	41
6.4 Modelos base.	42
6.4.1 Modelo base para grupo 1.....	42
6.5 Susceptibilidad de los grupos a los contaminantes	44
7. Discusión.....	47
8. Conclusión.....	49
9. Bibliografía	50

Resumen

La contaminación atmosférica juega un importante rol en la salud humana, especialmente en niños, quienes por una condición natural consumen mayor volumen de aire por unidad corporal, y por lo tanto se ven más afectados frente a la contaminación. El óxido nítrico exhalado es un marcador de inflamación bronquial usado en pacientes con asma que pudiera servir para identificar otros grupos susceptibles a contaminación atmosférica.

Objetivos: Determinar si la variación intraclase del óxido nítrico determinada por modelos multinivel es un predictor de susceptibilidad a contaminación atmosférica, identificar posibles subgrupos susceptibles a contaminación, determinar la variabilidad del óxido nítrico exhalado para los subgrupos identificados según su coeficiente de correlación intraclase, determinar subgrupos candidatos a ser más susceptibles a contaminación atmosférica, determinar impacto de la contaminación atmosférica en cuanto a incremento en el óxido nítrico de los subgrupos estudiados, comparar la susceptibilidad a la contaminación de los subgrupos con un grupo de referencia (participantes con asma).

Metodología: A partir del uso de métodos multinivel, los que permiten evaluar variabilidad dentro de los sujetos a través del coeficiente de correlación intraclase.

Resultados: Se compararon los niveles de óxido nítrico exhalado de tres subgrupos con los participantes con asma (grupo de referencia) y de estos se observó que el subgrupo 3 no presentó ningún efecto significativo en los niveles de óxido nítrico, los subgrupos 2 y 4 presentaron niveles de óxido nítrico similares al de los participantes con asma.

Conclusión: El método propuesto no permitió identificar subgrupos con mayor susceptibilidad a la contaminación que el grupo de referencia aunque el dióxido de azufre (SO₂) generó respuestas similares a las de participantes asmáticos en los grupos 2 y 4.

Abstract

Air pollution plays an important role in human health, especially in children, who by a natural condition consume more air per unit body, and therefore are more affected in the face of pollution. Exhaled nitric oxide is a marker of bronchial inflammation used in patients with asthma that could serve to identify other groups susceptible to air pollution.

Objectives: To determine if the intra-class variation of nitric oxide determined by multilevel models is a predictor of susceptibility to air pollution, to identify possible subgroups susceptible to contamination, to determine the variability of exhaled nitric oxide for the subgroups identified according to their intra-class correlation coefficient, to determine subgroups, to determine the impact of air pollution on the increase of nitric oxide in the subgroups studied, to compare the susceptibility to contamination of the subgroups with a reference group (participants with asthma).

Methodology: Based on the use of multilevel methods, which allow to evaluate variability within the subjects through the intra-class correlation coefficient.

Results: The exhaled nitric oxide levels of three subgroups were compared with the participants with asthma (baseline group) and from these it was observed that subgroup 3 had no significant effect on nitric oxide levels, subgroups 2 and 4 presented levels of nitric oxide similar to that of participants with asthma.

Conclusion: The proposed method did not allow the identification of subgroups with greater susceptibility to contamination than the reference group, although sulfur dioxide (SO₂) generated responses similar to those of asthmatic participants in groups 2 and 4.

CAPITULO 1

1. Introducción.

La contaminación atmosférica juega un rol importante en la salud respiratoria de los niños, dado que su sistema respiratorio se comporta como un blanco primario para los contaminantes, ya sean emitidos por fuentes móviles, fijas o al interior de los hogares (Gavidia et al., 2009). Los niños son, en general, más susceptibles a la contaminación atmosférica dado que poseen un estado anabólico activo lo que significa que respiran más aire, beben más agua e ingieren más comida por unidad corporal que los adultos, lo que redundaría en una mayor dosis ante una exposición ambiental dada (Gavidia et al., 2009).

Como producto de esto y otras patologías como asma, los niños pueden sufrir exacerbaciones, presentando síntomas respiratorios como tos y sibilancias, disminución de función pulmonar o inflamación de las vías respiratorias. Estas exacerbaciones se asocian a exposiciones tales como contaminación atmosférica, alérgenos entre otros (Fabian et al., 2012). Por ejemplo, aumentos a corto plazo en la contaminación del aire exterior se han asociado con síntomas y disminución temporal de la función pulmonar (Hoek et al., 2012).

En estudios previos se encontró que los niveles de contaminación por material particulado 2,5 fueron superiores a los niveles permitidos a nivel nacional e internacional. Tanto síntomas respiratorios como marcadores de inflamación de vías aéreas se vieron afectadas por el aumento de contaminantes (Yohannessen et al., 2015).

Sin embargo, no todos los niños responden en forma similar a la contaminación. En un nivel más básico se espera que niños asmáticos tengan mayor respuesta a este factor de riesgo, pero se postula que muchos otros factores pueden intervenir. Por ejemplo existe evidencia

que en niños con obesidad y sobrepeso existe interacción significativa entre niveles de contaminación y síntomas de problemas respiratorios (Dong et al., 2012). En este punto identificar factores que aumenten la susceptibilidad es de gran interés. Se ha postulado como posibles factores que aumentan esta susceptibilidad la atopia, el grado de control de asma, sexo, presencia de obesidad, entre otros. Dado que son muchos factores potenciales, y que se quisiera explorar la interacción entre ellos, es que planteamos la necesidad de evaluar un método de identificar *a priori* grupos susceptibles para luego realizar los estudios epidemiológicos de exposición respuesta propiamente tales.

Como una respuesta a esta necesidad, planteamos el uso de modelos multinivel. Estudios típicos de contaminación atmosférica, como los de panel, trabajan con medidas repetidas por cada sujeto. Estas agrupaciones hacen que observaciones que corresponden al mismo grupo reciban una serie de influencias comunes; por lo tanto, no existe independencia entre ellas y si se aplican métodos de regresión se viola un principio básico de la estadística, la independencia de los datos (Goldstein et al., 2010). Los métodos multinivel permiten desagregar la variabilidad de la respuesta de variación entre sujetos y dentro de los sujetos. Dado que el foco es la respuesta de la contaminación en el tiempo, el impacto de ésta debe afectar la variación dentro del sujeto. Por lo tanto, planteamos que la determinación de la **variabilidad dentro de los sujetos** puede usarse como tamizaje para identificar subgrupos de niños con una característica en común (obesidad, atopia, etc) que presenten una mayor susceptibilidad. Posteriormente, estos subgrupos pueden ser estudiados en relación a la contaminación atmosférica usando los métodos epidemiológicos usuales, y en caso de encontrar una mayor susceptibilidad pueden ser objeto de medidas de salud pública para protegerlos.

CAPITULO 2

2. Marco teórico.

2.1 Salud respiratoria infantil y contaminación atmosférica.

El óxido nítrico exhalado (ONe) es una molécula radical libre producido en las células endoteliales de la pared bronquial como parte intrínseca del proceso inflamatorio bronquial(Meave-Cueva et all, 2013).

El ONe corresponde a una herramienta diagnóstica en la práctica clínica diaria en pediatría. En este sentido la medición del ONe proporciona una valiosa información acerca de la inflamación subyacente, puede tener utilidad tanto para el diagnóstico de asma como para distinguir qué pacientes se pueden beneficiar más del tratamiento antiinflamatorio(Álvarez et all, 2010).

Los efectos inflamatorios, medidos a través del ONe, en niños asmáticos y atópicos de Chile por efectos de contaminantes atmosféricos han sido reportados previamente (Vidal et all, 2013).

Los efectos de la contaminación del aire sobre la salud se han estudiado extensamente. La contaminación intradomiciliaria y del aire ambiental están asociadas al desarrollo e incremento de la severidad de infecciones respiratorias en la infancia pudiendo producir exacerbaciones agudas en pacientes con asma (Goldizen et al., 2016).El diseño de estudios de datos de panel es una herramienta popular para el estudio de los efectos a corto plazo de la contaminación atmosférica sobre la salud humana. Los sujetos se observan repetidamente en el tiempo, facilitando así la evaluación de los efectos en la salud de los cambios en la exposición a través del tiempo(Janes et al., 2008).

Los factores que pueden desencadenar o agravar los síntomas asmáticos son infecciones víricas, alérgenos domésticos o laborales (por ejemplo, ácaros del polvo doméstico, polen o cucarachas), humo del tabaco, ejercicio y estrés, etc. Estas respuestas son más probables cuando el asma no está controlada (Reddel & Bateman, 2016).

2.2 Modelos Mixtos.

2.2.1 Introducción.

Son una metodología de investigación cuantitativa, fundamentada en la idea de que considerables grupos de datos, incluyendo los datos de observación recogidos en ciencias humanas y biológicas, tienen una estructura jerárquica o multinivel.

Para entender la lógica de ésta propuesta, imaginemos nuestra sociedad agrupada en niveles jerárquicos; por ejemplo un individuo pertenece a un barrio, ese barrio a un pueblo, ese pueblo a una provincia, esa provincia a una región y así sucesivamente la sociedad se va edificando sobre diferentes niveles jerárquicos o multinivel. Dichas agrupaciones se ordenan desde lo individual hasta lo colectivo, permaneciendo los primeros con todas sus características, e inmersos en el contexto y las actividades de los segundos(Acevedo, 2008).

