



UNIVERSIDAD
NACIONAL
DE COLOMBIA

Mortalidad por desnutrición en menores de cinco años, aproximación estadística de los determinantes, Colombia 2009-2012

Luis Carlos Forero Ballesteros

Universidad Nacional de Colombia
Facultad de Ciencias, Departamento de Estadística
Bogotá, Colombia
2017

Mortalidad por desnutrición en menores de cinco años, aproximación estadística de los determinantes, Colombia 2009-2012

Luis Carlos Forero Ballesteros

Trabajo de grado presentado como requisito parcial para optar al título de:
Magister en Ciencias, Bioestadística

Directora:

B. Piedad Urdinola Contreras Ph.D.

Línea de Investigación:

Bioestadística

Universidad Nacional de Colombia
Facultad de Ciencias, Departamento de Estadística
Bogotá, Colombia
2017

Dedicatoria

A mi familia, padres, hermanos y esposa, que han tenido que vivir este proceso como si fuera suyo.

Agradecimientos

A mi directora de trabajo de grado, B. Piedad Urdinola Contreras, profesora asociada del Departamento de Estadística de la Universidad Nacional de Colombia, por aceptar dirigir este camino, su paciencia, confianza y continua disposición, sus valiosas y oportunas orientaciones y estrategias de trabajo, que permiten siempre avanzar y abordar las preguntas de investigación, desde una mirada interdisciplinaria.

Resumen

La mortalidad por desnutrición es una situación influenciada por diferentes factores, por lo cual es importante identificar variables asociadas a este evento en la población, que permitan establecer los principales determinantes (sociales, demográficos, económicos, individuales, de la Seguridad Alimentaria y Nutricional), que explican esta mortalidad evitable. El objetivo fue identificar los determinantes de la mortalidad por desnutrición en población colombiana menor de cinco años para los años 2009 a 2012. Se exploró la relación entre la tasa de mortalidad por desnutrición y algunas variables del certificado de defunción, así como otras variables de diferentes fuentes, en una base de datos de panel construida con información de los departamentos y distrito capital de Colombia entre los años 2009 y 2012. Se estimaron los modelos de regresión realizando ajustes para cumplir los supuestos de correlación serial y heterocedasticidad con el método de errores estándar corregidos para panel PCSE. Se encontró que las variables socioeconómicas para todos los grupos de edad presentaron los coeficientes más altos y positivos, específicamente las de pobreza e indigencia. También se resaltan otras variables significativas como la afiliación en salud y la etnia indígena, que se asocian a un incremento de la tasa de mortalidad por desnutrición. Esta mortalidad no parece estar relacionada con la cobertura de vacunación y los programas de asistencia social, así como el área, sexo y nivel educativo que aunque en algunos casos fueron significativas, el coeficiente de cambio porcentual es muy bajo.

Palabras clave: Mortalidad, Desnutrición, Determinantes sociales de la salud, Modelos de regresión, Datos tipo panel.

Abstract

Mortality due to malnutrition is a situation influenced by different factors. Therefore, it is important to identify variables associated with this event in the population, which allow establishing the main determinants (social, demographic, economic, individual, Food and Nutrition Security) which explain this avoidable cause of mortality. The objective was to identify the determinants of mortality due to malnutrition in the Colombian population under five years of age for the years 2009 to 2012. The relationship between the mortality rate due to malnutrition and some variables of the death certificate, as well as other variables of different sources in a panel database constructed with information from the departments and capital district of Colombia between 2009 and 2012. The regression models were estimated by making adjustments to meet the assumptions of serial correlation and heteroscedasticity using the standard error method corrected for (PCSE) panel. We found that the socioeconomic variables for all age groups had the highest and positive coefficients, specifically those of poverty and indigence. It is also highlighted other significant variables such as affiliation to the health system, and indigenous ethnicity, which are associated with an increase in the mortality rate due to malnutrition. This mortality does not seem to be related to vaccination coverage and social assistance programs, as well as the area, sex and educational level that, although in some cases were significant, the percentage change coefficient is very low.

Keywords: Mortality, Malnutrition, Social determinants of health, Regression models, Panel data.

Contenido

	Pág.
Resumen	VII
Lista de figuras	XI
Lista de tablas	XII
Lista de símbolos y abreviaturas	XIV
Introducción	1
1. Objetivos	5
1.1 General.....	5
1.2 Específicos	5
2. Marco teórico	7
2.1 La Mortalidad	7
2.2 La Desnutrición.....	10
2.3 La Mortalidad por desnutrición	11
2.4 Los Determinantes.....	13
2.4.1 Determinantes sociales de la salud	14
2.4.2 Determinantes sociales de la Seguridad Alimentaria y Nutricional	15
2.4.3 Determinantes propuestos	16
2.5 Estadísticas vitales	18
2.5.1 Certificado de defunción.....	18
2.5.2 Clasificación internacional de enfermedades	20
2.6 Modelos estadísticos	20
2.6.1 Regresión lineal	21
2.6.2 Modelos lineales generalizados	22
2.6.3 Estimación de parámetros en el modelo lineal general.....	24
2.6.4 Modelos para datos de conteo	24
2.6.5 Sobredispersión	27
2.7 Datos Panel	27
2.7.1 Aspectos conceptuales	29
2.7.2 Modelos de datos panel estáticos	36
2.7.3 Test o pruebas recomendadas.....	43
2.7.4 Modelos de datos panel dinámicos	47
2.8 Variables expresadas en logaritmos	49

3. Metodología.....	51
3.1 Búsqueda y revisión de literatura	51
3.2 Fuente de los datos.....	51
3.3 Clasificación de la mortalidad por desnutrición	52
3.4 Estimación de la mortalidad.....	52
3.5 Variables seleccionadas para el análisis	53
3.5.1 Descripción de las variables	54
3.6 Procesamiento y análisis de datos	57
3.6.1 Depuración y ajuste de la base de datos	57
3.6.2 Metodología para datos de panel.....	58
3.6.3 Modelación estadística	60
4. Resultados	62
4.1 Análisis exploratorio de la información	62
4.1.1 Proporción de muertes.....	63
4.1.2 Tasas de mortalidad	63
4.1.3 Variables principales.....	65
4.2 Modelación.....	72
4.2.1 Modelos para el grupo de edad menor de 5 años.....	73
4.2.2 Modelos para el grupo de edad de uno a cuatro años	77
4.2.3 Modelos para el grupo de edad menor de 1 año.....	81
4.2.4 Consideraciones generales.....	86
5. Conclusiones y recomendaciones	87
5.1 Conclusiones.....	87
5.2 Recomendaciones.....	90
A. Anexo: Fuentes de información.....	92
B. Anexo: Descripción de las variables	95
C. Anexo: Comandos Stata.....	99
D. Anexo: Agrupación de departamentos por región	101
E. Anexo: Mortalidad por desnutrición, según el departamento y la región.....	103
F. Anexo: Mapas de Mortalidad por desnutrición.....	113
G. Anexo: Gráficos exploratorios de los paneles de datos	118
Bibliografía	121

Lista de figuras

	Pág.
Figura 2-1: Tasa de mortalidad en menores de cinco años por 1.000 nacidos vivos. Tendencia 1970 – 2010.....	9
Figura 2-2: Modelo de Determinantes Sociales de la Salud de la OMS	15
Figura 4-1: Proporción de muertes por desnutrición en menores de cinco años, según la región, período 2009 - 2012.	63
Figura 4-2: Tasa de mortalidad por desnutrición en menores de cinco años, según la región, período 2009 - 2012.	64
Figura 4-3: Tasa de mortalidad por desnutrición, según el grupo de edad, período 2009 - 2012	68
Figura F-1: Tasa de mortalidad por desnutrición en menores de 5 años, según departamento, Colombia 2009.	114
Figura F-2: Tasa de mortalidad por desnutrición en menores de 5 años, según departamento, Colombia 2010.	115
Figura F-3: Tasa de mortalidad por desnutrición en menores de 5 años, según departamento, Colombia 2011.	116
Figura F-4: Tasa de mortalidad por desnutrición en menores de 5 años, según departamento, Colombia 2012.	117
Figura G-1: Gráfico exploratorio del panel de datos para los menores de 5 años, según departamento, Colombia 2009-2012.	118
Figura G-2: Gráfico exploratorio del panel de datos grupo de 0 a 4 años, según departamento, Colombia 2009-2012.	119
Figura G-3: Gráfico exploratorio del panel de datos para los menores de 1 año, según departamento, Colombia 2009-2012.	120

Lista de tablas

	Pág.
Tabla 2-1: Transformaciones logarítmicas.....	49
Tabla 4-1: Defunciones de mortalidad por desnutrición, según el año.....	62
Tabla 4-2: Proporción total y por grupo de edad, de variables seleccionadas en menores de 5 años.....	66
Tabla 4-3: Proporción total y por grupo de edad de variables seleccionadas, después de la depuración, menores de 5 años.....	69
Tabla 4-4: Proporción total y por sexo de variables seleccionadas, después de la depuración, grupo de 1 a 4 años.....	70
Tabla 4-5: Proporción total y por sexo de variables seleccionadas, después de la depuración, menores de 1 año.....	70
Tabla 4-6: Coeficientes de regresión de los modelos tipo panel doble logarítmicos, para la tasa de mortalidad por desnutrición, en menores de cuatro años. Colombia, 2009 – 2012.....	75
Tabla 4-7: Coeficientes de regresión de los modelos tipo panel doble logarítmicos, para la tasa de mortalidad por desnutrición en individuos de uno a cuatro años. Colombia, 2009 - 2012.....	79
Tabla 4-8. Coeficientes de regresión de los modelos tipo panel doble logarítmicos, para la tasa de mortalidad por desnutrición en menores de 1 año. Colombia, 2009 -2012.....	83
Tabla B-1: Variables de la base de datos.....	96
Tabla C-1: Comandos Stata.....	99
Tabla D-1: Agrupación de departamentos por región.....	101
Tabla E-1: Proporciones y tasas de mortalidad por desnutrición, según el departamento. Menores de 5 años.....	103
Tabla E-2: Proporciones y tasas de mortalidad por desnutrición, según la región. Menores de 5 años.....	105

Tabla E-3: Proporciones y tasas de mortalidad por desnutrición, según el departamento. Grupo de 1 a 4 años.....	106
Tabla E-4: Proporciones y tasas de mortalidad por desnutrición, según la región. Grupo de 1 a 4 años.	108
Tabla E-5: Proporciones y tasas de mortalidad por desnutrición, según el departamento. Menores de 1 año.	109
Tabla E-6: Proporciones y tasas de mortalidad por desnutrición, según la región. Menores de 1 año.	111

Lista de símbolos y abreviaturas

Símbolos con letras latinas

Símbolo	Término
X_{it}	Vector de dimensión $(1 \times K)$ que contiene los valores de las covariables
u_{it}	Término de error. Idiosincrático o error variable con el tiempo
y_{it}	Es el valor de la variable respuesta
ε_{it}	Es el término del error puramente aleatorio.

Símbolos con letras griegas

Símbolo	Término
α_i	Efectos individuales
β_{it}	Es el vector $(k \times 1)$ de parámetros del modelo
δ_t	Hace referencia a los efectos no cuantificables
θ_i	Se llama parámetro natural de localización
μ_i	Representa los efectos no observables
σ^2	Varianza
ϕ	Es un parámetro de escala

Subíndices

Subíndice	Término
i	$1, \dots, N$ unidades muestrales (departamentos)
k	$1, \dots, K$ covariables o variables explicativas
t	$1, \dots, T$ periodos (años)

Superíndices

Superíndice	Término
-1	El inverso
2	El cuadrado
n	Exponente, potencia
θ	Parámetro canónico

Abreviaturas

Abreviatura	Término
CIE	Código Internacional de Enfermedades
ANDA	Archivo Nacional de Datos
DANE	Departamento Administrativo Nacional de Estadística
DNP	Departamento Nacional de Planeación
DNT	Desnutrición
EDA	Enfermedad Diarreica Aguda
EEVV	Estadísticas vitales
FAO	Organización de las Naciones Unidas para la Alimentación y la Agricultura (sigla en inglés)
FGLS	Estimador Mínimos Cuadrados Generalizados Factibles (sigla en inglés)
GMM	Método de los momentos generalizado (sigla en inglés)
i.i.d.	Independientes e idénticamente distribuidos
ICBF	Instituto Colombiano de Bienestar Familiar
IGME	Estimaciones de la Mortalidad Infantil (sigla en inglés)
INML-CF	Instituto Nacional de Medicina Legal y Ciencias Forenses
INS	Instituto Nacional de Salud
IRA	Infecciones Respiratorias Agudas
IRLS	Mínimos cuadrados iterativamente reponderados (sigla en inglés)
MANA	Plan de Mejoramiento Alimentario y Nutricional de Antioquia
MCO	Mínimos Cuadrados Ordinarios
MELI	Mejores Estimadores Lineales Insesgados
MINSALUD	Ministerio de Salud y Protección Social
MLG	Modelos Lineales Generalizados
MSPS	Ministerio de Salud y Protección Social
OLS	Ordinary Least Squares
OMS	Organización Mundial de la Salud
OPS	Organización Panamericana de la Salud
PAHO	Organización Panamericana de la Salud (sigla en inglés)
PCSE	Errores Estándar Corregidos para Panel (sigla en inglés)
PIB	Producto Interno Bruto
RUAF	Registro Único de Afiliados
SAN	Seguridad Alimentaria y Nutricional
SEN	Sistema Estadístico Nacional
SGD	Sistema de Gestión de Datos
SISBEN	Sistema de identificación y clasificación de potenciales beneficiarios para programas sociales
SISPRO	Sistema Integral de Información de la Protección Social
UNICEF	Fondo de las Naciones Unidas para la infancia
WHO	Organización Mundial de la Salud (sigla en inglés)

Introducción

La mortalidad se mide desde hace siglos como lo propuso John Graunt, a partir de tasas de mortalidad, que se definen como el número de defunciones en un período de tiempo determinado por cada mil habitantes, lo cual ha permitido identificar además de los patrones de mortalidad general, por edades y causas, las condiciones de desarrollo y pobreza de las poblaciones y la construcción de tablas de vida que traducen estas tasas en probabilidades (Alkema et al., 2014) (Acosta & Romero, 2014).

El propósito de esta investigación es identificar variables asociadas a la mortalidad por desnutrición en la población colombiana menor de cinco años, que permitan establecer los principales determinantes (sociales, demográficos, económicos, individuales, de la Seguridad Alimentaria y Nutricional), que explican esta mortalidad evitable. A partir de la información generada por los registros de defunción nacional, así como otras fuentes de información exógena, y la aplicación de métodos estadísticos y demográficos que permitan dar robustez a la información y los resultados; así como plantear un modelo estadístico, que permita identificar dichos determinantes de la mortalidad por desnutrición en los menores de cinco años.

Las tasas más altas de muerte en la población se evidencian en los adultos mayores, seguido por los menores de cinco años. En el primer caso está asociado a enfermedades crónicas resultado del proceso de envejecimiento, así como los hábitos y estilos de vida acumulados durante los años. Sin embargo, para el segundo caso, la enfermedad y muertes se asocian principalmente a procesos infecciosos, resultado de diferentes condiciones como las ambientales, la atención en salud, la influencia de la alimentación y el estado nutricional de los individuos (UNICEF et al., 2013) (Lisowska, 2013).

Esta última condición sobre el estado nutricional, está influenciado no sólo por factores biológicos, sino también por determinantes ambientales y psicosociales, entre otros, que

pueden llevar a un estado de malnutrición. De esta forma la desnutrición que corresponde a un tipo de malnutrición por deficiencia de calorías y nutrientes, está asociada a múltiples factores que hacen de su caracterización un ejercicio complejo; así mismo es origen de variadas enfermedades que no permiten identificarla como una de las causas básicas en las mediciones de la morbilidad y la mortalidad (Arias, Tarazona, Lamus, & Granados, 2013).

Las variables generalmente utilizadas para la estimación de la mortalidad incluyen la enfermedad diagnosticada o causa de defunción, la edad, el sexo y la ubicación geográfica. Sin embargo se hace necesario asociar otros aspectos que puedan tener influencia en este desafortunado desenlace, en muchas ocasiones evitable.

El principal parámetro para identificar la causa básica de las defunciones corresponde a la clasificación internacional de enfermedades en su décima versión con el acrónimo CIE10 (OPS, 2008). Para especificar los diagnósticos asociados a la situación de estudio de mortalidad por desnutrición se tiene en cuenta la clasificación en la lista 667-OPS CIE-10 con las defunciones asociadas a los códigos E40 - E64, D50 - D53 definidos como “deficiencias nutricionales y anemias nutricionales” (OPS, 1995).

En Colombia la estimación de mortalidad o el conteo de los casos de defunción se realizan directamente de la información de los registros de defunciones no fetales (registros individuales), producido anualmente por el sistema de estadísticas vitales del Departamento Administrativo Nacional de Estadística DANE, que para el objeto de estudio corresponde al periodo comprendido entre los años 2009 a 2012.

Dentro de las limitaciones del estudio se identifica el sub-registro de la información de defunción, que en Colombia se estimaba en 20,3% durante 2011, estos valores posicionan al país como uno de los más altos niveles de subregistro entre los principales países de Latinoamérica (Acosta & Romero, 2014). Otra limitación es la selección de diagnósticos asociados a deficiencias y anemias nutricionales, como causas principales o secundarias de muerte. Con lo anterior se espera que el número de casos reportados por las condiciones específicas de estudio, sea muy bajo.

El documento está integrado por seis secciones, sin contar la presente introducción; en la primera se relacionan los objetivos, seguidos del marco teórico que refiere los temas de importancia para el desarrollo del trabajo, la tercera sección describe la metodología desarrollada, desde la construcción de la base de datos y su procesamiento, hasta el planteamiento del análisis; la cuarta es la presentación de resultados y su análisis y se finaliza con las conclusiones, recomendaciones y referencias bibliográficas.

1.Objetivos

1.1 General

- Identificar los determinantes de la mortalidad por desnutrición en población colombiana menor de cinco años para los años 2009 a 2012.

1.2 Específicos

- Realizar un análisis exploratorio de la información sobre mortalidad por desnutrición, para el periodo de estudio.
- Plantear un modelo estadístico con las principales variables de estudio, para identificar su efecto estadístico en la explicación de la mortalidad.
- Establecer las variables asociadas con las defunciones por desnutrición.
- Evaluar el efecto conjunto de las variables seleccionadas sobre las defunciones por desnutrición.

2. Marco teórico

Para el desarrollo del componente teórico, se parte de las definiciones de mortalidad, desnutrición y la descripción alrededor de los determinantes, como ejes centrales del trabajo; posteriormente se mencionan las estadísticas vitales que hacen parte de las fuentes principales de información y finalmente el componente estadístico que definirá la metodología, donde se destaca el análisis de datos tipo panel.

2.1 La Mortalidad

Según el DANE la mortalidad se refiere a la cantidad de personas que fallecen en un país y en un período determinado (DANE, 2008).

Dichos fallecimientos se conocen como defunciones y son considerados registros demográficos; DANE (2012) lo define como la desaparición permanente de todo signo de vida, cualquiera que sea el tiempo transcurrido desde el nacimiento con vida (cesación postnatal de las funciones vitales sin posibilidad de resucitar). Esta definición excluye las defunciones fetales (DANE, 2012).

En menores de cinco años la mortalidad hace referencia a las muertes antes de cumplir cinco años de edad (Alkema et al., 2014). Se explica por factores endógenos al niño (deterioro biológico y genético) y a los exógenos a él y su familia, como condiciones sociales, económicas y ambientales (UNICEF et al., 2011).

Las nuevas estimaciones de mortalidad incluidas en el informe¹: “*Niveles y tendencias en la mortalidad infantil 2014*” del Grupo Interinstitucional de las Naciones Unidas sobre Estimaciones de la Mortalidad Infantil (IGME, por sus siglas en inglés²), revelan que en el año 2013 en el mundo murieron 6,3 millones de niños y niñas menores de cinco años, en su mayor parte por causas prevenibles, cifra que representa un descenso de 200.000 muertes respecto a 2012, que se traduce en alrededor de 17.000 muertes infantiles diarias (Acosta & Romero, 2014).

En 2013 las principales causas de muertes de los menores de cinco años en el mundo, son las complicaciones derivadas del nacimiento prematuro (17%), la neumonía (15%), las complicaciones en el parto y el alumbramiento (11%), la diarrea (9%) y el paludismo (7%), La subnutrición origina casi la mitad de todas las muertes de menores de 5 años (Acosta & Romero, 2014).

Los niños y niñas nacidos en Angola, país que presenta la tasa de mortalidad de menores de cinco años más elevada del mundo (167 muertes por cada 1.000 nacidos vivos), tienen 84 veces más probabilidades de morir antes de cumplir los cinco años que los que nacen en Luxemburgo, país que presenta la tasa más baja (2 muertes por cada 1.000). Dentro de un mismo país son factores clave la riqueza relativa, el nivel educativo y la ubicación: el riesgo de un niño o niña de morir aumenta si nace en una zona rural remota, en un hogar pobre o de una madre que no ha recibido educación (Acosta & Romero, 2014).

Por otro lado, aunque la tasa de mortalidad derivada de enfermedades y lesiones entre los niños menores de cinco años es casi 20 veces mayor en los países de ingreso bajo

¹ El informe Niveles y tendencias de la mortalidad infantil de 2014, es una publicación anual del Grupo Interinstitucional de las Naciones Unidas sobre Estimaciones de la Mortalidad Infantil, que encabeza UNICEF y en el que participan la Organización Mundial de la Salud, el Grupo del Banco Mundial y la División de Población del Departamento de Asuntos Económicos y Sociales de las Naciones Unidas.

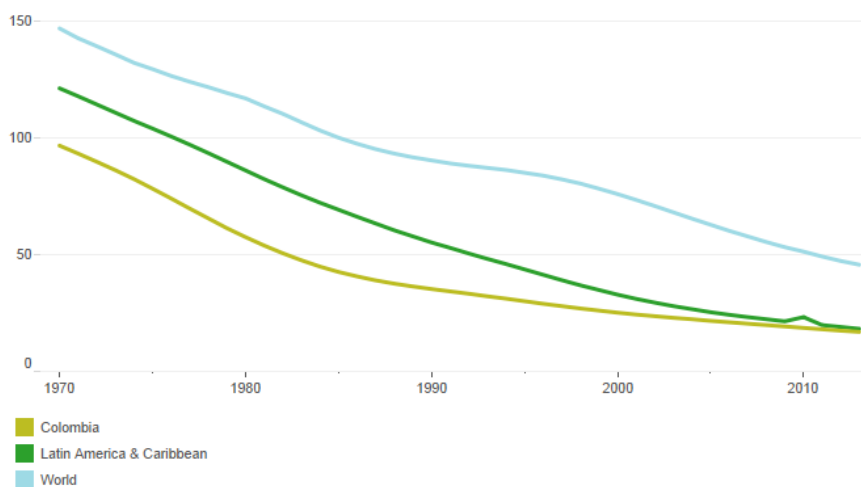
² El IGME, se constituyó en 2004 con el objeto de intercambiar datos sobre la mortalidad infantil, de armonizar las estimaciones en todo el sistema de las Naciones Unidas, perfeccionar los métodos que se emplean en el cálculo de la mortalidad infantil para informar de los avances logrados en la consecución de los Objetivos de Desarrollo del Milenio, y reforzar la capacidad de los países para elaborar cálculos de la mortalidad infantil oportunos y basados en valoraciones correctas.

que en los de ingreso alto; la mejora en el saneamiento y el aumento de la vacunación, junto con otras medidas de salud preventivas, han contribuido a reducir las tasas de mortalidad infantil y materna (UNICEF et al., 2013).

En Colombia, la distribución de la mortalidad por curso de vida muestra grandes diferencias, una de ellas se evidencia en la primera infancia en donde las afecciones originadas por el período perinatal representan entre 38% y 41% del total de grandes causas. Dentro de este subgrupo, se evidencia los trastornos respiratorios específicos del periodo perinatal como primera causa de mortalidad, observándose de 35 a 49 muertes por cada 100.000 niños menores de cinco años, entre el 2008 y 2010 (Hoz et al., 2014).

En la Figura 2-1 se presenta el cambio de la tasa de mortalidad en menores de cinco años en las últimas décadas, de forma comparativa para el Mundo, Latinoamérica y Colombia. Se destaca el estancamiento de Colombia en la disminución de la mortalidad, aproximadamente en los últimos 25 años, lo cual representa un retraso con otros países, incluyendo los de la región.

Figura 2-1: Tasa de mortalidad en menores de cinco años por 1.000 nacidos vivos. Tendencia 1970 – 2010.



Fuente: Banco Mundial <http://datos.bancomundial.org>

2.2 La Desnutrición

La desnutrición es el resultado de una ingestión alimentaria reducida en un tiempo prolongado o la absorción deficiente de los nutrientes consumidos como resultado de una enfermedad. Dicha reducción en la ingesta generalmente está asociada a su vez a una carencia de energía, de vitaminas y minerales (micronutrientes) (FAO, 2014).

Para la medición de la desnutrición se establecen indicadores como el bajo peso para la edad (desnutrición global), la baja talla para la edad (retraso en talla) y el bajo peso para la talla (desnutrición aguda), los cuales se construyen con las variables peso, talla o longitud, sexo y edad (FAO, 2014).

Partiendo de los indicadores anteriores, la prevalencia de la desnutrición en la niñez representa el porcentaje de niños menores de cinco años de edad con un peso para la edad o talla para la edad, por debajo de dos desviaciones estándar de la media de la población de referencia entre 0 y 59 meses de edad. Los datos están basados en los estándares mundiales de desarrollo infantil publicados por la OMS en 2006 (Minsalud, 2010).

Además la desnutrición es el resultado de múltiples factores asociados: económicos, políticos y sociales, y es, a la vez, la causante de múltiples problemas de salud, como la vulnerabilidad ante las infecciones recurrentes, que contribuyen a las altas tasas de mortalidad en la población menor de cinco años. Para el año 2007, 143 millones de niños menores de cinco años del mundo en desarrollo siguen sufriendo de desnutrición (Quiroga, 2012).

Según la Encuesta Nacional de la Situación Nutricional en Colombia 2010 ENSIN, el porcentaje de desnutrición crónica o retraso en crecimiento se redujo en los últimos 5 años de 16,0% a 13,2%, quedando a 5,2 puntos porcentuales de cumplir con la meta propuesta para el 2015 por el país, en el objetivo de desarrollo del milenio de erradicar la pobreza extrema y el hambre (Minsalud et al., 2011).

El porcentaje de desnutrición global es de 3,4%. Colombia se encuentra a 0,8 puntos porcentuales de cumplir con la meta propuesta para el 2015 por el país, en el objetivo de

desarrollo del milenio de erradicar la pobreza extrema y el hambre (Minsalud et al., 2011).

Al igual que el retraso en crecimiento, este tipo de desnutrición global presenta mayor prevalencia en los últimos años, en las familias con mayor número de hijos (6 o más) (8,9%), en aquellos cuya diferencia de edad con su hermano siguiente es menor a dos años (5,8%), en los hijos de mujeres sin educación (12,2%) y con menor nivel de SISBEN (4,7%) (Minsalud et al., 2011).

Por lugar de residencia, el área rural casi duplica la prevalencia frente a la urbana 4,7% frente a 2,9%, respectivamente. Al igual que en la desnutrición crónica, las regiones más afectadas son Atlántica (4,9%); y Amazonia y Orinoquia (3,6%). Comparado con 12 países latinoamericanos de los que se dispone de información, Colombia (Minsalud et al., 2011).

Las cifras de retraso en crecimiento en Colombia se encuentran en el tercer lugar de las prevalencias más bajas entre 12 países latinoamericanos de los que se dispone información, mientras presenta la menor cifra de desnutrición global (Minsalud et al., 2011).

2.3 La Mortalidad por desnutrición

En el estudio ecológico para el análisis del comportamiento de la mortalidad por desnutrición en menores de cinco años en las regiones de Antioquia, antes y después de la creación del programa MANA³, se menciona que en el año 2006, por primera vez en la historia reciente, la cifra total de muertes anuales por desnutrición entre los niños y niñas menores de cinco años fue inferior a los 10 millones. Esto representa una reducción de un 60% de la tasa de mortalidad en la infancia desde 1960 (Toro & Arango, 2013).

³ MANA: Plan de Mejoramiento Alimentario y Nutricional de Antioquia. Este programa trabaja por la seguridad alimentaria y nutricional de los antioqueños, principalmente de los niños, a través de diversos proyectos: brindando complementos alimenticios a menores de 6 años; aportando a la permanencia escolar mediante la entrega de víveres para los restaurantes escolares, y formando a la comunidad educativa en hábitos alimentarios saludables.

Un estudio mexicano refiere que durante los últimos 20 años en dicho país, han ocurrido cerca de 35 mil muertes de niños menores de cinco años por uno de los problemas más característicos de las sociedades en vías de desarrollo: la desnutrición infantil. Además la desnutrición, como causa de morbilidad y mortalidad, constituye un gran problema de carácter social, siendo la consecuencia más directa de una dieta insuficiente originada por bajos ingresos económicos en el seno familiar. Por otra parte, en cuanto al peso relativo que implica la mortalidad por desnutrición, este también se ha modificado en el mismo sentido, ya que en 1990 esta causa era la responsable de casi 5% de las muertes de menores de cinco años, en tanto que en el último año disponible dicho porcentaje ha caído por debajo de 2% (Fernández Cantón & Viguri Uribe, 2010).

Ruíz (2007) afirma que, para el periodo entre 1998 y 2002, en Colombia ocurrieron 39.197 defunciones en menores de 5 años, en las que estuvo presente la desnutrición ya fuera como causa directa, como causa antecedente o como otro estado patológico importante en el momento de la defunción. Además, demostró que en los niños menores de cinco años, existe la probabilidad que una de cada cinco muertes sea por desnutrición o esté asociada a esta, como sucede en la mayor parte de los municipios del país (Ruiz & Ruiz, 2007).

La investigación realizada por Quiroga (2012), que tenía como objetivo caracterizar la mortalidad por desnutrición utilizando un método de ajuste para la tasa, dadas las dificultades en su estimación por las limitaciones en la identificación de la desnutrición como causa básica de muerte, destaca en los resultados que la desnutrición, como antecedente, tiene asociación con las enfermedades infecciosas; la frecuencia de la enfermedad infecciosa como causa básica de muerte fue siete veces mayor cuando coexistió la desnutrición como causa antecedente. Una vez ajustada la tasa de mortalidad, se encontró que aumentó 5 veces el valor inicial. El método de “análisis de causas múltiples” establece su efectividad como metodología en el ajuste de este tipo de mortalidad (Quiroga, 2012).

Según resultados obtenidos del análisis de las estadísticas vitales de las fuentes DANE de las defunciones por desnutrición en Colombia para los menores de cinco años en el período 2002 a 2009 y de acuerdo con los datos disponibles para análisis, las defunciones por desnutrición en Colombia para menores de cinco años fueron 4.354,

esto representa 4% del total de muertes en menores de cinco años en el mismo período (Hoz et al., 2014).

2.4 Los Determinantes

Establecer el análisis desde los determinantes es fundamental para comprender y explicar los factores causales de problemas nutricionales, en especial los asociados a muertes evitables, por este tipo de problemas. Muchos factores han sido identificados asociados a la mortalidad infantil, Urdinola en el 2011, refiere que se destacan la educación materna, la calidad de la vivienda, el acceso a servicios públicos y a servicios sanitarios como los principales determinantes de la mortalidad infantil (Urdinola, 2011); sin embargo la mortalidad por desnutrición es un evento que requiere un análisis particular que permita identificar los factores con más peso.

Abordar el tema de los determinantes reconoce el desarrollo de diferentes corrientes y postulados, por ejemplo Mosley y Chen (1984) en sus estudios clásicos ofrecen un esquema integral de la muerte y la supervivencia infantil, concebida como un proceso que incorpora determinantes macro estructurales de distintos niveles (individual, del hogar y de la comunidad), dentro de los cuales se analizan el nivel educativo de la madre, las tradiciones, las normas y actitudes, relaciones de poder, el ingreso, la alimentación, la política económica, los sistemas de salud, entre otros. Y los determinantes próximos, que refieren los factores de la maternidad y del patrón reproductivo, el medio ambiente y la contaminación, las deficiencias nutricionales, lesiones y control personal de las enfermedades, entre otros (Longhi, 2013).

De esta forma la revisión se centrará en dos enfoques principalmente, el de determinantes sociales de la salud y el de los determinantes sociales de la Seguridad Alimentaria y Nutricional, este último basado en el primero y que se encuentra en construcción como es referido en el modelo conceptual y diseño metodológico de la próxima Encuesta Nacional de Situación Nutricional en Colombia ENSIN.

