

ISSN 0120-4157

Biomédica

Revista del Instituto Nacional de Salud

PUBLICACIÓN ANTICIPADA EN LINEA

El Comité Editorial de *Biomédica* ya aprobó para publicación este manuscrito, teniendo en cuenta los conceptos de los pares académicos que lo evaluaron. Se publica anticipadamente en versión pdf en forma provisional con base en la última versión electrónica del manuscrito pero sin que aún haya sido diagramado ni se le haya hecho la corrección de estilo.

Siéntase libre de descargar, usar, distribuir y citar esta versión preliminar tal y como lo indicamos pero, por favor, recuerde que la versión impresa final y en formato pdf pueden ser diferentes.

Citación provisional:

Mera-Mamián AY, Moreno-Montoya J, Rodríguez-Villamizar LA, Muñoz DI,

Segura AM, García HI. Fundamento y generalidades de la construcción de modelos estadísticos multinivel en el ámbito de la investigación en salud.

Biomédica. 2023;43 (4).

Recibido: 24-03-23

Aceptado: 25-09-23

Publicación en línea: 27-09-23

Fundamento y generalidades de la construcción de modelos estadísticos multinivel en el ámbito de la investigación en salud

Fundamentals and generalities of the construction of multilevel statistical models in health research

Modelos estadísticos multinivel en investigación en salud

Andry Yasmid Mera-Mamián ^{1,4}, José Moreno-Montoya ^{2,3}, Laura Andrea Rodríguez-Villamizar ³, Diana Isabel Muñoz ⁴, Ángela María Segura ¹, Héctor Iván García ⁵

¹ Escuela de Graduados, Universidad CES, Medellín, Colombia

² División de Estudios Clínicos y Epidemiología Clínica, Hospital Universitario de la Fundación Santa Fe de Bogotá, Bogotá, D.C., Colombia

³ Departamento de Salud Pública, Escuela de Medicina, Universidad Industrial de Santander, Bucaramanga, Colombia

⁴ Facultad de Fisioterapia, Universidad CES, Medellín, Colombia

⁵ Instituto de Cancerología Las Américas, Medellín, Colombia

Correspondencia:

Andry Yasmid Mera, Facultad de Fisioterapia, Universidad CES, Calle 10A #22-04, Altos de El Poblado, Medellín, Colombia

Tel.: +573013542030

amera@ces.edu.co

Contribución de los autores:

Andry Mera y José Moreno participaron en todas las etapas de la revisión.

Laura Rodríguez y Diana Muñoz: redacción.

Todos los autores participaron en la revisión crítica.

Este trabajo tiene como objetivo presentar una mirada global de la aplicabilidad de los modelos de análisis multinivel en el ámbito de la investigación sanitaria. Ofrece información sobre los fundamentos teóricos, metodológicos y estadísticos, además de los pasos básicos para la construcción de estos modelos, y ejemplos de su uso según la estructura jerárquica de los datos. Se destaca que antes de utilizar estos modelos se requiere contar con un soporte teórico sobre la necesidad de uso y una valoración estadística que dé cuenta del porcentaje de varianza explicada por el efecto de agrupación de las observaciones. Los requerimientos para llevar a cabo este tipo de análisis dependerán de condiciones especiales como el tipo de variables, la cantidad de unidades por nivel o el tipo de estructura jerárquica. Se concluye que los modelos de análisis multinivel son una herramienta útil para lograr la integración de información considerando la complejidad de las relaciones y las interacciones que determinan la mayoría de las condiciones de salud, incluida la pérdida de independencia entre las unidades de observación.

Palabras clave: análisis multinivel, investigación sobre servicios de salud; bioestadística; sesgo.

The objective is to present a global vision of the applicability of multilevel analysis models in health research. Information is presented on the theoretical, methodological, and statistical foundations, as well as the basic steps for the construction of these models, and examples of their use according to the hierarchical structure of the data. It should be noted that before using these models it is necessary to have a theoretical support on the need for use it and a statistical evaluation that accounts for the percentage of variance explained by the grouping effect of the observations. The requirements to conduct this type of analysis will depend on special conditions such as the type of variables, the number of units per level or the type of hierarchical structure. It is concluded that multilevel analysis models are a useful tool to achieve information integration considering the complexity of relations and interactions that determine most health conditions, including the loss of independence between observation units.