En la observación de las estructuras sociales, se afirma que un individuo interactúa permanentemente con su contexto. Las personas están influenciadas por los grupos de pertenencia y por las acciones de estos para transformar la sociedad. En este sentido, existe una interrelación entre ambos y esto se puede estudiar solamente por medio de estructuras jerárquicas en dos niveles: individuos, el primero y el grupo, el segundo(Hox et al., 1995).

Una jerarquía consiste en un conjunto de observaciones que conforman el nivel individual, primer nivel o micronivel, que se encuentran anidadas dentro de un nivel superior, segundo nivel o nivel macro. También se suele referir a los niveles macro como grupos o contextos. Ejemplo de estas estructuras son pacientes anidados en médicos o medidas repetidas anidadas dentro de personas(Perez, 2013). Un ejemplo es la estructura jerárquica de dos niveles. Las estructuras jerárquicas aparecen cuando tenemos una unidad del primer nivel o

del nivel más bajo anidada en una y solo una unidad de segundo nivel o nivel más alto. En el caso de los estudios de panel los datos se visualizan como muestra la figura 1.

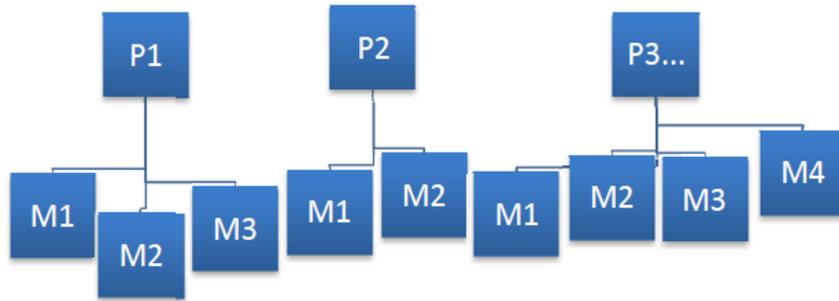


Fig. 1 Clasificación y diagrama de unidades de un diseño de medidas repetidas, mediciones (M1, M2,..., Mn) anidadas en pacientes (P1, P2,..., Pn).

2.2.2. Modelo Mixto de Dos Niveles.

Si se considera un modelo de dos niveles con la variable respuesta continua la ecuación se puede representar así:

$$y_{ij} = \beta_0 + \beta_1 X_{ij} + (\mu_{0j} + \mu_{1j} X_{ij} + e_{oij})$$

$$var(e_{oij}) = \sigma_{e0}^2$$

Donde:

y_{ij} : Variable respuesta medida en la i -ésima unidad del nivel educacional anidada en la j -ésima unidad del nivel contextual.

β_0 : Parámetro fijo, intercepto del nivel contextual.

β_1 : Parámetro fijo, coeficiente del regresor x en el nivel individual.

μ_{0j} : Variable aleatoria con parámetro $E(\mu_{0j}) = 0$ y $var(\mu_{0j}) = \sigma_{\mu_0}^2$

μ_{1j} : Variable aleatoria con parámetros $E(\mu_{1j}) = 0$ y $var(\mu_{1j}) = \sigma_{\mu_1}^2$ además se considera $cov(\mu_{0j}, \mu_{1j}) = \sigma_{\mu_{01}}^2$

Es así que la variable respuesta y_{ij} se expresa como la suma de una parte aleatoria y una fija y las variables aleatorias μ_{0j} y μ_{1j} hacen referencia al término residual.

Se requerirá estimar los dos coeficientes β además de los parámetros $\sigma_{\mu_0}^2$, $\sigma_{\mu_1}^2$, $\sigma_{\mu_{01}}^2$ y $\sigma_{e_0}^2$, siendo estas varianzas y covarianzas los parámetros aleatorios (Goldstein, 1999).

Así, considerando un modelo simple de dos niveles donde solo se considera el intercepto aleatorio dejando la pendiente fija, el modelo se podrá representar con la siguiente ecuación:

$$y_{ij} = \beta_0 + \beta_1 X_{ij} + \mu_{0j} + e_{0ij}$$

Con los supuestos:

$$e_{0ij} \sim N(0, \sigma_{e_0}^2)$$

$$\mu_{0j} \sim N(0, \sigma_{\mu_0}^2)$$

Bajo los supuestos mencionados se puede concluir que:

$$Var(Y_{ij} | \beta_0, \beta_1, X_{ij}) = Var(\mu_{0j}, e_{0ij}) = \sigma_{\mu_0}^2 + \sigma_{e_0}^2$$

Lo que representa la suma de las varianzas en el nivel contextual e individual y la covarianza entre dos medidas en el mismo sujeto estará dado por:

$$Cov(\mu_{0j} + e_{0i_1j}, u_{0j} + e_{0i_2j}) = cov(u_{0j}, u_{0j}) = \sigma_{u0}^2$$

2.2.3 Coeficiente de Correlación Intra Clase.

Este coeficiente es definido como la varianza entre los grupos dividida en la varianza total, donde la varianza total es la suma de la varianza entre los grupos y la varianza dentro de cada grupo (Twisk, 2006). Cuanto menor es la varianza dentro de los grupos mayor es el valor del índice de correlación intraclase.

Por lo tanto la correlación que existe entre dos mediciones en el mismo sujeto es:

$$\rho = \frac{\sigma_{u0}^2}{(\sigma_{u0}^2 + \sigma_{e0}^2)}$$

Lo que constituirá la correlación dentro del nivel contextual o intrasujetos, también conocida como coeficiente de correlación intraclase.

El coeficiente de correlación intraclase puede ser estimado sobre la varianza de los interceptos y sobre la varianza de los residuos. Este coeficiente es un indicador de la correlación existente dentro de las observaciones que pertenecen a un mismo sujeto (Goldstein, 1999).

2.2.4 Estimación de parámetros.

Si se conoce los valores de la varianza, se puede construir la matriz diagonal de varianzas covarianzas "V" y aplicar el procedimiento de estimación de Mínimos Cuadrados Generalizados para estimar los coeficientes fijos:

$$\hat{\beta} = (X^T V^{-1} X)^{-1} (X^T V^{-1} Y)$$

$$Var(\hat{\beta}) = (X^T V^{-1} X)^{-1}$$

Una alternativa en caso de que las varianzas no sean homogéneas a través del nivel contextual es:

$$Var(\hat{\beta}) = (X^T V^{-1} X)^{-1} X^T V^{-1} (Y - X\hat{\beta})$$

Donde “V” es una matriz de varianzas covarianzas, cuando los errores tienen una distribución normal, los estimadores de β son obtenidos por el método de Máxima Verosimilitud, pero esta estimación entrega estimadores sesgados de los parámetros aleatorios porque no toma en cuenta la variación muestral de los parámetros fijos. Esto puede resultar importante en muestras pequeñas por lo que para obtener estimadores no sesgados se puede acceder a una modificación conocida como método de Máxima Verosimilitud Restringida.

El procedimiento de estimación es iterativo, se comienza estimando los parámetros fijos usando el método de Mínimos Cuadrados el cual supone $\sigma_{u0}^2 = 0$ para obtener los residuos de la parte fija. Posteriormente se estiman los coeficientes σ_{u0}^2 , σ_{e0}^2 a través del método iterativo de Mínimos Cuadrados Generalizados usando la matriz de covarianzas estimadas donde el supuesto de normalidad permite expresar esta matriz como una función de los parámetros aleatorios (Goldstein & Rasbash, 1992).

2.2.5 Efectos Fijos y Aleatorios.

Un factor compuesto por diferentes tratamientos se dice que es un factor de efectos fijos si todos los posibles efectos en los que está interesado el investigador están presentes en el experimento. Un efecto aleatorio se refiere a un factor compuesto por una muestra aleatoria

de entre todos los tratamientos relevantes. Por ejemplo, los efectos fijos solo permiten que se hagan inferencias acerca de los tratamientos usados en el experimento (Kreft & De Leeuw, 2002).

2.2.6 Modelo mixto básico o modelo nulo.

El modelo nulo no tiene predictores, únicamente cuenta con la variable dependiente y una constante, a partir de estas se estiman los parámetros. Este modelo se puede representar a partir de dos ecuaciones que representan nuestros niveles de trabajo, o sea dos.

$$Y_{ij} = \beta_{0j} + e_{ij}$$

$$\beta_{0j} = \beta_{00} + \mu_{0j}$$

La primera ecuación representa al primer nivel o micro nivel, donde:

y_{ij} : Es la variable dependiente observada en el nivel inferior o micronivel, por ejemplo para una medición i dentro de un sujeto j .

β_{0j} : Representa al intercepto, o sea, al punto de partida de la recta en el plano, por ejemplo la media general de las mediciones en el sujeto j .

e_{ij} : Representa a los residuos, cuya media es cero y varianza σ_e^2 . Este modelo asume la varianza de error al azar, la misma para todos los sujetos.

La segunda ecuación representa al segundo nivel o macro nivel.

β_{00} : Representa al intercepto del nivel dos, esta correspondería a la gran media general de todos los sujetos j incluidos en la muestra.

μ_{0j} : Representa a los residuos del nivel dos. Tiene media cero y varianza $\sigma_{\mu_0}^2$

Cuando el puntaje de $\sigma_{\mu_0}^2$ es alto indica que existe una marcada diferencia entre los sujetos j del segundo nivel.