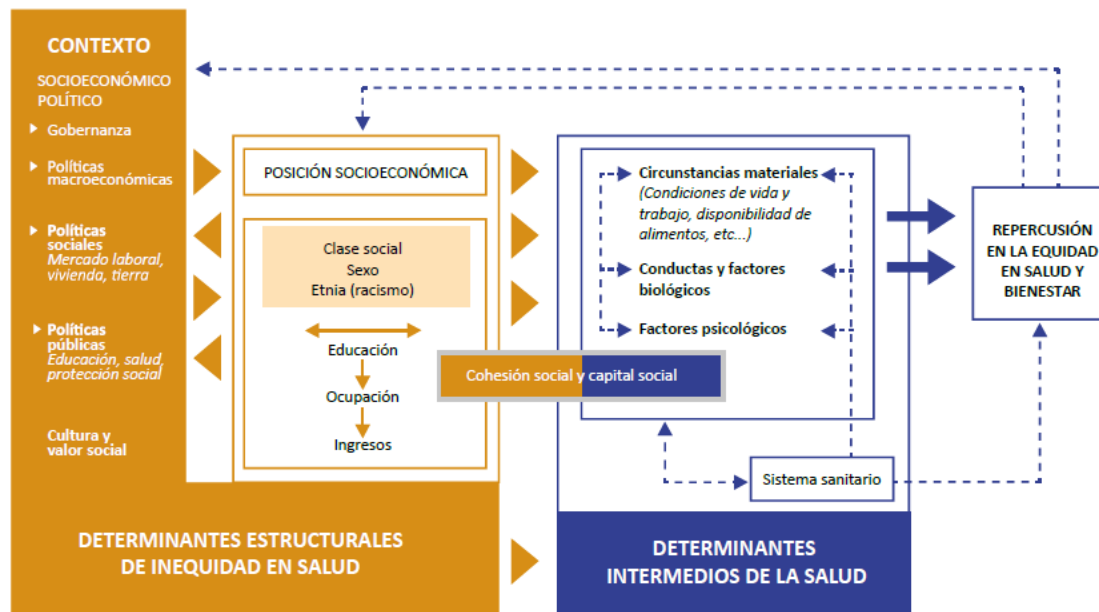
2.4.1 Determinantes sociales de la salud

Como se describe en el documento “Profundización del análisis de situación de las niñas, niños y adolescentes; en el ejercicio del derecho a la salud y el derecho a la alimentación y nutrición” (Peñuela et al., 2014), los determinantes sociales de la salud son las circunstancias en las que las personas nacen, crecen, viven, trabajan y envejecen, incluido el sistema de salud; es decir, las características sociales donde el curso de la vida tiene lugar y que apuntan a las características específicas del contexto social que influyen en la salud, las vías por las cuales las condiciones sociales se traducen en efectos sanitarios y los diferentes recursos a nivel mundial, nacional y local, siendo estas dependientes a su vez de las políticas adoptadas.

La OMS agrega que esas circunstancias son el resultado de la distribución del dinero, el poder y los recursos a nivel mundial, nacional y local, que depende a su vez de las políticas adoptadas. Los determinantes sociales de la salud explican la mayor parte de las inequidades sanitarias, esto es, de las diferencias injustas y evitables observadas en y entre los países en lo que respecta a la situación sanitaria (Cabezas, 2014).

La Comisión de los Determinantes Sociales de Salud describe unos determinantes estructurales y otros intermedios, en torno a un contexto socioeconómico y político (Hurtado, 2011), además plantea que los determinantes estructurales son aquellos que generan la estratificación social. Éstos incluyen los factores tradicionales de ingreso y educación, aunque actualmente es vital reconocer el género, la pertenencia étnica, la sexualidad y los recursos de capital social como los estratificadores sociales. Por otra parte, los determinantes intermedios fluyen de la configuración de estar bajo una estratificación social y, a su vez, determinan las diferencias en la exposición y vulnerabilidad de la salud que compone las condiciones. Se incluye en la categoría de determinantes intermedios: las condiciones de vida, las condiciones de trabajo, la disponibilidad de alimentos, comportamiento de la población y barreras para adoptar estilos de vida saludables y los servicios de salud, en la Figura 2-2 se representa el modelo de Determinantes de la OMS (Sexto informe ONS-Observatorio Nacional de salud, 2015).

Figura 2-2: Modelo de Determinantes Sociales de la Salud de la OMS



Fuente: Sexto informe ONS-Observatorio Nacional de Salud, 2015.

Los determinantes sociales de la salud de cada país tienen una caracterización muy particular, dependiendo del sistema social que se vive, de la ideología que impera y de la cultura dominante. Dentro de los determinantes más comunes se alistan: las redes de apoyo social, empleo y condiciones de trabajo y los estilos de vida. Las políticas sociales y económicas que existen en y entre los países influyen en el estado de salud de la población y determinan la desigualdad de las condiciones de vida, no siendo éstas resultado de la libre elección de los individuos. La voluntad política y la cooperación entre sectores a diferentes niveles son imprescindibles en la producción social de salud (González et al., 2012).

2.4.2 Determinantes sociales de la Seguridad Alimentaria y Nutricional

Son las características específicas y las vías mediante las cuales las condiciones sociales afectan la Seguridad Alimentaria y Nutricional - SAN de las poblaciones. Condiciones sociales en las cuales las personas viven y trabajan y que definen su posición diferencial en las jerarquías de poder, prestigio y recursos (Jiménez, 2016).

El enfoque de determinantes de la SAN establece que el ciclo de producción, consumo y utilización biológica de los alimentos, no es la principal fuerza que determina la buena o mala condición nutricional de las personas, sino que es un determinante más; siendo también importantes los factores ambientales y socioculturales que permiten a los grupos poblacionales y a las personas mejorar o mantener su Seguridad Alimentaria y Nutricional (Jiménez, 2016).

Este enfoque resalta lo social, intenta explicar cómo diferentes condiciones de disponibilidad de alimentos, acceso y consumo de ellos, estado de salud, ingresos económicos, trabajo y oportunidades sociales, marcan una diferencia en la condición nutricional de los individuos y la población.

2.4.3 Determinantes propuestos

Pensar en un enfoque de determinantes no implica realizar el análisis con todos o solo con los determinantes estructurales, intermedios o los de la SAN mencionados, se debe tener en cuenta el aporte de diferentes fuentes de información, como encuestas y registros administrativos que puedan aportar al análisis.

Partiendo de lo anterior y de los enfoques mencionados, se realiza una aproximación empírica, en la cual se establece el siguiente abordaje de los determinantes desde aspectos sociodemográficos, socioeconómicos, ambientales, de prestación de servicios de salud, individuales y política y protección social, que busca definir un conjunto de variables que puedan dar respuesta a los aspectos señalados.

Determinantes sociodemográficos

- Edad
- Edad de la madre
- Sexo
- Raza
- Estado civil
- Estrato socio-económico
- Área de residencia urbana/Rural
- Tamaño y estructura de la familia
- Educación
- Estado civil de la madre
- Nivel educativo de la madre

Determinantes socio-económicos

- Ocupación
- Ingreso
- Pobreza
- Desempleo
- Línea de indigencia
- Coeficiente de Gini
- Producto Interno Bruto (PIB) per cápita

Determinantes de tipo ambiental

- Tabaquismo
- Contaminación
- Condiciones de la vivienda
- Pavimentación de las vías de transporte
- Condiciones de los servicios públicos

Determinantes de prestación de servicios de salud

- Oferta de los servicios
- Afiliación en salud o seguridad social
- Controles prenatales

Determinantes individuales

- Peso al nacer (<2500g) reportado en certificado de defunción
- Peso al nacer (<2500g) según certificado de nacimiento
- Semanas de gestación
- Número de personas atendidas por ciertas enfermedades infecciosas y parasitarias
- Número de personas atendidas por deficiencias nutricionales
- Tasa mortalidad por Enfermedad Diarreica Aguda EDA
- Tasa de mortalidad por Infecciones Respiratorias Agudas IRA
- Comportamiento y estilos de vida

Determinantes de política y protección social

- Número de personas vinculadas a programas de asistencia social
- Coberturas de vacunación

Es de aclarar que las variables propuestas relacionadas con los diferentes determinantes dependen de la disponibilidad de la información, para las unidades de análisis (departamentos) y el período de tiempo establecido.

2.5 Estadísticas vitales

Según el documento metodológico de estadísticas vitales del DANE, las estadísticas vitales – EEVV, son un proceso que acopia información de la frecuencia de los hechos vitales (nacimientos y defunciones) y sus características para luego compilar, analizar, evaluar, presentar y difundir esos datos (DANE, 2012).

La información de Estadísticas Vitales está en un subsistema que tiene el propósito de recopilar la información de los nacimientos y defunciones ocurridos en todo el territorio nacional, el DANE con esta información produce y difunde la información estadística oficial de nacimientos y defunciones fetales y no fetales.

La información de defunciones producida por el DANE con fines estadísticos es captada desde su creación en 1953 y partir del año 1998 se estableció en el país el Sistema de Registro Civil y Estadísticas Vitales, a través de los certificados de nacido vivo y defunción; los cuales, a su vez, están constituidos por dos partes: un certificado antecedente y uno estadístico (DANE, 2012).

En el año 2005, se inició, de manera conjunta entre DANE y Ministerio de la Protección Social, hoy Ministerio de Salud y Protección Social, el proceso de modernización tecnológica del Subsistema de Estadísticas Vitales, mediante la implementación de una plataforma electrónica, perteneciente al Sistema Integral de Información de la Protección Social (SISPRO), a través del módulo de nacimientos y defunciones del Registro Único de Afiliados (RUAF), para captar de manera sincrónica los eventos vitales, capturando en línea la información de los nacimientos y defunciones.

2.5.1 Certificado de defunción

El certificado de defunción es la fuente fundamental de información de la mortalidad, y es el registro que acredita la muerte de una persona. Para el caso del certificado en medio electrónico RUAF, existe un certificado antecedente para el registro civil en medio físico con numeración impresa que se entrega a la familia y del cual se toma el número para asignárselo al certificado de defunción que se diligencia electrónicamente y cuyo formato es exactamente igual al certificado en medio físico (DANE, 2011).

El certificado además de servir de instrumento legal, también se utiliza como un instrumento de salud pública; la información estadística obtenida del certificado de defunción posibilita identificar casos de enfermedades objeto de vigilancia epidemiológica, para medir la magnitud de problemas en salud, evaluar programas y para estudios de factores de riesgo. La información de defunciones también es utilizada como insumo demográfico; la información estadística obtenida del certificado de defunción permite conocer el volumen de muertes y categorizarlas por edad, sexo y demás variables que caracterizan la dinámica de la población (DANE, 2011).

Los certificados de nacido vivo y de defunción son diligenciados por personal médico, enfermeras o personal de salud autorizado que atienda el hecho vital; cuando la causa de la defunción es un evento externo y/o violento o existan dudas en cuanto al diagnóstico inicial que generó la muerte, los certificados debe ser diligenciados por los médicos del Instituto Nacional de Medicina Legal y Ciencias Forenses (INML-CF). Para los casos en que el hecho vital ha sucedido sin tener contacto con el sector salud, los formatos de los certificados deberán ser diligenciados por los funcionarios de las notarías y registradurías (DANE, 2012).

El certificado de defunción se divide en seis partes:

- La primera (mediciones 2 a 24) corresponde a la información general de la defunción.
- La segunda parte (mediciones 25 a 36) incluye aspectos relacionados con las defunciones fetales o de menores de un año.
- La tercera parte (mediciones 37 a 39) contempla aspectos relacionados con las defunciones de mujeres en edad fértil, con el fin de captar la mortalidad materna durante el embarazo, el parto, el puerperio y mortalidad materna tarda.
- La cuarta parte (mediciones 40 a 42) es de uso específico para los casos de muertes violentas.
- La quinta parte (mediciones 43 a 46) está destinada a consignar las causas de la defunción o en los casos diligenciados por personal no médico (personal de salud autorizada y funcionarios de Registro Civil), la causa probable de la defunción.
- La sexta parte (mediciones 47 a 53) contiene los datos de quien certifica la defunción.

2.5.2 Clasificación internacional de enfermedades

La Clasificación estadística Internacional de Enfermedades y problemas relacionados con la salud, décima revisión (CIE-10), puede definirse como un sistema de categorías mutuamente excluyentes a las cuales se asignan enfermedades, lesiones y motivos de consulta de acuerdo con criterios previamente establecidos. La clasificación abarca todo el rango de enfermedades existentes en la terminología médica (Gómez, 2015).

Constituye uno de los estándares internacionales más usados para elaborar estadísticas de morbilidad y mortalidad en el mundo. El propósito de la CIE es permitir el registro sistemático, el análisis, la interpretación y la comparación de los datos de mortalidad y morbilidad recolectados en diferentes países o áreas, y en diferentes momentos. La clasificación permite la conversión de los términos diagnósticos y de otros problemas de salud, de palabras a códigos alfanuméricos que facilitan su almacenamiento y posterior recuperación para el análisis de la información (Arango, Cabrera, & De Mendoza, 2015).

En la actualidad se utiliza internacionalmente la CIE-10, para clasificar enfermedades y problemas relacionados y, en este sentido, tomar decisiones administrativas y epidemiológicas a partir de los datos diligenciados por los profesionales de salud, en 117 países de todo el mundo en registros clínicos y certificados de defunción (Gómez, 2015).

La CIE-10, es el instrumento utilizado por el codificador para asignar el código correspondiente a las causas de defunción anotadas por el médico para posteriormente seleccionar y codificar la causa básica de defunción (DANE, 2011).

Existe además una lista de agrupación de las causas de defunción elaborada por la Organización Panamericana de la Salud OPS, denominada lista 667.

2.6 Modelos estadísticos

El modelamiento estadístico ofrece la posibilidad de encontrar metodologías de análisis en la medida en que la construcción rigurosa de un modelo adecuado para el problema conducirá a una mejor comprensión del mismo (Torres, 2012).

Ortiz y Montenegro (2005) mencionan que el propósito central de la estadística es la representación simplificada de la realidad, con el propósito de hacer descripciones parciales y predicciones. Es así como el estadístico, haciendo uso de sofisticadas herramientas matemáticas, construye modelos matemáticos de distinta naturaleza dependiendo del tipo de problema estudiado y de la forma como las mediciones pueden o deben ser hechas sobre los objetos estudiados (Ortiz & Montenegro, 2005).

Desde un punto de vista muy general y moderno, el modelamiento puede ser clasificado en dos grandes áreas: modelamiento supervisado y modelamiento no supervisado. Cuando el propósito es usar las entradas para predecir los valores de la salida se habla de modelamiento supervisado. En el modelamiento no supervisado las salidas se auto organizan según algunos criterios a partir de las entradas. Este es el caso de del descubrimiento de patrones latentes en los datos, área que ha cobrado mucho auge en los últimos años (Ortiz & Montenegro, 2005).

Un modelo estadístico para datos observados en n unidades de análisis ($i = 1, \dots, n$) a lo largo de un conjunto de periodos determinados ($t = 1, \dots, n_i$), con k variables explicativas, tiene la siguiente forma (Castro, 2012):

$$y_{it} = \alpha_{it} + X_{it}\beta_{it} + u_{it} \quad (2-1)$$

Donde y_{it} es el valor de la variable respuesta, X_{it} es un vector de dimensión $(1 \times k)$ que contiene los valores de las covariables, β_{it} es el vector $(k \times 1)$ de parámetros del modelo y u_{it} el error. Los subíndices i y t asociados a los parámetros pueden ser incluidos o no dependiendo de los objetivos del modelo (Castro, 2012).

A continuación se describen algunos aspectos teóricos de los modelos que podrían resultar adecuados para el tema de estudio.

2.6.1 Regresión lineal

El análisis de regresión es útil para averiguar la forma probable de las relaciones entre variables, y el objetivo final, cuando se emplea es predecir o estimar el valor de una variable que corresponde al valor de otra u otras variables (Wayne, 2002).

Una regresión lineal queda representada como aparece en la siguiente ecuación:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \varepsilon_i \quad (2-2)$$

Donde:

Y_i la variable dependiente o respuesta.

$\beta_0, \beta_1, \dots, \beta_k$ son los coeficientes de la regresión.

X_1, X_2, \dots, X_k las variables independientes.

ε_i es un término de error aleatorio

$i = 1, \dots, n$ (registros) $k = 1, \dots, p$ (covariables)

Matricialmente, el modelo 2-2 se escribe como⁴:

$$Y = X\beta + \xi \quad (2-3)$$

Donde:

Y es el vector de respuesta de orden $n \times 1$.

X de orden $n \times (p + 1)$ matriz estocástica, es decir de valores fijos.

β de orden $(p + 1) \times 1$ vector de $p + 1$ parámetros: constantes a estimar.

ξ es el vector de errores aleatorios no observados de orden $n \times 1$.

Condiciones

Media del error es nula: $E(\mu_i) = 0 \quad i = 1 \dots n$

Varianza constante: $Var(\mu_1) = \vartheta^2 \quad i = 1 \dots n$

Covarianza nula: $Cov(\mu_1 \mu_j) = 0 \quad i \neq j$

2.6.2 Modelos lineales generalizados

Esta sección sigue de cerca el texto de Montgomery 2006. El modelo lineal generalizado MLG (GLM por sus siglas en inglés, *Generalized Linear Model*), es una unificación de los modelos de regresión lineal y no lineal, que también permite incorporar distribuciones de respuesta no normales. En un MLG la distribución de la variable de respuesta solo necesita ser un miembro de la familia exponencial, que comprende las distribuciones

⁴ Notas de clase, curso "Modelación y Regresión Aplicada". Luis Fernando Grajales PhD. Universidad Nacional de Colombia, Sede Bogotá. 2015.

Normal, Poisson, Binomial, Exponencial y Gamma, entre sus miembros, además, el modelo lineal con error normal no es más que un caso especial del modelo lineal generalizado, por lo que en muchos aspectos se puede considerar que el lineal generalizado es un método unificador de muchos aspectos del modelado y análisis empírico de datos (Montgomery, Peck, & Vining, 2006). Las distribuciones que son miembros de la familia exponencial tiene la forma general:

$$f(y_i, \theta_i, \phi) = \exp \{ [y_i \theta_i - b(\theta_i)] / a(\phi) + h(y_i, \phi) \} \quad (2-4)$$

Donde ϕ es un parámetro de escala, y θ_i se llama parámetro natural de localización.

Componentes de un modelo lineal generalizado

Componente aleatoria: identifica la variable respuesta Y y su distribución de probabilidad.

$$\exp \left\{ \frac{y^\theta - b(\theta)}{a(\theta)} + c(y, \phi) \right\} \quad (2-5)$$

Donde θ es el parámetro canónico, ϕ es un parámetro de perturbación y las funciones $a()$, $b()$ y $c()$ son conocidas. Además, se cumple que

$$\mu_i = E(Y) = b'(\theta) \text{ y } Var(Y) = a(\phi)b''(\theta) \quad (2-6)$$

Componente sistemática: especifica las variables explicativas utilizadas en la función predictora lineal. Es decir el vector de covariables $X' = (x_1, x_2, \dots, x_p)$

$$n_i = \sum_{j=1}^p x_j \beta_j = X' \quad (2-7)$$

Siendo β el vector a estimar.

Función de enlace: relaciona las dos componentes anteriores (μ_i y n_i). Una función $g(.)$ diferenciable y monótona tal que:

$$E(Y_i) = \mu_i = g^{-1}(n_i) \quad (2-8)$$

Para el análisis de la mortalidad algunos estudios han utilizado modelos log-lineales que asumen una distribución de Poisson. Es decir, modelos lineales generalizados con distribución de Poisson (Frome, 1983) (Puig, Ginebra, & Gispert, 2005).

2.6.3 Estimación de parámetros en el modelo lineal general

El método de máxima verosimilitud es la base teórica de la estimación de parámetros en el modelo lineal general, sin embargo, la implementación real de la máxima verosimilitud da como resultado un algoritmo basado en los mínimos cuadrados iterativamente reponderados (IRLS)⁵.

Casi siempre, los estimadores de máxima verosimilitud tienen mejores propiedades estadísticas que los estimadores por mínimos cuadrados. Los de máxima verosimilitud son insesgados y tienen varianza mínima comparados con otros estimadores insesgados. También son estimadores consistentes y un conjunto de estadísticos suficientes. Por otro lado, la estimación por máxima verosimilitud requiere hipótesis estadísticas más estrictas que las de los estimadores de mínimos cuadrados. Los estimadores de mínimos cuadrados requieren solo hipótesis de segundo momento (que son hipótesis sobre el valor esperado, las varianzas y las covarianzas entre los errores aleatorios). Los mínimos cuadrados ponderados también son una forma potencialmente útil para manejar el problema de varianza no constante (Montgomery et al., 2006).

2.6.4 Modelos para datos de conteo

Son aquellos que tienen como variable dependiente una variable discreta de conteo que toma valores no negativos. En este tipo de modelos en general se identifican pocos valores distintos y abundancia de ceros (Hachuel et al., 2010).

Variable de interés:

$$Y \in \{0, 1, 2, \dots\} \quad (2-9)$$

El objetivo es analizar Y como función de variables explicativas

⁵ Para más detalles ver Montgomery, Douglas C. 2006.

$$E(Y|X_1, X_2, \dots, X_k) \quad (2-10)$$

$$E(Y|X_1, X_2, \dots, X_k) = \beta_0 + \beta_1 X_1 + \dots + \beta_K X_K \quad (2-11)$$

Para datos de conteo se suele utilizar la distribución Poisson como componente aleatorio en el proceso de ajuste de un modelo lineal generalizado. El fenómeno por el cual un modelo lineal generalizado tiene mayor variabilidad que la presupuesta por el componente aleatorio del mismo se denomina sobredispersión. Los métodos para tratar la sobredispersión se basan en modelos que complementan los métodos más convencionales que se concentran solamente en modelar correctamente la relación media-variancia (Hachuel et al., 2010).

Los datos Poisson se dicen sobredispersos si la variancia excede la media, por lo que un indicador simple de la magnitud de la sobredispersión la da la comparación entre la media muestral y la variancia de la variable de conteo en estudio. Si bien el modelo Poisson es el paradigma o el modelo básico para respuestas tipo conteo, la regresión Binomial Negativa es casi siempre pensada como el modelo alternativo al Poisson cuando hay sobredispersión en los datos (Hachuel et al., 2010).

2.6.4.1 Regresión de Poisson

En este escenario de modelado, la variable respuesta de interés no tiene distribución normal. En este caso, la variable respuesta representa un conteo de algún elemento relativamente raro. Se desea modelar la relación entre los conteos observados y variables regresoras o predictoras potencialmente útiles (Montgomery et al., 2006).

Se supondrá que y_i , la variable de respuesta, es un conteo, como por ejemplo la observación $y_i = 0, 1, \dots$. Un modelo probabilístico razonable para los datos de conteo es, con frecuencia, la distribución de Poisson:

$$f(y) = \frac{e^{-\mu} \mu^y}{y!}, \quad y = 0, 1, \dots \quad (2-12)$$

Donde el parámetro $\mu > 0$. La distribución de Poisson es otro ejemplo de distribuciones de probabilidad en donde se relaciona el promedio y la variancia.

El modelo de regresión de Poisson se puede escribir en la forma

$$y_i = E(y_i) + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (2-13)$$

Una de las principales razones por las que el modelo Poisson falla es la heterogeneidad no observada. Esto significa que hay factores no observados, en especial características de los individuos, que ejercen alguna influencia sobre la variabilidad relacionada con la variable de respuesta (Salinas Rodríguez et al., 2009).

2.6.4.2 Regresión Binomial Negativa

La heterogeneidad que se ignora en el modelo Poisson, puede modelarse de manera explícita mediante el uso de la regresión binomial negativa. De esta forma la distribución Binomial Negativa es una alternativa de modelamiento de conteos cuando la varianza de los datos es mayor que su media; es decir cuando hay problemas de sobredispersión en relación al modelo Poisson (Berdugo, 2010).

Este procedimiento está diseñado para ajustar un modelo de regresión en el cual la variable dependiente Y consiste de conteos. El modelo de regresión ajustado relaciona Y con una o más variables predictoras X , que pueden ser cuantitativas o categóricas. El procedimiento ajusta un modelo usando máxima verosimilitud o mínimos cuadrados ponderados.

La densidad de la distribución binomial negativa es

$$f(y; K, \mu) = \frac{\Gamma(y + k)}{\Gamma(k)\Gamma(y + 1)} \left(\frac{k}{\mu + k}\right)^k \left(1 - \frac{k}{\mu + k}\right)^y \quad (2-14)$$

Con $y = 0, 1, 2, \dots$ donde k y μ son los parámetros.

Se tiene que

$$E(Y) = \mu, \quad var(Y) = \mu + \mu^2 / k. \quad (2-15)$$

2.6.5 Sobredispersión

Es un fenómeno que sucede a veces cuando se modelan datos de respuesta con una distribución binomial P de Poisson, básicamente quiere decir que la varianza de la respuesta es mayor de la que se podría esperar para esa elección de distribución estadística de la respuesta (Montgomery et al., 2006).

2.7 Datos Panel

Las definiciones y aspectos teóricos presentados a continuación siguen de cerca los textos de Wooldridge (2002; 2009), Baronio & Vianco (2010) y Labra & Torrecillas (2014).

Los datos en panel son un conjunto de datos que combina una dimensión temporal (serie de tiempo) y otra transversal (individuos). Permiten trabajar simultáneamente varios periodos de tiempo y los efectos individuales, y a su vez, tratar el problema de la endogeneidad. Esta técnica tiene en cuenta los efectos fijos de los individuos que pueden ocasionar comportamientos no aleatorios de las variables, y las series de tiempo cuyos datos tienen su propia dinámica que debe ser estudiada (Labra & Torrecillas, 2014).

Labra & Torrecillas (2014), refiere además que existen dos tipos de análisis con datos de panel, los Estáticos y Dinámicos. En el caso de los estáticos tienen una aplicación sencilla a través de software y evalúan un conjunto de variables como explicativas de algún fenómeno en estudio y determinar así, si el conjunto de datos presenta efectos individuales fijos o variables; sin embargo, esta metodología tiene deficiencias, como el tratamiento adecuado de la endogeneidad. En cuanto al análisis dinámico permiten incorporar en el modelo una estructura endógena, mediante la integración de efectos pasados a través de variables instrumentales (Labra & Torrecillas, 2014).

La principal diferencia entre paneles dinámicos y estáticos está en la capacidad y forma de tratar la endogeneidad de las variables.

El desarrollo de técnicas de datos de panel, puede tratar en forma independiente el conjunto de datos de un individuo, en este caso departamento, en el tiempo, lo que se conoce como efectos individuales (α_i).

El principal objetivo de los datos en panel, es capturar la heterogeneidad no observable, entre agentes de estudio y el tiempo, dado que esta heterogeneidad no se puede detectar ni con estudios de series temporales ni tampoco con los de corte transversal (Mayorga & Muñoz, 2000). Esta técnica permite un análisis más dinámico al incorporar la dimensión temporal de los datos, lo que enriquece el estudio.

La aplicación de esta metodología permite analizar dos aspectos de suma importancia cuando se trabaja con este tipo de información y que forman parte de la heterogeneidad no observable: i) los efectos individuales específicos y ii) los efectos temporales.

Los efectos individuales son los que afectan de manera desigual a cada unidad de estudio (departamentos), los cuales son invariables en el tiempo; los efectos temporales son los que afectan por igual a todas las unidades de estudio (Baronio & Vianco, 2010).

El modelo general de datos panel está definido:

$$Y_{it} = \alpha_{it} + b_1X_{1it} + b_2X_{2it} + \dots + b_kX_{kit} + U_{it} \quad (2-16)$$

con $i = 1, \dots, n$ y $t = 1, \dots, T$

Donde i se refiere a la unidad de estudio y t a la dimensión en el tiempo.

α es un vector de interceptos que puede contener entre 1 y $n + t$ parámetros.

b es un vector de K parámetros.

X_{1it} es la i -ésima observación al momento t para las K variables explicativas X_1, X_2, \dots, X_k
 α_{it} recoge la heterogeneidad provocada por los efectos de los individuos y/o tiempo y por variables no observables.

La muestra total de las observaciones en el modelo viene dado por $n \times T$.

Los modelos de datos de panel se pueden interpretar a través de sus componentes de errores. El término de error u_{it} puede expresarse:

$$u_{it} = \mu_i + \delta_t + \varepsilon_{it} \quad (2-17)$$

Donde,

μ_i representa los efectos no observables que difieren entre las unidades de estudio pero no en el tiempo.

δ_t hace referencia a los efectos no cuantificables que varían en el tiempo pero no entre las unidades de estudio.

ε_{it} es el término del error puramente aleatorio.

Con base en este modelo general, y en ciertos supuestos y restricciones se derivan algunas variantes de modelos de datos de panel. Dependiendo de la consideración que se le dé al término independiente se distinguen tres enfoques (Salmerón, 2012) que se mencionan a continuación pero describirán con detalle más adelante:

Modelo agrupado: Es constante para todos los individuos y en los periodos (es decir, $\alpha_{it} = \alpha$).

Modelo de efectos fijos: El término independiente puede ser distinto para cada individuo (es decir, $\alpha_{it} = \alpha_i$), cada período (es decir, $\alpha_{it} = \alpha_t$) o ambos.

Efectos variables o aleatorios: El término independiente, α_{it} , es una variable aleatoria.

2.7.1 Aspectos conceptuales

A continuación se realiza una descripción de los conceptos estadísticos relacionados con los modelos de panel, que serán de utilidad para el desarrollo metodológico.

2.7.1.1 Heterocedasticidad

La heterocedasticidad hace referencia a que la varianza de los errores de la muestra no es constante, de esta forma se entiende por homocedasticidad que la varianza de los errores es la misma para toda la muestra. Una forma de saber si nuestra estimación tiene problemas de heterocedasticidad es a través de la prueba del Multiplicador de Lagrange de Breusch y Pagan. Sin embargo, de acuerdo con Greene, ésta y otras pruebas son sensibles al supuesto sobre la normalidad de los errores; afortunadamente, la prueba Modificada de Wald para Heterocedasticidad funciona aun cuando dicho supuesto es violado.

Supuesto de homocedasticidad: para cualquier valor de las variables explicativas, el error u tiene la misma varianza. En otras palabras,

$$\text{Var}(u|x_1, \dots, x_k) = \sigma^2 \quad (2-18)$$

Significa que la varianza en el término del error u_i condicional en las variables explicativas, es la misma para todas las combinaciones de valores de las variables explicativas. Si este supuesto no se satisface, entonces el modelo muestra heterocedasticidad.

2.7.1.2 Endogeneidad

Se refiere a la correlación entre la variable dependiente y el error, es una relación causal entre las variables, las cuales quedan explicadas dentro del modelo (Labra & Torrecillas, 2014).

Se pueden considerar dos tipos de endogeneidad:

1. Endogeneidad del modelo: Está relacionada con el efecto de la variable dependiente sobre sí misma, es decir el modelo está determinado por su pasado.
2. Endogeneidad de las variables independientes: se produce porque X_{it} está determinada por su condición pasada (X_{it-n}).

Puede ocurrir que exista relación entre las variables independientes, lo que se identifica mediante un coeficiente de correlación alto entre ellas y representaría multicolinealidad.

2.7.1.3 Multicolinealidad

El concepto de colinealidad se refiere a las relaciones que existen entre las covariables, y no entre éstas y la variable dependiente (Celis de la Rosa, 2008). Si existe algún tipo de relación entre las covariables o cuando hay dependencias casi lineales entre las covariables, se dice que existe el problema de multicolinealidad. La presencia de multicolinealidad tiene una gran cantidad de efectos graves sobre los estimados de coeficientes de regresión por mínimos cuadrados (Montgomery et al., 2006).

2.7.1.4 Autocorrelación

El problema de la correlación serial o “autocorrelación” se da cuando los errores e_{it} no son independientes con respecto al tiempo. Aunque en lo descrito anteriormente se

intenta modelar la heterogeneidad temporal y espacial en un modelo, las ecuaciones pueden estar mal especificada en otros aspectos; De acuerdo con los supuestos de Gauss-Markov, los estimadores de OLS son los Mejores Estimadores Lineales Insesgados (MELI) siempre y cuando los errores e_{it} sean independientes entre si y se distribuyan idénticamente con varianza constante σ^2 .

Desafortunadamente, con frecuencia estas condiciones son violadas en datos panel: la independencia se viola cuando los errores de diferentes unidades están correlacionados (correlación contemporánea), o cuando los errores dentro de cada unidad se correlacionan temporalmente (correlación serial), o ambos. A su vez, la distribución “idéntica” de los errores es violada cuando la varianza no es constante (heteroscedasticidad) (Aparicio & Márquez, 2005).

Existen muchas maneras de diagnosticar problemas de autocorrelación. Sin embargo, cada una de estas pruebas funciona bajo ciertos supuestos sobre la naturaleza de los efectos individuales. Wooldridge desarrolló una prueba muy flexible basada en supuestos mínimos.

2.7.1.5 Correlación Contemporánea

Resulta conveniente analizar otro problema que surge de la estimación con datos tipo panel. Las estimaciones en datos panel pueden tener problemas de correlación contemporánea si las observaciones de ciertas unidades están correlacionadas con las observaciones de otras unidades en el mismo período de tiempo.

El problema de correlación contemporánea se refiere a la correlación de los errores de al menos dos o más unidades en el mismo tiempo t . En otras palabras, tenemos errores contemporáneamente correlacionados si existen características inobservables de ciertas unidades que se relacionan con las características inobservables de otras unidades. Por ejemplo, los errores de dos territorios pueden relacionarse pero mantenerse independientes de los errores de los demás territorios. La prueba de Breusch y Pagan identifica problemas de correlación contemporánea en los residuales de un modelo de efectos fijos.

2.7.1.6 Variables exógenas, endógenas, predeterminadas e instrumentales

Esta clasificación de las variables se refiere por que puede ser de utilidad para procedimientos que se realizan con datos de panel.