Keywords: multilevel analysis, health services research; biostatistics; bias.

En investigación sanitaria las estructuras jerárquicas o anidadas aparecen cuando las unidades de observación en las que se mide el desenlace de interés se anidan naturalmente dentro de otras que las agrupan. Esto sucede, por ejemplo, con individuos humanos que se agrupan en unidades superiores, como familias, barrios o escuelas (1). La misma agrupación puede encontrarse al considerar como unidades de observación el conjunto de mediciones repetidas en el tiempo para un mismo sujeto (2), o cuando las observaciones se recopilan con metodologías cuya selección se basa en el emparejamiento (3).

Dada la frecuencia con la que el fenómeno ocurre en el ámbito de la salud, aprovechar la estructura jerarquizada puede resultar útil para ampliar el alcance de la investigación científica, a la vez que mejora la comprensión del efecto que exposiciones a diferentes niveles pueden provocar en los individuos. No obstante, el uso de modelos para datos jerarquizados debe obedecer a una necesidad analítica preestablecida desde la formulación de la pregunta y la hipótesis de trabajo. En ese sentido, es necesario que el supuesto de que la asociación de interés entre una o varias exposiciones y un desenlace específico está razonablemente mediada por el hecho de que las unidades de análisis se agrupan, se encuentre biológica, social o clínicamente soportado.

Una forma de evaluar estadísticamente si existe un efecto de agrupación entre las unidades de análisis es el coeficiente de correlación intraclase (ICC), una medida que hace referencia a la fuerza de la asociación lineal existente entre las mediciones del desenlace de interés sobre los sujetos individuales, una vez comparada con la que deriva de la comparación entre grupos (4).

Los análisis multinivel surgen en el área de la psicometría y la educación como alternativa para explorar el efecto que tienen las estructuras de relaciones jerárquicas o de datos anidados en la valoración de las exposiciones y los desenlaces de interés en el ámbito epidemiológico, y consideran la pérdida de independencia de los sujetos de observación debido a su cercanía en cada nicho agrupador. Estos, al igual que los modelos de regresión de un único nivel, dependen para su especificación de la escala de medida de la variable de respuesta y del tipo de función de enlace usado en su construcción (1).

Este tipo de análisis aporta en el control de la denominada *falacia ecológica*, o sesgo de agregación, que se produce cuando se infiere inadecuadamente acerca de la variabilidad interindividual a partir de información captada a nivel grupal, cuando las variables individuales y grupales tiene efectos independientes o miden conceptos teóricos diferentes (5), lo que puede generar que se asuma erróneamente que la correlación estadística entre dos variables a nivel agregado es igual a la correlación entre las correspondientes variables a nivel individual (6). Por tanto, considerando su aporte en el análisis de datos anidados en diferentes escenarios, el presente trabajo tiene como propósito ofrecer una mirada global acerca de la construcción, el uso y las particularidades metodológicas de este tipo de modelos en el ámbito de la investigación sanitaria.

Fundamento del uso de análisis multinivel

La falta de independencia entre las unidades de observación respecto a la oportunidad de considerarse expuestas o a riesgo constituye un desafío común en la investigación en salud y su omisión puede implicar la sobrestimación de la significación estadística (7,8) y de los tamaños de efecto (9), lo que lleva a la

obtención de estimaciones con intervalos de confianza inapropiadamente estrechos y al aumento del riesgo de error de tipo I (10). Para enfrentar el problema, los modelos multinivel toman en cuenta el nivel de agregación de las unidades de observación (11,12).