La información contenida en las ecuaciones mencionadas se puede presentar en una sola ecuación que integra los elementos de ambas, así se obtiene:

$$y_{ij} = \beta_{00} + \mu_{0j} + e_{ij}$$

Donde:

y_{ij} :Corresponde a la variable dependiente, el valor de la medición i en el sujeto j

β_{00} :Identifica la media general de mediciones entre los sujetos.

μ_{0j} :Es el residuo o varianza residual del segundo nivel (los sujetos) y representa la desviación del rendimiento de la media entre los sujetos (diferencia de estimación entre el valor esperado y el valor observado).

e_{ij} : Corresponde a la varianza residual del micro nivel, las mediciones i dentro de cada sujeto.

El coeficiente β_{00} corresponde a la parte fija del modelo.

Los coeficientes μ_{0j} y e_{ij} representan las varianzas residuales de la parte aleatoria del modelo. Como se puede apreciar, la ecuación de regresión del modelo nulo brinda información de dos elementos, la parte fija y la parte aleatoria de una estructura anidada de datos.

En la parte aleatoria del modelo solo están contenidas las varianzas de los residuos, estos residuos no corresponden a parámetros del modelo, pero las varianzas de estos sí(Acevedo, 2008).

2.3 Contraste de hipótesis para los coeficientes de modelos anidados.

La función de verosimilitud permite el cálculo de la *desviación*, esta *desviación*mos indica qué tan bien se ajusta un determinado modelo a los datos de los cuales proviene. La *desviación*

se define como d y es igual a $-2\log L$, en donde L es el valor de la función de verosimilitud obtenido como resultado de la convergencia del proceso de iteraciones llevado a cabo durante el cálculo y ajuste del modelo. Para efectos prácticos se considera que a menor *desviación* el modelo presenta un mejor ajuste a los datos.

Cuando a un determinado modelo se le añaden parámetros, se dice que este aumenta su complejidad, para este caso se usa el concepto de modelo anidado, esto hace referencia a que el modelo resultante proviene de un modelo con menos parámetros (Modelling, 2017), pero ambos modelos comparten la misma base de datos. Estos modelos son comparables entre sí usando sus *desviaciones*. Este contraste estadístico se llama test de razón de verosimilitud o LR test y está dado por:

$$LR = d_1 - d_2 = -2\log \frac{L_2}{L_1}$$

Donde estando el modelo 1 anidado en el modelo 2:

d_1 : corresponde a la *desviación* del modelo 1

d_2 : corresponde a la *desviación* del modelo 2

L_1 : es el valor de la verosimilitud del modelo 1

L_2 : es el valor de la verosimilitud del modelo 2

El estadístico LR sigue una distribución Ji-cuadrado con grados de libertad que corresponden al número extra de parámetros del modelo más complejo, o sea, del modelo 2.

También es posible usar el LR test para comparar un modelo anidado nulo (Hox, 2002), que solo contiene efectos fijos, con un modelo de mayor complejidad en cuanto a la incorporación

de parámetros (con efectos aleatorios además de los fijos). Por lo mencionado es posible el contraste de la hipótesis nula:

$$H_0 : \sigma_0^2 = \sigma_{uk}^2 = 0$$

2.4 Comparación de modelos no anidados.

Existen dos indicadores que permiten la comparación de modelos no anidados, el criterio de información de Akaike (Akaike's Information Criterion, *AIC*) y el criterio de información Bayesiano (Bayesian Information Criterion, *BIC*).

El criterio de información de Akaike incorpora un ajuste por el número de parámetros estimados. Este índice asume que los modelos comparados se han ajustado sobre el mismo conjunto de datos.

El *AIC* se calcula a partir de la desviación d y el número q de parámetros estimados como se muestra a continuación (Hox, 2002):

$$AIC = d + 2q$$

El criterio de información Bayesiano además de la desviación y el número de parámetros considera el número de observaciones usado para realizar la estimación (N), esto corresponde al número de términos independientes para el cálculo de la verosimilitud. El cálculo del *BIC* se muestra a continuación (Scharz, 1978):

$$BIC = d + q \log N$$

Ambos índices, tanto el *AIC* y el *BIC* incorporan una penalización de su puntaje en la medida que el número de parámetros incorporados al modelo aumenta. Para efectos prácticos los dos indicadores disminuyen su valor cuando el modelo se ajusta mejor a los datos(Hox, 2002).

2.5 Modelos Auto-regresivos

Se puede especificar un modelo de respuesta auto-regresiva de primer orden incluyendo la respuesta anterior $Y_{i-1,j}$ como variable en un modelo de regresión lineal para Y_{ij} :

$$Y_{ij} = \beta_1 + \gamma Y_{i-1,j} + \beta_2 x_{2ij} + \dots + \beta_8 x_{8ij} + \epsilon_{ij}, \epsilon_{ij} | Y_{i-1,j}, x_{ij} \sim N(0, \sigma^2)$$

La respuesta anterior se denomina a veces "respuesta de lag-1".

El modelo de retraso-respuesta es sólo razonable si las ocasiones están igualmente espaciadas en el tiempo. De lo contrario, sería extraño asumir que la respuesta rezagada tiene el mismo efecto en la respuesta actual independientemente del intervalo de tiempo entre ellos(Rabe-Hesketh & Skrondal, 2008).

2.5.1 Respuesta auto-regresiva y efectos aleatorios

Podríamos especificar un modelo de intercepto aleatorio que incluya a las respuestas retrasadas como variables. Una característica útil de estos modelos es que pueden usarse para distinguir entre dos explicaciones competitivas de la dependencia dentro del sujeto: la heterogeneidad no observada (representada por los efectos aleatorios) o la experiencia previa (representada por las respuestas rezagadas). Por ejemplo, la dependencia dentro de los sujetos de los salarios a lo largo del tiempo, por encima de la expuesta por las variables observadas, puede deberse a que algunos sujetos son especialmente dotados y por lo tanto

más valorados y/o algunos sujetos que han experimentado altos salarios en el pasado, lo que potencialmente podría colocarlos en una buena posición de negociación(Rabe-Hesketh & Skrondal, 2008).

CAPITULO 3

3. Objetivos de la investigación.

3.1 Objetivo General:

Determinar si la variación intraclase en ONE determinada por modelos multinivel es un predictor de susceptibilidad a contaminación atmosférica para subgrupos de participantes evaluados.

3.2 Objetivos Específicos:

1. Identificar posibles subgrupos susceptibles a contaminación.
2. Determinar la variabilidad del ONE para los subgrupos identificados según su coeficiente de correlación intraclase.
3. Determinar subgrupos candidatos a ser más susceptibles a contaminación atmosférica.
4. Determinar impacto de la contaminación atmosférica en cuanto a incremento en el ONE de los subgrupos estudiados.
5. Comparar la susceptibilidad a la contaminación de los subgrupos con un grupo de referencia (participantes con asma).

CAPITULO 4

4. Hipótesis.

Se postula que es posible identificar grupos de alta variabilidad en ONE usando el método de screening y que estos subgrupos serán más susceptibles a la contaminación atmosférica que el grupo de referencia (participantes con asma).

CAPITULO 5

5. Metodología.

5.1 Diseño de estudio.

Estudio observacional usando un panel de niños, con mediciones longitudinales, esto es mediciones repetidas para cada sujeto.

5.2 Universo y muestra.

La muestra a utilizar corresponde a una cohorte de 220 niñas y niños entre seis y catorce años, asmáticos y no asmáticos seguidos durante los meses de invierno del 2010 y 2011 en las Comunas de Independencia y Recoleta de Santiago, Chile. Los niños asmáticos fueron reclutados en el Hospital Roberto del Río y Consultorio Cruz Meló; mientras que los niños no asmáticos se reclutaron en el Consultorio mencionado. A todos los participantes se les explicó los objetivos del estudio y se les solicitó la firma del consentimiento informado a los padres (o adulto responsable del niño) y del asentimiento a los niños. Se excluyó a los niños que reportaban haber fumado. Cada niño recibió visitas cada 15 días donde se hicieron evaluaciones de función pulmonar e inflamación de las vías respiratorias. Del total de niños y niñas seguidos 149 de ellos presentó 2 o más mediciones de ONE y de estos 149 niños y niñas se extrajeron los subgrupos definidos según su coeficiente de correación intraclase.

5.3 Métodos o procedimientos de recolección de información e instrumentos usados.

Para minimizar molestias al niño y evitar sesgos por traslados, a cada niño se le visitó en domicilio por un profesional de la salud cada dos semanas durante los tres meses de seguimiento (alrededor de seis mediciones por niño). Para asegurar la presencia del niño y evitar variaciones circadianas, las mediciones se realizaron entre las 16:00 y las 20:00 horas.

5.4 Variables.

5.4.1 Variable Dependiente.

Óxido Nítrico exhalado (ONe) de los niños y niñas. El ONe se obtuvo a partir de la medición del aire espirado de un subconjunto de niños y niñas incluidos en el estudio.