Variables exógenas: son las que se encuentran fuera del modelo, no tienen relación con el resto de independientes y por esta razón no existe correlación entre los errores de la variable y los del modelo.

$$Cov(x_1, \varepsilon) = 0 \quad (2-19)$$

Siendo:

x_1 : la variable exógena, ε : el error del modelo

Variables endógenas: son las que están determinadas dentro del modelo, es decir existe causalidad en ambos sentidos ($X_i \rightarrow Y_i$; $Y_i \rightarrow X_i$).

$$Cov(x_1, \varepsilon) \neq 0 \quad (2-20)$$

Siendo:

x_1 : la variable endógena, ε : el error del modelo

Variable predeterminada: son variables fuera del modelo y que se determinan con anterioridad al actual momento. El valor futuro de la variable puede estar correlacionado con el término de error del modelo pero no con su retardo. Este tipo de variables están relacionadas con la variable independiente.

$$Cov(x_{1s}, \varepsilon_t) \neq 0 \text{ si } S < T \quad (2-21)$$

Siendo:

x_1 : la variable endógena, ε : los errores del modelo, $s - t$: distintos períodos de tiempo.

Variables instrumentales: Son proxy de variables que no se pueden observar o utilizar directamente por diferentes problemas, como la falta de datos, mala calidad de la información, imposibilidad de cuantificarlas o series incompletas.

2.7.1.7 Paneles balanceados y desbalanceados

Los paneles pueden ser balanceados (*strongly balanced*) cuando el número de períodos es igual para todos los individuos o unidades de análisis o desbalanceados (*unbalanced panel data*) cuando el panel está incompleto, es decir no se encuentran completas toda la serie de datos para un individuo o unidad de análisis, o faltan unidades para ciertos años de una misma variable. En los datos panel balanceados ($T_i = T$ para todo i) y en los no balanceados ($T_i \neq T$ para algún i).

Cuando los *missing* (datos faltantes) son elevados, pueden surgir limitaciones en el análisis, generando inconsistencia en los resultados, o simplemente impidiendo realizar la estimación del modelo o función (Labra & Torrecillas, 2014).

2.7.1.8 Errores Estándar Corregidos para Panel (PCSE⁶)

Calcula las estimaciones de error estándar corregido por el panel para modelos de series de tiempo lineales donde los parámetros son estimados por regresión OLS o Prais-Winsten. Cuando se calculan los errores estándar y las estimaciones de varianza-covarianza, el método asume que las perturbaciones son, por defecto, heterocedásticas y correlacionadas simultáneamente entre paneles.

Es una alternativa a los Mínimos Cuadrados Generalizados Factibles (*Feasible Generalized Least Squares* ó FGLS) para ajustar modelos de series de tiempo lineales cuando no se supone que las perturbaciones son independientes e idénticamente distribuidas (i.i.d.). En su lugar, se supone que las alteraciones son heterocedásticas a través de paneles o heterocedásticas y correlacionadas simultáneamente entre paneles.

También se puede suponer que las perturbaciones están autocorrelacionadas dentro del panel y el parámetro de autocorrelación puede ser constante a través de paneles o diferente para cada panel.

Partiendo del modelo:

⁶ PCSE por sus siglas en inglés (Panel Corrected Standard Error).

$$y_{it} = x_{it}\beta + \epsilon_{it} \quad (2-22)$$

Donde $i = 1, \dots, m$ es el número de unidades (o paneles); $t = 1, \dots, T_i$; T_i es el número de períodos en el panel i ; y ϵ_{it} es una perturbación que puede estar autocorrelacionada a lo largo de t o correlacionada simultáneamente a través de i .

Este modelo también se puede escribir panel por panel como:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \beta + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_m \end{bmatrix} \quad (2-23)$$

Para un modelo con perturbaciones heterocedásticas y correlación contemporánea pero sin autocorrelación, se supone que la matriz de covarianza de perturbación es:

$$E[\epsilon\epsilon'] = \Omega = \begin{bmatrix} \sigma_{11}I_{11} & \sigma_{12}I_{12} & \cdots & \sigma_{1m}I_{1m} \\ \sigma_{21}I_{21} & \sigma_{22}I_{22} & \cdots & \sigma_{2m}I_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{m1}I_{m1} & \sigma_{m2}I_{m2} & \cdots & \sigma_{mm}I_{mm} \end{bmatrix} \quad (2-24)$$

Donde σ_{ii} es la varianza de las perturbaciones para el panel i , σ_{ij} es la covarianza de las perturbaciones entre el panel i y el panel j cuando los períodos de los paneles se corresponden, y I es una matriz de identidad $T_i \times T_i$ con paneles balanceados. Los paneles no necesitan ser balanceados, pero la expresión para la covarianza de los disturbios será más general si están desbalanceados.

Esto también podría escribirse como:

$$E[\epsilon\epsilon'] = \sum_{m \times m} \otimes I_{T_i \times T_i} \quad (2-25)$$

Dónde \sum es la matriz de covarianza panel por panel y I es una matriz de identidad.

Partiendo de lo anterior PCSE produce estimaciones OLS de los parámetros cuando no se especifica ninguna autocorrelación, o estimaciones Prais-Winsten cuando se

especifica la autocorrelación. Si se especifica la autocorrelación, las estimaciones de los parámetros están condicionadas a las estimaciones de los parámetros de autocorrelación.

2.7.1.9 Regresión Prais-Winsten

Prais utiliza el método de mínimos cuadrados generalizado para estimar los parámetros en un modelo de regresión lineal en el que los errores están correlacionados en serie. Específicamente, se supone que los errores siguen un proceso autorregresivo de primer orden.

El estimador de Prais puede usar cualquier estimación consistente de ρ (para el parámetro de autocorrelación), para transformar la ecuación, y cada una de estas estimaciones cumple con ese requisito.

Prais se ajusta a una regresión lineal de variables dependientes o independientes que se corrige para los residuos secuencialmente correlacionados de primer orden usando el estimador de regresión transformado Prais-Winsten.

El proceso de error autocorrelacionado más común es el proceso autorregresivo de primer orden. Bajo este supuesto, el modelo de regresión lineal se puede escribir como:

$$y_t = \mathbf{x}_t\boldsymbol{\beta} + u_t \quad (2-26)$$

Donde los errores satisfacen,

$$u_t = \rho u_{t-1} + e_t \quad (2-27)$$

Y los e_t son independientes e idénticamente distribuidos $N(0, \sigma^2)$.

El estimador de Prais-Winsten es un estimador generalizado de mínimos cuadrados (GLS). El método Prais-Winsten se deriva del modelo AR (1) para el término de error descrito anteriormente.

2.7.2 Modelos de datos panel estáticos

Se hará referencia a dos técnicas de análisis en el uso de panel de datos, ya que los efectos individuales (α_i) pueden ser tratados como efectos fijos y efectos aleatorios. Para llevar a cabo la estimación, se asume que los α_i son constantes a lo largo del tiempo.

Sin embargo antes de profundizar en los modelos según los efectos, se hará mención a la variación (ya que las variables pueden tener variación tanto en el tiempo como entre individuos) y al modelo agrupado.

2.7.2.1 Variabilidad *within* y *between*

Para la variación en el tiempo para un individuo dado, se tiene variabilidad *within* (s_w^2) y la variación entre individuos variabilidad *between* (s_B^2).

Variabilidad “*within*” (en torno a la media individual $\bar{x}_i = 1/T \sum_t x_{it}$).

$$s_w^2 = \frac{1}{NT - 1} \sum_i \sum_t (x_{it} - \bar{x}_i)^2 = \frac{1}{NT - 1} \sum_i \sum_t (x_{it} - \bar{x}_i + \bar{x})^2 \quad (2-28)$$

Variabilidad “*between*” (variación de \bar{x}_i en torno a \bar{x})

$$s_B^2 = \frac{1}{N - 1} \sum_i (\bar{x}_i - \bar{x})^2 \quad (2-29)$$

NT se entiende como el total de observaciones. Para paneles no balanceados debe ser:

$$\sum_{i=1}^N T_i \quad (2-30)$$

Las estadísticas pueden describir los datos “*within*”: $x_{it} - \bar{x}_i + \bar{x}$ y “*between*”: \bar{x}_i

2.7.2.2 Modelo agrupado

El vector de interceptos puede contener entre 1 y $n + t$ parámetros y ser constante para todos los individuos y períodos. De esta forma se tiene el modelo agrupado o *pooled* que será estimado por MCO:

$$Y_{it} = \alpha + b_1 X_{1it} + b_2 X_{2it} + \dots + b_k X_{kit} + U_{it} \quad (2-31)$$

El inconveniente de este modelo es que ignora la estructura de panel de los datos e incumple la hipótesis de no autocorrelación entre las perturbaciones. Además, se puede estar interesado en determinar si se estima un modelo agrupado o T modelos de series temporales. Si las perturbaciones cumplen las hipótesis básicas, se puede aplicar un test de Chow:

$$F = \frac{(\vec{e}'_r \vec{e}_r - \vec{e}' \vec{e}) / (N - 1)k}{\vec{e}' \vec{e} / N(T - k)} \sim F_{(N-1)k, N(T-k)} \quad (2-32)$$

Donde \vec{e}_r son los residuos del modelo agrupado y \vec{e}_r son los residuos de estimar los modelos por separado. Este test es similar a realizar un contraste mediante una F usando variables ficticias para la constante y la pendiente.

2.7.2.3 Modelo con efectos individuales

La estructura general se establece así:

$$y_{it} = \beta_1 x_{1it} + \dots + \beta_k x_{kit} + U_{it} \quad (2-33)$$

$$= \beta_1 x_{1it} + \dots + \beta_k x_{kit} + \alpha_i + \varepsilon_{it} \quad (2-34)$$

Donde,

x_{1it}, \dots, x_{kit} : variables explicativas (observables)

$U_{it} = \alpha_i + \varepsilon_{it}$: término de error compuesto (inobservado)

α_i : efectos individuales (heterogeneidad inobservada permanente en el tiempo)

ε_{it} : error idiosincrásico

Como se mencionó al inicio de esta sección existen dos modelos sustancialmente diferentes según el tratamiento de α_i :

1. Modelos de efectos fijos
2. Modelo de efectos aleatorios

2.7.2.4 Modelo de Efectos Fijos (Fixed Effects)

Una manera de modelar el carácter “individual” de cada estado es a través del modelo de efectos fijos. Este modelo no supone que las diferencias entre estados sean aleatorias, sino constantes o fijas y por ello debemos estimar cada intercepto u_i . Considera que existe un término constante diferente para cada individuo, y supone que los efectos individuales son independientes entre sí.

Con este modelo se considera que las variables explicativas afectan por igual a las unidades de corte transversal y que éstas se diferencian por características propias de cada una de ellas, medidas por medio del intercepto.

Modelo de efectos fijos individuales

En este modelo tenemos N términos independientes que recogen las diferencias entre los distintos individuos y que se conocen como efectos fijos individuales. La variación de los efectos fijos individuales proviene de las variables omitidas que varían entre los distintos individuos pero no en el tiempo.

$$y_{it} = \alpha_i + X_{it}\beta + u_{it} \quad (2-35)$$

Modelo de efectos fijos temporales

Un modelo con efectos fijos temporales es aquel que permite tener en cuenta las variables que son constantes entre individuos, pero evolucionan en el tiempo.

$$y_{it} = \alpha_t + X_{it}\beta + u_{it} \quad (2-36)$$

En este caso se tiene T interceptos que recogen los efectos fijos temporales que provienen de las variables omitidas que varían en el tiempo, pero no entre los individuos.

Modelo de efectos fijos individuales y temporales

Si las variables omitidas son constantes en el tiempo pero varían entre los individuos, mientras que otras son constantes para los individuos pero varían en el tiempo, entonces se habla de efectos fijos individuales y temporales:

$$y_{it} = \alpha_{it} + X_{it}\beta + u_{it} \quad (2-37)$$

Para tratar los efectos fijos se emplea el estimador intragrupos (*within*), el cual asume que el efecto individual está correlacionado con las variables explicativas. De esta forma:

$$\text{corr}(\alpha_i, X) \neq 0 \quad (2-38)$$

Considerando el caso del modelo de efectos fijos individuales, además del término independiente, se puede recoger mediante variables ficticias la variación entre los individuos y se supone que las pendientes permanecen constantes. En tal caso se estimaría el siguiente modelo:

$$y_{it} = \alpha_1 + \sum_{j=2}^N \alpha_j d_j + X_{it}\beta + u_{it} \quad (2-39)$$

Donde:

α_1 es el efecto del primer individuo, el cual se utiliza como categoría base.

d_j son variables binarias que toman el valor 1 si el dato corresponde al individuo j y cero en otro caso.

α_j son los coeficientes de las variables ficticias y representan el grado en que los valores de los interceptos del resto de individuos difieren respecto del intercepto base α_1 (por ejemplo, $\alpha_1 + \alpha_2$ representa el efecto individual del individuo 2).

En el caso de efectos fijos temporales la variable binaria d_j tomaría el valor 1 si el dato corresponde al tiempo j y cero en otro caso. Mientras que bajo ambos supuestos se tiene que:

$$y_{it} = \alpha_1 + \sum_{j=2}^N \alpha_j d_j + \delta_1 + \sum_{l=2}^T \delta_l d'_l + X_{it}\beta + u_{it} \quad (2-40)$$

Permite que los regresores x_{1it}, \dots, x_{kit} estén correlacionados con α_i . Todo el análisis será condicional en α_i .

El supuesto fundamental es:

$$E[\varepsilon_{it} | \alpha_i, x_{1it}, \dots, x_{kit}] = 0 \quad (2-41)$$

Los regresores deben seguir no correlacionados con ε_{it} esto implica:

$$E[y_{it} | \alpha_i, x_{1it}, \dots, x_{kit}] = \beta_1 x_{1it} + \dots + \beta_k x_{kit} + \alpha_i y$$

$$\frac{\delta E[y_{it} | \alpha_i, x_{1it}, \dots, x_{kit}]}{\delta x_{j,it}} = \beta_j \quad (2-42)$$

Se puede identificar el efecto marginal β_j aunque el regresor es endógeno, respecto al término de error compuesto U_{it} . Los regresores pueden estar correlacionados U_{it} , pero sólo con su parte constante en el tiempo.

Estimadores de efectos fijos

Sea un modelo de efectos individuales

$$y_{it} = \beta_1 x_{1it} + \dots + \beta_k x_{kit} + \alpha_i + \varepsilon_{it} \quad (2-43)$$

Se supone:

$$E[\varepsilon_{it} | \alpha_i, x_{1it}, \dots, x_{kit}] = 0 \quad (2-44)$$

La estimación de los parámetros β requiere la eliminación de α_i . Estos estimadores sólo utilizan variación “*within*” de los datos.

Son consistentes tanto si los regresores están correlacionados con la heterogeneidad permanente como si no.

2.7.2.5 Modelo de Efectos Aleatorios (Random Effects)

A diferencia del modelo de efectos fijos, el modelo de efectos aleatorios considera que los efectos individuales no son independientes entre sí, sino que están distribuidos aleatoriamente alrededor de un valor dado (Mayorga & Muñoz, 2000).

Para tratar los efectos aleatorios se emplea el Método Generalizado de Momentos (MGM), que es una extensión más eficiente de MCO. Este estimador asume la condición de que los efectos individuales no están correlacionados con las variables explicativas del modelo:

$$\text{corr}(\alpha_i, X) = 0 \quad (2-45)$$

Siendo:

α_i Los efectos individuales.

X Las variables explicativas.

El modelo de efectos aleatorios permite suponer que cada unidad transversal tiene un intercepto diferente. Este modelo se expresa como:

$$Y_{it} = \alpha_i + \beta_1 X_{1it} + e_{it} \quad (2-46)$$

Donde $\alpha_i = \alpha + u_i$. Lo que representa que no se considera a α como fija, se supone que es una variable aleatoria con un valor medio α y una desviación aleatoria u_i de este valor medio. Reemplazando α en (2-46) se tendría:

$$Y_{it} = \alpha_i + \beta_1 X_{1it} + u_i + e_{it} \quad (2-47)$$

Si se analiza esta última ecuación, se observa que si la varianza de u_i es igual a cero $\sigma_u^2 = 0$, entonces no existe diferencia relevante entre las ecuaciones (2-46) y (2-47).

Este estimador tiene la ventaja de que permite conocer los α_i separadamente, lo que contribuye a entender de mejor forma el modelo. Además, evita una sobrestimación del parámetro β , lo que ocurre cuando se aplica el estimador de efectos fijos (Labra & Torrecillas, 2014).

El efecto individual α_i se trata como puramente aleatorio, debe especificarse su distribución, condicional en los regresores.

Supuesto habitual: α_i no está correlacionado con los regresores

$$\alpha_i | X_{it} \sim N(0, \sigma_\alpha^2) \quad (2-48)$$

Estimadores de efectos aleatorios

Sea un modelo de efectos individuales

$$y_{it} = \beta_1 x_{1it} + \dots + \beta_k x_{kit} + \alpha_i + \varepsilon_{it} \quad (2-49)$$

Donde

$$\begin{aligned} E[\alpha | X_{it}] &= 0; & \text{Var}[\alpha_i | X_{it}] &= \sigma_\alpha^2 \\ E[\varepsilon_{it} | X_{it}] &= 0; & \text{Var}[\varepsilon_{it} | X_{it}] &= \sigma_\varepsilon^2 \end{aligned} \quad (2-50)$$

Esto implica que los regresores son exógenos respecto al término de error compuesto

$$U_{it} = \alpha_i + \varepsilon_{it} \quad (2-51)$$

$$E[U_{it} | X_{it}] = 0 \quad (2-52)$$

Además, se tiene una estructura de correlación particular

$$\text{Corr}(U_{it}, U_{is}) = \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_\varepsilon^2}, \quad t \neq s \quad (2-53)$$

Por tanto, se puede estimar eficientemente mediante el estimador de mínimos cuadrados generalizados factibles FGLS. El estimador de efectos aleatorios se obtiene estimando por MCO el modelo transformado.

$$(y_{it} - \hat{\theta}_i \bar{y}_i) = \alpha (1 - \hat{\theta}_i) + (X_{it} - \hat{\theta}_i \bar{X}_i)' \beta + \alpha_i (1 - \hat{\theta}_i) + (\varepsilon_{it} - \hat{\theta}_i \bar{\varepsilon}_i) \quad (2-54)$$

$\hat{\theta}_i$ es un estimador consistente de

$$\theta_i = 1 - \sqrt{\sigma_\varepsilon^2 / (T_i \sigma_\alpha^2 + \sigma_\varepsilon^2)} \quad (2-55)$$

El estimador de efectos aleatorios usa tanto variación “*within*” como “*between*”

2.7.2.6 Modelo de Efectos Fijos vs Modelo de Efectos aleatorios

Resulta muy importante conocer si el modelo adecuado para analizar nuestros datos es el de efectos fijos o el de efectos aleatorios. El modelo de efectos aleatorios supone que esta correlación es igual a cero. Si las u_i y las variables X están correlacionadas, no incluir u_i en el modelo producirá un sesgo de variable omitida en los coeficientes de X .

Hausman demostró que la diferencia entre los coeficientes de efectos fijos y aleatorios ($\beta_{ef} - \beta_{ea}$) puede ser usada para probar la hipótesis nula de que u_i y las variables X no están correlacionadas.

Para eso se realiza la prueba de Hausman para verificar que los estimadores de efectos aleatorios y de efectos fijos no difieren sustancialmente.

2.7.3 Test o pruebas recomendadas

A continuación, se presenta una serie de pruebas que son recomendables realizar para detectar si existe heterocedasticidad y autocorrelación en el modelo, para ver si el modelo de datos de panel estáticos se prefiere al pool de datos, o para incorporar variables dicotómicas de tiempo en el modelo.

2.7.3.1 Test de efectos fijos

Considerando el modelo de efectos fijos en forma matricial:

$$y = X\beta + Z\mu + \varepsilon \quad (2-56)$$

En donde Z es una matriz de $N - 1$ variables binarias por intercepto. Bajo el supuesto de normalidad de ε , la hipótesis nula: $H_0: \mu_1 = \mu_2 = \dots = \mu_{N-1} = 0$, ausencia de efectos fijos, pueden evaluarse con test F estándar de significancia conjunta de las $N - 1$ variables binarias.

2.7.3.2 Test de efectos aleatorios (Breusch-Pagan).

Para contrastar la hipótesis de que no hay efectos aleatorios se puede usar la prueba de Breusch-Pagan, que es un test de multiplicadores de Lagrange. Las pruebas de Breusch y Pagan para efectos aleatorios, y la prueba F de significancia de los efectos fijos indican que tanto el modelo de efectos aleatorios como el de efectos fijos son mejores que el modelo agrupado. Para decidir cuál depende de la posible correlación entre el componente de error individual u_i y las variables X .

Bajo la hipótesis de normalidad, el estadístico se define como:

$$LM = \frac{NT}{2(T-1)} \left(\frac{\sum_{i=1}^N (\sum_{t=1}^T e_{it})^2}{\sum_{i=1}^N \sum_{t=1}^T e_{it}^2} - 1 \right) \sim \chi_1^2 \quad (2-57)$$

2.7.3.3 Test de Hausman

Esta prueba ya mencionada permite determinar qué modelo es el más adecuado para el panel de datos que se está analizando, si el de efectos fijos o de efectos aleatorios. Utiliza para ello una prueba Chi-cuadrado con la hipótesis nula de que el modelo de efectos aleatorios es el que mejor explica la relación de la variable dependiente con las explicativas, y por tanto se tiene la hipótesis alternativa de que el mejor método que se ajusta es el de efectos fijos (Mayorga & Muñoz, 2000).

Este test compara los β obtenidos por medio del estimador de efectos fijos y efectos aleatorios, identificando si las diferencias entre ellos son o no significativas. Por tanto, primero se debe estimar por el método menos eficiente pero consistente (efectos fijos) y posteriormente por el estimador eficiente y consistente (efectos aleatorios). En ambos casos la matriz de pesos debe ser homocedástica.

El estadístico de contraste:

$$(\hat{\beta}_{EF} - \hat{\beta}_{EA})' [Var(\hat{\beta}_{EF}) - Var(\hat{\beta}_{EA})]^{-1} (\hat{\beta}_{EF} - \hat{\beta}_{EA}) \sim X_k^2 \quad (2-58)$$

El subíndice *EF* hace referencia a efectos fijos y *EA* efectos aleatorios.

El contraste compara los coeficientes estimables de los regresores que varían con el tiempo, de esta forma el estadístico de contraste mide la “distancia” entre ambas estimaciones.

Si se rechaza la Hipótesis Nula (H_0), los estimadores sí difieren, y la conclusión es que el procedimiento de efectos fijos es más conveniente que el de efectos aleatorios (ya que es probable que las perturbaciones estén correlacionadas con alguna variable explicativa). Si no se rechaza la Hipótesis Nula, no hay sesgo de qué preocuparse y se prefiere los efectos aleatorios que, al no estimar tantas variables dicotómicas, es un modelo más eficiente.

2.7.3.4 Prueba de autocorrelación de Wooldridge

El método de Wooldridge utiliza los residuales de una regresión de primeras diferencias, observando que si u_{it} no está serialmente correlacionado, entonces la correlación entre los errores u_{it} diferenciados para el período t y $t - 1$ es igual a -0.5. En realidad, la prueba de Wooldridge consiste en probar esta igualdad, donde la hipótesis nula indica que no hay autocorrelación de primer orden. Es una prueba muy flexible basada en supuestos mínimos. En la primera diferencia entre los datos del modelo elimina el efecto de nivel individual, el término basado en las covariables invariantes en el tiempo y la constante.

$$y_{it} - y_{it-1} = (\mathbf{X}_{it} - \mathbf{X}_{it-1})\boldsymbol{\beta}_1 + \varepsilon_{it-1}$$

$$\Delta y_{it} = \Delta \mathbf{X}_{it}\boldsymbol{\beta}_1 + \Delta \varepsilon_{it} \quad (2-59)$$

Donde Δ es el primer operador de diferencia.

Wooldridge (2002), propone probar la correlación con la siguiente estadística de prueba:

$$w = \frac{\sum_{i=1}^N \sum_{t=1}^{T-1} \sum_{s=t+1}^T \hat{u}_{it} \hat{u}_{is}}{[\sum_{i=1}^N (\sum_{t=1}^{T-1} \sum_{s=t+1}^T \hat{u}_{it} \hat{u}_{is})^2]^{1/2}} \rightarrow N(0,1) \quad (2-60)$$

Donde \hat{u}_{it} son los residuos OLS agrupados. La estadística de prueba w puede detectar muchos tipos de correlación serial en el término de error u por lo que tiene poder contra la especificación de efectos aleatorios unidireccionales y la correlación serial en términos de error.

2.7.3.5 Prueba de Durbin-Watson

Otra prueba para la correlación serial AR(1) es la de Durbin-Watson. El estadístico de Durbin-Watson (DW) se basa en los residuales de MCO:

$$DW = \frac{\sum_{t=2}^n (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^n \hat{u}_t^2} \quad (2-61)$$

2.7.3.6 Prueba de heterocedasticidad de Wald

Este test permite comprobar la heterocedasticidad del modelo. La hipótesis nula de esta prueba es que no existe problema de heteroscedasticidad, es decir $\sigma_i^2 = \sigma^2$ para toda $i = 1, \dots, N$, donde N es el número de unidades transversales. Funciona aun cuando el supuesto sobre la normalidad de los errores es violado (Aparicio & Márquez, 2005).

2.7.3.7 Test de endogeneidad Durbin Wu Hausman (DWH)

Compara estimaciones efectuadas mediante diferentes métodos: MCO y modelo endógeno utilizando variables instrumentales. La prueba evalúa la consistencia de un estimador cuando se compara con un estimador alternativo, menos eficiente, que ya se sabe que es consistente. Ayuda a evaluar si un modelo estadístico corresponde a los datos.

2.7.4 Modelos de datos panel dinámicos

Este tipo de modelos se mencionan de forma general ya que no se utilizarán en la modelación. El objetivo es estimar un modelo dinámico del tipo:

$$y_{it} = \delta y_{i, t-1} + x'_{it}\beta + u_{it} \quad (2-62)$$

Con $u_{it} = \mu_i + v_{it}$

2.7.4.1 Autocorrelación

Para que la estimación sea consistente y se justifique la utilización de modelos dinámicos, que empleen los retardos en diferencias o niveles como instrumentos, se requiere que los errores no estén serialmente correlacionados, lo que se comprueba con el test de Arellano y Bond.

2.7.4.2 Heterocedasticidad

La heterocedasticidad es un problema recurrente en los modelos dinámicos. Para probar la existencia de la misma, disponemos del test de Breusch-Pagan, que es aplicable solo a regresiones lineales simples que no consideran los efectos fijos.

Este tipo de modelos han sido desarrollados con el propósito de incorporar en la estimación las relaciones de causalidad que se generan en el interior del modelo, como una forma de tratar los problemas de endogeneidad (Labra & Torrecillas, 2014).

2.7.4.3 Tratamiento de la Endogeneidad

La endogeneidad puede ser tratada a través de diferentes vías, sin embargo una de las formas más habitualmente empleada es a través de variables instrumentales expresadas como retardos de la variable endógena. Dependiendo del estimador que se emplee, los retardos pueden ser formulados como diferencias o niveles.

2.7.4.4 Estimador de Arellano y Bond (1991) *Difference GMM*.

Utiliza como instrumentos las diferencias de los retardos. Realiza la regresión con variables endógenas utilizando sus diferencias.

2.7.4.5 Estimador de Arellano-Bover (1995) *System GMM*.

Fue desarrollado como un estimador que incluye los retardos en niveles de las variables como instrumentos con el objetivo de hacer frente a paneles con un t (período de tiempo) pequeño, y por tanto con un reducido número de instrumentos. Al incorporar las variables en niveles conforma un sistema de ecuaciones, dando lugar a su nombre *System GMM*.

2.7.4.6 Test de Sargan

Este test es adecuado cuando la estimación se ha hecho considerando la matriz de pesos homocedástica, como es el caso de One step. Verifica la validez de los instrumentos, pero es vulnerable a la proliferación de éstos y requiere además errores homocedásticos (One step) para ser consistente (Roodman, 2008). Por ello, el test de Sargan es más robusto para estimaciones One step y en aquellas donde no hay riesgo de sobreidentificación (Labra & Torrecillas, 2014).

2.7.4.7 Test de Hansen

Este test permite detectar la sobreidentificación del modelo cuando se ha empleado la matriz de pesos heterocedástica en la estimación, es decir es válido para estimaciones con Two step. Analiza la validez de los instrumentos, por lo que es posible deducir si nos encontramos ante un modelo endógeno.

En el caso de que teniendo en cuenta las consideraciones anteriores, no detectemos endogeneidad y no sospechemos de ella, entonces el camino más recomendado es la utilización de paneles estáticos o MCO. Si por el contrario, la literatura o los test dan indicios o confirman la presencia de endogeneidad, necesariamente deberemos utilizar un método que permita tratarla (Labra & Torrecillas, 2014).

2.7.4.8 One step y Two step

Todos los estimadores expuestos anteriormente pueden efectuar el análisis a través de dos mecanismos: One step: Utiliza solo la matriz de pesos homocedástica para la estimación. Two step: Utiliza para la estimación la matriz de pesos heterocedástica. La literatura indica que los estimadores Two step son más eficientes.

2.8 Variables expresadas en logaritmos

El uso de las variables transformadas a logaritmos se debe a varias situaciones: a) Una expresión no lineal se quiere trabajar mediante métodos lineales, por lo tanto se linealiza generando una expresión en logaritmos. b) Como estrategia matemática para disminuir la dispersión original de una serie, a través de una compresión de los valores originales dentro de un rango menor que el original; esto disminuye la posibilidad de que se presente heterogeneidad (varianza no constante de la perturbación aleatoria) un problema que afecta la eficiencia de los estimadores. c) El uso de logaritmos modifica el significado e interpretación de los parámetros obtenidos. Cuando ambas variables dependiente (y) e independiente (x) están escritas en logaritmos la interpretación de los parámetros de un modelo de regresión se basa en la magnitud del cambio porcentual en “ y ” ante una variación del 1% de la variable “ x ” (De Arce & Mahía, 2012).

Las transformaciones logarítmicas permiten modelizar relaciones en términos de “porcentajes” (como elasticidades), en vez de linealmente.

$$\ln(x + \Delta x) - \ln(x) = \ln\left(1 + \frac{\Delta x}{x}\right) \cong \frac{\Delta x}{x} \quad (2-63)$$

Cálculo

$$\frac{d\ln(x)}{dx} = \frac{1}{x} \quad (2-64)$$

Pueden ser tres casos de transformación, dependiendo de si es Y y/o X es la variable transformada tomando logaritmos.

Tabla 2-1: Transformaciones logarítmicas.

Caso	Función de regresión
1. lineal – log	$Y_i = \beta_0 + \beta_i \ln(X_i) + u_i$
2. log – lineal	$\ln(Y_i) = \beta_0 + \beta_i X_i + u_i$
3. log-log	$\ln(Y_i) = \beta_0 + \beta_i \ln(X_i) + u_i$

Fuente: Interpretación de los parámetros de un modelo básico de regresión lineal (De Arce & Mahía, 2012).

3. Metodología

En este capítulo se presenta el proceso metodológico establecido para la búsqueda de literatura, las fuentes de información utilizadas para obtener los datos, las variables definidas, así como los procedimientos para realizar las estimaciones y el análisis.

3.1 Búsqueda y revisión de literatura

Se realizó una búsqueda en la cual se incluyeron documentos como artículos y libros de investigaciones sobre mortalidad en menores de cinco años publicados tanto en revistas nacionales e internacionales desde el año 2005, también literatura estadística sobre análisis de regresión, modelación estadística, metodología de datos panel, entre otros. La consulta se realizó en diferentes buscadores como *Scielo*, *PubMed*, *google scholar*, el SINAB de la Universidad Nacional del Colombia; la búsqueda fue hecha utilizando palabras clave o términos *MeSH* como “mortalidad”, “desnutrición”, “malnutrición” “infantil”, “Colombia” así como “regresión”, “modelos” “análisis multivariado” “análisis categórico” “datos panel” o una combinación de los anteriores, tanto en español como inglés.

3.2 Fuente de los datos

Los datos corresponden a los registros individuales de mortalidad, que hacen parte del sistema de estadísticas vitales, que como ya se mencionó reúnen las frecuencias de los hechos vitales (nacimientos y defunciones); de esta forma la principal fuente de información se obtiene de los microdatos para defunciones DANE, de los datos registrados a partir de los certificados de defunción; sin embargo para tener más información que dé respuesta a los determinantes que se establecen desde la construcción teórica, se obtuvo también información de las publicaciones y bases de datos del Sistema Estadístico Nacional SEN del DANE, y del Sistema de Gestión de

Datos SGD de SISPRO (Sistema de Información del Ministerio de Salud y Protección Social) que proporciona datos e indicadores agrupados en el país.

A partir de dichas fuentes de información se seleccionaron las variables que hacen parte de la base de datos, cumpliendo en lo posible el requerimiento de obtener información para los diferentes territorios y períodos de análisis, que permitan la construcción de paneles balanceados. En el anexo A, se describen las fuentes de datos utilizadas.

3.3 Clasificación de la mortalidad por desnutrición

En la valoración de la desnutrición como causa de muerte, se tiene en cuenta la desnutrición o alguna deficiencia nutricional relacionada en los códigos de la Clasificación Internacional de Enfermedades (CIE 10) y las enfermedades asociadas a desnutrición.

Para especificar los diagnósticos asociados a la desnutrición se partió de la clasificación de mortalidad 602, en la lista 667-OPS CIE-10 de la Organización Panamericana de la Salud que corresponde a las defunciones asociadas con los códigos E40 - E64, D50 - D53 definidos como “deficiencias nutricionales y anemias nutricionales” (OPS, 1995), que en los indicadores nacionales se identifica como “mortalidad por desnutrición”.

3.4 Estimación de la mortalidad

Las estimaciones se realizan por cada año, para los menores de 5 años y por grupos de edad, en menores de un año y de 1 a 4 años. Las tasas se calculan con base en las proyecciones de población DANE 2005-2020. Es de aclarar que no se realiza ningún ajuste de sub-registro o aplicación de metodologías de cálculo indirectas. El indicador de mortalidad se puede expresar como Wachter (2014) Wachter, K. W. (2014). Essential demographic methods. Harvard University Press.

$$nMx = \frac{nDx}{nNx} * 100.000 \quad (3-1)$$

Dónde:

nMx : Es la tasa de mortalidad del grupo de edad x a $x + n - 1$ en el año z

nDx : Es el número de defunciones ocurridas en el año z a personas con edades cumplidas entre x y $x + n - 1$

nNx : Es la población del año z en el grupo de edad x a $x + n - 1$

3.5 Variables seleccionadas para el análisis

El primer aspecto para modelar algún fenómeno es considerar el tipo de variable que está bajo estudio (Pérez et al., 2012). El conocimiento de las variables contribuye a determinar un modelo estadístico que se ajuste a las características de los datos.

De esta forma existe un conjunto de variables demográficas, socioeconómicas e individuales, que pueden explicar el fenómeno de la mortalidad. La selección es de acuerdo con las exigencias del modelo y la disponibilidad de información proveniente de las estadísticas vitales y de otras fuentes externas⁷.

Para el presente trabajo la variable dependiente es la tasa estandarizada de mortalidad por desnutrición, que representan la muerte de los individuos menores de cinco años, y las sub agrupaciones que se definan, para una población y período de tiempo específico, clasificados con diagnósticos de deficiencia y anemia nutricional.

Las independientes son aquellas que dan respuesta a los determinantes y que se consideran relevantes; es de aclarar que solo se incluye las que tienen más información para el período y departamentos de estudio.

Partiendo de lo anterior las variables consideradas son:

- Variables relacionadas a los determinantes sociodemográficos: edad, sexo, etnia, área de residencia, estado civil de la madre, nivel educativo de la madre, edad de la madre.
- Variables relacionadas a los determinantes socioeconómicos: tasa de desempleo, pobreza, línea de indigencia, coeficiente de Gini, Producto Interno Bruto (PIB) per cápita.

⁷ En el presente trabajo se consideran “*exógenas*”, todas las variables de fuentes de información diferentes al certificado de defunción. De esta forma las “*endógenas*” son las explícitas al microdato de defunción. Esta especificación aunque se denomina igual, no se refiere a lo planteado en la sección 2.7.1.6. “Variables exógenas, endógenas, predeterminadas e instrumentales”.

- Variables relacionadas a los determinantes de prestación de servicios de salud: Afiliación en salud.
- Variables relacionadas a los determinantes individuales: peso al nacer reportado en certificado de defunción, bajo peso al nacer (<2500g) según certificado de nacimiento, semanas de gestación, proporción de atendidos por enfermedades infecciosas y parasitarias, proporción de atendidos por deficiencias nutricionales, tasa de mortalidad por Enfermedad Diarreica Aguda EDA en menores de 5 años, tasa de mortalidad por Infección Respiratoria Aguda IRA en menores de 5 años.
- Variables relacionadas a los determinantes de política y protección social: Tasa de personas vinculadas a programas de asistencia social, coberturas de vacunación.

No fue posible incluir variables de determinantes de tipo ambiental, como contaminación, condiciones de la vivienda o de los servicios públicos, así como de comportamiento y estilos de vida, ya que no se encontró información suficiente para el período y las unidades de análisis.

3.5.1 Descripción de las variables

A continuación se describen las variables utilizadas en el análisis. En el anexo B se incluye además la denominación de las variables en la base de datos y la fuente de información.

3.5.1.1 Variables obtenidas del certificado de defunción.

Las siguientes variables son obtenidas de los microdatos de defunciones del DANE y agrupadas por departamento.

- Departamento de residencia: Departamento de residencia habitual del fallecido (para muerte fetal o de menor de un año el de la madre).
- Año: Representa el año en que ocurrió la defunción, para el presente trabajo se utilizó información del período 200 a 2012.
- Sexo: Identifica el sexo del fallecido, sus categorías son femenino y masculino.
- Grupo de edad: Agrupación de edades, según la edad del fallecido en menores de 1 año, de 1 a 4 años y de 0 a 4 años.

- Etnia: Define la pertenencia étnica de acuerdo con la cultura, pueblo o rasgos físicos del fallecido o como era reconocido, se agrupo en las siguientes categorías: Indígena, Afrocolombiano, Rom y Ninguna etnia.
- Área de residencia: Corresponde al área de residencia habitual del fallecido (para muerte de menor de un año la el de la madre), sus categorías son rural y urbano.
- Afiliación en salud: Hace referencia al régimen de seguridad social del fallecido (para muerte de menor de un año el de la madre), se agrupo en las siguientes categorías: contributivo, subsidiado y no asegurado.
- Semanas de gestación: Identifica el número de semanas o tiempo de gestación en las cuales se tuvo el menor, se agrupo en las siguientes categorías: menor de 38 semanas, de 38 a 41 semanas y mayor o igual a 42 semanas.
- Peso al nacer reportado en certificado de defunción: Indica el peso al nacer del menor de un año, se agrupo en las siguientes categorías: mayor o igual a 2500 gramos y menor de 2500 gramos.
- Edad de la madre: Define la edad de la madre del menor de un año fallecido, se agrupo en las siguientes categorías: 10 a 14 años, 15 a 49 años y de 50 o más años.
- Nivel educativo de la madre: Indica el último año de estudios aprobado de la madre, se agrupo en las siguientes categorías: primaria, secundaria, técnico o profesional y ningún nivel educativo.
- Estado civil de la madre: Indica el estado conyugal de la madre, se agrupo en las siguientes categorías: Casada, Soltera, Unión Libre y Otro estado civil.

3.5.1.2 Variables obtenidas de otras fuentes diferentes al certificado de defunción.

Las variables incluidas se obtuvieron de diferentes fuentes y la condición para tenerlas en venta e incluirlas era tener información desagregada por departamento y para el período de análisis.

- Producto Interno Bruto (PIB) per cápita: Es un indicador macroeconómico de productividad y desarrollo económico.
- Tasa de desempleo: Expresa el nivel de desocupación entre la población económicamente activa.
- Coeficiente de Gini: corresponde a una medida de la desigualdad.

- Incidencia de la Pobreza: representa el número de personas pobres expresado como porcentaje del total de la población en un determinado año.
- Línea de indigencia: Establece si los hogares cuentan con ingresos suficientes como para cubrir una canasta de alimentos capaz de satisfacer un umbral mínimo de necesidades.
- Cobertura de vacunación: Corresponde al porcentaje de las coberturas de vacunación por departamento y año de forma general sin desagregar por el biológico.
- Tasa de mortalidad por EDA (Enfermedad Diarreica Aguda): Es el número de muertes por EDA en menores de 5 años por 100.000 personas.
- Tasa de mortalidad por IRA (Infección Respiratoria Aguda): Es el número de muertes por IRA en menores de 5 años por 100.000 personas.
- Nacimientos: Información de nacimientos que se obtiene a partir de los Certificados de Nacido Vivo.
- Proyecciones de población: proyecciones realizadas por el DANE tomando como base los resultados ajustados de población del Censo 2005.
- Proporción de bajo peso al nacer: nacidos vivos con peso menor a 2500 gramos, se obtiene del certificado de nacimiento.
- NBI: Población en Necesidades Básicas Insatisfechas. Busca determinar, con ayuda de algunos indicadores simples, si las necesidades básicas de la población se encuentran cubiertas.

3.5.1.3 Variables calculadas

Las siguientes variables fue necesario calcularlas.

- Tasa estandarizada de mortalidad por desnutrición: representa el número de muertes por deficiencias nutricionales y anemias nutricionales por cada 100.000 habitantes, según el grupo de edad.
- Proporción de atendidos por deficiencias nutricionales: representa las personas atendidas en los servicios de salud por diagnósticos asociados a deficiencias nutricionales.
- Proporción de atendidos por enfermedades infecciosas y parasitarias: representa las personas atendidas en los servicios de salud por diagnósticos asociados a enfermedades infecciosas y parasitarias.

- Tasa de personas vinculadas a programas de asistencia social: representa el número de personas vinculadas a programas de asistencia social por cada 100.000 habitantes, según el grupo de edad.

3.6 Procesamiento y análisis de datos

Lo primero que se plantea es un análisis de la base de datos y las variables de utilidad que harán parte del modelo estadístico. De esta forma se parte de un análisis descriptivo de los datos, para identificar el comportamiento de la información. Posteriormente se desarrollara la metodología para Datos de Panel.

3.6.1 Depuración y ajuste de la base de datos

Al realizar la agrupación de la información de las variables por departamentos, de la base de datos de defunción inicial, las opciones de respuesta o categorías de las variables originales, se convierten cada una en una nueva variable, debido a lo anterior es necesario agrupar las opciones de respuesta para limitar la cantidad de variables que se generan. Las nuevas variables son proporciones que al sumarlas completan la unidad e identifican las diferentes categorías de la variable cualitativa.

Dentro de los principales ajustes a la base de datos se destaca:

- En el área de residencia, “urbano” hace referencia a la cabecera y se integró la información de “centro poblado” y “rural disperso” como la categoría “rural”.
- Para el nivel educativo de la madre, los registros de la opción “preescolar” (8 registros) se integraron a la categoría “ninguno”; los registros de “básica secundaria”, “media académica”, “media técnica” y “normalista” se agruparon en la categoría “secundaria”; se definió una categoría denominada “técnico o profesional”, que agrupa los estudios técnicos, tecnológicos, profesionales y un registro de postgrado (maestría).
- Para la variable reconocimiento de acuerdo con la cultura o rasgos físicos, denominada como “etnia” se suma a la categoría “ninguna etnia”, la opción “Raizal” que tan solo sumaba dos (2) registros.
- Para la afiliación a salud, los registros de la categoría excepción y del régimen especial (11 registros), se agrupan con la categoría “ningún régimen”.

- Para la variable departamento de residencia se identifican doce (12) registros que se clasifican sin información del departamento y tres registros (3) con el código 75 que no corresponde a ningún departamento en Colombia, sino a residencia en el extranjero; por lo tanto se agrupan en la categoría “otro”. Esta situación es de importancia porque los quince (15) registros referidos son definidos en el análisis exploratorio con el nombre de categoría mencionado y descartados en el resto de análisis.

Lo anterior permite también identificar los registros sin información, que se convierten en una nueva variable al realizar la agrupación por departamentos; esta situación hace necesario plantear un procedimiento de imputación o eliminación de los registros que no tienen información completa, para las variables de interés que se utilizan en el análisis, con el propósito de obtener una base de datos consistente. Para esta aproximación de los determinantes se realizó eliminación de registros sin información completa, pero solo para las variables consideradas, según el grupo de edad.

Es de aclarar que en la base agrupada construida se identificaron nueve (9) departamentos sin casos de mortalidad que difieren de los que no cuentan con información completa, por lo tanto son incluidos en el análisis.

3.6.2 Metodología para datos de panel

El análisis utilizando la metodología de regresión para datos de panel, requiere de un procedimiento determinado para la construcción de la base de datos, la validación de supuestos y la selección del modelo. A continuación se menciona el proceso desarrollado.

1. Base de datos

La base de datos tiene una construcción específica, en las filas de la base se incluyeron las unidades de análisis (departamentos) y en las columnas las variables de interés, que son objeto del análisis. Se utilizó el formato para bases de datos longitudinales, tipo panel, que se aplica para el análisis en Stata. Se incluyó una variable que hace referencia al año (*year*) y otra que identifique a los departamentos (*id*), esta última para el

trabajo en Stata debe ser continua, por lo cual se utilizan los códigos DANE de cada departamento.

2. Declaración de la base

Es necesario indicarle al programa que se trabaja con datos de panel, por lo cual se utiliza el comando *xtset*, seguido de las variables mencionadas *id* y *year* (*xtset id year*), la salida reportada posterior al comando indica los años del período y si el panel es balanceado o desbalanceado.

3. Cálculo de los efectos fijos y aleatorios

Se utiliza en primer lugar el comando del estimador *xtreg* seguido de la variable dependiente y las independientes, además se especifica el tratamiento de los efectos individuales “fe” (*fixed effects*) o de efectos aleatorios “re” (*random effects*). También se puede utilizar el método robusto que realiza la estimación teniendo en cuenta la heterocedasticidad; ambas opciones son consistentes, pero la opción robusta evita el sesgo y la ineficiencia de los β .

4. Probar el test de Hausman

Después de realizar las estimaciones de efectos fijos y aleatorios se realiza este test el cual no se calcula con la opción robusta. La hipótesis nula (H_0), que se plantea para este test: “ H_0 . Los coeficientes de modelos fijos y modelos aleatorios son exactamente los mismos” que tiene como criterio de rechazo: $\text{Prob} > \chi^2$ (mayor a 0,05), es decir, no hay correlación entre los efectos individuales y las variables explicativas, lo que indica que el estimador aleatorio debe ser utilizado. En caso contrario, $\text{Prob} < \chi^2$ (menor a 0,05), se emplea el estimador de efectos fijos.

Lo que hace Hausman es revisar si los estimadores son iguales en un modelo y el otro. La prueba asume que los estimadores son los mismos y no existe diferencia ($p > 5\%$), pero si son diferentes los estimadores, se deben calcular los estimadores con el método de efectos fijos.

5. Pruebas para validar supuestos

Se realizan las pruebas que son recomendables para detectar si existe heteroscedasticidad y autocorrelación en el modelo.

- Prueba de autocorrelación de Wooldridge

Si se rechaza la hipótesis nula, ($F < 0.000$) hay un problema de autocorrelación que es necesario corregir, una de las soluciones es utilizar el comando *xtregar* el cual corrige la autocorrelación de primer orden de los residuos; otra posibilidad es utilizar el comando *xtpcse* que realiza estimaciones con errores estándar corregidos para panel (PCSE).

- Prueba de heteroscedasticidad de Wald

Si se rechaza la hipótesis nula, indica que hay heteroscedasticidad. Para solucionar este problema se puede utilizar el comando *xtgls* que realiza estimaciones de mínimos cuadrados generalizados factibles (FGLS) o el comando *xtpcse*, que como se mencionó también corrige la autocorrelación. El primero requiere de paneles balanceados mientras el segundo se aplica tanto a balanceados como desbalanceados, además asume que las perturbaciones son por defecto heterocedásticas y correlacionadas contemporáneamente entre paneles.

3.6.3 Modelación estadística

La modelación se realiza a partir de técnicas estadísticas y variables socioeconómicas, demográficas e individuales del certificado de defunción. Para la construcción del modelo se tomó la información de las defunciones ocurridas en los años 2009, 2010, 2011 y 2012.

Con el propósito de estandarizar la unidad de medida de los datos se calcula el logaritmo para cada una de las variables, tanto la dependiente como las independientes. La ventaja práctica del logaritmo natural además de que la transformación logarítmica genera linealidad en los parámetros, es que la interpretación de los coeficientes de regresión es sencilla; ya que el coeficiente es el cambio porcentual en la variable dependiente para un cambio porcentual en la variable independiente.

Las variables en el modelo se introducen según las categorías que tengan menos una, es decir si una variable cualitativa tiene k categorías, se introducen en el modelo $k-1$ variables, para evitar perfecta multicolinealidad.

Se especifica el siguiente modelo de datos de panel:

$$\begin{aligned}
 & \text{Log (tasa Def}_{it}) \\
 & = \alpha_{it} + \beta_1 \text{Log}(\text{área de residencia}_{it}) + \beta_2 \text{Log}(\text{Sexo}_{it}) \\
 & + \beta_3 \text{Log}(\text{Etnia}_{it}) + \beta_4 \text{Log}(\text{afiliación en salud}_{it}) \\
 & + \beta_5 \text{Log} (\text{Edad de la madre}_{it}) \\
 & + \beta_6 \text{Log} (\text{Estado civil de la madre}_{it}) \\
 & + \beta_7 \text{Log} (\text{Nivel educativo de la madre}_{it}) + \dots + \beta_k \text{Log}(X_{it}) + U_{it}
 \end{aligned} \tag{3-1}$$

$$i = 1, \dots, 33 \quad t = 2009, \dots, 2012 \quad k = 1, \dots, n$$

Donde,

i 32 departamentos y un Distrito (Bogotá) de Colombia.

t período de análisis (cuatrienio 2009, 2010, 2011 y 2012).

X_{it} corresponde a las demás variables independientes incluidas en el modelo como: peso al nacer reportado en certificado de defunción⁸, proporción de bajo peso al nacer⁹, semanas de gestación, proporción de atendidos por deficiencias nutricionales, proporción de atendidos por enfermedades infecciosas y parasitarias, Producto Interno Bruto (PIB) per cápita, tasa de desempleo, coeficiente de Gini, incidencia de la pobreza, línea de indigencia, cobertura de vacunación, tasa de personas vinculadas a programas de asistencia social, tasa de mortalidad por EDA (Enfermedad Diarreica Aguda) en menores de 5 años, tasa de mortalidad por IRA (Infección Respiratoria Aguda) en menores de 5 años.

U_{it} término del error conjunto.

⁸ El peso al nacer se obtiene del certificado de defunción solo para los menores de 1 año y se agrupo en las categorías: mayor o igual a 2500 gramos y menor de 2500 gramos.

⁹ La proporción de bajo peso al nacer se establece para un peso por debajo de 2500 gramos y se obtiene del certificado de nacimiento, esta variable hace parte de las independientes para los grupos de edad de 0 a 4 y 1 a 4 años.

4. Resultados

El análisis de información se realizó a partir de los microdatos de defunciones DANE, considerados en el trabajo como información “endógena”, así como de datos y variables de otras fuentes que se definen como la información “exógena” y se refirieron en la metodología. Aunque la información de defunciones es muy criticada por el subregistro y los problemas de calidad (OPS, 2012), representa la principal fuente de información relacionada a la mortalidad.

4.1 Análisis exploratorio de la información

Para comprender el comportamiento de los datos se realiza un análisis exploratorio de las principales variables que se obtienen de los microdatos de defunción, que además de caracterizar el evento, en este caso la mortalidad, oriente los procedimientos estadísticos. Es básicamente un paso previo a la aplicación del análisis de datos panel y de orientación para el manejo de las variables.

La población obtenida según la causa básica de muerte por deficiencias nutricionales y anemias nutricionales o mortalidad por desnutrición como es analizada en los indicadores de mortalidad, corresponde a 1.354 registros para el período entre los años 2009 a 2012, en la tabla 4-1 se muestran los registros de defunción por año.

Tabla 4-1: Defunciones de mortalidad por desnutrición, según el año.

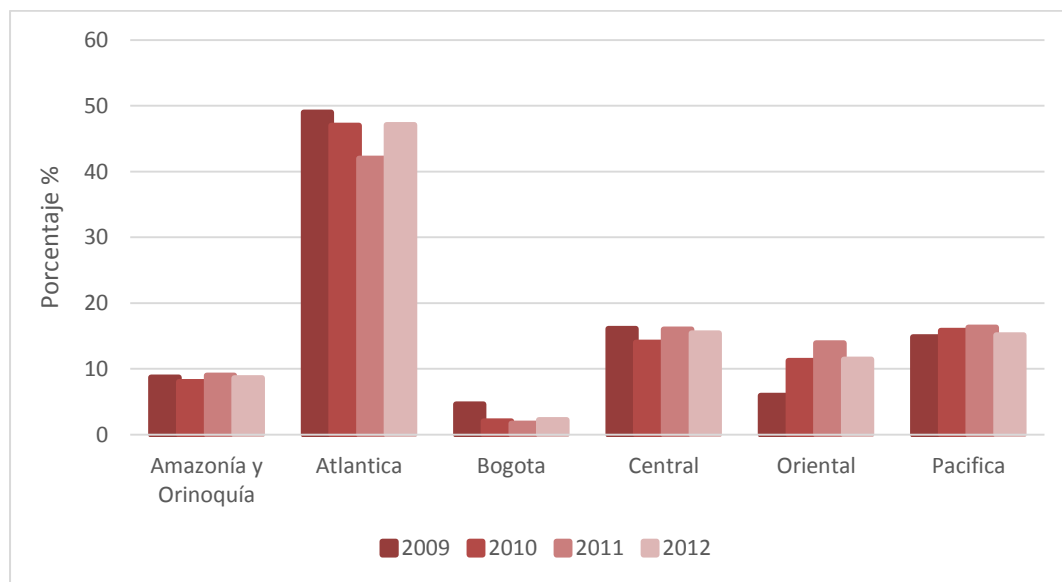
Año	Defunciones	Porcentaje
2009	392	29,0
2010	349	25,8
2011	288	21,3
2012	325	24,0
Total	1354	100

Fuente: Elaboración propia.

4.1.1 Proporción de muertes

Se agrupó la información por regiones¹⁰ y se identificó que para cada una de estas el comportamiento por año es similar; sin embargo entre ellas si hay diferencias notorias y como se observa en la Figura 4-1 la región Atlántica tiene la proporción más alta de mortalidad, cercana a la mitad de los casos reportados. Información más detallada con las proporciones por año y departamento, así como los intervalos de confianza del 95%, se presentan en el Anexo E.

Figura 4-1: Proporción de muertes por desnutrición en menores de cinco años, según la región, período 2009 - 2012.



Fuente: Elaboración propia.

4.1.2 Tasas de mortalidad

Se estimó la tasa de mortalidad¹¹ (Figura 4-2), las cuales permiten un análisis ajustado al tamaño poblacional, por lo tanto se hará una descripción más detallada que el caso de las proporciones; para el año 2009 la tasa nacional se estimó en 9,2 por cada cien mil habitantes y representa la más alta para los cuatro años de análisis; aunque ha

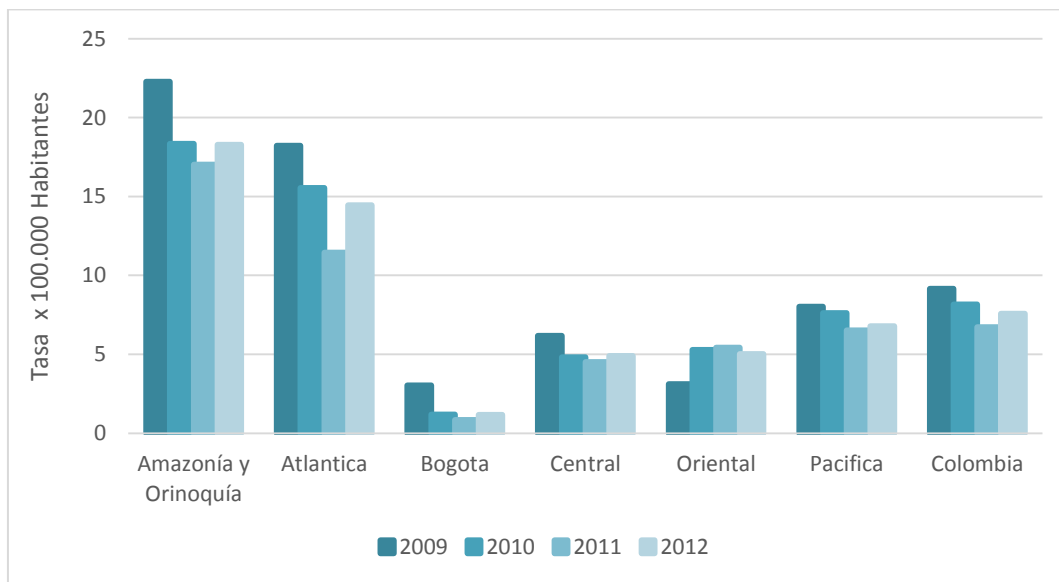
¹⁰ La agrupación por regiones se realizó como se establece en la ENSIN 2010. Ver Anexo D.

¹¹ La estimación de las tasas se realizaron sin ajustar por sub-registro.

disminuido no se identifica una tendencia a la reducción, ya que en el 2012 aumento 0,9 por cien mil habitantes respecto al 2011.

En cuanto a las regiones el comportamiento de la mortalidad es similar para todos los años, ya que las menores tasas de mortalidad por cada cien mil habitantes son para Bogotá, con unas tasas entre 0,8 y 3,0 y las mayores tasas en todo los años son para Amazonía y Orinoquía con tasas entre 17 y 22,3, frente a las demás regiones que se concentran con tasas por debajo de 8.

Figura 4-2: Tasa de mortalidad por desnutrición en menores de cinco años, según la región, período 2009 - 2012.



* Las tasas presentadas son por cada cien mil habitantes.

Fuente: Elaboración propia.

En cuanto a las tasas de mortalidad por departamento, para el año 2009 la tasa de mortalidad infantil más baja es la de Cundinamarca con 0,9 y la mayor es para Guainía con una tasa 102,6 por cada cien mil habitantes, una tasa muy alejada del total nacional, y de la concentración de tasas para la mayoría de departamentos (24 territorios) que está por debajo de 20. Seguido de Guainía, con las mayores tasas se resalta Amazonas, Vichada y La Guajira con tasas por encima de 40.

En el año 2010 Guaviare y Quindío no presentaron muertes y seguido a ellos con la menor tasa se encuentra Cundinamarca que al igual que el 2009 continua con 0,9; hacia el extremo con mayor mortalidad se encuentra Vichada con una tasa de 183,3 y seguido a este Guainía y La Guajira con tasas sobre 30; cifras que como en el 2009 están alejadas de la tasa nacional y del resto de departamentos (30), los cuales están por debajo de 20 por cada cien mil habitantes.

Para el 2011 la mayor tasa es para Guainía con 138,9 por cada cien mil habitantes, seguido de Vichada con una tasa de 84,8, el resto de departamentos (30) se encuentran en tasas por debajo de 20. En este año San Andrés y Arauca no presentaron casos de mortalidad por desnutrición.

En el caso del 2012 Guainía nuevamente representa el departamento con la tasa más alta en este caso de 137,6 por cien mil habitantes, muy alejado de los siguientes departamentos que corresponden a Amazonas y Vichada con tasas de 68 y 52,1 respectivamente. Quindío y San Andrés para el 2012 no presentaron muertes y 26 departamentos tienen tasas por debajo de 20 por cada cien mil habitantes.

En general para el período de análisis (2009- 2012) la tasa más alta registrada es para el departamento de Vichada en el 2010, además este departamento y Guainía han estado entre las tasas más altas en los cuatros años, y los departamentos de Amazonas y La Guajira en tres de estos años; de esta forma estos cuatro departamentos son los más afectados por altas tasas de mortalidad por desnutrición.

4.1.3 Variables principales

Teniendo en cuenta las variables principales descritas en la metodología se realiza un análisis descriptivo de estas, que además se diferenciaron por grupo de edad teniendo en cuenta que algunos autores (Urdinola 2011, Ruíz 2003, Quiroga 2012), resaltan las diferencias en la mortalidad según la edad específica. De esta forma se observa en la Tabla 4-2 que la mayor proporción de muertes corresponde a los menores de 1 año (66,9%), se destaca además que para este grupo se diligencia en el certificado variables que para otros no, como es el caso del peso al nacer, las semanas de gestación y para la madre: la edad, el estado civil y el nivel educativo.

Para el total de los datos no se identifica diferencias importantes según el área de residencia y el sexo, tan solo alrededor de 3 puntos porcentuales es superior la cantidad de registros para el área urbana y el sexo masculino.

Con respecto a la etnia, la mayor proporción no refiere ningún grupo étnico (52,8%), seguido de la clasificación afrocolombiano(a) (9,9%).

En cuanto a la afiliación en salud o seguridad social, al régimen subsidiado son reportados más de la mitad de los fallecidos (63,5%), lo cual es un aspecto de resaltar y de importancia, ya que puede ser asociado a familias de bajos recursos, teniendo en cuenta que según el análisis de situación de salud (ASIS, 2014), la población subsidiada pertenece a los estratos 1 y 2.

Tabla 4-2: Proporción total y por grupo de edad, de variables seleccionadas en menores de 5 años

Variable	Total % (IC 95%)	Grupo de edad		<i>p</i> ¹
		Menor de 1 año % (IC 95%)	De 1 a 4 años % (IC 95%)	
<i>n</i>	1354	906	448	
Sexo				<0,05 ²
Masculino	51,1 (48,3 - 53,7)	54,3 (51,0 - 57,6)	44,4 (39,8 - 49,2)	
Femenino	48,9 (46,2 - 51,6)	45,6 (42,3 - 48,9)	55,6 (50,8 - 60,2)	
Indeterminado	0,1 (0,0 - 0,5)	0,1 (0,0 - 0,7)	0	
Área de residencia				<0,05
Urbano	50 (47,3 - 52,7)	52,9 (49,6 - 56,2)	44,2 (39,6 - 48,9)	
Rural	46,4 (43,7 - 49,1)	43,3 (40,0 - 46,6)	52,7 (47,9 - 57,4)	
Sin información	3,6 (2,7 - 4,8)	3,9 (2,7 - 5,4)	3,1 (1,8 - 5,3)	
Etnia (%)				<0,05
Indígena	25,5 (23,2 - 27,9)	20,3 (17,8 - 23,1)	35,9 (31,5 - 40,6)	
ROM (Gitano)	1,4 (0,9 - 2,2)	0,2 (0,0 - 0,9)	3,8 (2,3 - 6,1)	
Afrocolombiano ³	9,9 (8,4 - 11,6)	10,6 (8,7 - 12,8)	8,5 (6,2 - 11,6)	
Ninguno	52,8 (50,1 - 55,5)	57,1 (53,8 - 60,3)	44,2 (39,6 - 48,9)	
Sin información	10,4 (8,9 - 12,2)	11,8 (9,8 - 14,1)	7,6 (5,4 - 10,5)	

Tabla 4-2: (Continuación)

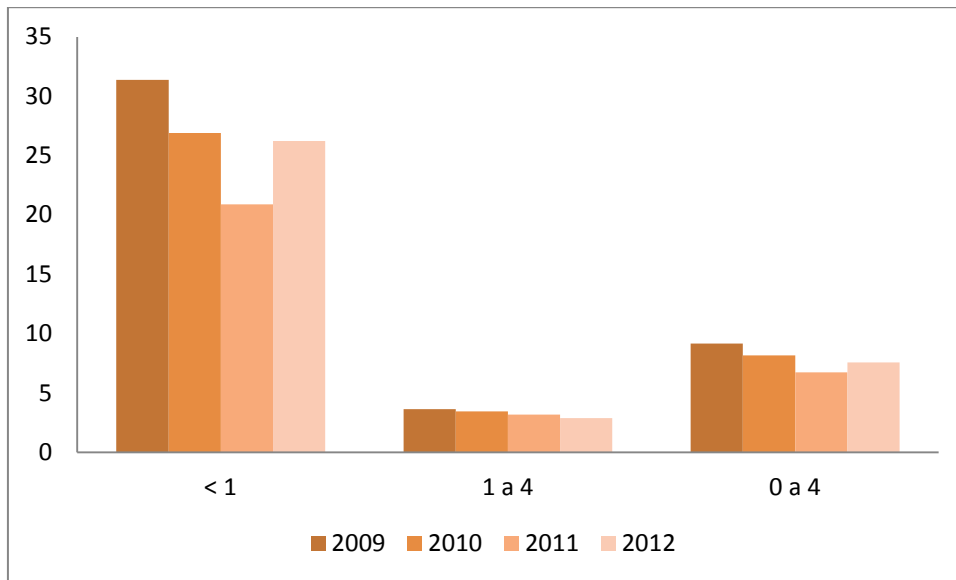
Variable	Total % (IC 95%)	Grupo de edad		p ¹
		Menor de 1 año % (IC 95%)	De 1 a 4 años % (IC 95%)	
Afiliación en salud				0,36
Contributivo	5,8 (4,6 - 7,2)	6,1 (4,7 - 7,9)	5,1 (3,4 - 7,7)	
Subsidiado	63,5 (60,9 - 66,1)	65 (61,8 - 68,1)	60,5 (55,8 - 65)	
Otro	0,8 (0,4 - 1,5)	0,8 (0,3 - 1,7)	0,9 (0,3 - 2,4)	
No asegurado	26,7 (24,3 - 29,1)	25,2 (22,4 - 28,2)	29,7 (25,5 - 34,2)	
Sin información	3,2 (2,4 - 4,4)	3 (2,0 - 4,4)	3,8 (2,3 - 6,1)	
Peso al nacer				<0,05
Mayor Igual 2500	53,8 (51,1 - 56,5)	80,5 (77,7 - 83)		
Menor 2500	13,1 (11,4 - 15,0)	19,5 (17,0 - 22,3)		
Sin información	33,1 (30,6 - 35,7)	0		
Edad de la madre				<0,05 ²
10 a 14	0,5 (0,2 - 1,1)	0,8 (0,3 - 1,7)		
15 a 49	58,7 (56,0 - 61,4)	87,7 (85,4 - 89,8)		
50 o mas	0,2 (0,1 - 0,7)	0,3 (0,1 - 1,1)		
Sin información	40,5 (37,9 - 43,2)	11,1 (9,2 - 13,4)		
Estado civil de la madre				<0,05
Casada	4 (3,0 - 5,2)	6 (4,6 - 7,8)		
Soltera	8,6 (7,2 - 10,3)	12,9 (10,8 - 15,3)		
Unión libre	34,7 (32,2 - 37,3)	51,9 (48,6 - 55,2)		
Otro	1,8 (1,2 - 2,7)	2,6 (1,7 - 4)		
Sin información	50,9 (48,2 - 53,6)	26,6 (23,8 - 29,6)		
Nivel educativo de la madre				<0,05
Ninguno	9,2 (7,8 - 10,9)	13,8 (11,7 - 16,3)		
Primaria	19,6 (17,6 - 21,9)	29,4 (26,4 - 32,5)		
Secundaria	12,9 (11,2 - 14,9)	19,3 (16,8 - 22,1)		
Técnico o profesional	0,9 (0,5 - 1,6)	1,3 (0,7 - 2,4)		
Sin información	57,3 (54,6 - 60,0)	36,2 (33,1 - 39,4)		
Tiempo de gestación				<0,05
Menor de 38	12,3 (10,6 - 14,2)	18,3 (15,9 - 21,0)		
De 38 a 41	26,1 (23,8 - 28,6)	39,1 (35,9 - 42,4)		
Mayor Igual a 42	28,5 (26,1 - 31,0)	42,6 (39,4 - 45,9)		
Sin información	33,1 (30,6 - 35,7)	0		

¹ X² Pearson o Fisher para variables categóricas.² Más del 20% de las casillas de la sub-tabla esperaban frecuencias de casilla inferiores a 5. Puede que los resultados de chi-cuadrado no sean válidos.³ Aquí y en todo el documento afrocolombiano Incluye la información de: negro, mulato y afrodescendiente.