5.4.2 Variables de Formación de Subgrupos:

Variabes consideradas para la formación de subgrupos de la muestra (objetivos 1, 2 y 3). Se utilizarán las variables de características físicas o condiciones propias de los niños y niñas (tanto endógenas como exógenas). Entre estas serán consideradas:

- a. Asma. Trastorno inflamatorio crónico de las vías aéreas caracterizada por episodios recurrentes de sibilancias, dificultad respiratoria, opresión torácica y tos especialmente nocturna o durante la madrugada (Reddel & Bateman, 2016). Para el presente estudio se considera paciente con asma todo aquel paciente que presenta diagnóstico médico de dicha patología
- b. Atopia. *Estado de hipersensibilidad anómala que presentan ciertos individuos ante la presencia de sustancias o condiciones que para el resto de la población son inocuas* (Pepper, s.f.). Para identificar la presencia de atopia, se realizó a todos los sujetos del estudio un test cutáneo de alergias según método estandarizado para los siguientes aeroalérgenos: polvo de habitación, mezcla de dermatophagoides, mezcla de plumas, mezcla de caspas de animales, mezcla de hongos, mezcla de polen de malezas, pastos y árboles. Se consideró reacción positiva cuando el diámetro de la pápula fue ≥ 3 mm. Se definió como atópico a aquellos que presentaron al menos un resultado positivo en el test cutáneo.

- c. Sexo. Diferencias y características biológicas, anatómicas, fisiológicas y cromosómicas que definen los seres humanos como hombres y mujeres. Se utilizará la variable de control sexo a toda la población que tenga diferentes características masculinas o femeninas, considerándolo masculino o femenino de acuerdo a la clasificación de la base de datos.
- d. Edad. La edad se define como el tiempo que ha vivido una persona desde su nacimiento. La variable de control edad será categorizada en dos periodos o rangos, siendo el primero de 6 a 9 años de edad y el segundo de 10 a 14 años de edad, los cuales nos permitirán poder segmentar nuestra muestra para someterla a un análisis estadístico.
- e. Obesidad. Corresponde a una categoría del estado nutricional y es la condición física que presenta una persona, como resultado del balance entre el aporte, absorción y utilización de los nutrientes en el organismo.

Para realizar la evaluación nutricional de los participantes se utilizará el índice de masa corporal (IMC) según edad y sexo, comparado con la población de referenciapropuesto por el “*Center forDisease Control*”y el “*National Center forHealthStatistics*” (CDC-NCHS)de EEUU para niños y niñas de 6 a 18 años (Pizarro, 2004) en anexos 1 y 2.

Se utilizará el percentil del Índice de Masa Corporal entendido como la razón entre el peso (expresado en kilos) y el cuadrado de la estatura (expresada en metros), según edad y sexo de la siguiente forma:

$$\text{IMC} = \text{peso}/\text{talla}^2$$

IMC < p10: Bajo peso

IMC entre p10 y < p85: Normal

IMC entre p 85 y < a 95: Riesgo de Obesidad

IMC ≥ a p95: Obesidad.

- f. Consulta en urgencia. Corresponde al antecedente de haber consultado en servicio de urgencia por causa respiratoria. En términos operacionales el familiar responsable del

participante responde afirmativamente a la consulta de si “*su hijo ha sido atendido en el servicio de urgencias por silbidos en el pecho o algún problema en la respiración*” Este antecedente se considerará como una variable categórica de valor “*cero*” para quienes no han consultado en urgencia y “*uno*” para quienes sí lo han hecho.

- g. *Ingreso Familiar inferior a trescientos mil pesos.* Corresponde al reporte de ingreso familiar inferior a trescientos mil pesos. Esta variable se consideró en términos operacionales como una variable categórica de valor de valor “*cero*” para quienes reportaron ingreso familiar sobre trescientos mil pesos y “*uno*” para quienes reportaron ingreso familiar inferior o igual a trescientos mil pesos.

5.4.3 Variables Independientes:

Las siguientes variables se consideraron como variables independientes en los modelos de contaminación atmosférica.

- a. *Día de la semana.* Corresponde al día de la semana en la que se realizó la medición y va de lunes a domingo en escala ordinal de “*uno a siete*” de lunes a domingo respectivamente.
- b. *Año.* Corresponde al año de evaluación de los participantes y corresponden a los años 2010 y 2011.
- c. *Carga viral.* Virus influenza a, virus influenza b, virus influenza, virus parainfluenza, virus respiratorio sincicial, virus metaneumovirus y virus total. Cada uno de estos virus fue medido como variables continuas de manera semanal.
- d. *Humedad relativa.* Valor de humedad ambiental relativa con valores de cero a 100% categorizada de la siguiente manera:

Categoría 1: de 0 a 24,99%

Categoría 2: de 25 a 49,99%

Categoría 3: de 50 a 74,99%

Categoría 4: de 75 a 100%

e. Temperatura media. Valor medio de las temperaturas registradas durante el día categorizadas de la siguiente manera:

Categoría 1: temperatura mínima a 8,99°C

Categoría 2: 9 a 11,99°C

Categoría 3: 12 a 14,99°C

Categoría 4: 15 a máxima temperatura media

f. Temperatura mínima. Valor mínimo de las temperaturas registradas durante cada día categorizados de la siguiente manera:

Categoría 1: temperatura mínima a 3,99°C

Categoría 2: 4 a 6,99°C

Categoría 3: 7 a 9,99°C

Categoría 4: 10 a máxima temperatura mínima

g. Material particulado 2,5 (PM 2,5): El MP2,5 corresponde al material particulado menor a 2,5 micrones y corresponde a una variable de tipo continua.

h. Material particulado 10 (PM 10): Corresponde al material particulado sobre 10 micrones y corresponde a una variable de tipo continua.

i. Material particulado 2,5 – 10 (MPCO): Corresponde a material particulado entre 2,5 y 10 micrones.

j. Monóxido de carbono (CO): Corresponde a un gas tóxico y corresponde a una variable continua.

k. Dióxido de nitrógeno (NO₂): Gas contaminante medido como variable continua.

l. Ozono (O₃): Corresponde a un gas contaminante medido como una variable continua.

m. Dióxido de azufre (SO₂): Corresponde a un gas contaminante medido como variable continua.

5.5 Técnicas de procesamiento de la información.

A través del uso de modelos multinivel y el coeficiente de correlación intraclase, determinar la variabilidad intraclase de los subgrupos de la muestra para posteriormente estudiar el efecto de la contaminación atmosférica para discriminar la posible existencia de algunos grupos más susceptibles.

Se utilizó el software estadístico STATA 14.0

5.6 Organización de las Variables.

5.6.1 Definición de los subgrupos.

En una primera etapa se buscará identificar posibles subgrupos de participantes susceptibles a contaminación. Los criterios para la formación de subgrupos será que:

1° Los subgrupos creados serán a partir de niños que comparten la presencia de dos variables de formación de subgrupo, por ejemplo, asma-atopía, asma-obesidad, etc.

No se consideraran los subgrupos formados por la presencia de una característica y su propio complemento, es decir, asma-no asma, atopía-no atopía, etc.

No se considerarán subgrupos formados por ausencia de variables de formación de subgrupos, es decir, no asma-no atopía, no asma-no obesidad, etc.

2° Los subgrupos seleccionados deberán presentar mayor variabilidad en las mediciones de ONe que el subgrupo definido como de referencia (participantes con asma).

3° Será criterio para seleccionar un subgrupo como candidato a ser evaluado que cuente con treinta o más participantes.

5.6.2 Jerarquización de las variables en la etapa de modelación.

Las variables serán jerarquizadas o anidadas de la siguiente manera:

Las variables consideradas en el micro nivel (o nivel 1) serán las correspondientes al rendimiento en las pruebas de ONe de cada niño (medido de dos a siete ocasiones). Esta variable en el modelo corresponderá a y_{ij} . Esto corresponde a la medición “ j ” dentro del sujeto “ i ”.

También serán considerados dentro del primer nivel los errores e_{0ij} producto de la diferencia entre la estimación y el valor observado de y_{ij} .

Las variables consideradas en el nivel contextual, macro nivel, o nivel 2, serán las correspondientes a las variables de carga viral, humedad relativa temperatura, día de la semana y año la caracterización de los sujetos, que además serán las variables explicativas de los modelos. Estas variables serán consideradas del segundo nivel, ya que estas son las que dan el contexto a las mediciones en las pruebas de función pulmonar dentro de cada niño o niña.

También serán consideradas dentro del segundo nivel los errores μ_{0j} que corresponden al residuo o varianza residual del segundo nivel (los sujetos) y representa la desviación del rendimiento de la media entre los sujetos (diferencia de estimación entre el valor esperado y el valor observado) según la variable de caracterización, por ejemplo sexo.

5.7 Proceso de modelamiento.

Para evaluar el efecto de pertenecer a cada grupo en los niveles de Óxido Nítrico exhalado (ONe), que corresponde a la variable dependiente y es continua, se implementaron modelos de regresión multinivel con el fin de observar el efecto específico de cada contaminante dentro en los niveles de ONe de los participantes.

Correspondiendo y_{ij} al valor de la variable dependiente ONe para la i -ésima medición en el j -ésimo participante, el **modelo nulo** para cada grupo de participantes observados presenta la forma $y_{ij} = \beta_{00} + \mu_{0j} + e_{ij}$, donde β_{00} corresponde a la media general de mediciones de ONe entre los sujetos y además se identifica como la parte fija del modelo, correspondiendo μ_{0j} al residuo del nivel de los participantes y representa la desviación del rendimiento de la media entre los participantes, e_{ij} corresponde a la varianza residual de las mediciones i dentro de cada participante. Los coeficientes μ_{0j} y e_{ij} representan las varianzas residuales y representan la parte aleatoria del modelo.