Fuente: Elaboración propia.

La literatura (Ruiz, 2003) y el comportamiento de los datos demuestran una diferencia en la mortalidad para los menores de un año y el grupo entre 1 a 4 años, esto hace necesario realizar un análisis diferenciado, desde lo descriptivo hasta la modelación estadística. En la Figura 4-3 se presenta la tasa de mortalidad por grupos de edad y es claro que el más afectado es el menor de 1 año; además la tasa para estese ha incrementado con el transcurrir de los años.

Figura 4-3: Tasa de mortalidad por desnutrición, según el grupo de edad, período 2009 - 2012



* Las tasas presentadas son por cada cien mil habitantes.

Fuente: Elaboración propia.

Partiendo de lo anterior se dispusieron 3 bases de datos según la edad: 1. Base total con los registros de los menores de 5 años, 2. Base con la información del grupo de 1 a 4 años y 3. Base con los registros de los menores de 1 año.

En el proceso de depuración mencionado en la metodología se realiza eliminación de los registros sin información que para la base total de la población menor de 5 años representa una reducción de 199 registros, quedando una nueva base con 1.155 registros (763 de menores de 1 año y 392 de 1 a 4 años).

Las variables consideradas para la eliminación de registros corresponden al sexo, área de residencia, etnia y la afiliación en salud; que además son variables en las que se recolecta información para todos los grupos de edad; ya que como se mencionó hay algunas variables en las que solo se diligencia información para los menores de 1 año.

En la Tabla 4-3, se observa la nueva distribución de las variables al eliminar los registros.

Tabla 4-3: Proporción total y por grupo de edad de variables seleccionadas, después de la depuración, menores de 5 años

Variable	Total	Grupo de edad	
		Menor de 1 año	De 1 a 4 años
<i>n</i>	1155	763	392
Sexo (%)			
Masculino	51,5	55,2	44,4
Femenino	48,5	44,8	55,6
Área de residencia (%)			
Urbano	51,3	54,7	44,6
Rural	48,7	45,3	55,4
Etnia (%)			
Indígena	28,1	22,5	38,8
ROM (Gitano)	1,5	0,1	4,1
Afrocolombiano	11,3	12,3	9,4
Ninguno	59,1	65	47,7
Afiliación en salud (%)			
Contributivo	6,4	6,8	5,6
Subsidiado	65,6	67,2	62,5
No asegurado	28	26	31,9

Fuente: Elaboración propia.

En la depuración de la base de datos del grupo entre 1 y 4 años, se tuvieron en cuenta las mismas variables del grupo anterior y la reducción fue de 56 registros, en la Tabla 4-4 se presenta la distribución de las variables.

Tabla 4-4: Proporción total y por sexo de variables seleccionadas, después de la depuración, grupo de 1 a 4 años

Variable	Total	Sexo	
		Femenino	Masculino
<i>n</i>	392	218	174
Área de residencia (%)			
Urbano	44,6	41,3	48,9
Rural	55,4	58,7	51,1
Etnia (%)			
Indígena	38,8	40,4	36,8
ROM (Gitano)	4,1	1,4	7,5
Afrocolombiano	9,4	9,2	9,8
Ninguno	47,7	49,1	46
Afiliación en salud (%)			
Contributivo	5,6	5,5	5,7
Subsidiado	62,5	61,5	63,8
No asegurado	31,9	33	30,5

Fuente: Elaboración propia.

Para los menores de 1 año el proceso de depuración y eliminación de registros fue el más exigente, ya que se redujo en alrededor de 48% la cantidad de registros iniciales (906), debido a que se consideraron más variables; sin embargo este grupo de edad es el que cuenta con más variables endógenas propias del certificado de defunción. De esta forma este grupo de edad no solo tiene las más altas tasas de mortalidad, sino que además cuenta con más variables para el análisis.

Tabla 4-5: Proporción total y por sexo de variables seleccionadas, después de la depuración, menores de 1 año

Variable	Total	Sexo	
		Femenino	Masculino
<i>n</i>	468	216	252
Área de residencia (%)			
Urbano	55,1	50,9	58,7
Rural	44,9	49,1	41,3

Tabla 4-5: (Continuación)

Variable	Total	Sexo	
		Femenino	Masculino
Etnia (%)			
Indígena	18,4	20,4	16,7
ROM (Gitano)	0,2	0,5	0
Afrocolombiano	12,2	11,6	12,7
Ninguno	69,2	67,6	70,6
Afiliación en salud (%)			
Contributivo	8,3	8,8	7,9
Subsidiado	62,8	59,7	65,5
No asegurado	28,8	31,5	26,6
Peso al nacer (%)			
Mayor Igual 2500	77,8	78,7	77
Menor 2500	22,2	21,3	23
Edad de la madre (%)			
10 a 14	0,9	0	1,6
15 a 49	98,9	99,5	98,4
50 o mas	0,2	0,5	0
Estado civil de la madre (%)			
Casada	7,3	7,9	6,7
Soltera	17,3	18,1	16,7
Unión libre	71,6	68,5	74,2
Otro	3,8	5,6	2,4
Nivel educativo de la madre (%)			
Ninguno	18,6	19,4	17,9
Primaria	47	46,8	47,2
Secundaria	32,1	30,6	33,3
Técnico o profesional	2,4	3,2	1,6
Tiempo de gestación (%)			
Menor de 38	21,8	22,2	21,4
De 38 a 41	48,1	46,3	49,6
Mayor Igual a 42	30,1	31,5	29

Fuente: Elaboración propia.

4.2 Modelación

Un propósito de plantear un análisis de datos panel es aumentar la precisión de los parámetros estimados, gracias al aumento del número de observaciones al combinar datos de corte seccional con los de series de tiempo, es decir los datos de cada año unidos por un período de un cuatrienio, que permite aumentar la cantidad de registros para un evento de baja prevalencia como el de estudio.

Se logró construir un panel de datos, que tiene como variable dependiente la tasa de mortalidad por desnutrición según los grupos de edad mencionados. Para cada grupo se construyeron tres modelos; el primero incluye solo las variables propias del certificado de defunción, el segundo incluye estas sumadas a las externas de las cuales se tiene información para todos los departamentos y un tercer modelo, con adición de otras variables también externas pero socioeconómicas de importancia, para las cuales solo tienen información de 24 territorios.

Es así como se considera un modelo con una variable dependiente y_{it} la cual depende de n variables independientes endógenas definidas como X_{it} y m variables exógenas como Z_{it}

$$y_{it} = X_{it}\beta_1 + Z_{it}\beta_2 \quad (4-1)$$

Se exploró la relación entre la tasa de mortalidad por desnutrición y algunas variables del certificado de defunción, así como otras variables de diferentes fuentes, en una base de datos de panel construida con información de los departamentos y distrito capital de Colombia entre los años 2009 y 2012. Se estimaron los modelos de regresión realizando ajustes para cumplir los supuestos de correlación serial y heterocedasticidad con el método de errores estándar corregidos para panel PCSE y siguiendo las indicaciones de (Wooldridge, 2002), (Labra & Torrecillas, 2014) y (Aparicio & Márquez, 2005), con transformación doblemente logarítmica como quedó expresado en la ecuación (3-2).

En cuanto a la prueba para comprobar la presencia de correlación contemporánea no fue posible realizarla, debido a que se tiene un panel desbalanceado como consecuencia de la depuración de información, sin embargo el método para panel PCSE la tiene en cuenta y ajusta para la estimación.

Ya se mencionó que al realizar la agrupación de los datos por departamento, algunas de las variables iniciales se subdividieron en sus categorías, sin embargo continúan representando una sola variable original. Para definir la categoría de referencia no se estableció alguna instrucción específica como el valor alto o bajo de su proporción, en este caso se debió más a las categorías que desde la teoría son importantes de tener en cuenta.

Es así como las variables finales se clasificaron de acuerdo a los determinantes que se esperaba dieran respuesta al evento en estudio, de la siguiente forma: variables sociodemográficas, socioeconómicas, de prestación de servicios de salud, individuales y variables de política y protección social.

Las categorías de referencia se definieron de la siguiente forma: para el área de residencia la zona “rural”, para el sexo la categoría “femenino”, en la etnia se seleccionó de referencia “ninguna etnia”, en la edad de la madre el rango entre “15 a 49 años”, el estado civil de referencia es la opción “otro”, “ningún nivel educativo” es la referencia para nivel educativo de la madre, “no asegurado” para la afiliación en salud, el “peso mayor o igual a 2500 gramos” para la variable de peso al nacer y de “38 a 41 semanas” para la edad gestacional.

Como en los tradicionales análisis de regresión el signo de los coeficientes estimados permite ayuda a describir si el riesgo de muerte aumenta o disminuye respecto a la categoría de referencia.

Se reitera que probablemente las cifras finales tengan subregistro, por lo que los resultados de estos modelos pueden ser interpretados como una cota inferior del evento bajo estudio.

4.2.1 Modelos para el grupo de edad menor de 5 años

Los datos integran el período de 2009 a 2012 para 33 desagregaciones territoriales de Colombia que corresponden a 32 departamentos y un distrito (Bogotá). El panel no es balanceado, debido a la eliminación de registros sin información, en el cual se perdieron 3 departamentos: Norte de Santander 2009, Sucre 2009 y Sucre 2011.

- Modelo 1.

El primer modelo que se considera, es en el que solo se incluyen las variables endógenas o propias del certificado de defunción. Es así como en el modelado se examina el efecto en la tasa de defunción de las variables independientes: área de residencia, sexo, etnia y régimen de afiliación a salud.

- Modelo 2.

Este modelo incluye además de las variables endógenas mencionadas en el modelo 1, las definidas como exógenas traídas de otras fuentes de información como son la cobertura de vacunación, la proporción de bajo peso al nacer, la proporción de atendidos por enfermedades infecciosas y parasitarias, la proporción de atendidos por deficiencias nutricionales, las tasas de mortalidad por IRA y EDA para menores de 5 años, el PIB per cápita y la tasa de personas vinculadas a programas de asistencia social. De esta forma en el modelado se examina el efecto de dichas variables endógenas y exógenas en la tasa de defunción.

- Modelo 3.

En este modelo además de las variables mencionadas en el modelo 2 se incluyen las variables exógenas tasa de desempleo, línea de indigencia, coeficiente de Gini y la incidencia de pobreza; que no fueron tenidas en cuenta en el modelo 2, ya que no tienen información para 9 departamentos¹².

Selección de efectos fijos o aleatorios

Para los tres modelos anteriores, en la estimación del modelo de efectos fijos, la prueba F de significancia de dichos efectos (Prob > F = 0.0000) indica que es preferible usar el método de efectos fijos al modelo agrupado. Así mismo en la estimación del modelo de efectos aleatorios y la prueba de Breusch and Pagan¹³ (Prob > chi2 = 0.0000) indica que es preferible usar la estimación de efectos aleatorios en lugar del agrupado.

¹² De estas variables no se tiene información para los departamentos de Arauca, Casanare, Putumayo, Amazonas, Guainía, Guaviare, Vaupés, Vichada y San Andrés islas.

¹³ Breusch and Pagan Lagrangian multiplier test for random effects

Se aplicó la prueba de Hausman para identificar si se utiliza el modelo de efectos fijos o el aleatorio, para los modelos 2 y 3, la hipótesis nula se rechaza ($p=0,0000$) y ($p=0,0025$) respectivamente, es decir, la diferencia entre los coeficientes de efectos aleatorios y fijos sí es sistemática. Por lo tanto, es apropiado usar el método de efectos fijos.

Validación de supuestos

Se aplicó la prueba de autocorrelación de Wooldridge y la prueba indica que existe autocorrelación, que es necesario corregir (modelo 1: p-valor 0,0259; modelo 2: p-valor 0,0206; modelo 3: 0,0164). Se realiza la prueba de heteroscedasticidad de Wald y con p-valor 0,0000, para los tres modelos la prueba indica que la varianza no es constante y existe heteroscedasticidad.

Para solucionar la violación al supuesto se utiliza el estimador con Errores Estándar Corregidos para Panel PCSE ante la presencia de problemas de autocorrelación y heteroscedasticidad.

La ecuación de los modelos 1, 2 y 3 se estimó mediante una regresión Prais Winsten de datos de panel empleando el método PCSE (*Prais-Winsten regression, heteroskedastic panels corrected standard errors*), en la cual el panel es heteroscedástico desbalanceado y autocorrelacionado. Las estimaciones de los modelos se presentan en la tabla 4-6.

Tabla 4-6: Coeficientes de regresión de los modelos tipo panel doble logarítmicos, para la tasa de mortalidad por desnutrición, en menores de cuatro años. Colombia, 2009 – 2012.

Variable	Modelo 1	Modelo 2	Modelo 3
VARIABLES SOCIODEMOGRÁFICAS			
<u>Área de residencia¹</u>			
Urbano	0.0630	0.0671	0.0139
<u>Sexo²</u>			
Masculino	0.0818	0.0909	-0.0064
<u>Etnia³</u>			
Indígena	0.2434*	0.1998*	0.1812*
Afrocolombiano	0.0693	0.1086*	0.0415
Rom	0.0296	0.0041	0.0442
VARIABLES DE PRESTACIÓN DE SERVICIOS DE SALUD			
<u>Afiliación en salud⁴</u>			
Contributivo	-0.0044	0.6833	0.0972*
Subsidiado	0.2134*	0.2173*	0.2765*

Tabla 4-6: (Continuación)

Variable	Modelo 1	Modelo 2	Modelo 3
VARIABLES INDIVIDUALES			
% Bajo peso al nacer (<2500 gr.)		-1.7407*	-0.4195
Proporción de atendidos por deficiencias nutricionales		0.2306	0.1919
Proporción de atendidos por enfermedades infecciosas y parasitarias		-0.5279*	-0.3315
Tasa de mortalidad por EDA		0.0726	0.2752*
Tasa de mortalidad por IRA		-0.0479	-0.065
VARIABLES SOCIOECONÓMICAS			
Producto Interno Bruto (PIB) per cápita.		-0.5183*	0.3728
Tasa de desempleo			0.0216
Coficiente de Gini			-2.8379*
Incidencia de la Pobreza			1.6091*
Línea de indigencia			3.9768*
VARIABLES DE POLÍTICA Y PROTECCIÓN SOCIAL			
Cobertura de vacunación		-0.2089	0.2630
Tasa de personas vinculadas a programas de asistencia social		-0.0217	-0.0211
ESTADÍSTICOS DE LA REGRESIÓN			
Número de departamentos	129	129	93
Constante	0.0508	6.1047*	-25.3501*
R-cuadrado	0.5277	0.6703	0.7459
Prob > chi2	2.804e-22	5.344e-42	3.44E-74
Wald chi2	117.2927	237.1924	405.3136

*Variable significativa (0,05%)

¹ Categoría de referencia: Rural² Categoría de referencia: Femenino³ Categoría de referencia: Ninguna etnia⁴ Categoría de referencia: No asegurado

Fuente: Elaboración propia.

En la etnia para los tres modelos la población indígena es significativa con respecto a no pertenecer a ninguna etnia y con coeficientes cercanos a 0,2 en las tres especificaciones. De igual forma el régimen de afiliación a salud subsidiado, respecto a la población no afiliada; en dichas variables al incrementar su proporción aumenta la proporción de mortalidad¹⁴. De esta forma por ejemplo en el modelo 1, se estima que un incremento de 1% en la etnia indígena se asocia con un incremento del 0,24% en la tasa de muertes por desnutrición y en el modelo 3, un aumento de 1% en el régimen subsidiado se asocia a un incremento de la tasa de mortalidad en 0,28%.

¹⁴ En este análisis, siempre que se mencione mortalidad se hace referencia a la mortalidad por desnutrición.

Además de lo anterior se resalta en el modelo 2, que aumentar el PIB per cápita disminuye la mortalidad y para el modelo 3, se identifica que la variable de mayor coeficiente, en valor absoluto y significativa, incluso comparada con las variables de los otros modelos, resulta ser la línea de indigencia, que indica que con un incremento de 1% de esta, la tasa de mortalidad por desnutrición se asocia a un aumento de 3,98%.

Se resalta que todos los coeficientes se comportan como se espera y son consistentes con la literatura para las variables socioeconómicas incidencia de pobreza, línea de indigencia, afiliación en salud y etnia, ya que el aumento de estas, incrementa también la mortalidad.

Las únicas variables con un comportamiento contra intuitivo fueron el índice de Gini y la proporción de nacidos con Bajo Peso al Nacer, en el modelo 2, lo que puede deberse a una alta correlación con otras variables en el modelo; ya que al aumentar 1% el bajo peso al nacer la tasa de mortalidad se asocia a disminución de 1,74%, lo cual no coincide con lo descrito en la literatura; aunque se presenten estos resultados las variables se mantienen por su relevancia en la literatura. Algo similar sucede en el modelo 3 con el coeficiente de Gini, en el cual el aumento de este en 1% estima que la tasa de mortalidad desciende en 2,8%, lo cual no sería congruente con un coeficiente de desigualdad.

La R^2 más alta es para el modelo 3 (0,746) que indica que el modelo explica 74.6% de la variación total en la variable dependiente, siendo el de mejor ajuste al modelo, como se espera en estos modelos de mortalidad que sufren de dimensionalidad.

4.2.2 Modelos para el grupo de edad de uno a cuatro años

El panel para este grupo también es no balanceado, por la misma situación descrita. En el proceso de eliminación de variables sin información se perdieron 3 departamentos: Amazonas 2009, Caldas 2009 y Norte de Santander 2010.

- **Modelo 1.**

Al igual que en el primer grupo de edad presentado este primer modelo que se considera, solo incluye las variables endógenas o propias del certificado de defunción. Es así como

en el modelado se examina el efecto del área de residencia, el sexo, la etnia y el régimen de afiliación a salud; sobre la tasa de defunción.

- Modelo 2.

Este modelo incluye además de las variables endógenas mencionadas en el modelo 1, las definidas como exógenas traídas de otras fuentes de información como son la cobertura de vacunación, la proporción de bajo peso al nacer, la proporción de atendidos por enfermedades infecciosas y parasitarias, la proporción de atendidos por deficiencias nutricionales, el PIB per cápita y la tasa de personas vinculadas a programas de asistencia social. De esta forma en el modelado se examina el efecto de dichas variables endógenas y exógenas en la tasa de defunción. A diferencia del modelo 2 del grupo de 0 a 5 años en este no se incluyen las tasas de mortalidad por IRA y EDA por que no están definidas para el rango de edad.

- Modelo 3.

En este modelo además de las variables mencionadas en el modelo 2 se incluyen las variables exógenas tasa de desempleo, línea de indigencia, coeficiente de Gini y la incidencia de pobreza; que no fueron tenidas en cuenta en el modelo 2, ya que no tienen información para 9 departamentos¹⁵.

Selección de efectos fijos o aleatorios

Para los tres modelos anteriores, en la estimación del modelo de efectos fijos, la prueba F de significancia de dichos efectos ($\text{Prob} > F = 0.0000$) indica que es preferible usar el método de efectos fijos al modelo agrupado. Así mismo en la estimación del modelo de efectos aleatorios y la prueba de Breusch and Pagan ($\text{Prob} > \text{chi}^2 = 0.0000$) indica que es preferible usar la estimación de efectos aleatorios en lugar del agrupado.

Se aplicó la prueba de Hausman para identificar si se utiliza el modelo de efectos fijos o el aleatorio, para el modelo 1 y 2, la hipótesis nula se rechaza ($p=0,0497$) y ($p=0,0000$) respectivamente, es decir, la diferencia entre los coeficientes de efectos aleatorios y fijos

¹⁵ De estas variables no se tiene información para los departamentos de Arauca, Casanare, Putumayo, Amazonas, Guainía, Guaviare, Vaupés, Vichada y San Andrés islas.

sí es sistemática. Por lo tanto, se concluye que es conveniente utilizar el método de efectos fijos.

En el caso de modelo 1 la prueba arroja ($p=0,5698$) que se interpreta como una evidencia de que no puede rechazarse la hipótesis, por lo tanto se prefiere el modelo de efectos aleatorios.

Validación de supuestos

Se aplicó la prueba de autocorrelación de Wooldridge y la prueba indica que no existe autocorrelación (modelo 1: p-valor 0,9681; modelo 2: p-valor 0,5691; modelo 3: 0,6648). Se realiza la prueba de heteroscedasticidad de Wald y con p-valor 0,0000, para los tres modelos la prueba indica que la varianza no es constante y existe heteroscedasticidad. Para corregir la heteroscedasticidad se utiliza el estimador con Errores Estándar Corregidos para Panel PCSE.

De esta forma la ecuación de los modelos 1, 2 y 3 se estimó mediante una regresión lineal de datos de panel empleando el método PCSE (*Linear regression, heteroskedastic panels corrected standard errors*), en la cual el panel es heteroscedástico desbalanceado y no autocorrelacionado. Las estimaciones de los modelos se presentan en la tabla 4-7.

Tabla 4-7: Coeficientes de regresión de los modelos tipo panel doble logarítmicos, para la tasa de mortalidad por desnutrición en individuos de uno a cuatro años. Colombia, 2009 - 2012.

Variable	Modelo 1	Modelo 2	Modelo 3
Variables sociodemográficas			
<u>Área de residencia¹</u>			
Urbano	0.0804*	0.0729*	0.0362
<u>Sexo²</u>			
Masculino	-0.0212	0.0057	0.0663
<u>Etnia³</u>			
Indígena	0.3124*	0.2620*	0.2145*
Afrocolombiano	-0.0247	0.0091	-0.0024
Rom	0.0370	0.0254	-0.0319
Variables de prestación de servicios de salud			
<u>Afiliación en salud⁴</u>			
Contributivo	-0.1232*	-0.0218	0.0196
Subsidiado	0.1783*	0.1717*	0.1137*

Tabla 4-7: (Continuación)

Variable	Modelo 1	Modelo 2	Modelo 3
VARIABLES INDIVIDUALES			
% Bajo peso al nacer (<2500 gr.)		-1.1409*	-0.7440
Proporción de atendidos por deficiencias nutricionales		0.3378*	0.2432
Proporción de atendidos por enfermedades infecciosas y parasitarias		-0.5243*	-0.4673
VARIABLES SOCIOECONÓMICAS			
Producto Interno Bruto (PIB) per cápita.		-0.3534*	0.0055
Tasa de desempleo			-0.4648
Coefficiente de Gini			-2.6849*
Incidencia de la Pobreza			0.4798
Línea de indigencia			-4,3854*
VARIABLES DE POLÍTICA Y PROTECCIÓN SOCIAL			
Cobertura de vacunación		-0.1027	-0.1608
Tasa de personas vinculadas a programas de asistencia social		-0.0549	-0.0205
ESTADÍSTICOS DE LA REGRESIÓN			
Número de departamentos	128	128	93
Constante	0.0008	4.3392*	22,0354
R-cuadrado	0.5907	0.6925	0.6966
Prob > chi2	2.286e-33	3.686e-47	4.425e-36
Wald chi2	170.1924	255.9669	213.9457

*Variable significativa (0,05%)

¹ Categoría de referencia: Rural

² Categoría de referencia: Femenino

³ Categoría de referencia: Ninguna etnia

⁴ Categoría de referencia: No asegurado

Fuente: Elaboración propia.

Al igual que el grupo de edad anterior la población indígena es significativa con respecto a no pertenecer a ninguna etnia, así como el régimen de afiliación a salud subsidiado, respecto a la población no afiliada; esto se debe a que este grupo de edad hace parte del anterior. Sin embargo, vale la pena resaltar que los tamaños de los coeficientes cambian. Por ejemplo, para la variable etnia crece en casi 0.1 con respecto al modelo del grupo de edad de 0 a 4 años.

En cuanto a la incidencia de la pobreza para este grupo no fueron significativamente diferente de cero, lo cual plantea una diferencia en los grupos, que también se debe revisar en el análisis de los menores de 1 año.

En este grupo de edad se identifica además que al incrementar la proporción de atendidos por enfermedades infecciosas y parasitarias disminuye la tasa de mortalidad

(0,5%); mientras aumentar la proporción de atenciones por deficiencias nutricionales se asocia a un incremento de la tasa de mortalidad por desnutrición (0,34%).

Para el modelo 3, llama la atención la variable de línea de indigencia, ya que al igual que el grupo anterior se esperaría un incremento directamente proporcional con la tasa de mortalidad, sin embargo pasa lo contrario, en donde por cada incremento de 1% de la línea de indigencia disminuye en 4,39% la tasa de mortalidad. También sorprende el coeficiente de Gini que seguido a la línea de indigencia tiene el coeficiente más alto, e indica que con un incremento de 1% de este, la tasa de mortalidad por desnutrición se asocia a una disminución de 2,68%.

Además de lo anterior se resalta en el modelo 2, que aumentar el PIB per cápita disminuye la mortalidad.

La R^2 más alta es para el modelo tres (0,6966) que indica que el modelo explica 69.7% de la variación total en la variable dependiente.

4.2.3 Modelos para el grupo de edad menor de 1 año

El panel no es balanceado, debido a la eliminación de registros sin información, en el cual se perdieron 9 departamentos: Arauca 2012, Casanare 2011, Guainía 2010, Norte de Santander 2009, Putumayo 2010, Sucre 2009, Sucre 2011, Vaupés 2009 y Vaupés 2012. El período continúa siendo 2009 a 2012 para 33 desagregaciones territoriales de Colombia.

- **Modelo 1.**

El primer modelo que se considera, es en el que solo se incluyen las variables endógenas o propias del certificado de defunción y de los modelos planteados hasta el momento es el que más variables incluye porque en el certificado de defunción existen algunas que solo se recolectan para este grupo de edad. Es así como además del área de residencia, el sexo, la etnia y el régimen de afiliación a salud; se tiene algunas de la madre (edad, el nivel educativo, el estado civil) las semanas de gestación y el peso al nacer.

- Modelo 2.

Este modelo incluye además de las variables endógenas mencionadas en el modelo 1, las definidas como exógenas traídas de otras fuentes de información como son la cobertura de vacunación, la proporción de bajo peso al nacer, la proporción de atendidos por enfermedades infecciosas y parasitarias, la proporción de atendidos por deficiencias nutricionales, el PIB per cápita y la tasa de personas vinculadas a programas de asistencia social. De esta forma en el modelado se examina el efecto de dichas variables endógenas y exógenas en la tasa de defunción.

- Modelo 3.

En este modelo además de las variables mencionadas en el modelo 2 se incluyen las variables exógenas tasa de desempleo, línea de indigencia, coeficiente de Gini y la incidencia de pobreza; que no fueron tenidas en cuenta en el modelo 2, ya que no tienen información para 9 departamentos¹⁶.

Selección de efectos fijos o aleatorios

Para los tres modelos anteriores, en la estimación del modelo de efectos fijos, la prueba F de significancia de dichos efectos ($\text{Prob} > F = 0.0000$) indica que es preferible usar el método de efectos fijos al modelo agrupado. Así mismo en la estimación del modelo de efectos aleatorios y la prueba de Breusch and Pagan ($\text{Prob} > \text{chi}^2 = 0.0000$) indica que es preferible usar la estimación de efectos aleatorios en lugar del agrupado.

Se aplicó la prueba de Hausman para identificar si se utiliza el modelo de efectos fijos o el aleatorio, para el modelo 2 y 3, la hipótesis nula se rechaza ($p=0,000$) y ($p=0,000$) respectivamente, es decir, la diferencia entre los coeficientes de efectos aleatorios y fijos sí es sistemática. Por lo tanto, conviene usar el método de efectos fijos.

En el caso de modelo 1 la prueba arrojó un resultado negativo (valor de la $\text{Chi}^2 -1375$) lo cual no debe suceder, pero que a los efectos de la prueba se interpreta como una fuerte evidencia de que no puede rechazarse la hipótesis, por lo tanto se prefiere el modelo de efectos aleatorios.

¹⁶ De estas variables no se tiene información para los departamentos de Arauca, Casanare, Putumayo, Amazonas, Guainía, Guaviare, Vaupés, Vichada y San Andrés islas.

Validación de supuestos

Se aplicó la prueba de autocorrelación de Wooldridge y la prueba indica que existe autocorrelación para el modelo 1, que es necesario corregir (modelo 1: p-valor 0,0482; mientras para los modelos 2 y 3 no se evidencia autocorrelación (modelo 2: p-valor 0,1954, modelo 3: p-valor 0,1135). Se realiza la prueba de heteroscedasticidad de Wald y con p-valor 0,0000, para los tres modelos la prueba indica que la varianza no es constante y existe heteroscedasticidad.

Al igual que en los modelos anteriores se utiliza la estimación PCSE para corregir la autocorrelación y heteroscedasticidad.

La ecuación del modelo 1 se estimó mediante una regresión Prais Winsten de datos de panel empleando el método PCSE (*Prais-Winsten regression, heteroskedastic panels corrected standard errors*), en la cual el panel es heteroscedástico desbalanceado y autocorrelacionado.

La ecuación de los modelos 2 y 3 se estimó mediante una regresión lineal de datos de panel empleando el método PCSE (*Linear regression, heteroskedastic panels corrected standard errors*), en la cual el panel es heteroscedástico desbalanceado y no autocorrelacionado.

Las estimaciones de los modelos se presentan en la tabla 4-8.

Tabla 4-8. Coeficientes de regresión de los modelos tipo panel doble logarítmicos, para la tasa de mortalidad por desnutrición en menores de 1 año. Colombia, 2009 -2012.

Variable	Modelo 1	Modelo 2	Modelo 3
Variables sociodemográficas			
<u>Área de residencia</u> ¹			
Urbano	0.0827	0.1104*	0.05832
<u>Sexo</u> ²			
Masculino	0.0445	0.0264	0.0458
<u>Etnia</u> ³			
Indígena	.2870*	0.2316*	0.1457*
Afrocolombiano	-0.1018*	-0.1043*	-0.0184
Rom	0.2831*	0.2757*	0.2945

Tabla 4-8: (Continuación)

Variable	Modelo 1	Modelo 2	Modelo 3
Edad de la madre⁴			
10 a 14 años	-0.2608*	-0.1932*	-0.0935
50 o más años	0.1279	0.2000	-0.0457
Estado civil de la madre⁵			
Casada	0.0485	0.0615	0.1075*
Soltera	0.1095*	0.1008	0.1548*
Unión Libre	0.2616*	0.2301*	0.1922*
Nivel educativo de la madre⁶			
Primaria	0.0415	0.0789	-0.0132
Secundaria	0.0973*	0.0989*	0.0361
Técnico o profesional	-0.0419	-0.0459	-0.0542
VARIABLES DE PRESTACIÓN DE SERVICIOS DE SALUD			
Afiliación en salud⁷			
Contributivo	-0.0259	-0.0508	0.0345
Subsidiado	0.1236*	0.1205*	0.1163*
VARIABLES INDIVIDUALES			
Peso al nacer reportado en certificado de defunción⁸			
Menor de 2500 gramos	0.0445	0.0105	0.0730*
Semanas de gestación⁹			
Menor de 38 semanas	-0.0822*	-0.0485	-0.0469
Mayor o igual a 42 semanas	0.0892	0.0873	0.0466
Proporción de atendidos por deficiencias nutricionales			
		0.2452	-0.0848
Proporción de atendidos por enfermedades infecciosas y parasitarias			
		-0.5482*	0.4158
VARIABLES SOCIOECONÓMICAS			
Producto Interno Bruto (PIB) per cápita.		-0.6185*	0.0890
Tasa de desempleo			-0.2313
Coefficiente de Gini			-3.1096*
Incidencia de la Pobreza			1.6222*
Línea de indigencia			4.0249*
VARIABLES DE POLÍTICA Y PROTECCIÓN SOCIAL			
Cobertura de vacunación		-0.0116	0.4327
Tasa de personas vinculadas a programas de asistencia social		-0.1081	-0.0060
ESTADÍSTICOS DE LA REGRESIÓN			
Número de departamentos	123	123	93
Constante	-0.0558	5.0180*	-24.9267*
R-cuadrado	0.7299	0.7616	0.8319
Wald chi2	323,0914	516,3649	900,5774
Prob > chi2	8.40E-55	1.382E-94	2.45E-172

*Variable significativa (0,05%)

Categorías de referencia:¹ Rural² Femenino³ Ninguna etnia⁴ 15 a 49 años⁵ Otro estado civil⁶ Ningún nivel educativo⁷ Categoría de referencia: No asegurado⁸ Mayor o igual a 2500 gramos.⁹ 38 a 41 semanas de gestación

Fuente: Elaboración propia.