Se comienza con el modelo nulo con el fin de poder compararlo con el modelo siguiente, en el que a partir de la modalidad forward, se le sumaran parámetros, que corresponderán a las variables explicativas, de manera secuencial para obtener así los **modelos base** a los que posteriormente les agregará niveles de distintos contaminantes ambientales medidos durante el período de recolección de datos. Para la comparación de los modelos se usará el *LR test* y

los índices *AIC* y *BIC*. Los índices *AIC* y *BIC* son obtenidos inmediatamente después de corrido tanto el modelo nulo como cada modelo anidado posterior.

Los modelos nulos para cada grupo incorporan, en primer lugar las variables día de la semana y año de medición esto ya que según estudios previos se encuentran relacionados con la variable dependiente (Yohannessen et al, 2015) son variables que pueden jugar un rol en la variación de los niveles medidos de óxido nítrico exhalado.

Según lo mencionado al modelo nulo se le agrega la variable día de la semana y en este momento se formula el **modelo día de la semana**, que es el modelo nulo más el primer parámetro agregado.

$$y_{ij} = \beta_{00} + \beta_1 x_{1ij} + \mu_{0j} + e_{ij}$$

Donde x_{1ij} corresponde al día de la semana en que se realiza la *i-ésima* medición de la variable dependiente en el *j-ésimo* participante del estudio y β_1 corresponde al coeficiente fijo de la variable x_1 .

Corrido el **modelo día de la semana** se obtiene sus índices *AIC* y *BIC*.

Luego se procede a la comparación del modelo nulo con el **modelo día de la semana** a través del *LR test*.

Posteriormente, y de manera general, si el coeficiente β_1 es significativo en el modelo esta variable se incluye y se procede con las siguientes variables que corresponden a cargas virales y variables meteorológicas (humedad relativa y temperatura (media y mínima)). Como ya se mencionó el criterio para incluir las variables día de la semana y año de medición corresponde a un criterio de inclusión por presencia en estudios previos dado que ambas variables pueden jugar un rol.

Dado que las variables meteorológicas habitualmente presentan un efecto autoregresivo, es decir, que el efecto observado en el nivel ONE en el *j-ésimo* participante puede presentar una

relación más fuerte con los niveles observados en días previos de las variables meteorológicas, es que para la definición de los modelos base se consideró el ajuste de cada modelo con los valores de uno a siete días previos de las variables meteorológicas y secuencialmente con las medias de uno, tres, cinco y siete días previos con el fin de determinar el mejor modelo base posible considerando el efecto autoregresivo de las variables meteorológicas.

El procedimiento de comparación de modelos no anidados consideró los índices de *AIC* y *BIC* para determinar el ajuste de estos a los datos.

De los resultados obtenidos en cada paso se decide mantener o excluir cada variable en el modelo, para luego repetir el procedimiento incorporando otra variable hasta completar el **modelo base** con los mejores indicadores de ajuste. El *p valor* considerado para la incorporación de los parámetros a cada modelo fue de 0,1.

Ya definido un modelo base para cada grupo de participantes se le agregará la medición de cada contaminante y las respectivas medias de los lags de uno, tres, cinco y siete días para apreciar el efecto auto-regresivo en los niveles de *ONe* en cada modelo base.

5.8 Aspectos éticos.

La investigación es observacional, se aplica sobre datos recolectados durante los años 2010 y 2011, oportunidad en la cual a todos los participantes se les explicó los objetivos del estudio y se les solicitó la firma del consentimiento informado a los padres (o adulto responsable del niño) y del asentimiento a los niños.

Los datos que se utilizarán fueron recolectados en el marco del proyecto FONDECYT 11090309 y el protocolo fue aprobado por el comité de ética en la investigación con seres humanos de la Facultad de Medicina.

CAPITULO 6

6. Análisis y resultados.

6.1 Etapa de selección y caracterización de subgrupos.

Los grupos de comparación provienen del total de sujetos con mediciones de ONe que estuvo conformado por un total de 149 participantes.

Del total de participantes se seleccionó cuatro sub-grupos. El primer subgrupo se compuso por niños y niñas con diagnóstico de asma. Los siguientes tres grupos fueron formados a partir de la combinación de pares de variables de formación de subgrupos mencionadas en capítulo 5.

La **tabla 1** presenta las variables con que se formó cada subgrupo y los coeficientes de correlación intraclase (CCI) para cada uno.

Los subgrupos seleccionados presentan CCI inferior al grupo de referencia, se seleccionaron los dos primeros por presentar los CCI más bajo y el quinto por interés en observar a niños y niñas con consulta en servicios de urgencia y de bajo ingreso económico.

Los cuatro subgrupos se encuentran definidos a continuación:

Grupo 1: Niñas y niños con diagnóstico de asma.

Grupo 2: Niñas con reporte de consultas en urgencia.

Grupo 3: Niñas y niños obesos con ingreso familiar inferior a trescientos mil pesos.

Grupo 4: Niñas y niños con consulta en urgencia e ingreso familiar inferior a trescientos mil pesos.

Tabla 1. Variables de formación de subgrupos y CCI.

Variables	CCI
Asma *	0,629
Asiste a urgencia - sexo femenino **	0,47
Obesidad - IFI a trescientos mil pesos **	0,488
Familiar fuma adentro - sexo femenino	0,534
IFI a trescientos mil pesos - sexo femenino	0,539
Asiste a urgencia - IFI a trescientos mil pesos **	0,546
Asiste a urgencia - Familiar fuma adentro	0,557
Familiar fuma adentro - Obeso	0,56
Asiste a urgencia - Antecedente de hospitalización	0,587
Asiste a urgencia - Obeso	0,591
Asiste a urgencia - Alergia al pasto	0,595
Sexo masculino - IFI a trescientos mil pesos	0,607
Familiar fuma adentro - IFI a trescientos mil pesos	0,614
Atopia - IFI a trescientos mil pesos	0,621
Alergia al pasto - IFI a trescientos mil pesos	0,624
Asiste a urgencia - Atopía	0,625
Asiste a urgencia - Sexo masculino	0,628
Atopía - Sexo femenino	0,577

* Grupo de referencia

** Grupos seleccionados para medir efecto de contaminantes

IFI: Ingreso Familiar Inferior (a trescientos mil pesos)

6.1.1 Caracterización de los subgrupos en relación a su CCI.

A modo de caracterización inicial de los subgrupos seleccionados se presenta la **tabla 2**, se detalla el número de sujetos, número de observaciones y CCI calculado para cada subgrupo.

Tabla 2. Características de los subgrupos en cuanto a CCI, número de participantes y número de observaciones.

Grupo	n° sujetos	n° observaciones	CCI
Grupo 1	72	350	0,629
Grupo 2	57	269	0,472
Grupo 3	34	152	0,468
Grupo 4	86	417	0,564

Los grupos 1 y 4 son los que contienen un mayor número de participantes y también un mayor número de observaciones. También es posible observar que los grupos 1 y 4 poseen un CCI mayor. Los grupos 2 y 3 son los que poseen un menor número de sujetos y un CCI de menor valor. Es importante destacar que los grupos 2, 3 y 4 cumplen la condición de presentar un CCI menor que el grupo de niñas y niños con asma.

6.1.2 Características generales de los subgrupos.

Tal como se aprecia en la **tabla 3**, la composición por sexo dentro de los grupos es asimétrica ya que el grupo 2 se encuentra formado exclusivamente por mujeres.

En cuanto a la distribución de pacientes según diagnóstico de asma se observa que el grupo 1 está compuesto exclusivamente por pacientes con asma y los grupos 2, 3 y 4 por un 49, 44 y 45% cada uno.

Sobre el antecedente de haber consultado en servicios de urgencia por causa respiratoria el grupo 1 presenta un 86%, el grupo 3 presenta un 76%, pero los grupos 2 y 4 están completamente compuestos por sujetos con antecedente de consulta en servicios de urgencia por causa respiratoria.

La proporción de obesidad para cada grupo fue de 32% para el grupo 1, de 23% para el grupo 2, de 100% para el grupo 3 y de 30% para el grupo 4. Finalmente sobre la proporción de participantes con ingreso familiar declarado bajo trescientos mil pesos el grupo 1 presenta

63%, el grupo 2 presenta 75% y los grupos 3 y 4 presentan 100% de su composición con esta característica.

Tabla 3. Características de los grupos en cuanto a composición por las principales variables de formación de subgrupos exploradas.

Característica	Grupo 1	Grupo 2	Grupo 3	Grupo 4
Número participantes	72	57	34	86
Observaciones	350	269	152	417
Sexo, M/F (%M)	33/39 (46%)	0/57 (0%)	19/15 (56%)	43/43 (50%)
Edad en años, media (\pm DS)	10,5 \pm 2	10,11 \pm 2,3	9,7 \pm 2,2	10 \pm 2
Edad en categorías (% Cat. 1)*	29/43 (40%)	29/28 (51%)	17/17 (50%)	43/43 (50%)
Asma (%)	72 (100%)	28 (49%)	15 (44%)	39 (45%)
Asma leve (%)	33 (45%)	10 (18%)	6 (17%)	18 (21%)
Asma moderada (%)	39 (54%)	18 (32%)	9 (27%)	21 (24%)
Atopia (%)	44 (63%)	30 (55%)	17 (55%)	40 (49%)
Consulta a servicio de urgencia (%)	62 (86%)	57 (100%)	26 (76%)	86 (100%)
Hospitalización (%)	32 (45%)	22 (38%)	9 (27%)	32 (38%)
Tabaco intradomiciliario (%)	17 (24%)	15 (26%)	8 (24%)	22 (26%)
Tabaco extradomiciliario (%)	51 (71%)	42 (74%)	30 (88%)	68 (79%)
Obesidad (%)	23 (32%)	13 (23%)	34 (100%)	26 (30%)
IFI a trescientosmil pesos (%)**	45 (63%)	43 (75%)	34 (100%)	86 (100%)

*Categoría 1 “de 6 a 9 años” y categoría 2 “de 10 a 14 años”

** IFI: Ingreso familiar inferior a trescientos mil pesos

Tabla 4. Distribución de Óxido Nítrico exhalado por grupo en cuanto a su media y desviación estándar de la media.

Grupo	Media \pm DS	Mínimo	Máximo
Grupo 1	17,4 \pm 1,49	4	83
Grupo 2	15,8 \pm 1,26	4	68
Grupo 3	13,9 \pm 1,42	4	50
Grupo 4	15,5 \pm 1,10	4	83

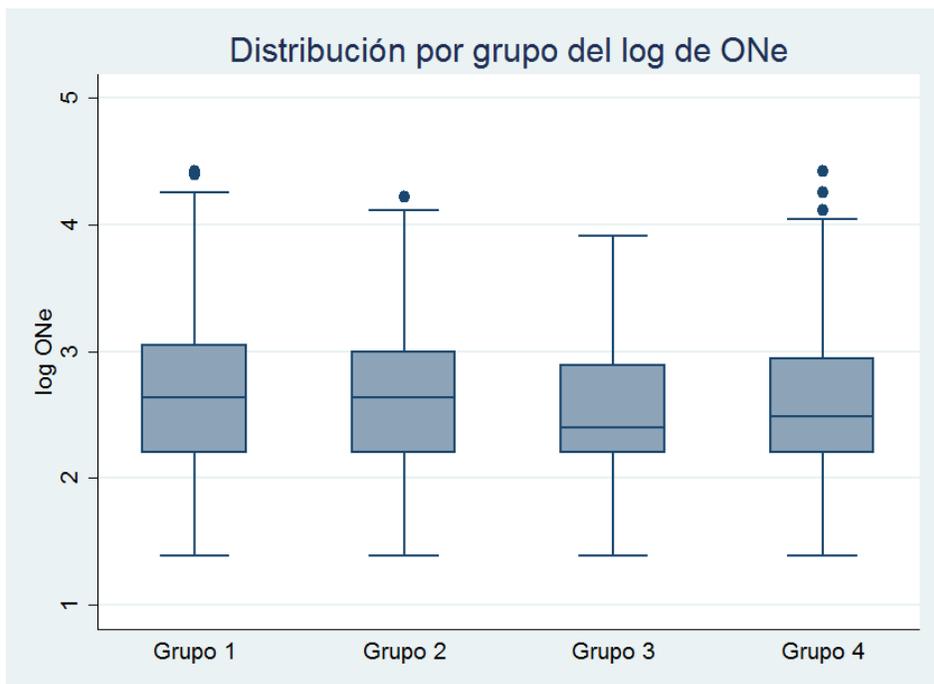
Según se observa en la **tabla 4** no se aprecian diferencias para las medias y desviación estándar de la media para los grupos 1, 2 y 4.

Según se observa en la **tabla 4** el grupo 3 presenta una media de ONe más baja que el grupo de referencia.

Previo al proceso de modelamiento se realizó una transformación logarítmica de los valores de ONe, esto debido a una característica desviación a la derecha de las mediciones de esta variable. En el **gráfico 1** se presenta el diagrama de cajas para apreciar la distribución de los valores de logaritmo de ONe.

Es posible apreciar que los valores medios se encuentran cercanos entre sí tomando como referencia los cuartiles superiores e inferiores de cada grupo. La media más baja corresponde al grupo 3 y la media más alta corresponde al grupo 1.

Gráfico 1. Distribución por grupos del logaritmo de óxido nítrico exhalado.



6.2 Exposición a contaminantes por subgrupo.

Según se observa en la **tabla 5** la media y la desviación estándar de la media los contaminantes se comportan de manera homogénea para los cuatro subgrupos en estudio.

En relación al PM 2,5 ($\mu\text{g}/\text{m}^3$) los grupos presentaron una media entre 31,5 y 32,6 ($\mu\text{g}/\text{m}^3$) con desviaciones estándar de la media entre 1,27 y 2,06 ($\mu\text{g}/\text{m}^3$). Los valores mínimos de concentración de este contaminante oscilaron entre 8,4 y 9,3 ($\mu\text{g}/\text{m}^3$) y los valores máximos oscilaron entre 56,2 y 59,3 ($\mu\text{g}/\text{m}^3$).

En relación a la distribución del PM 10 dentro de los grupos se puede observar que la media varía entre 77,6 y 80,1 ($\mu\text{g}/\text{m}^3$) y las desviaciones estándar de la media van entre 3,19 y 5,23 ($\mu\text{g}/\text{m}^3$), el valor mínimo para el grupo 1 fue de 21,5 ($\mu\text{g}/\text{m}^3$) y para los grupos 2, 3 y 4 fue de 17,2. Los valores máximos registrados fueron entre 147,6 y 147,7 ($\mu\text{g}/\text{m}^3$). Para los valores registrados de MPCO los valores medios oscilaron entre 46,1 y 47,5 ($\mu\text{g}/\text{m}^3$) con desviaciones estándar de la media entre 2,12 y 3,43 ($\mu\text{g}/\text{m}^3$). Para los valores registrados de CO se registraron valores medios entre 1,28 y 1,3 (ppm) con desviaciones estándar de la media entre 0,05 y 0,09 (ppm).

Para los valores de NO₂ se registraron valores medios entre 43,8 y 44,6 con desviaciones estándar de la media entre 1,51 y 2,49 (ppb).

Para los valores de O₃ se observaron medias entre 3,85 y 4,4 (ppb) con desviaciones estándar de la media entre 0,29 y 0,46 (ppb).

Finalmente para los valores de exposición a SO₂ se registraron valores medios entre 1,8 y 1,9 (ppb) y desviaciones estándar de la media entre 0,39 y 0,44.

Tabla 5. Media y desviación estándar de la media de los niveles de contaminantes a los que estuvo expuesto cada grupo.

Contaminante	Grupo	Media ± DS	Mínimo	Máximo
PM 2,5 (µg/m ³)	Grupo1	31,7 ± 1,33	8,4	59,3
	Grupo2	32 ± 1,54	8,4	59,3
	Grupo3	31,5 ± 2,06	9,3	56,2
	Grupo4	32,6 ± 1,27	9,3	59,3
PM 10 (µg/m ³)	Grupo1	77,6 ± 3,38	21,5	147,7
	Grupo2	79,1 ± 3,84	17,2	147,7
	Grupo3	78,5 ± 5,23	17,2	147,7
	Grupo4	80,1 ± 3,19	17,2	147,6
MPCO (µg/m ³)	Grupo1	46,1 ± 2,26	11,1	101
	Grupo2	47,1 ± 2,53	7,9	101
	Grupo3	47 ± 3,43	7,9	101
	Grupo4	47,5 ± 2,12	7,9	101
CO (ppm)	Grupo1	1,24 ± 0,06	0,37	2,77
	Grupo2	1,29 ± 0,06	0,37	2,77
	Grupo3	1,28 ± 0,09	0,37	2,77
	Grupo4	1,3 ± 0,05	0,37	2,77
NO2 (ppb)	Grupo1	44 ± 1,61	21,8	87,3
	Grupo2	44,5 ± 1,84	21,8	87,3
	Grupo3	43,8 ± 2,49	22,2	87,3
	Grupo4	44,6 ± 1,51	22,2	87,3
O3 (ppb)	Grupo1	4,4 ± 0,39	0,31	15,6
	Grupo2	3,85 ± 0,33	0,31	15,6
	Grupo3	3,98 ± 0,46	0,31	15,3
	Grupo4	3,86 ± 0,29	0,31	15,6
SO2 (ppb)	Grupo1	1,87 ± 0,11	0,44	4,5
	Grupo2	1,9 ± 0,12	0,39	4,5
	Grupo3	1,8 ± 0,13	0,39	4,1
	Grupo4	1,9 ± 0,09	0,39	4,5

6.3 Exposición a variables meteorológicas por subgrupo.

En la **tabla 6** se presentan los valores de las variables meteorológicas medidas para cada uno de los grupo de participantes. Según la media y la desviación estándar de la media las variables meteorológicas se comportan de manera homogénea para los cuatro subgrupos en estudio.

En relación a los valores de humedad relativa para los cuatro grupos observados se registró que los valores de humedad relativa presentaron medias entre 66,2 y 66,9% con desviación estándar de sus medias entre 1,22 y 1,85%. Las medias de temperaturas mínimas presentaron una oscilación entre 6,05 y 6,27°C con desviaciones estándar de sus medias de 0,24 y 0,4°C. Las medias de las temperaturas medias registradas oscilaron entre 10,47°C y 10,68°C con desviaciones estándar de sus medias entre 0,29 y 0,44°C.