Las nuevas variables incluidas para este grupo, nivel educativo y semanas de gestación presentan significancia para alguna de sus categorías; sin embargo los coeficientes de cambio porcentual son pequeños.

En el modelo 1 que no se incluyen variables exógenas se identifican que ser indígena o Rom incrementan el riesgo de mortalidad, mientras ser afrocolombiano la disminuye, con respecto a no pertenecer a ninguna etnia. Ser soltero o estar en unión libre incrementa el riesgo de mortalidad comparado con otro estado civil. En el modelo 2 se observa que el incremento de 1% en el PIB se asocia a disminución de la tasa de mortalidad en 0,62%. En el modelo 3 el estado civil de unión libre representa un riesgo más alto de mortalidad que estar soltero o casado con respecto a otro estado civil.

Tanto en el modelo 2 como el 3 se identifica que la unión libre incrementa más el riesgo que ser soltera o casada con respecto a otro estado civil de la madre. Para todos los modelos nuevamente el régimen subsidiado representa un incremento en la mortalidad respecto a no estar afiliado.

Para este grupo de edad vuelven a cobrar importancia las variables socioeconómicas donde el incremento de la incidencia de la pobreza o la línea de indigencia en 1% se asocian a un incremento de 1,6% y 4,0% la tasa de mortalidad por desnutrición.

Se encuentran resultados no esperados como en los grupos anteriores en cuanto al coeficiente de Gini que al incrementarse disminuye la mortalidad.

La R^2 más alta, aunque no con mucha diferencia para los otros modelos es para el modelo tres (0,8319) que indica que el modelo explica 83.2% de la variación total en la variable dependiente.

4.2.4 Consideraciones generales

Los mejores modelos especificados son los que incluyen variables exógenas, ya que obtuvieron los mejores estimadores estadísticamente significativos.

Los resultados muestran unas variables más significativas para el grupo de edad de menores de 1 año; además para este grupo la variación total en la variable dependiente explicada por los modelos expresada en el R^2 es más alta.

Hay resultados relacionados con disminución de la tasa de mortalidad que puede ser inconsistentes porque no es lo que se espera, como es el aumento de la tasa de desempleo, el bajo peso al nacer o la desigualdad entre los ingresos de los hogares medida por el coeficiente de Gini que indica que al aumentar este disminuye la mortalidad por desnutrición

Aunque como se mencionó para el grupo de 1 a 4 años el coeficiente para la línea de indigencia no es consistente con lo esperado, para los otros grupos corresponde a la estimación más alta en la cual la elasticidad¹⁷ estimada de esta respecto a la tasa de mortalidad es 4,02 y es estadísticamente significativa ($p < 0,05$). Por consiguiente, como se ha mencionado para los grupos de 0 a 4 años y los menores de 1 año, al incrementar la línea de indigencia aumenta la tasa de mortalidad por desnutrición.

Aunque en modelos de regresión siempre se busca como mejor modelo el más parsimonioso, es decir el que incluya la menor cantidad de variables explicativas, en este caso para todos los grupos de edad estudiados el mejor modelo es el que incluyo todas las variables, ya que se pueden identificar asociaciones de la mortalidad con variables adicionales a las obtenidas del certificado de defunción.

Finalmente los resultados muestran que no todas las variables con coeficientes significativos presentaron un efecto esperado sobre la mortalidad.

¹⁷ La elasticidad se refiere a que el coeficiente es el previsto cambio porcentual en la variable dependiente para un cambio porcentual en la variable independiente.

5. Conclusiones y recomendaciones

5.1 Conclusiones

Muchos de los análisis en salud se realizan con el enfoque de riesgo, sin embargo este se queda corto comparado con el de determinantes, el de riesgo va más hacia lo individual, mientras el de determinantes tiene en cuenta las personas y su entorno, plantea ver más allá de los factores de riesgo. De aquí la importancia de analizar la mortalidad desde los determinantes, que en este caso se apoya de la aplicación de métodos demográficos y estadísticos.

Definir un enfoque desde los determinantes de la Seguridad Alimentaria y Nutricional es un esfuerzo necesario pero muy reciente que se está abordando en la Encuesta Nacional de la Situación Nutricional en Colombia ENSIN, que se encuentra en ejecución; dicho enfoque requiere de información e indicadores de la situación del país y sus territorios que orienten la influencia de diferentes factores en el estado de salud y nutrición de la población.

Establecer determinantes desde el certificado de defunción es limitado ya que las variables que incluye no dan respuesta a todos los factores que pueden causar una muerte, sumado a esto la falta calidad en el diligenciamiento del certificado hace que se identifiquen diferentes variables sin completar; lo anterior hace necesario traer información exógena de otras fuentes.

Sin embargo las fuentes de información exógena también pueden ser limitadas, ya que no es posible encontrar indicadores y variables para todo el período de estudio que fortalezca el análisis; esto impide dar una respuesta completa a los determinantes que puedan explicar la mortalidad por desnutrición.

El análisis exploratorio mostró que el total de muertes por desnutrición en menores de 5 años para el período estudiado fue de 1354, lo que representa 3,3% del total de muertes en el país en este rango de edad; aunque esta cifra parece una proporción pequeña, es de recordar que es una mortalidad evitable que no se debería presentar y menos iniciando el curso de vida.

La literatura y el comportamiento de los datos permiten identificar una diferencia en la mortalidad para los menores de un año y el grupo entre 1 a 4 años, esto hace necesario realizar un análisis diferenciado, desde lo descriptivo hasta la modelación estadística.

Para el período de análisis el departamento con a mayor proporción de mortalidad para la población menor de 5 años corresponde a la Guajira (12,8 – año 2009), al analizar de forma diferenciada por grupo de edad se identifica que para el caso de menores de 1 año la mortalidad predomina en el departamento del Atlántico (11,9 – año 2012) y en los individuos entre 1 a 4 años corresponde a La Guajira (20,2 – año 2009).

En cuanto a las tasas de mortalidad por cien mil nacidos vivos las más altas para el período de análisis en el grupo menor de 5 años corresponde a Vichada (183,3 – año 2010), para los menores de un año Guainía representa la tasa más alta (656 - año 2012) y en los individuos de 1 a 4 años corresponde a Vichada (95 – año 2010). De esta forma se resalta que los departamentos de Guainía, Vichada, Amazonas y la Guajira son los departamentos que ocupan los primeros lugares con las tasas de mortalidad por cada cien mil habitantes, más altas.

Lo anterior no es un resultado del todo desconocido o no esperado, ya que para otros indicadores relacionados con el estado nutricional de la población, como en el retraso en talla, la ENSIN 2010 para los menores de 5 años, reporto que la prevalencia más alta se presentó en los departamentos de Vaupés (34,6), Amazonas (28,6), La Guajira (27,9) y Guainía (23,4).

El uso de datos estructurados en panel cobra importancia cuando se quieren tener en cuenta los efectos individuales, es decir el conjunto de datos de un individuo en el tiempo. Para el ejercicio realizado los individuos corresponden a los departamentos del país y el tiempo al cuatrienio de estudio.

Se exploró la relación entre la tasa de mortalidad por desnutrición con una serie de variables propias del certificado de defunción y otras externas, sobre una base de datos de panel de 32 departamentos y el distrito capital de Colombia. Se estimaron varios modelos de regresión de efectos fijos y aleatorios según el grupo de edad y la cantidad de variables independientes incluidas en estos. Se aplicaron transformaciones logarítmicas tanto a la variable dependiente como las independientes, para mejorar el ajuste, estandarizar la unidad de medida y facilitar la interpretación.

Para el grupo de menores de 5 años se identificó que los resultados son consistentes con la literatura para las variables socioeconómicas: incidencia de pobreza y línea de indigencia, así como para la afiliación en salud de las variables de prestación de servicios y etnia en las variables sociodemográficas; ya que al incrementar su proporción, también aumenta la proporción de la tasa de mortalidad por desnutrición. Sin embargo se observaron resultados no esperados con el bajo peso al nacer y el coeficiente de Gini que plantean que al aumentarse, se asocian con disminución de la mortalidad por desnutrición.

En el grupo de edad de 1 a 4 años, sucede un comportamiento similar al grupo anterior para la afiliación en salud y la etnia, pero se identifica además que al incrementar en 1% la proporción de atendidos por enfermedades infecciosas y parasitarias disminuye la tasa de mortalidad (0,5%), mientras aumentar las atenciones por deficiencias nutricionales en 1% se asocia a un incremento de la tasa de mortalidad por desnutrición (0,34%). Nuevamente se encuentran resultados no esperados para el coeficiente de Gini, la línea de indigencia y la tasa de desempleo que indica que al aumentar estos disminuye la tasa mortalidad por desnutrición.

Para el grupo de menores de 1 año vuelven a cobrar importancia las variables socioeconómicas, donde el incremento de la pobreza o la indigencia en 1% se asocian a un incremento de 1,6% y 4,0% respectivamente, de la tasa de mortalidad por desnutrición. Dentro de las variables que solo hace parte de este grupo, el estado civil de unión libre representa un riesgo más alto de mortalidad que estar soltero o casado con respecto a otro estado civil. Se observa también que el incremento de 1% en el PIB se asocia a disminución de la tasa de mortalidad en 0,62%

En respuesta a la pregunta de establecer las variables asociadas con las defunciones por desnutrición se encuentra que las socioeconómicas para todos los grupos de edad presentaron los coeficientes más altos y positivos, específicamente las de pobreza e indigencia. También se resaltan aunque no con los coeficientes más altos pero si significativas, la afiliación en salud dentro de las variables de prestación de servicios de salud y la etnia indígena dentro de las variables sociodemográficas, que se asocian a un incremento de la tasa de mortalidad por desnutrición.

La mortalidad por desnutrición no parece estar relacionada con la cobertura de vacunación y los programas de asistencia social, así como el área, sexo y nivel educativo que aunque en algunos casos fueron significativas, el coeficiente de cambio porcentual es muy bajo.

5.2 Recomendaciones

Este trabajo tuvo en cuenta para definir los casos de mortalidad por desnutrición solo los definidos por causa básica, se puede ampliar el análisis teniendo en cuenta las muertes asociadas a desnutrición, es decir los casos de muerte que no son clasificados como causa básica, pero están como otro tipo de causa en el certificado de defunción, que puede permitir identificar la verdadera magnitud del problema.

El certificado de defunción, el cual es el principal instrumento para recolectar la información del hecho vital estudiado, no es diligenciado de manera completa lo cual afecta el análisis de la información. Es necesario informar al DANE el no diligenciamiento de información en los certificados de defunción que permita establecer estrategias o fortalecer las existentes.

Se puede plantear un análisis organizando y agrupando las variables por ejemplo de forma dicotómica, que permita extender la utilización de métodos estadísticos para este tipo de variables.

Se puede realizar el análisis de la mortalidad por desnutrición planteando una metodología de imputación de la información no diligenciada en los certificados de defunción.

Dado el subregistro que se presenta en la información de defunciones es posible realizar un análisis ajustando las tasas de mortalidad con algún método indirecto.

Cabe mencionar que los resultados obtenidos deben ser interpretados con cuidado en la toma de decisiones ya que los resultados pueden variar según las metodologías y variables explicativas incluidas.

A. Anexo: Fuentes de información

- Archivo Nacional de Datos – ANDA

Un catálogo en el que los usuarios pueden explorar, buscar, comparar, solicitar acceso y descargar información relacionada con censos, encuestas por muestreo y uso estadístico de registros administrativos. El ANDA contiene metadatos de operaciones estadísticas que producen el DANE y otras entidades del Sistema Estadístico Nacional – SEN. Asimismo, algunos de los archivos tienen disponibles microdatos de acceso público https://formularios.dane.gov.co/Anda_4_1/index.php/home

- Microdatos DANE

Son los datos suministrados al Departamento Administrativo Nacional de Estadística DANE, en el desarrollo de los censos, encuestas y registros administrativos, en resúmenes numéricos, que no hagan posible deducir de ellos información alguna de carácter individual. El acceso a los microdatos anonimizados es de uso público y de carácter gratuito y está disponible en la página Web del DANE.

http://formularios.dane.gov.co/Anda_4_1/index.php/catalog/MICRODATOS

- Estadísticas Vitales - EEVV

El Sistema de Registro Civil y Estadísticas Vitales está orientado a recoger información sobre algunos hechos que ocurren a lo largo de la vida de las personas, tales como nacimiento, matrimonio, divorcio, anulación, adopción, legitimación, reconocimiento y defunción. Dentro de este sistema se identifican dos subsistemas, registro civil y estadísticas vitales. El primero, en esencia, es una institución pública cuyo objetivo es registrar y almacenar información sobre el acaecimiento de hechos vitales y sus características con fines jurídicos, administrativos y de otra índole; y el segundo, es un proceso que acopia información de los hechos vitales, las características pertinentes a los propios sucesos, para luego compilar, sistematizar, analizar, evaluar, presentar y

difundir esos datos. Actualmente, el Sistema de Estadísticas Vitales en Colombia sólo incluye registros de nacimientos y defunciones fetales y no fetales.

http://formularios.dane.gov.co/Anda_4_1/index.php/catalog/375

- SISPRO

Sistema de Gestión de Datos SGD de SISPRO (Sistema de Información del Ministerio de Salud y Protección Social). El Sistema Integral de Información de la Protección Social es una herramienta que permite obtener, procesar y consolidar la información necesaria para la toma de decisiones que apoyen la elaboración de políticas, el monitoreo regulatorio y la gestión de servicios en cada uno de los niveles y en los procesos esenciales del sector: aseguramiento, financiamiento, oferta, demanda y uso de servicios. Suministra información para toda la ciudadanía. Su diseño se basa en una Bodega de Datos en la cual se concentra la información necesaria para la construcción de indicadores y reportes. Los datos provienen de fuentes de información internas y externas al Ministerio. En el SISPRO se consolida y dispone la información a través de los siguientes componentes: RUAF, RIPS, PILA, SIHO, SGD, SISMED, entre otros.

<https://www.minsalud.gov.co/salud/Paginas/SistemaIntegraldeInformaci%C3%B3nSISPRO.aspx>

- Cuentas Nacionales Departamentales

Las Cuentas Departamentales tienen como objetivo general conocer la estructura y el comportamiento económico de los diferentes Departamentos del país y suministrar elementos de análisis para la planeación y el desarrollo económico Departamental. Las Cuentas Departamentales son una desagregación geográfica de las Cuentas Nacionales, que toman las definiciones y conceptos del Sistema de Cuentas Nacionales y se actualizan en función de los cambios del año base. Sus inicios se remontan a la década de los 80, período en el cual se iniciaron los cálculos en términos de la base 1975.

<https://www.dane.gov.co/index.php/estadisticas-por-tema/cuentas-nacionales/cuentas-nacionales-departamentales>

- Gran Encuesta Integrada de Hogares - GEIH

La Gran Encuesta Integrada de Hogares (GEIH) tiene como objetivo principal proporcionar información básica sobre el tamaño y estructura de la fuerza de trabajo (empleo, desempleo e inactividad), así como de las características sociodemográficas de la población colombiana, permitiendo caracterizar a la población según el sexo, edad, el parentesco, nivel educativo, la afiliación al sistema de seguridad social en salud entre otros. Igualmente, a través de la encuesta se clasifica a las personas según su fuerza de trabajo en ocupadas, desocupadas o inactivas. De esta forma es posible estimar los principales indicadores del mercado laboral colombiano, como son la Tasa Global de Participación (TGP), la Tasa de Ocupación (TO) y la Tasa de Desempleo (TD) Se trata de una investigación continua, que se aplica en todo el territorio del país y que permite la desagregación de resultados para el total nacional, total cabeceras, total centros poblados y rural disperso, cada una de las 23 ciudades capitales y áreas metropolitanas, y San Andrés

<https://www.dane.gov.co/index.php/estadisticas-por-tema/mercado-laboral/empleo-y-desempleo/geih-historicos>

- Programa Ampliado de Inmunizaciones (PAI)

Es una acción conjunta de las naciones del mundo y de organismos internacionales interesados en apoyar acciones tendientes a lograr coberturas universales de vacunación, con el fin de disminuir las tasas de mortalidad y morbilidad causadas por las enfermedades inmunoprevenibles y con un fuerte compromiso de erradicar, eliminar y controlar las mismas. El PAI, es una prioridad política del Gobierno Nacional dentro del marco de las acciones de integralidad del plan decenal de salud pública 2012 - 2021, que desarrolla el Ministerio de Salud y Protección Social. La Subdirección de Enfermedades Transmisibles considera que es fundamental implementar y desarrollar estrategias y actividades que garanticen de manera permanente el mejoramiento continuo de las coberturas de vacunación en la población Colombiana, con mayor énfasis en la población menor de 6 años trascendiendo a la familia en general

<https://www.minsalud.gov.co/salud/publica/Vacunacion/Paginas/pai.aspx>

B. Anexo: Descripción de las variables

Como se describió en la metodología las variables son organizadas de la siguiente forma:

- VARIABLES RELACIONADAS A LOS DETERMINANTES SOCIODEMOGRÁFICOS: edad, sexo, etnia, área de residencia, estado civil de la madre, nivel educativo de la madre, edad de la madre.
- VARIABLES RELACIONADAS A LOS DETERMINANTES SOCIOECONÓMICOS: tasa de desempleo, pobreza, línea de indigencia, coeficiente de Gini, Producto Interno Bruto (PIB) per cápita.
- Variables relacionadas a los determinantes de prestación de servicios de salud: Afiliación en salud.
- VARIABLES RELACIONADAS A LOS DETERMINANTES INDIVIDUALES: peso al nacer reportado en certificado de defunción, bajo peso al nacer (<2500g) según certificado de nacimiento, semanas de gestación, proporción de atendidos por enfermedades infecciosas y parasitarias, proporción de atendidos por deficiencias nutricionales, tasa mortalidad por Enfermedad Diarreica Aguda EDA, tasa de mortalidad por Infecciones Respiratorias Agudas IRA.
- VARIABLES RELACIONADAS A LOS DETERMINANTES DE POLÍTICA Y PROTECCIÓN SOCIAL: Tasa de personas vinculadas a programas de asistencia social, coberturas de vacunación

En la tabla B-1, se realiza una descripción del nombre de las variables codificadas en la base de datos, su definición y fuente de información donde se obtuvieron.

Tabla B-1: Variables de la base de datos.

VARIABLE	DEFINICIÓN	UNIDAD DE MEDIDA	FUENTE
CODPTORE	Departamento de residencia habitual del fallecido (para muerte de menor de un año el de la madre).	Código Dane	EEVV ¹⁸
AÑO	Representa el año en que ocurrió la defunción, para el presente trabajo se utilizó información del período 200 a 2012.	Años	EEVV
TASA_DEF	Representa el número de muertes por deficiencias nutricionales y anemias nutricionales por cada 100.000 habitantes, según el grupo de edad.	Por 100.000	EEVV
SEXO	Identifica el sexo del fallecido, sus categorías son femenino y masculino.	Por 100 (%)	EEVV
GRU_ED2	Realiza una agrupación según la edad del fallecido en menores de 1 año, de 1 a 4 años y de 0 a 4 años.	Años	EEVV
IDPERTET	Define la pertenencia étnica de acuerdo con la cultura, pueblo o rasgos físicos del fallecido o como era reconocido, se agrupo en las siguientes categorías: Indígena, Afrocolombiano, Rom y Ninguna etnia.	Por 100 (%)	EEVV
AREA_RES	Corresponde al área de residencia habitual del fallecido (para muerte de menor de un año la el de la madre), sus categorías son rural y urbano.	Por 100 (%)	EEVV
SEG_SOCIAL	Hace referencia a la afiliación en salud o régimen de seguridad social del fallecido (para muerte de menor de un año el de la madre), se agrupo en las siguientes categorías: contributivo, subsidiado y no asegurado.	Por 100 (%)	EEVV
T_GES	Identifica el número de semanas o tiempo de gestación en las cuales se tuvo el menor, se agrupo en las siguientes categorías: menor de 38 semanas, de 38 a 41 semanas y mayor o igual a 42 semanas.	Por 100 (%)	EEVV
PESO_NAC	Indica el peso al nacer del menor de un año, se agrupo en las siguientes categorías: mayor o igual a 2500 gramos y menor de 2500 gramos.	Por 100 (%)	EEVV
EDAD_MADRE	Define la edad de la madre del menor de un año fallecido, se agrupo en las siguientes categorías: 10 a 14 años, 15 a 49 años y de 50 o más años.	Por 100 (%)	EEVV

¹⁸ Ver abreviaturas al inicio del documento.

Tabla B-1: (Continuación)

VARIABLE	DEFINICIÓN	UNIDAD DE MEDIDA	FUENTE
NIV_EDUM	Indica el último año de estudios aprobado de la madre, se agrupo en las siguientes categorías: primaria, secundaria, técnico o profesional y ningún nivel educativo.	Por 100 (%)	EEVV
EST_CIVM	Indica el estado conyugal de la madre, se agrupo en las siguientes categorías: Casada, Soltera, Unión Libre y Otro estado civil.	Por 100 (%)	EEVV
VACUNACIÓN	Corresponde al porcentaje de las coberturas de vacunación por departamento y año de forma general sin desagregar por el biológico.	Por 100 (%)	Cubo Vacunación SISPRO
NACIMIENTOS	Información de nacimientos que se obtiene a partir de los Certificados de Nacido Vivo.	Personas	DANE
PROYECCIONES	Proyecciones de población realizadas por el DANE tomando como base los resultados ajustados de población del Censo 2005.	Personas	DANE
BAJO_PESO_PRO	Proporción de bajo peso al nacer (<2500g) según el certificado de nacimiento.	Por 100 (%)	EEVV
RIPS_INFECPARA	Proporción de individuos atendidos por enfermedades infecciosas y parasitarias.	Por 100 (%)	Cubo reporte RIPS SISPRO
RIPS_DEF_NUTR	Proporción de individuos atendidos por deficiencias nutricionales.	Por 100 (%)	Cubo reporte RIPS SISPRO
GINI	Coefficiente de GINI, que corresponde a una medida de la desigualdad.	Proporción	DNP
PIB	Producto Interno Bruto (PIB) per cápita. Es un indicador macroeconómico de productividad y desarrollo económico.	PIB / número de habitantes	DANE
POBREZA	Incidencia de la Pobreza. Número de personas pobres expresado como porcentaje del total de la población en un determinado año.	Por 100 (%)	DANE
LINEA_INDIGENCIA	Línea de indigencia. Establece si los hogares cuentan con ingresos suficientes como para cubrir una canasta de alimentos capaz de satisfacer un umbral mínimo de necesidades.	Índice	DANE
TASA_DESEMPLEO	Tasa de desempleo. Expresa el nivel de desocupación entre la población económicamente activa.	Por 100 (%)	DANE
A_SOCIAL	Tasa de personas vinculadas a programas de asistencia social.	Por 100.000	Cubo Asistencia Social SISPRO

Tabla B-1: (Continuación)

VARIABLE	DEFINICIÓN	UNIDAD DE MEDIDA	FUENTE
TASA_EDA	Tasa de mortalidad por EDA (Enfermedad Diarreica Aguda).	Por 100.000	EEVV
TASA_IRA	Tasa de mortalidad por IRA (Infección Respiratoria Aguda).	Por 100.000	EEVV

Fuente: Elaboración propia.

C. Anexo: Comandos Stata

Para el procesamiento y análisis se utilizó el programa estadístico stata, muy utilizado para el manejo grandes archivos de datos y en este caso para el análisis de datos de panel. Se crearon archivos en formato *.do* que corresponde a la sintaxis de stata (lenguaje de comandos), donde se relacionan los comandos utilizados en el trabajo realizado con la base de datos. Entre los comandos y subcomandos básicos se destacan:

Tabla C-1: Comandos Stata

COMANDOS	DESCRIPCIÓN
Comandos para organizar los datos	
<i>clear</i>	Limpia los datos de la memoria
<i>display</i>	Muestra los valores de las funciones
<i>search</i>	Obtienes ayuda (sin precisar comando)
<i>log</i>	Guarda en un archivo todos los comandos y resultados
<i>save</i>	Guarda los datos en un archivo Stata
<i>use</i>	Usa una base de datos de Stata
<i>sort</i>	Clasifica los datos en función a una variable
Comandos para el manejo de los datos	
<i>codebook</i>	Obtener información de las variables, etiquetas y su resumen
<i>count</i>	Contar número de observaciones
<i>drop</i>	Elimina variables
<i>generate</i>	Crea una nueva variable
<i>keep</i>	Mantener subconjunto de variables
<i>label</i>	Manejo de las etiquetas de las variables o base de datos
<i>recode</i>	Recodificar una variable numérica en una categórica
<i>rename</i>	Cambia el nombre de una variable existente
<i>replace</i>	Cambia el valor de una variable en todos o en algunos de los registros
Comandos para procesos estadísticos	
<i>summarize</i>	Muestra un resumen de estadísticos descriptivos
<i>table</i>	Tablas de frecuencias 2x2 o más y resumen de estadísticos descriptivos
<i>tabulate</i>	Tablas de frecuencias de una y dos vías
<i>anova</i>	Análisis de varianza
<i>correlate</i>	Correlación entre variables
<i>glm</i>	Modelos lineales generalizados
<i>logit</i>	Regresión logística (salida en escala logarítmica: coeficientes)
<i>logistic</i>	Regresión logística (salida en escala antilogarítmica: Odds ratio)
<i>regress</i>	Regresión lineal

Tabla C-1: (Continuación)

COMANDOS	DESCRIPCIÓN
Comandos para análisis de datos panel	
<i>xtset</i>	Comando para declarar la base de datos
<i>xtsum</i>	Calcula estadísticas resumidas
<i>xtline</i>	Realiza exploración grafica de datos panel
<i>xtreg</i>	Para regresiones usando datos de panel estáticos
<i>stimates store</i>	Guarda los coeficientes estimados
<i>hausman</i>	Realiza el test que compara entre efectos fijos y efectos aleatorios
<i>xttest0</i>	Breusch-Pagan prueba de LM Lagrange Multiplicador
<i>xtserial</i>	Prueba de autocorrelación de Wooldridge
<i>Xttest2</i>	Prueba de correlación transversal / contemporánea Breusch-Pagan LM Indep.
<i>xttest3</i>	Prueba de heterocedasticidad
<i>xtpcse</i>	Estimadores con Errores Estándar Corregidos para Panel PCSE
<i>xtregar</i>	Regresión controlando autocorrelación con AR(1)
<i>robust</i>	Errores estándar robustos permite la estimación a pesar heterocedasticidad
<i>xtgls</i>	Estimadores de Mínimos Cuadrados Generalizados Factibles FGLS
<i>testparm</i>	Para ver si los efectos de tiempo fijo son necesarios

Fuente: Microglosario de STATA, (Universidad de Navarra, 2014)

D. Anexo: Agrupación de departamentos por región

Tabla D-1: Agrupación de departamentos por región

Región	Código DANE	Departamento
Amazonía y Orinoquía	91	Amazonas
	81	Arauca
	85	Casanare
	94	Guainía
	95	Guaviare
	86	Putumayo
	97	Vaupés
	99	Vichada
Atlántica	8	Atlántico
	13	Bolívar
	20	Cesar
	23	Córdoba
	44	La Guajira
	47	Magdalena
	70	Sucre
	88	Archipiélago de San Andrés, Providencia y Santa Catalina
Bogotá	11	Bogotá
Central	5	Antioquia
	17	Caldas
	18	Caquetá
	41	Huila
	63	Quindío
	66	Risaralda
	73	Tolima

Tabla D-1: (Continuación)

Región	Código DANE	Departamento
Oriental	15	Boyacá
	25	Cundinamarca
	50	Meta
	54	Norte de Santander
	68	Santander
Pacífica	19	Cauca
	27	Choco
	52	Nariño
	76	Valle del Cauca

Fuente: Elaboración propia.

E. Anexo: Mortalidad por desnutrición, según el departamento y la región

Tabla E-1: Proporciones y tasas de mortalidad por desnutrición, según el departamento. Menores de 5 años.

Departamento	2009					2010					2011					2012				
	Casos	%	IC 95%		Tasa	Casos	%	IC 95%		Tasa	Casos	%	IC 95%		Tasa	Casos	%	IC 95%		Tasa
			LI	LS				LI	LS				LI	LS				LI	LS	
Antioquia	23	5,9	3,8	8,8	4,5	20	5,7	3,6	8,9	3,9	19	6,6	4,0	10,1	3,7	13	4,0	2,2	6,9	2,5
Atlántico	27	6,9	4,7	10,0	12,4	26	7,5	5,0	10,9	12,0	18	6,3	3,8	9,7	8,4	12	3,7	2,0	6,5	5,6
Bogotá D.C.	18	4,6	2,8	7,3	3,0	7	2,0	0,9	4,3	1,2	5	1,7	0,6	4,0	0,8	7	2,2	1,0	4,6	1,2
Bolívar	29	7,4	5,1	10,6	14,0	25	7,2	4,8	10,5	12,1	15	5,2	2,9	8,4	7,3	34	10,5	7,5	14,4	16,5
Boyacá	5	1,3	0,5	3,1	4,2	10	2,9	1,5	5,4	8,5	3	1,0	0,2	3,0	2,6	8	2,5	1,2	5,0	7,0
Caldas	2	0,5	0,1	2,0	2,4	2	0,6	0,1	2,3	2,4	2	0,7	0,1	2,5	2,4	2	0,6	0,1	2,5	2,5
Caquetá	11	2,8	1,5	5,1	20,4	10	2,9	1,5	5,4	18,6	9	3,1	1,4	5,9	16,8	5	1,5	0,6	3,8	9,3
Cauca	7	1,8	0,8	3,8	5,4	13	3,7	2,1	6,4	10,0	12	4,2	2,2	7,2	9,2	11	3,4	1,8	6,2	8,4
Cesar	24	6,1	4,0	9,1	21,6	21	6,0	3,9	9,2	19,0	15	5,2	2,9	8,4	13,6	28	8,6	5,9	12,4	25,4
Córdoba	27	6,9	4,7	10,0	15,3	17	4,9	3,0	7,8	9,6	21	7,3	4,6	10,9	11,8	19	5,9	3,7	9,1	10,6
Cundinamarca	2	0,5	0,1	2,0	0,9	2	0,6	0,1	2,3	0,9	9	3,1	1,4	5,9	3,8	3	0,9	0,2	2,9	1,3
Choco	20	5,1	3,2	7,9	30,1	12	3,4	1,9	6,1	18,1	11	3,8	1,9	6,7	16,6	13	4,0	2,2	6,9	19,6
Huila	11	2,8	1,5	5,1	9,8	2	0,6	0,1	2,3	1,8	4	1,4	0,4	3,5	3,5	4	1,2	0,4	3,3	3,5
La Guajira	50	12,8	9,7	16,6	44,9	40	11,5	8,4	15,4	34,9	23	8,0	5,1	11,7	19,6	30	9,2	6,4	13,1	24,9
Magdalena	29	7,4	5,1	10,6	20,6	27	7,7	5,3	11,2	19,3	27	9,4	6,3	13,4	19,4	16	4,9	2,9	8,0	11,5
Meta	9	2,3	1,1	4,5	10,1	11	3,2	1,7	5,7	12,3	14	4,9	2,7	8,0	15,5	15	4,6	2,7	7,7	16,4
Nariño	9	2,3	1,1	4,5	5,5	18	5,2	3,2	8,2	11,0	17	5,9	3,5	9,3	10,4	10	3,1	1,6	5,8	6,1
Norte de Santander	3	0,8	0,2	2,4	2,3	6	1,7	0,7	3,9	4,7	6	2,1	0,8	4,5	4,7	9	2,8	1,4	5,4	7,1

Tabla E-1: (Continuación)

Departamento	2009					2010					2011					2012				
	Casos	%	IC 95%		Tasa	Casos	%	IC 95%		Tasa	Casos	%	IC 95%		Tasa	Casos	%	IC 95%		Tasa
			LI	LS				LI	LS				LI	LS				LI	LS	
Quindío	4	1,0	0,3	2,8	8,7	0	0,0	0,0	0,0	0,0	2	0,7	0,1	2,5	4,3	0	0,0	0,0	0,0	0,0
Risaralda	6	1,5	0,6	3,5	7,8	7	2,0	0,9	4,3	9,1	5	1,7	0,6	4,0	6,5	16	4,9	2,9	8,0	20,9
Santander	4	1,0	0,3	2,8	2,3	10	2,9	1,5	5,4	5,9	8	2,8	1,2	5,4	4,7	2	0,6	0,1	2,5	1,2
Sucre	5	1,3	0,5	3,1	5,9	7	2,0	0,9	4,3	8,3	2	0,7	0,1	2,5	2,4	14	4,3	2,5	7,3	16,7
Tolima	6	1,5	0,6	3,5	4,5	8	2,3	1,1	4,6	6,1	5	1,7	0,6	4,0	3,9	10	3,1	1,6	5,8	7,8
Valle del Cauca	22	5,6	3,6	8,5	6,1	12	3,4	1,9	6,1	3,3	7	2,4	1,0	4,9	1,9	15	4,6	2,7	7,7	4,1
Arauca	2	0,5	0,1	2,0	5,9	2	0,6	0,1	2,3	6,0	0	0,0	0,0	0,0	0,0	2	0,6	0,1	2,5	6,1
Casanare	1	0,3	0,0	1,6	2,8	2	0,6	0,1	2,3	5,7	2	0,7	0,1	2,5	5,6	1	0,3	0,0	2,0	2,8
Putumayo	7	1,8	0,8	3,8	17,6	2	0,6	0,1	2,3	5,1	4	1,4	0,4	3,5	10,1	3	0,9	0,2	2,9	7,6
A. de San Andrés	1	0,3	0,0	1,6	15,6	1	0,3	0,0	1,8	15,7	0	0,0	0,0	0,0	0,0	0	0,0	0,0	0,0	0,0
Amazonas	8	2,0	1,0	4,1	79,1	1	0,3	0,0	1,8	9,8	2	0,7	0,1	2,5	19,5	7	2,2	1,0	4,6	68,0
Guainía	5	1,3	0,5	3,1	102,6	3	0,9	0,2	2,7	60,0	7	2,4	1,0	4,9	138,9	7	2,2	1,0	4,6	137,6
Guaviare	2	0,5	0,1	2,0	14,4	0	0,0	0,0	0,0	0,0	2	0,7	0,1	2,5	14,3	1	0,3	0,0	2,0	7,1
Vaupés	2	0,5	0,1	2,0	33,4	1	0,3	0,0	1,8	16,8	1	0,4	0,0	1,9	16,8	2	0,6	0,1	2,5	33,5
Vichada	7	1,8	0,8	3,8	76,7	17	4,9	3,0	7,8	183,3	8	2,8	1,2	5,4	84,8	5	1,5	0,6	3,8	52,1
Otro	4	1,0	0,3	2,8	*	7	2,0	0,9	4,3	*	3	1,0	0,2	3,0	*	1	0,3	0,0	2,0	*
Total Colombia	392	100	-	-	9,2	349	100	-	-	8,2	288	100	-	-	6,7	325	100	-	-	7,6

* Otro: corresponde a 15 registros que no tenían información del departamento de residencia o la residencia es en el extranjero.