Tabla 6. Valores medios y desviación estándar de la media de humedad relativa y temperatura a los que estuvo expuesto cada grupo.

V. Meteorológicas	Grupo	Media \pm DS	Mínimo	Máximo
Humedad relativa (%)	Grupo1	66,9 \pm 1,31	35,5	87,3
	Grupo2	66,4 \pm 1,51	35,5	87,3
	Grupo3	66,2 \pm 1,85	35,5	86,8
	Grupo4	66, 8 \pm 1,22	35,5	87,3
Temperatura mínima (°C)	Grupo1	6,15 \pm 0,28	0,62	11,81
	Grupo2	6,05 \pm 0,3	0,62	11,81
	Grupo3	6,27 \pm 0,4	0,62	10,88
	Grupo4	6,05 \pm 0,24	0,61	11,81
Temperatura media (°C)	Grupo1	10,68 \pm 0,33	5,36	19,58
	Grupo2	10, 49 \pm 0,35	5,36	19,58
	Grupo3	10,63 \pm 0,44	5,36	17,56
	Grupo4	10,47 \pm 0,29	5,36	19,58

6.4 Modelos base.

Los modelos base para los grupos se encuentran expresados en **tabla 7**, se incluye día de la semana (día s.) y año. El VRS se repite en los cuatro modelos, también se encuentra el metaneumovirus (MNV). Para los modelos base se incluyeron algunas variables meteorológicas que se mencionan en la tabla.

Tabla 7. Modelos base para los subgrupos.

Grupo	Variables
Grupo 1	día semana, año, VRS, media 3 días hr, media 7 días temp media
Grupo 2	día semana, año, VRS, MNV, media 5 días temp. mínima
Grupo 3	día semana, año, VRS, media 7 días temp. Mínima
Grupo 4	día semana, año, VRS, MNV, humedad relativa lag 6, temp mínima lag 3

6.4.1 Modelo base para grupo 1.

A continuación, en las tablas 8, 9, 10 y 11 se presentan la tabla de resumen del modelo base para el grupo 1 con los parámetros incluidos, sus coeficientes, error estándar y valor p.

Tabla 8. Modelo base para grupo 1.

Parámetro	Coefficiente	Error estándar	p-valor
día de la semana			
categoría 2	0,0527	0,0943	0,5760
categoría 3	0,0313	0,0990	0,7520

categoría 4	-0,1088	0,1090	0,3180
categoría 5	0,0304	0,1140	0,7900
año	0,5391	0,1087	<0.001
virus sincicial	0,0079	0,0017	<0.001
media 3 días hr			
categoría 3	0,2727	0,0975	0,005
categoría 4	0,2438	0,1311	0,063
media 7 días t. media			
categoría 2	-0,0637	0,0527	0,227
categoría 3	-0,1491	0,0570	0,009
categoría 4	0,2280	0,1421	0,109

En negrita se encuentran los valores p significativos a un 0,05.

Tabla 9. Modelo base para grupo 2.

Parámetro	Coefficiente	Error estándar	p-valor
día de la semana			
categoría 2	0,1654	0,1041	0,112
categoría 3	0,0888	0,1073	0,408
categoría 4	-0,0200	0,1324	0,88
categoría 5	0,2477	0,1296	0,056
año	0,3723	0,1425	0,009
virus sincicial	0,0066	0,0019	0,001
metaneumovirus	0,0632	0,0333	0,058
media 5 días t. mínima			
categoría 2	-0,0901	0,0765	0,239
categoría 3	-0,1859	0,0832	0,025

En negrita se encuentran los valores p significativos a un 0,05.

Tabla 10. Modelo base para grupo 3.

Parámetro	Coefficiente	Error estándar	p-valor
día de la semana			
categoría 2	0,3106	0,1435	0,03
categoría 3	0,1537	0,1209	0,204
categoría 4	0,2336	0,1455	0,108

categoría 5	0,3295	0,1664	0,048
año	0,1007	0,1474	0,495
virus sincial	0,0091	0,0024	<0.001
media 7 días t. mínima			
categoría 2	-0,1519	0,0794	0,056
categoría 3	-0,3250	0,0844	<0.001

En negrita se encuentran los valores p significativos a un 0,05.

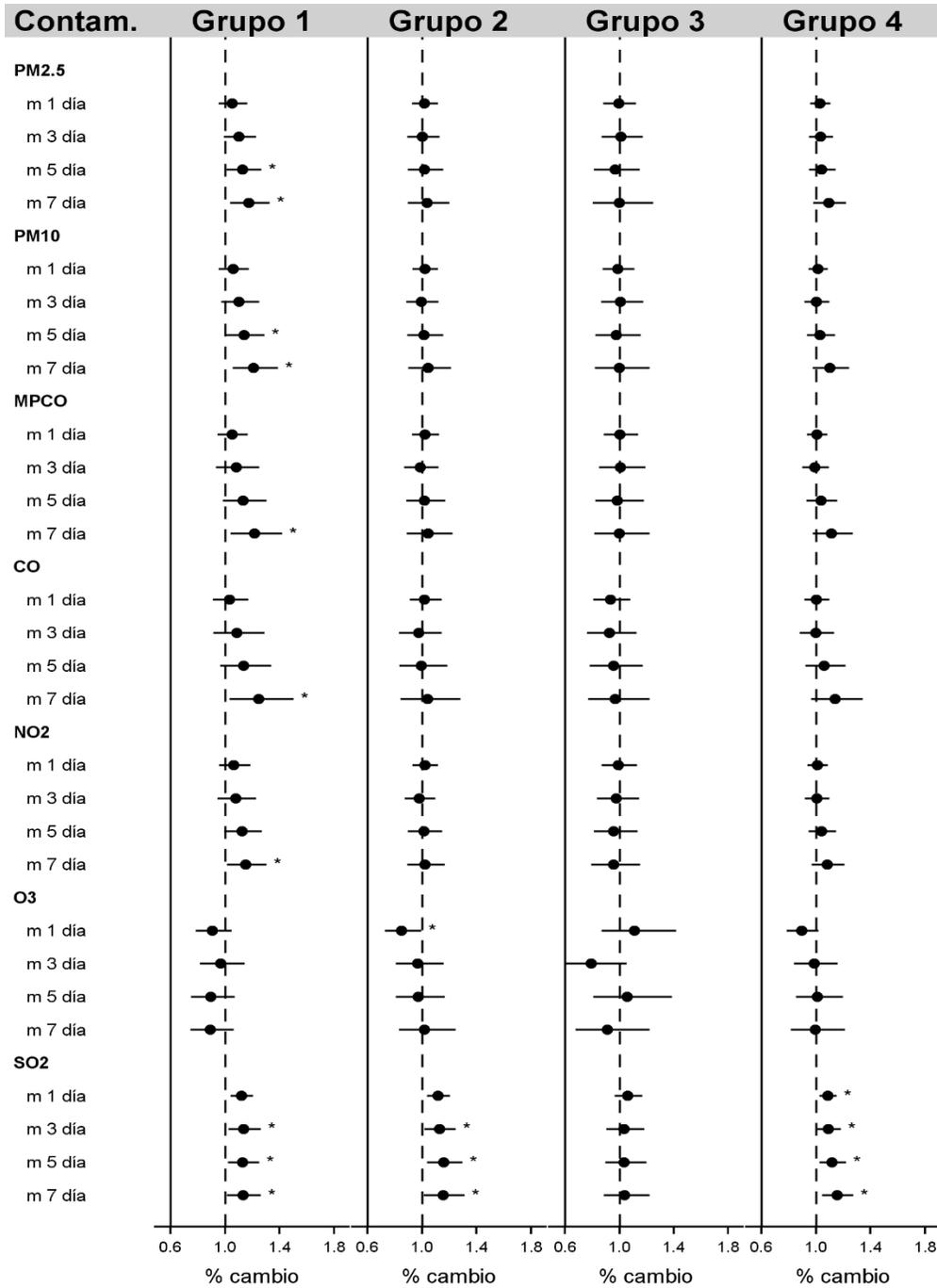
Tabla 11. Modelo base para grupo 4.

Parámetro	Coefficiente	Error estándar	p-valor
día de la semana			
categoría 2	0,2885	0,0942	0,002
categoría 3	0,2095	0,0935	0,025
categoría 4	0,1870	0,0989	0,059
categoría 5	0,2842	0,1086	0,009
año	0,3720	0,1093	<0,001
virus sincial	0,0084	0,0015	<0.001
metaneumovirus	0,0398	0,0232	0,087
hrlag 6			
categoría 3	0,0638	0,0579	0,27
categoría 4	0,0581	0,0684	0,396
t. min lag 3			
categoría 2	0,0485	0,0619	0,433
categoría 3	-0,1066	0,0630	0,09
categoría 4	-0,0738	0,1105	0,504

En negrita se encuentran los valores p significativos a un 0,05.

6.5 Susceptibilidad de los grupos a los contaminantes

Gráfico 2. Efecto de los contaminantes en el nivel de ONE para los modelos base definidos para cada grupo.



*Significancia estadística al 0,05

En el **gráfico 2**, para el grupo 1 se aprecia un incremento del ONE para la mayoría de los contaminantes evaluados. En PM 2,5 y PM 10 para la respuesta de las medias 5 y 7 días se observan incrementos significativos.

Se observa respuesta significativa en MPCO, CO y NO₂ para las medias de 7 días. Para el O₃ no se observa respuesta significativa.