** Tasa por 100.000 habitantes

Fuente: Elaboración propia.

Tabla E-2: Proporciones y tasas de mortalidad por desnutrición, según la región. Menores de 5 años.

Región	2009					2010					2011					2012				
	Casos	%	IC 95%		Tasa	Casos	%	IC 95%		Tasa	Casos	%	IC 95%		Tasa	Casos	%	IC 95%		Tasa
			LI	LS				LI	LS				LI	LS				LI	LS	
Amazonía y Orinoquía	34	8,7	6,2	12,0	22,3	28	8,0	5,5	11,5	18,3	26	9,0	6,0	13,0	17,0	28	8,6	5,9	12,4	18,3
Atlántica	192	49,0	43,9	54,0	18,2	164	47,0	41,7	52,4	15,5	121	42,0	36,3	48,0	11,4	153	47,1	41,6	52,7	14,4
Bogotá D.C.	18	4,6	2,8	7,3	3,0	7	2,0	0,9	4,3	1,2	5	1,7	0,6	4,0	0,8	7	2,2	1,0	4,6	1,2
Central	63	16,1	12,7	20,2	6,2	49	14,0	10,7	18,2	4,8	46	16,0	11,9	20,7	4,5	50	15,4	11,7	19,9	4,9
Oriental	23	5,9	3,8	8,8	3,1	39	11,2	8,2	15,1	5,3	40	13,9	10,1	18,4	5,4	37	11,4	8,2	15,5	5,0
Pacífica	58	14,8	11,5	18,8	8,0	55	15,8	12,2	20,1	7,6	47	16,3	12,2	21,1	6,5	49	15,1	11,5	19,5	6,8
Otro	4	1,0	0,3	2,8	*	7	2,0	0,9	4,3	*	3	1,0	0,2	3,0	*	1	0,3	0,0	2,0	*
Total general	392	100	-	-	9,2	349	100	-	-	8,2	288	100	-	-	6,7	325	100	-	-	7,6

* Otro: corresponde a 15 registros que no tenían información del departamento de residencia o la residencia es en el extranjero.

** Tasa por 100.000 habitantes

Fuente: Elaboración propia.

Tabla E-3: Proporciones y tasas de mortalidad por desnutrición, según el departamento. Grupo de 1 a 4 años.

Departamento	2009					2010					2011					2012				
	Casos	%	IC 95%		Tasa	Casos	%	IC 95%		Tasa	Casos	%	IC 95%		Tasa	Casos	%	IC 95%		Tasa
			LI	LS				LI	LS				LI	LS				LI	LS	
Antioquia	8	6,5	2,8	12,3	1,9	10	8,5	4,1	15,0	2,4	9	8,3	3,9	15,2	2,2	5	5,1	1,7	11,5	1,2
Atlántico	2	1,6	0,2	5,7	1,1	5	4,2	1,4	9,6	2,9	3	2,8	0,6	7,9	1,7	4	4,1	1,1	10,1	2,3
Bogotá D.C.	6	4,8	1,8	10,2	1,3	3	2,5	0,5	7,3	0,6	0	0,0	0,0	0,0	0,0	3	3,1	0,6	8,7	0,6
Bolívar	11	8,9	4,5	15,3	6,6	5	4,2	1,4	9,6	3,0	5	4,6	1,5	10,5	3,0	7	7,1	2,9	14,2	4,3
Boyacá	2	1,6	0,2	5,7	2,1	2	1,7	0,2	6,0	2,1	0	0,0	0,0	0,0	0,0	6	6,1	2,3	12,9	6,6
Caldas	1	0,8	0,0	4,4	1,5	1	0,9	0,0	4,6	1,5	2	1,9	0,2	6,5	3,0	0	0,0	0,0	0,0	0,0
Caquetá	4	3,2	0,9	8,1	9,3	4	3,4	0,9	8,5	9,3	1	0,9	0,0	5,1	2,3	2	2,0	0,3	7,2	4,7
Cauca	2	1,6	0,2	5,7	1,9	3	2,5	0,5	7,3	2,9	4	3,7	1,0	9,2	3,8	2	2,0	0,3	7,2	1,9
Cesar	9	7,3	3,4	13,3	10,1	7	5,9	2,4	11,8	7,9	6	5,6	2,1	11,7	6,8	8	8,2	3,6	15,5	9,1
Córdoba	2	1,6	0,2	5,7	1,4	8	6,8	3,0	12,9	5,7	4	3,7	1,0	9,2	2,8	8	8,2	3,6	15,5	5,6
Cundinamarca	0	0,0	0,0	0,0	0,0	0	0,0	0,0	0,0	0,0	4	3,7	1,0	9,2	2,1	0	0,0	0,0	0,0	0,0
Choco	5	4,0	1,3	9,2	9,5	3	2,5	0,5	7,3	5,7	8	7,4	3,3	14,1	15,2	3	3,1	0,6	8,7	5,7
Huila	5	4,0	1,3	9,2	5,5	0	0,0	0,0	0,0	0,0	2	1,9	0,2	6,5	2,2	0	0,0	0,0	0,0	0,0
La Guajira	25	20,2	13,5	28,3	28,4	16	13,6	8,0	21,1	17,7	12	11,1	5,9	18,6	12,9	13	13,3	7,3	21,6	13,7
Magdalena	10	8,1	3,9	14,3	8,8	7	5,9	2,4	11,8	6,2	9	8,3	3,9	15,2	8,1	4	4,1	1,1	10,1	3,6
Meta	2	1,6	0,2	5,7	2,8	4	3,4	0,9	8,5	5,6	6	5,6	2,1	11,7	8,3	3	3,1	0,6	8,7	4,1
Nariño	3	2,4	0,5	6,9	2,3	7	5,9	2,4	11,8	5,3	10	9,3	4,5	16,4	7,6	2	2,0	0,3	7,2	1,5
Norte de Santander	0	0,0	0,0	0,0	0,0	1	0,9	0,0	4,6	1,0	1	0,9	0,0	5,1	1,0	2	2,0	0,3	7,2	2,0
Quindío	1	0,8	0,0	4,4	2,7	0	0,0	0,0	0,0	0,0	2	1,9	0,2	6,5	5,4	0	0,0	0,0	0,0	0,0
Risaralda	3	2,4	0,5	6,9	4,9	4	3,4	0,9	8,5	6,5	4	3,7	1,0	9,2	6,5	6	6,1	2,3	12,9	9,8
Santander	2	1,6	0,2	5,7	1,5	2	1,7	0,2	6,0	1,5	2	1,9	0,2	6,5	1,5	0	0,0	0,0	0,0	0,0
Sucre	1	0,8	0,0	4,4	1,5	0	0,0	0,0	0,0	0,0	0	0,0	0,0	0,0	0,0	4	4,1	1,1	10,1	6,0
Tolima	1	0,8	0,0	4,4	0,9	4	3,4	0,9	8,5	3,8	4	3,7	1,0	9,2	3,8	2	2,0	0,3	7,2	1,9
Valle del Cauca	5	4,0	1,3	9,2	1,7	4	3,4	0,9	8,5	1,4	2	1,9	0,2	6,5	0,7	2	2,0	0,3	7,2	0,7
Arauca	0	0,0	0,0	0,0	0,0	2	1,7	0,2	6,0	7,4	0	0,0	0,0	0,0	0,0	1	1,0	0,0	5,6	3,8
Casanare	1	0,8	0,0	4,4	3,6	2	1,7	0,2	6,0	7,1	1	0,9	0,0	5,1	3,5	1	1,0	0,0	5,6	3,5
Putumayo	2	1,6	0,2	5,7	6,3	1	0,9	0,0	4,6	3,2	1	0,9	0,0	5,1	3,2	1	1,0	0,0	5,6	3,2
A. de San Andrés	0	0,0	0,0	0,0	0,0	1	0,9	0,0	4,6	19,6	0	0,0	0,0	0,0	0,0	0	0,0	0,0	0,0	0,0
Amazonas	2	1,6	0,2	5,7	25,0	0	0,0	0,0	0,0	0,0	2	1,9	0,2	6,5	24,7	4	4,1	1,1	10,1	49,2
Guainía	2	1,6	0,2	5,7	51,4	1	0,9	0,0	4,6	25,4	0	0,0	0,0	0,0	0,0	3	3,1	0,6	8,7	74,9

Tabla E-3: (Continuación)

Departamento	2009					2010					2011					2012				
	Casos	%	IC 95%		Tasa	Casos	%	IC 95%		Tasa	Casos	%	IC 95%		Tasa	Casos	%	IC 95%		Tasa
			LI	LS				LI	LS				LI	LS				LI	LS	
Guaviare	2	1,6	0,2	5,7	18,1	0	0,0	0,0	0,0	0,0	0	0,0	0,0	0,0	0,0	0	0,0	0,0	0,0	0,0
Vaupés	0	0,0	0,0	0,0	0,0	1	0,9	0,0	4,6	21,1	0	0,0	0,0	0,0	0,0	1	1,0	0,0	5,6	21,1
Vichada	2	1,6	0,2	5,7	27,5	7	5,9	2,4	11,8	95,0	4	3,7	1,0	9,2	53,4	1	1,0	0,0	5,6	13,1
Otro	3	2,4	0,5	6,9	*	3	2,5	0,5	7,3	*	0	0,0	0,0	0,0	*	0	0,0	0,0	0,0	*
Total Colombia	124	100	-	-	3,62	118	100	-	-	3,449	108	100	-	-	3,156	98	100	-	-	2,861

* Otro: corresponde a 15 registros que no tenían información del departamento de residencia o la residencia es en el extranjero.

** Tasa por 100.000 habitantes

Fuente: Elaboración propia.

Tabla E-4: Proporciones y tasas de mortalidad por desnutrición, según la región. Grupo de 1 a 4 años.

Región	2009					2010					2011					2012				
	Casos	%	IC 95%		Tasa	Casos	%	IC 95%		Tasa	Casos	%	IC 95%		Tasa	Casos	%	IC 95%		Tasa
			LI	LS				LI	LS				LI	LS				LI	LS	
Amazonía y Orinoquía	11	8,9	4,5	15,3	9,0	14	11,9	6,6	19,1	11,5	8	7,4	3,3	14,1	6,6	12	12,2	6,5	20,4	9,8
Atlántica	60	48,4	39,3	57,5	7,1	49	41,5	32,5	51,0	5,8	39	36,1	27,1	45,9	4,6	48	49,0	38,7	59,3	5,7
Bogotá D.C.	6	4,8	1,8	10,2	1,3	3	2,5	0,5	7,3	0,6	0	0,0	0,0	0,0	0,0	3	3,1	0,6	8,7	0,6
Central	23	18,6	12,1	26,5	2,8	23	19,5	12,8	27,8	2,8	24	22,2	14,8	31,2	2,9	15	15,3	8,8	24,0	1,8
Oriental	6	4,8	1,8	10,2	1,0	9	7,6	3,6	14,0	1,5	13	12,0	6,6	19,7	2,2	11	11,2	5,7	19,2	1,9
Pacífica	15	12,1	6,9	19,2	2,6	17	14,4	8,6	22,1	2,9	24	22,2	14,8	31,2	4,2	9	9,2	4,3	16,7	1,6
Otro	3	2,4	0,5	6,9	*	3	2,5	0,5	7,3	*	0	0,0	0,0	0,0	*	0	0,0	0,0	0,0	*
Total general	124	100	-	-	3,6	118	100	-	-	3,4	108	100	-	-	3,2	98	100	-	-	2,9

* Otro: corresponde a 15 registros que no tenían información del departamento de residencia o la residencia es en el extranjero.

** Tasa por 100.000 habitantes

Fuente: Elaboración propia.

Tabla E-5: Proporciones y tasas de mortalidad por desnutrición, según el departamento. Menores de 1 año.

Departamento	2009					2010					2011					2012				
	Casos	%	IC 95%		Tasa	Casos	%	IC 95%		Tasa	Casos	%	IC 95%		Tasa	Casos	%	IC 95%		Tasa
			LI	LS				LI	LS				LI	LS				LI	LS	
Antioquia	15	5,6	3,2	9,1	14,6	10	4,3	2,1	7,8	9,6	10	5,6	2,7	10,0	9,6	8	3,5	1,5	6,8	7,6
Atlántico	25	9,3	6,1	13,5	58,3	21	9,1	5,7	13,6	49,3	15	8,3	4,7	13,4	35,4	8	3,5	1,5	6,8	18,9
Bogotá D.C.	12	4,5	2,3	7,7	10,1	4	1,7	0,5	4,4	3,3	5	2,8	0,9	6,4	4,1	4	1,8	0,5	4,5	3,3
Bolívar	18	6,7	4,0	10,4	43,6	20	8,7	5,4	13,1	48,1	10	5,6	2,7	10,0	24,0	27	11,9	8,0	16,8	64,6
Boyacá	3	1,1	0,2	3,2	12,9	8	3,5	1,5	6,7	35,0	3	1,7	0,4	4,8	13,4	2	0,9	0,1	3,2	9,1
Caldas	1	0,4	0,0	2,1	6,0	1	0,4	0,0	2,4	6,1	0	0,0	0,0	0,0	0,0	2	0,9	0,1	3,2	12,3
Caquetá	7	2,6	1,1	5,3	64,9	6	2,6	1,0	5,6	55,8	8	4,4	1,9	8,6	74,6	3	1,3	0,3	3,8	27,9
Cauca	5	1,9	0,6	4,3	19,5	10	4,3	2,1	7,8	38,4	8	4,4	1,9	8,6	30,5	9	4,0	1,8	7,4	34,0
Cesar	15	5,6	3,2	9,1	67,6	14	6,1	3,4	10,0	63,2	9	5,0	2,3	9,3	40,7	20	8,8	5,5	13,3	90,4
Córdoba	25	9,3	6,1	13,5	69,6	9	3,9	1,8	7,3	25,0	17	9,4	5,6	14,7	47,1	11	4,9	2,4	8,5	30,3
Cundinamarca	2	0,8	0,1	2,7	4,3	2	0,9	0,1	3,1	4,3	5	2,8	0,9	6,4	10,6	3	1,3	0,3	3,8	6,3
Choco	15	5,6	3,2	9,1	110,5	9	3,9	1,8	7,3	66,6	3	1,7	0,4	4,8	22,2	10	4,4	2,1	8,0	74,3
Huila	6	2,2	0,8	4,8	26,8	2	0,9	0,1	3,1	8,9	2	1,1	0,1	4,0	8,8	4	1,8	0,5	4,5	17,5
La Guajira	25	9,3	6,1	13,5	107,0	24	10,4	6,8	15,1	99,5	11	6,1	3,1	10,7	44,5	17	7,5	4,4	11,7	67,3
Magdalena	19	7,1	4,3	10,9	69,2	20	8,7	5,4	13,1	72,8	18	10,0	6,0	15,3	65,5	12	5,3	2,8	9,1	43,6
Meta	7	2,6	1,1	5,3	38,8	7	3,0	1,2	6,1	38,9	8	4,4	1,9	8,6	44,1	12	5,3	2,8	9,1	65,5
Nariño	6	2,2	0,8	4,8	18,4	11	4,8	2,4	8,4	33,7	7	3,9	1,6	7,9	21,4	8	3,5	1,5	6,8	24,5
Norte de Santander	3	1,1	0,2	3,2	12,0	5	2,2	0,7	5,0	19,6	5	2,8	0,9	6,4	19,5	7	3,1	1,3	6,3	27,2
Quindío	3	1,1	0,2	3,2	32,3	0	0,0	0,0	0,0	0,0	0	0,0	0,0	0,0	0,0	0	0,0	0,0	0,0	0,0
Risaralda	3	1,1	0,2	3,2	19,5	3	1,3	0,3	3,8	19,4	1	0,6	0,0	3,1	6,5	10	4,4	2,1	8,0	64,8
Santander	2	0,8	0,1	2,7	5,9	8	3,5	1,5	6,7	23,5	6	3,3	1,2	7,1	17,5	2	0,9	0,1	3,2	5,9
Sucre	4	1,5	0,4	3,8	23,8	7	3,0	1,2	6,1	41,4	2	1,1	0,1	4,0	11,8	10	4,4	2,1	8,0	58,9
Tolima	5	1,9	0,6	4,3	19,3	4	1,7	0,5	4,4	15,6	1	0,6	0,0	3,1	3,9	8	3,5	1,5	6,8	31,4
Valle del Cauca	17	6,3	3,7	10,0	23,5	8	3,5	1,5	6,7	11,0	5	2,8	0,9	6,4	6,8	13	5,7	3,1	9,6	17,6
Arauca	2	0,8	0,1	2,7	29,7	0	0,0	0,0	0,0	0,0	0	0,0	0,0	0,0	0,0	1	0,4	0,0	2,4	15,4
Casanare	0	0,0	0,0	0,0	0,0	0	0,0	0,0	0,0	0,0	1	0,6	0,0	3,1	14,1	0	0,0	0,0	0,0	0,0
Putumayo	5	1,9	0,6	4,3	62,8	1	0,4	0,0	2,4	12,5	3	1,7	0,4	4,8	37,6	2	0,9	0,1	3,2	25,1
A. de San Andrés	1	0,4	0,0	2,1	78,4	0	0,0	0,0	0,0	0,0	0	0,0	0,0	0,0	0,0	0	0,0	0,0	0,0	0,0
Amazonas	6	2,2	0,8	4,8	286,4	1	0,4	0,0	2,4	47,1	0	0,0	0,0	0,0	0,0	3	1,3	0,3	3,8	138,9
Guainía	3	1,1	0,2	3,2	305,2	2	0,9	0,1	3,1	190,1	7	3,9	1,6	7,9	656,0	4	1,8	0,5	4,5	370,4

Tabla E-5: (Continuación)

Departamento	2009					2010					2011					2012				
	Casos	%	IC 95%		Tasa	Casos	%	IC 95%		Tasa	Casos	%	IC 95%		Tasa	Casos	%	IC 95%		Tasa
			LI	LS				LI	LS				LI	LS				LI	LS	
Guaviare	0	0,0	0,0	0,0	0,0	0	0,0	0,0	0,0	0,0	2	1,1	0,1	4,0	70,9	1	0,4	0,0	2,4	35,3
Vaupés	2	0,8	0,1	2,7	161,9	0	0,0	0,0	0,0	0,0	1	0,6	0,0	3,1	81,0	1	0,4	0,0	2,4	80,8
Vichada	5	1,9	0,6	4,3	269,4	10	4,3	2,1	7,8	525,2	4	2,2	0,6	5,6	205,4	4	1,8	0,5	4,5	202,1
Otro	1	0,4	0,0	2,1		4	1,7	0,5	4,4		3	1,7	0,4	4,8		1	0,4	0,0	2,4	
Total Colombia	268	100	-	-	31,37	231	100	-	-	26,9	180	100	-	-	20,87	227	100	-	-	26,24

* Otro: corresponde a 15 registros que no tenían información del departamento de residencia o la residencia es en el extranjero.

** Tasa por 100.000 habitantes

Fuente: Elaboración propia.

Tabla E-6: Proporciones y tasas de mortalidad por desnutrición, según la región. Menores de 1 año.

Región	2009					2010					2011					2012				
	Casos	%	IC 95%		Tasa	Casos	%	IC 95%		Tasa	Casos	%	IC 95%		Tasa	Casos	%	IC 95%		Tasa
			LI	LS				LI	LS				LI	LS				LI	LS	
Amazonía y Orinoquía	23	8,6	5,5	12,6	74,9	14	6,1	3,4	10,0	45,5	18	10,0	6,0	15,3	58,5	16	7,1	4,1	11,2	51,8
Atlántica	132	49,3	43,1	55,4	62,5	115	49,8	43,2	56,4	54,2	82	45,6	38,1	53,1	38,5	105	46,3	39,6	53,0	49,2
Bogotá D.C.	12	4,5	2,3	7,7	10,1	4	1,7	0,5	4,4	3,3	5	2,8	0,9	6,4	4,1	4	1,8	0,5	4,5	3,3
Central	40	14,9	10,9	19,8	19,6	26	11,3	7,5	16,1	12,8	22	12,2	7,8	17,9	10,8	35	15,4	11,0	20,8	17,0
Oriental	17	6,3	3,7	10,0	11,6	30	13,0	8,9	18,0	20,4	27	15,0	10,1	21,1	18,3	26	11,5	7,6	16,3	17,6
Pacífica	43	16,0	11,9	21,0	29,8	38	16,5	11,9	21,9	26,2	23	12,8	8,3	18,6	15,8	40	17,6	12,9	23,2	27,3
Otro	1	0,4	0,0	2,1	*	4	1,7	0,5	4,4	*	3	1,7	0,4	4,8	*	1	0,4	0,0	2,4	*
Total general	268	100	-	-	31,4	231	100	-	-	26,9	180	100	-	-	20,9	227	100	-	-	26,2

* Otro: corresponde a 15 registros que no tenían información del departamento de residencia o la residencia es en el extranjero.

** Tasa por 100.000 habitantes

Fuente: Elaboración propia.

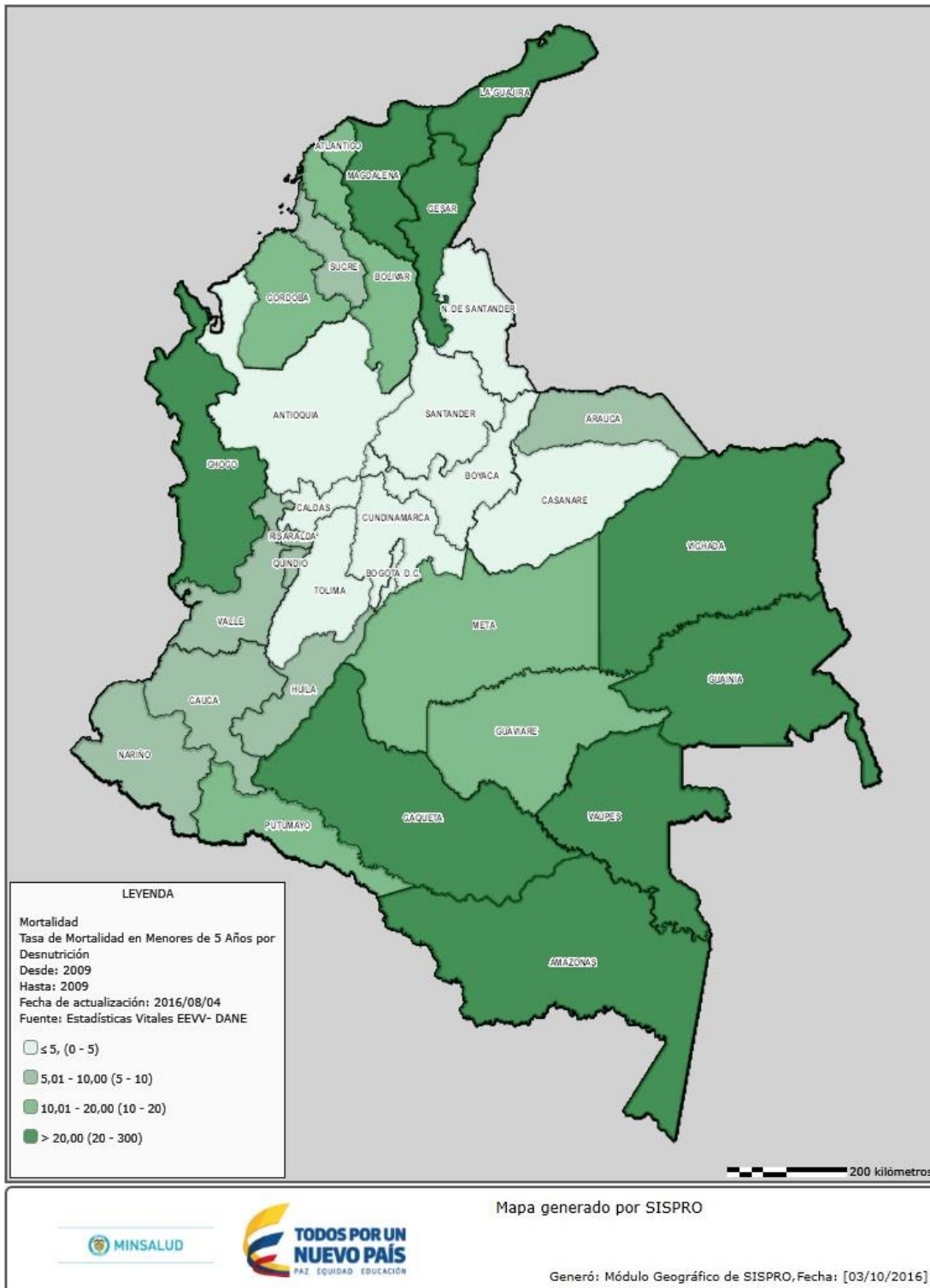
F. Anexo: Mapas de Mortalidad por desnutrición

Los mapas que se relacionan presentan la tasa de mortalidad por desnutrición en menores de 5 años por cada 100.000 habitantes, están elaborados por departamento según el año, para el período de 2009 a 2012.

Los mapas se generaron en el módulo geográfico del Sistema Integral de Información de Protección Social SISPRO, que se constituye en una herramienta útil que utiliza el Ministerio de Salud y Protección Social para analizar la información en salud desde cada región¹⁹.

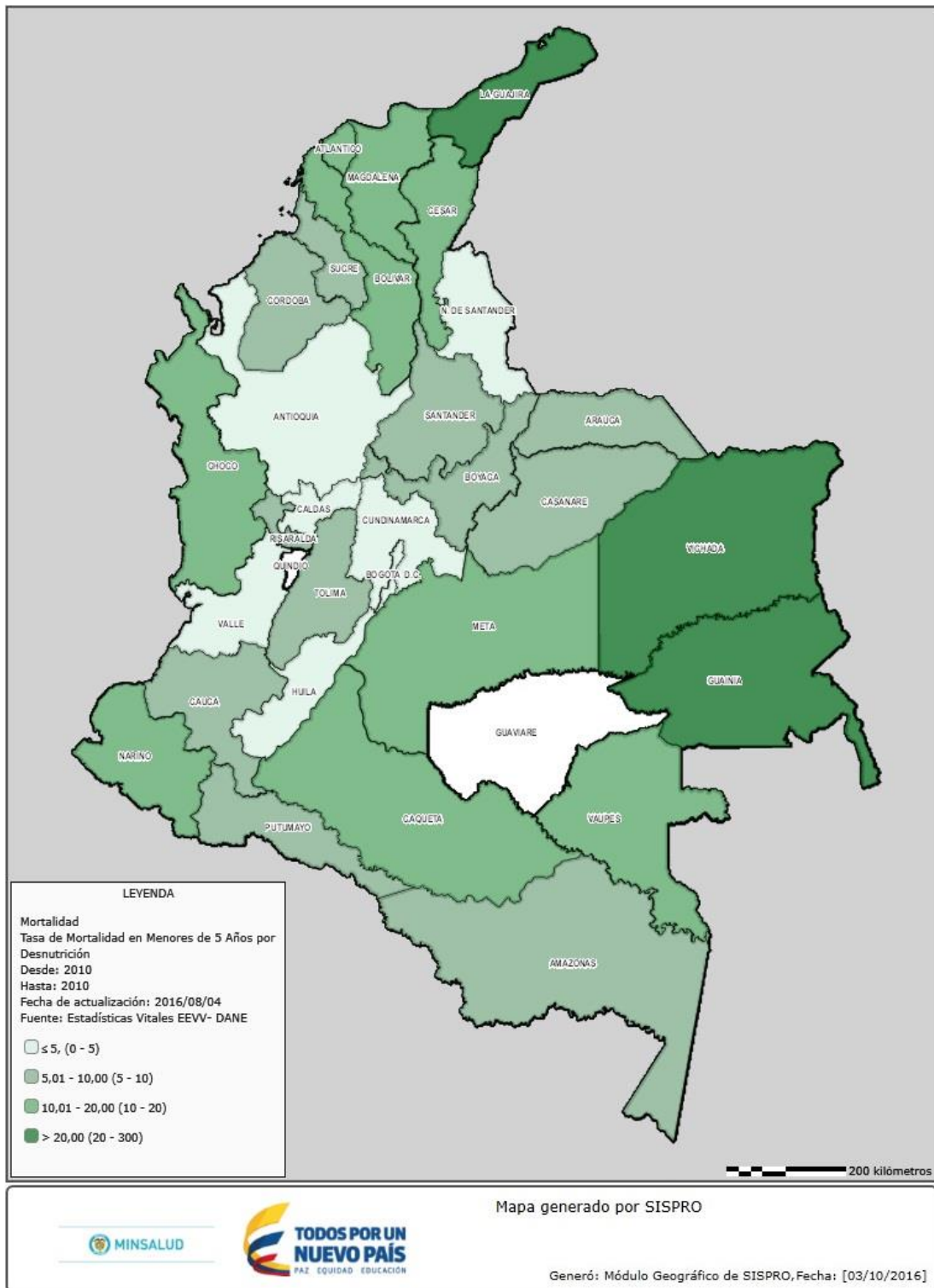
¹⁹ Para mayor información sobre esta herramienta, ingresar a los enlaces: <http://www.sispro.gov.co/> y <http://sig.sispro.gov.co/sigmisp/>

Figura F-1: Tasa de mortalidad por desnutrición en menores de 5 años, según departamento, Colombia 2009.



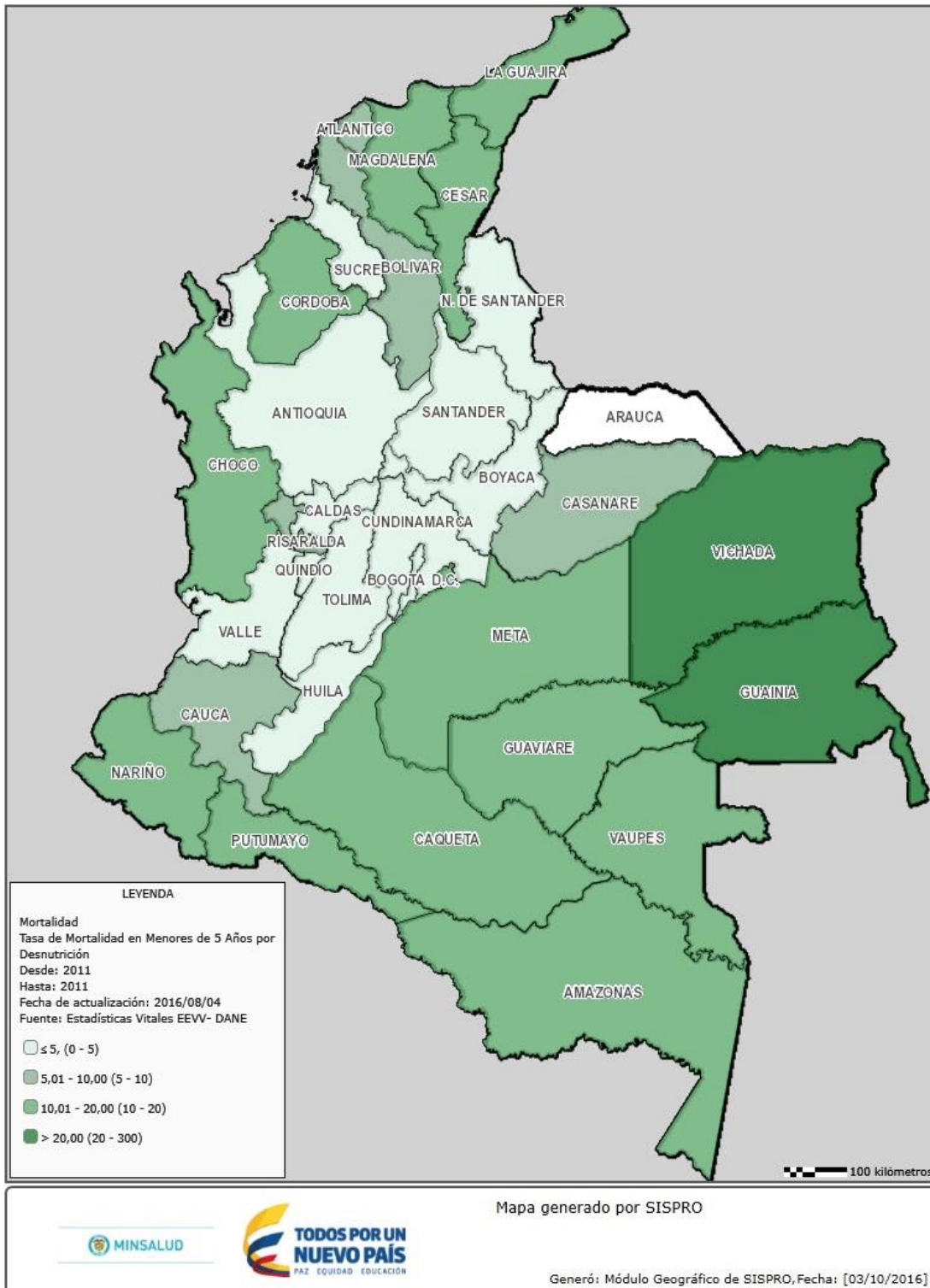
Fuente: Módulo geográfico de SISPRO

Figura F-2: Tasa de mortalidad por desnutrición en menores de 5 años, según departamento, Colombia 2010.



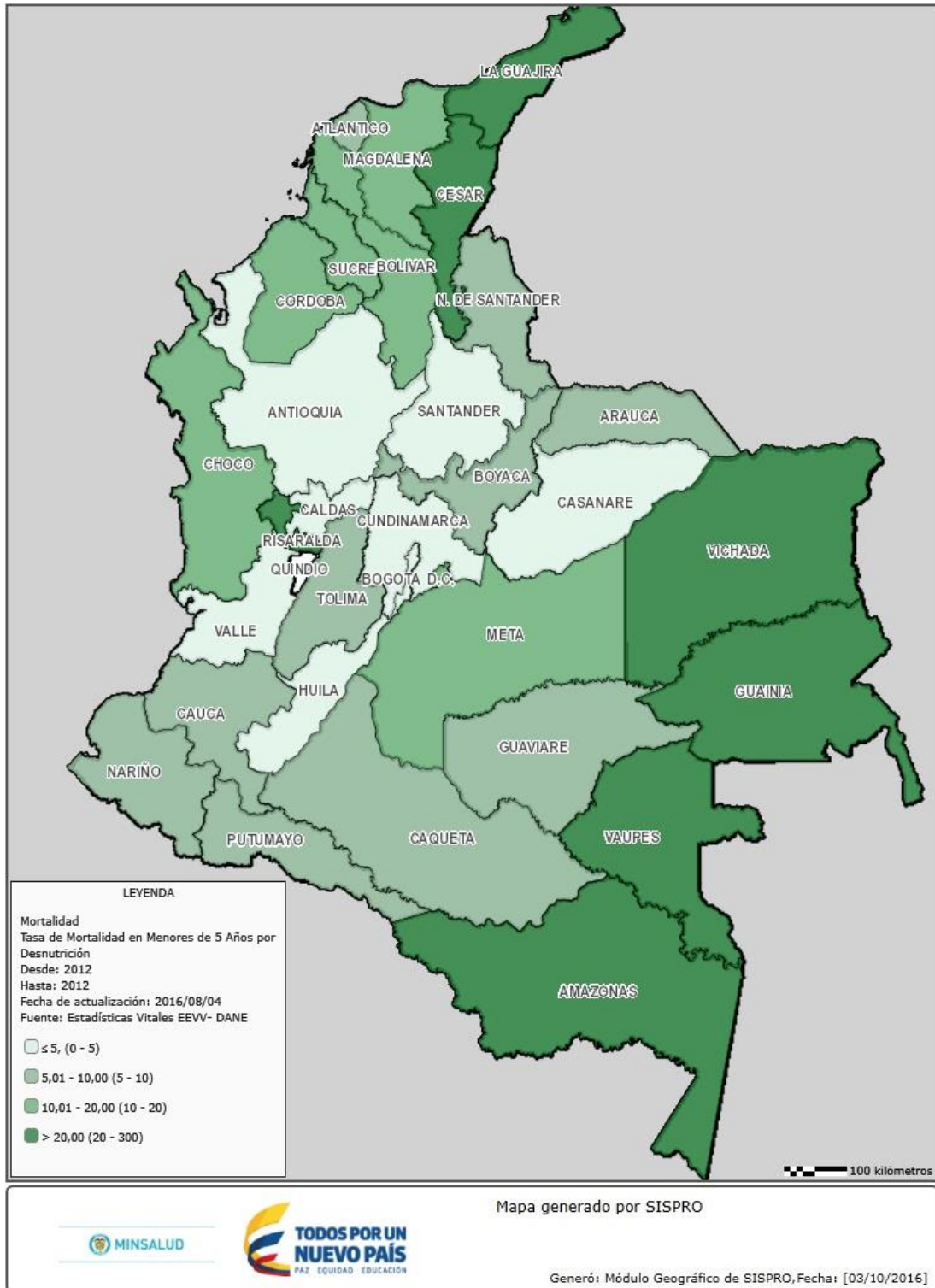
Fuente: Módulo geográfico de SISPRO

Figura F-3: Tasa de mortalidad por desnutrición en menores de 5 años, según departamento, Colombia 2011.



Fuente: Módulo geográfico de SISPRO

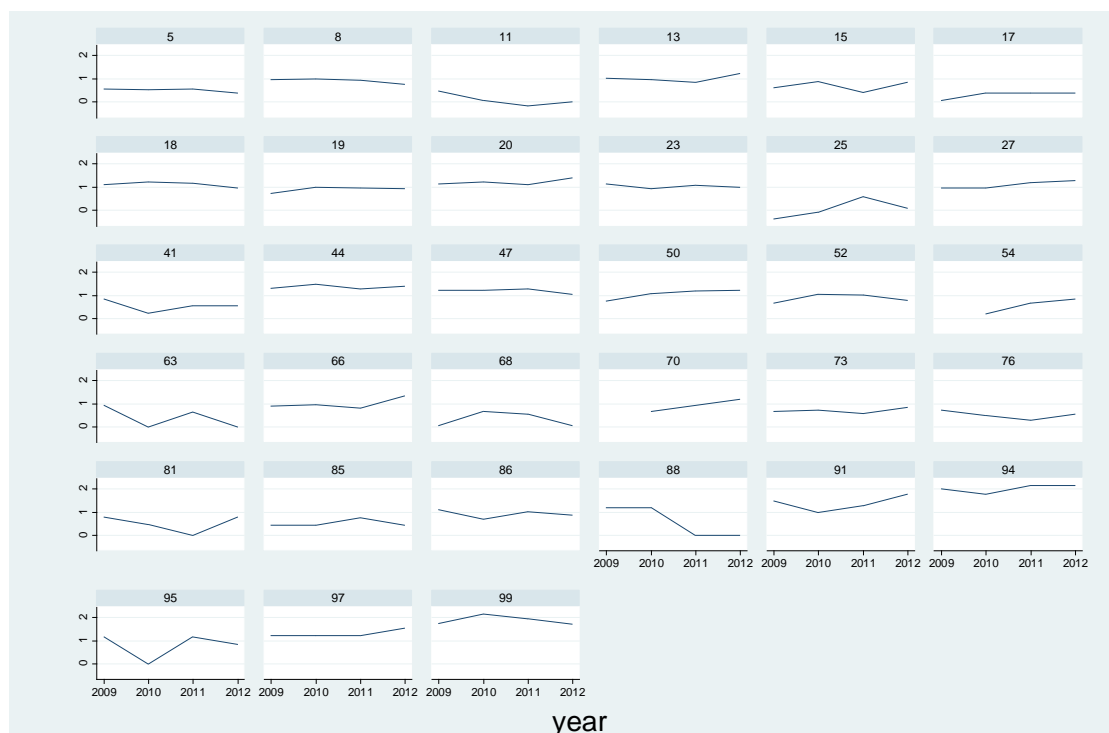
Figura F-4: Tasa de mortalidad por desnutrición en menores de 5 años, según departamento, Colombia 2012.



Fuente: Módulo geográfico de SISPRO

G. Anexo: Gráficos exploratorios de los paneles de datos

Figura G-1: Gráfico exploratorio del panel de datos para los menores de 5 años, según departamento, Colombia 2009-2012.



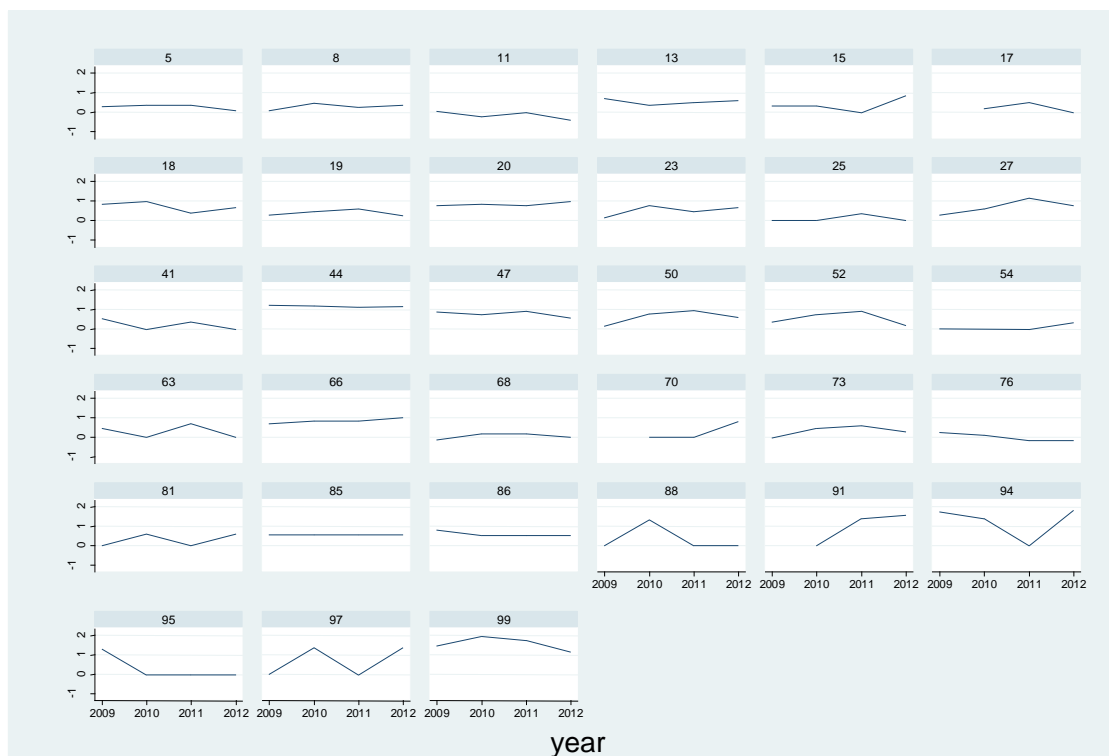
Fuente: Elaboración propia en Stata

Eje x: Log de la tasa de mortalidad por desnutrición

Eje y: Año (2009 – 2012)

Cada gráfica representa un departamento según el código DANE (ver anexo D).

Figura G-2: Gráfico exploratorio del panel de datos grupo de 0 a 4 años, según departamento, Colombia 2009-2012.



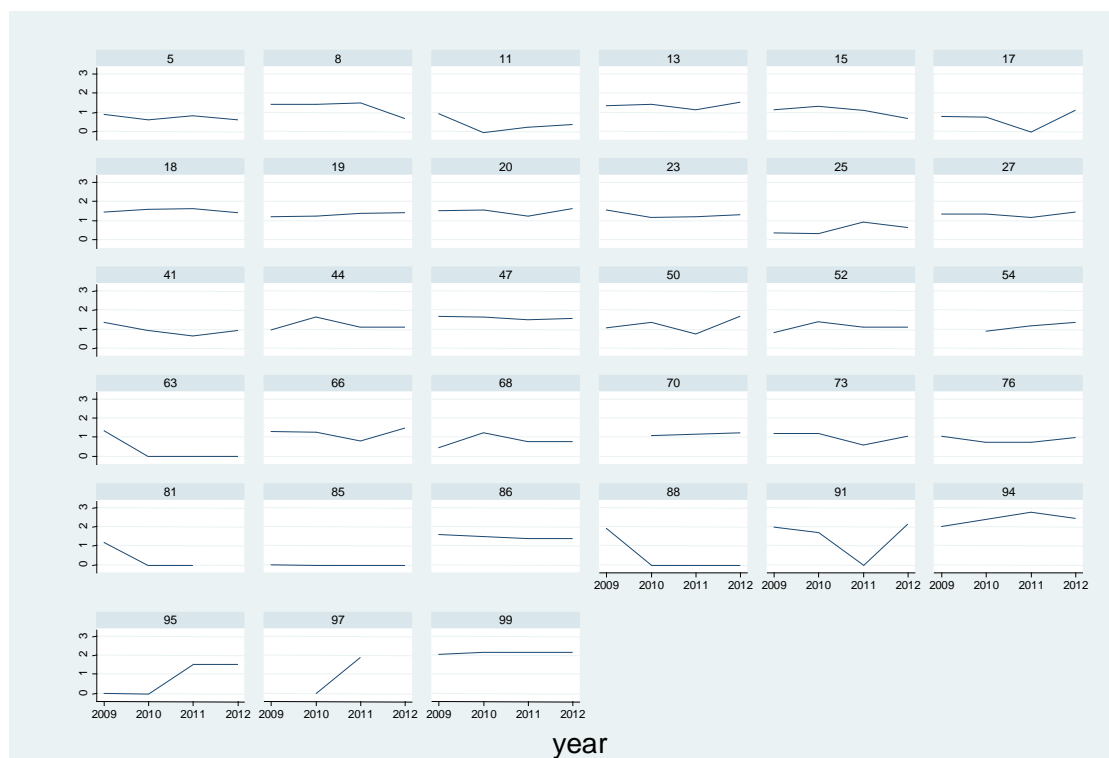
Fuente: Elaboración propia en Stata

Eje x: Log de la tasa de mortalidad por desnutrición

Eje y: Año (2009 – 2012)

Cada gráfica representa un departamento según el código DANE (ver anexo D).

Figura G-3: Gráfico exploratorio del panel de datos para los menores de 1 año, según departamento, Colombia 2009-2012.



Fuente: Elaboración propia en Stata

Eje x: Log de la tasa de mortalidad por desnutrición

Eje y: Año (2009 – 2012)

Cada gráfica representa un departamento según el código DANE (ver anexo D).

Códigos DANE por departamento

id	Departamento	id	Departamento	id	Departamento	id	Departamento
5	Antioquia	23	Córdoba	63	Quindío	88	Archipiélago de San
8	Atlántico	25	Cundinamarca	66	Risaralda	91	Amazonas
11	Bogotá	27	Choco	68	Santander	94	Guainía
13	Bolívar	41	Huila	70	Sucre	95	Guaviare
15	Boyacá	44	La Guajira	73	Tolima	97	Vaupés
17	Caldas	47	Magdalena	76	Valle del Cauca	99	Vichada
18	Caquetá	50	Meta	81	Arauca		
19	Cauca	52	Nariño	85	Casanare		
20	Cesar	54	Norte de Santander	86	Putumayo		

Bibliografía

- Acosta, K., & Romero, J. (2014). *Estimación indirecta de la tasa de mortalidad infantil en Colombia, 1964-2008*. Banco de la República-Economía Regional.
- Alkema, L., New, J. R., Pedersen, J., & You, D. (2014). Child mortality estimation 2013: an overview of updates in estimation methods by the United Nations Inter-agency Group for Child Mortality Estimation. *PloS One*, *9* (7).
- Amaya, A. C. (2014). Informe final mortalidad por y asociada a desnutrición en menores de cinco años, Colombia, 2014. INS, Instituto Nacional de Salud.
- Aparicio, J., & Márquez, J. (2005). Diagnóstico y especificación de modelos panel en STATA 8.0.
- Arango, A. P., Cabrera, N. L., & De Mendoza, A. J. H. (2015). Necesidad de un sistema informático de registro y control en Anatomía Patológica para la red hospitalaria en Cuba. *Revista Cubana de Informática Médica*, *7*(2), 185–195.
- Arias, M., Tarazona, M. C., Lamus, F., & Granados, C. (2013). Nutritional state associated with social determinants in Arhuaco children aged less than 5 years-old. *Revista de Salud Pública*, *15*(4), 613–625.
- Banco Mundial. (2014). *Informe sobre el desarrollo mundial 2014. Panorama general: Riesgo y oportunidad. La administración del riesgo como instrumento de desarrollo* (p. 66).
- Baronio, A., & Vianco, A. (2010). Manual de Econometría. *Primera Parte. Facultad de Ciencias Económicas. Universidad Nacional de Río Cuarto*.

- Bay, G., & Orellana, H. (2007). La calidad de las estadísticas vitales en la América Latina. *CEPAL, Comisión Económica Para América Latina Y El Caribe*.
- Berdugo, E. (2010). *Estudio comparativo entre los modelos Poisson-Normal y binomial negativo log-lineal mixto, para datos de conteos longitudinales*. Universidad Nacional de Colombia. Facultad de Ciencias. Departamento de Estadística.
- Cabezas, C. (2014). Enfermedades infecciosas desatendidas: un permanente reto para la salud pública y la equidad en el Perú. *Revista Peruana de Medicina Experimental Y Salud Pública*, 31(2), 326–335.
- Cárdenas, R. (2001). La medición de la mortalidad prematura: alcances y limitaciones de los indicadores. *Estudios Demográficos Y Urbanos*, 677–697.
- Castillo, A. J. S. (2010). Métodos Estadísticos con R y R Commander. *Recurso Electrónico Gratuito*, 1–157.
- Castro, A. F. (2012). *Modelamiento y cuantificación de los efectos del conflicto armado colombiano sobre las agendas de fecundidad de las mujeres 200-2010*. Universidad Nacional de Colombia. Facultad de Ciencias. Departamento de Estadística.
- Catalán, C. E. (2011). Series temporales. CSIC, Consejo Superior de Investigaciones Científicas Cataluña España.
- Celis de la Rosa, A. de J. (2008). *Bioestadística*.
- Correia, L. L., Campos, J. S., Andrade, F. M. de O., Machado, M. M. T., Lindsay, A. C., Leite, Á. J. M., ... Cunha, A. J. L. A. da. (2014). Prevalence and determinants of child undernutrition and stunting in semiarid region of Brazil. *Revista de Saude Publica*, 48(1), 19–28.
- DANE. (2008). Departamento Administrativo Nacional de Estadística. Estimación de la mortalidad, 1985-2005.

- DANE. (2010). Departamento Administrativo Nacional de Estadística. Metodología de estimación de la tasa de mortalidad infantil municipal 2005-2007. Retrieved from https://www.dane.gov.co/files/investigaciones/fichas/Tasa_mortalidad_infantil.pdf
- DANE. (2011). Departamento Administrativo Nacional de Estadística. Manual de crítica certificado de nacido vivo y defunciones EE.VV.
- DANE. (2012). Departamento Administrativo Nacional de Estadística. Metodología estadísticas vitales, EE.VV.
- DANE. (2013). Departamento Administrativo Nacional de Estadística. Ficha metodológica, Estadísticas vitales-(EEVV).
- DANE. (2013). Departamento Administrativo Nacional de Estadística. Resultados pobreza monetaria y desigualdad 2012, por departamentos.
- De Arce, R., & Mahía, R. (2001). Modelos Arima. *Econometría II, Modelos de Series Temporales*, 1–31.
- De Arce, R., & Mahía, R. (2012). Interpretación de los parámetros de un modelo básico de regresión lineal.
- Díaz, A. M. M., Ávila, T. P. N., & Gallo, A. D. V. (2005). Del análisis cluster a los mapas autoorganizados de Kohonen para segmentación psicográfica de mercados. *Simposio Estadística 2005*, 89–106.
- DNP, D. N. de P. (2014). Evolución de las coberturas de los servicios de acueducto y alcantarillado (1985-2013). *DNP, Departamento Nacional de Planeación*, 51–59.
- FAO, O. de las N. U. para la A. y la A. (2014). Mapa del Hambre 2014 de la FAO. Retrieved from <http://www.fao.org/hunger/es/>

- Fernández Cantón, S. B., & Viguri Uribe, R. (2010). La mortalidad por desnutrición en México en menores de cinco años, 1990-2009. *Boletín médico del Hospital Infantil de México*, 67(5), 471–473.
- Flórez, C. E. (2000). *Las transformaciones socio demográficas en Colombia*. Banco de la República. Retrieved from <http://www.banrep.gov.co/es/node/16591>
- Fonseca, Z., Ayala, D., Uribe, L. J., & Castaño, T. (2014). Aproximación a los Determinantes de la Doble Carga Nutricional en Colombia. *Bogotá, Observatorio de Seguridad Alimentaria Y Nutricional (OSAN)*, 52.
- Frome, E. L. (1983). The analysis of rates using Poisson regression models. *Biometrics*, 665–674.
- García, S., Sarmiento, O. L., Forde, I., & Velasco, T. (2013). Socio-economic inequalities in malnutrition among children and adolescents in Colombia: the role of individual-, household-and community-level characteristics. *Public Health Nutrition*, 16(09), 1703–1718.
- Gómez, A. (2015). Clasificación Internacional de Enfermedades (CIE): Descifrando la CIE-10 y esperando la CIE-11. *Monitor Estratégico*, 7, 66–73.
- González, E. C., Gelis, M. M., Cruz, M. E. S., Figueroa, E. M., Hernández, M. V., & Pérez, L. de A. C. (2012). Los determinantes sociales de la salud y sus diferentes modelos explicativos. *Infodir (Revista de Información Para La Dirección En Salud)*, 8(15).
- Granados, R. M. (2011). Efectos fijos o aleatorios: test de especificación. *Universidad de Granada*, 1–5.
- Hachuel, L., Boggio, G., & Harvey, G. (2010). Modelos alternativos para el análisis de datos de conteo con exceso de ceros.

- Härdle, W., Mori, Y., & Vieu, P. (2006). *Statistical methods for biostatistics and related fields*. Springer Science & Business Media.
- Hoechle, D. (2007). Robust standard errors for panel regressions with cross-sectional dependence. *The Stata Journal*, 7(3), 281–312. Retrieved from <http://www.stata-journal.com/sjpdf.html?articlenum=st0128>
- Horiuchi, S., Wilmoth, J. R., & Pletcher, S. D. (2008). A decomposition method based on a model of continuous change. *Demography*, 45(4), 785–801.
- Hoz, F. de la, Duran, M. E. M., García, O. E. P., & Bonilla, H. Q. (2014). Mortalidad por y asociada a desnutrición en < 5 años. *INS, Instituto Nacional de Salud*, 32.
- Hurtado, D. A. (2011). *Determinantes sociales y epidemiología del curso de la vida*.
- Hyndman, R. J., Booth, H., & Yasmineen, F. (2013). Coherent mortality forecasting: the product-ratio method with functional time series models. *Demography*, 50(1), 261–283.
- INS. (2014). Protocolo de Vigilancia en Salud Pública. Mortalidad por y asociada a desnutrición en < 5 años.
- INS. (2015). Informe del evento mortalidad por y asociada a desnutrición en menores de cinco años hasta el período epidemiológico V, Colombia, 2015.
- INS. (2016). Una Mirada al ASIS y análisis en profundidad; informe técnico ONS 2016-I. *Imprenta Nacional de Colombia*, 21–57. Retrieved from <http://www.ins.gov.co/lineas-de-accion/ons/SiteAssets/Paginas/publicaciones/Informe%20FINAL%20-%207.pdf>
- INS. (2014) ¿Por qué usar tasas ajustadas de mortalidad para comparar poblaciones?, 1–7.

- Jiménez, S. A. Z. (2016). *El reporte de la información de la ENSIN con enfoque de determinantes*.
- Labra, R., & Torrecillas, C. (2014). Guía CERO para datos de panel. Un enfoque práctico, 61.
- Lisowska, M. (2013). Asfixia perinatal: causa de muerte y daños neurológicos, un dilema de salud mundial. *Universidad, Ciencia Y Desarrollo, V. 8 Fascículo 5*.
- Londoño, D. A., Ortega Lenis, D., & Olaya Ochoa, J. (2009). Modelación del número de homicidios vía regresión de Poisson. *Heurística, 16*, 81–90. Retrieved from <http://bibliotecadigital.univalle.edu.co/handle/10893/6105>
- Longhi, F. (2013). Pobreza y mortalidad infantil. Una aproximación teórica al estudio de sus relaciones.
- Madrid, L. F. (2005). Deficiencias nutricionales en los adultos y adultos mayores. *An. Venez. Nutr, 18*(1), 82–89.
- Mayorga, M., & Muñoz, E. (2000). La técnica de datos de panel una guía para su uso e interpretación. *Banco Central de Costa Rica*, 1–20. f
- Mellado, M. O. (2004). Pronóstico de las defunciones por medio de los modelos autorregresivos integrados de promedios móviles. *Papeles de Población, 10*(42), 250–264. Retrieved from <http://www.redalyc.org/pdf/112/11204209.pdf>
- Minsalud, M. de S. y P. S. (2010). Resolución 2121 de 2010, “Por la cual se adoptan los Patrones de Crecimiento publicados por la Organización Mundial de la Salud-OMS en el 2006 y 2007 para los niños, niñas y adolescentes de 0 a 18 años de edad y se dictan otras disposiciones.”
- Minsalud, M. de S. y P. S. (2013). Análisis de situación de salud de poblaciones diferenciales Colombia 2012. Imprenta Nacional.

- Minsalud, M. de S. y P. S., INS, I. N. de S., & ICBF, I. C. de B. F. (2011). *Encuesta Nacional de la Situación Nutricional en Colombia (ENSIN) 2010* (Primera Edición). Bogotá, Colombia: Da Vinci Editores & Cia.
- Montgomery, D. C., Peck, E. A., & Vining, G. G. (2006). *Introducción al análisis de regresión lineal* (Tercera). Mexico: Compañía editorial continental.
- Monti, G. R. (2008). Desnutrición hospitalaria: una patología subdiagnosticada. *Revista de La Asociación Médica Argentina*, 121(4), 25–28.
- Mosley, W. H., & Chen, L. C. (1984). An analytical framework for the study of child survival in developing countries. *Population and development review*, 10, 25-45.
- MSPAS, M. de S. P. y A. S. G. (2009). Reducción de mortalidad por desnutrición aguda y las enfermedades asociadas.
- OMS, O. M. de la S. (2008a). *Comisión sobre Determinantes Sociales de la Salud - 2005-2008*. Retrieved from http://www.who.int/social_determinants/thecommission/es/
- OMS, O. M. de la S. (2008b). *Comisión sobre Determinantes Sociales de la Salud - ¿Qué es, por qué, y cómo?*.
- OMS, O. M. de la S. (2016). Reducción de la mortalidad en la niñez. *Nota descriptiva*, (No. 178). Retrieved from <http://www.who.int/mediacentre/factsheets/fs178/es/>
- ONS, O. N. de S. (2013). Metodologías de análisis de la información en salud, 47.
- OPS, O. P. de la S. (1995). *Clasificación estadística internacional de enfermedades y problemas relacionados con la salud: Décima Revisión* (2003rd ed., Vols. 1–3). Organización Panamericana de la Salud.
- Ortiz, J., & Montenegro, Á. (2005). Modelamiento estadístico. Facultad de Ciencias, Departamento de Estadística, Universidad Nacional de Colombia. Simposio de estadística XV. Bogotá.

- Parul, C. (2010). Impact of the economic crisis and increase in food prices on child mortality: exploring nutritional pathways. *The Journal of Nutrition*, 140(1), 177S–181S. Retrieved from <http://jn.nutrition.org/content/140/1/177S.short>
- Peñuela, A. M., Linero, M. I., Guiza, L. M., & Martínez, N. P. (2014). Profundización del análisis de la situación de las niñas, niños y adolescentes; en el ejercicio del derecho a la salud y el derecho a la alimentación y nutrición. Ministerio de Salud.
- Pérez, A. G. Á., Fariñas, A. G., Salvá, A. R., & Gorbea, M. B. (2007). Voluntad política y acción intersectorial: Premisas clave para la determinación social de la salud en Cuba. *Revista Cubana de Higiene Y Epidemiología*, 45(3), 1–16.
- Pérez, D. L. S., Santa, F., & López, H. J. F. (2012). Modelos de regresión espacial para el comportamiento de las enfermedades infecciosas dengue y malaria en Colombia para los años 2000, 2005 y 2010. *UD Y La Geomática*, (6), 110–128.
- Perozo, M. F. C. (2012). *Modelación espacial de la mortalidad infantil, el período postnatal de cribado y el volumen de atención hospitalaria en Colombia*. Universidad Nacional de Colombia. Retrieved from <http://www.bdigital.unal.edu.co/12918/>
- Puig, X., Ginebra, J., & Gispert, R. (2005). Análisis de la evolución temporal de la mortalidad mediante modelos lineales generalizados. *Gaceta Sanitaria*, 19(6), 481–485. Retrieved from <http://scielo.isciii.es/pdf/gv/v19n6/nota.pdf>
- Quiroga, E. F. (2012). Mortalidad por desnutrición en menores de cinco años, Colombia, 2003-2007. *Biomédica*, 32(4), 499–509.
- Ramirez Figueroa, J. A., & Constante Mendoza, K. M. (2009). Análisis Estadístico de la morbilidad y mortalidad infantil. Caso del área del Hospital del Niño Francisco De Ycaza Bustamante año 2000.

- Reyes, J. G. G., Zúñiga, A. S., & Cruz, M. G. (2007). Prevalencia de desnutrición del adulto mayor al ingreso hospitalario. *Nutrición Hospitalaria*, 22(6), 702–709.
- Reyna, O. T. (2007). Panel data analysis fixed and random effects using Stata (v. 4.2). *Data & Statistical Services, Princeton University*.
- Rice, A. L., Sacco, L., Hyder, A., & Black, R. E. (2000). Malnutrition as an underlying cause of childhood deaths associated with infectious diseases in developing countries. *Bulletin of the World Health Organization*, 78(10), 1207–1221.
- Roitman, M. (2005). *Aplicación de técnicas de datos en panel a la medición de eficiencia relativa entre empresas de distribución eléctrica reguladas: una guía para la práctica regulatoria*. Buenos Aires, Argentina: Instituto de Economía, Universidad Argentina de la Empresa.
- Ruiz, M., & Ruiz, N. (2007). La mortalidad por desnutrición en Colombia 1998-2002. *Bogotá*.
- Salinas-Rodríguez, A., Manrique-Espinoza, B., & Sosa-Rubí, S. G. (2009). Statistical analysis for count data: use of healthcare services applications. *Salud Publica de Mexico*, 51(5), 397–406.
- Setel, P. W., Sankoh, O., Rao, C., Velkoff, V. A., Mathers, C., Gonghuan, Y.,... Lopez, A. D. (2005). Sample registration of vital events with verbal autopsy: a renewed commitment to measuring and monitoring vital statistics. *Bulletin of the World Health Organization*, 83(8), 611–617.
- Solar, O., & Irwin, A. (2007). A conceptual framework for action on the social determinants of health., 1–77.
- Spedicato, G. A., & Clemente, G. P. (2015). Mortality projection with demography and lifecontingencies packages, 1–11.

- SPSS, I., & IBM, C. (2010). IBM SPSS Advanced statistics 19, Manual.
- Toro, J. H., & Arango, D. C. (2013). Mortality due to malnutrition among children aged five and younger: five years before and after the implementation of the MANA program. Antioquia, 1998-2007. *Revista Facultad Nacional de Salud Pública*, 31(1), 93–101.
- UNICEF, (2011). *Level & Trends in Child Mortality. Report 2011* (p. 24).
- UNICEF, (2013). Ending preventable child deaths from pneumonia and diarrhoea by 2025: the integrated global action plan for pneumonia and diarrhoea (GAPD).
- UNICEF, (2014). *Level & Trends in Child Mortality. Report 2014* (p. 36).
- Salmerón, R. (2012). Universidad de Granada. Modelos con datos de panel, Econometría II. Retrieved from <http://www.ugr.es/~romansg/material/WebEco/Eco2-Panel.pdf>
- Universidad de Navarra. (2014). Microglosario de STATA. Retrieved from <http://www.unav.edu/departamento/preventiva/index.php?section=2>
- Urdinola, B. P. (2011). Determinantes socioeconómicos de la mortalidad infantil en Colombia, 1993. *Revista Colombiana de Estadística*, 34(1), 39–72.
- Urdinola, B. P., Avilés, F. T., & Velasco, J. A. (2016). Mortality atlas of homicides in Colombia: contagion and under-registration for small areas.
- USAID, A. de E. U. para el D. I. (1995). *La desnutrición y la mortalidad infantil: repercusiones programáticas de nuevas pruebas* (Actualización investigación) (pp. 1–8). Retrieved from <http://www.basics.org/documents/pdf/MCM-Spanish.pdf>
- Wayne, D. (2002). Bioestadística. Base para el análisis de las ciencias de la salud. 4ta. Edición. Editorial Limusa. México., 501.
- WFP, W. F. P., & CDC, C. for D. C. and P. (2007). *A Manual: Measuring and Interpreting Malnutrition and Mortality*. UNHCR Operational publications. Retrieved from <http://www.unhcr.org/45f6abc92.pdf>

Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data*.

Cambridge, MA: Massachusetts Institute of Technology.

Wooldridge, J. M. (2009). *Introducción a la econometría: un enfoque moderno* (4th ed.).

Paraninfo.

Yoshioka, H., & Ochoa, J. C. E. (2009). Desigualdad socioeconómica y mortalidad infantil

en Nicaragua: ¿Una cuestión étnica? *Latin American Research Review*, 44(1),

199–217. Retrieved from <https://muse.jhu.edu/article/263733/summary>