Para el SO₂ se observa respuesta significativa en todas las medias de lags evaluados.

Para los grupos 2 y 4 no se observa respuesta significativa en los contaminantes evaluados a excepción del SO₂, en el cual se observa una respuesta significativa y de aumento en la medida que aumentan los lags evaluados.

Para el grupo 3 no se aprecia ninguna respuesta significativa en relación a los contaminantes evaluados.

Para el grupo 4 se observa respuesta significativa de incremento en el ONE solo en el SO₂, en el cual se aprecia respuesta significativa para todas las medias de lags evaluados.

CAPITULO 7

7. Discusión.

La respuesta inflamatoria bronquial ha sido descrita y explorada para la población de pacientes asmáticos y también para población con atopía (Vidal et al, 2013) sin embargo la respuesta inflamatoria a nivel bronquial es secundaria a diferentes estímulos y no es exclusiva de pacientes con diagnóstico de asma y/o atopía. Distintos contaminantes intradomiciliarios como el tabaco y extradomiciliarios como contaminantes ambientales de distintas fuentes (Oyarzún, 2010)(Yohannessen et al, 2015), presentan efectos inflamatorios en la población general. Por otro lado existen distintas condiciones presentes durante la infancia asociadas a procesos inflamatorios como por ejemplo la obesidad, la que también ha sido documentada (Castro-Rodriguez, 2007) previamente. En el presente estudio se buscó explorar si la agrupación de pacientes por estas características endógenas y exógenas, estudiados de forma longitudinal, presenta una variabilidad mayor en el óxido nítrico exhalado que la de pacientes con diagnóstico de asma. Esto efectivamente fue así y se encontraron grupos con mayor variabilidad y en ellos se observó el efecto de contaminantes en el OR de ONe.

A diferencia de estudios transversales (Vidal, 2013) en que no se encontraron diferencias significativas en el ONe en pacientes asmáticos y/o atópicos al emplear métodos multinivel para participantes del mismo estudio sí fue posible evidenciar un incremento estadísticamente significativo en el ONe para diferentes contaminantes coincidiendo con lo reportado por otros autores para el MP 2,5(Yohannessen et al, 2015).

Tal como se mencionó en el capítulo 1 existen estudios que dan cuenta de la variabilidad del ONe en pacientes con diagnóstico de asma o atopía y de pacientes con asma y atopía. En estos pacientes el incremento en los niveles de ONe es cercano al 20%, estos resultados son

comparables a los obtenidos en la presente investigación para el grupo 1 (pacientes con asma) para los contaminantes MP 10 sobre el 20%, MP 2,5 – 10 sobre el 21% y CO sobre el 24%(Brussee JE, 2005).

No se ha encontrado estudios previos que usen la variabilidad en las respuestas dentro de cada sujeto como herramienta de tamizaje para identificar grupos de niños susceptibles a algún contaminante ambiental específico distintos de los niños asmáticos o atópicos o asmáticos y atópicos. Pero para el grupo de participantes con consultas en urgencia e ingreso familiar inferior a trescientos mil pesos se encontraron incrementos superiores o equivalentes.

CAPITULO 8

8. Conclusión.

El grupo 1 (participantes asmáticos) presentó un comportamiento de incremento de respuesta para los niveles ONe en la mayoría de los contaminantes evaluados. Es llamativa la respuesta en dos contaminantes. En el O_3 en presencia del cual se observa una disminución del ONe, pero de manera no significativa. En el SO_2 se observa un incremento ascendente en la medida que aumentan las medias de lags evaluadas y todas presentan resultados significativos del punto de vista estadístico.

El grupo 3 no presentó respuesta significativamente mayor en los niveles de ONe para ninguno de los contaminantes evaluados.

Los grupos 2 y 4 presentaron respuestas significativamente mayores en los niveles de ONe para el SO_2 , no así para el resto de los contaminantes.

Los subgrupos evaluados no parecieron más susceptibles a la contaminación, por lo menos comparados con el grupo de participantes asmáticos.

9. Bibliografía

- Acevedo, Raziel. (2008). *Cuadernos Metodológicos. Los Modelos Jerárquicos Lineales: Fundamentos Básicos Para Su Uso y Aplicación*. San José: Instituto de Investigaciones Psicológicas, Universidad de Costa Rica.
- Álvarez Caro, F., Pérez Guirado, A., Ruiz del Árbol Sánchez, P., De Miguel Mallén, Á., & Álvarez Berciano, F. (diciembre de 2010). *Scielo*. Obtenido de http://www.scielo.org.ar/scielo.php?script=sci_arttext&pid=S0325-00752010000600011
- Brussee JE, S. (2005). Exhaled nitric oxide in 4-year-old children: relationship with asthma and atopy. *European Respiratory journal*, 455-461.
- Castro-Rodriguez. (Marzo de 2007). *ScienceDirect*. Obtenido de <http://www.sciencedirect.com/science/article/pii/S0300289607710463>
- Dong, G., Qian, Z., Liu, M.-M., Wang, D., Ren, W.-H., Fu, Q., . . . Trevathan, E. (31 de julio de 2012). Obesity enhanced respiratory health effects of ambient air pollution in Chinese children: the Seven Northeastern Cities study. China.
- Fabian, M., Stout, N., Adamkiewiks, G., Geggel, A., Ren, C., Sandel, M., & Levy, J. (2012). The effects of indoor environmental exposures on pediatric asthma: a discrete event simulation model. *Environmental Health*, 1-15.
- Gavidia, T., Pronkzuc, J., & Dsly, P. (2009). Impactos ambientales sobre la salud respiratoria de los niños. Carga global de las enfermedades respiratorias pediátricas ligadas al ambiente. *Revista Chilena de Enfermedades Respiratorias*, 99-108.
- Goldizen, F., Sly, P., & Knibbs, L. (2016). Respiratory effects of air pollution on children. *Pediatric Pulmonology*, 94-108.
- Goldstein, Harvey. (1999). *Multilevel Statistical Models*. Londres: Institute of Education.
- Goldstein, Harvey. (2010). *Multilevel Statistical Models*. Wiley.
- Goldstein, Harvey; Rasbash, Jon. (1992). Efficient computational procedures for the estimation of parameters in multilevel models based on iterative generalised least squares. *Computational Statistics & Data Analysis*, 63-71.
- Hoek, G., Pattenden, S., Willers, S., Antova, T., Fabianova, E., Braun-Fahrländer, C., . . . Fletcher, T. (2012). PM10, and children's respiratory symptoms and lung function in the PATY study. *European Respiratory Journal*, 538-547.
- Hox, J. (1995). *Applied Multilevel Analysis*. Amsterdam: TT-Publikaties.
- Hox, J. (2002). *Multilevel Analysis Techniques and Applications*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.

<http://www.who.int/respiratory/asthma/es/>. (s.f.). Recuperado el agosto de 2016

Janes, H., Sheppard, L., & Shepherd, K. (2008). Statistical Analysis of Air Pollution Panel Studies: An Illustration. *Elsevier*, 792-802.

Kreft, I., & De Leeuw, J. (2002). *Introducing Multilevel Modeling*. London: SAGE Publications Inc.

Meave-Cueva, L., Díaz-García, L., Llamosas-Gallardo, B., & Ortiz-Hernandez, A. (2013). Medición del óxido nítrico exhalado en niños con crisis asmática: Revisión sistemática. *Acta Pediátrica de México*, 247-252.

Ministerio de Salud. (2011).

<http://web.minsal.cl/portal/url/item/7222754637f08646e04001011f014e64.pdf>.

Recuperado el agosto de 2016

Modelling, C. f. (marzo de 2017). *LEMMA (Learning Environment for Multilevel Methods and Applications)*,. Obtenido de <https://www.cmm.bris.ac.uk/lemma/>

Morales, E., García-Esteban, R., de la Cruz, O., Basterrechea, M., Lertxundi, A., de Dicastillo, M., . . . Sunyer, J. (2015). Intrauterine and early postnatal exposure to outdoor air pollution and lung function at preschool age. *Thorax*, 64-73.

Oyarzún, M. (2010). Contaminación aérea y sus efectos en salud. *Revista Chilena de Enfermedades Respiratorias*, 16-25.

Pepper, I. (s.f.). <http://atlas.med.uchile.cl/35.htm>. Recuperado el agosto de 2016

Perez , V. (2013). Factores de Riesgo de Sibilancias Recurrentes en Lactantes. Enfoque Frecuentista y Bayesiano. *Tesis*.

Rabe-Hesketh, S., & Skrondal, A. (2008). *Multilevel and Longitudinal Modeling Using Stata*. Texas: Stata Press.

Reddel , H., & Bateman, E. (diciembre de 2016). *ginasthma.org*. Obtenido de <http://www.ginasthma.org/>

Scharz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 461-464.

Twisk, J. (2006). *Applied Multilevel Analysis*. Cambridge: Cambridge University Press.

Velástegui, C., Pérez-Canto, P., Zárate, V., Arenas, D., Salinas, P., & Moreno, G. (2010). Impacto del asma en escolares de dos centros de salud primaria. *Revista Médica de Chile*, 205-212.

Vidal, D., Yohannessen, K., Prieto, L., Ubilla, C., & Ruiz, P. (2013). Óxido nítrico exhalado, asma y atopia: análisis en un panel de niños en Santiago, Chile. *Revista Médica de Chile*, 743-750.