El análisis multinivel puede adaptarse a una estructura jerárquica clásica y a estructuras con jerarquías y relaciones complejas, como la clasificación cruzada o la membresía múltiple. La estructura jerárquica clásica se presenta cuando unidades de nivel 1 están anidadas en grupos que a su vez pueden pertenecer a otros de nivel superior, como individuos anidados en escuelas que a su vez se anidan en barrios, o el caso de los metaanálisis, en los que el desenlace se mide a nivel individual y agrupado a través de la medida de resumen obtenida para cada estudio (2). También ocurre en el escenario de las medidas repetidas, en el que diferentes mediciones a lo largo del tiempo pueden provenir del mismo individuo, quien representa un nivel superior (13,14).

Por su parte, el análisis de clasificación cruzada se presenta cuando unidades de nivel inferior corresponden a más de un nivel superior a la vez, sin que estos dos niveles superiores estén anidados entre sí (15). En el caso de la estructura no jerárquica de membresía múltiple, las unidades del primer nivel pueden pertenecer a varias unidades del nivel superior. Por ejemplo cuando se analizan datos de individuos respecto a su lugar de estudio en un periodo específico, en el que pudieron asistir a diferentes escuelas (4).

En los modelos multinivel, al igual que en los modelos de regresión convencionales, de un solo nivel, también son importantes los términos *intercepto* (valor de la variable dependiente cuando las variables explicativas valen 0) y

pendiente (cambio en el valor de la variable dependiente por cada unidad de cambio en una variable independiente, una vez consideradas constantes las demás). No obstante, cuando se trata de datos agrupados, la intersección y la pendiente pueden variar en cada grupo según el nicho agrupador de las observaciones o la magnitud de una variable medida a ese nivel, lo que convierte en un tema de interés la estimación de dichos parámetros y su variación en los modelos multinivel.

Para efectos de lo anterior, es posible construir dos tipos de modelos principales: de interceptos aleatorios (también denominados de efectos fijos) y de efectos aleatorios (también denominados de pendientes aleatorias). En el primero, el término de intersección varía aleatoriamente entre los grupos, pero los efectos/pendientes son los mismos. En este caso, la variable dependiente se ve afectada por un componente fijo que corresponde con la media general del intercepto común a todos los conglomerados y los efectos de las variables independientes, que también son comunes, además del componente aleatorio: la variación entre individuos y la variación del intercepto entre los conglomerados (16). Este tipo de modelos resulta de particular utilidad cuando se requiere valorar en qué medida las diferencias entre individuos se deben a su pertenencia a grupos después de ajustar por otras variables (17).

En el segundo, tanto interceptos como efectos/pendientes son aleatorios (2), lo que permite a las unidades de observación tener sus propias trayectorias, aspecto importante cuando los grupos de observación tienen características muy heterogéneas. A diferencia del modelo de efectos fijos, es razonable considerar que el efecto a nivel de grupo varía aleatoriamente entre estos, es decir, los

efectos de las variables explicativas varían entre grupos (17). La figura 1 compara un modelo de regresión de un solo nivel, un modelo de interceptos aleatorios y un modelo de efectos aleatorios.

La elección de uno u otro modelo dependerá de la pregunta de investigación, los objetivos de estudio, la escala de medida de la variable usada para conformar los grupos (discreta o continua) y el contexto (1).

Respecto al método de estimación de los coeficientes, los más utilizados para la regresión multinivel se basan en la máxima verosimilitud; en particular, el método de máxima verosimilitud restringida o residual (REML), que emplea una función de probabilidad calculada a partir de un conjunto transformado de datos cuando hay muchos parámetros no especificados o desconocidos, lo que ayuda a obtener estimadores menos sesgados de varianza y covarianza; además tiene la ventaja de que se puede usar para dar estimadores con datos no normales (18). Este método permite el uso de predictores o variables independientes en todos los niveles de análisis, debido a que se crean ecuaciones de predicción para cada nivel de la estructura anidada, y se presentan ventajas sobre otros modelos cuando existe una variación del parámetro de pendiente en el nivel superior. Al respecto se ha indicado que ante la sospecha de endogeneidad (falta de correlación entre los residuos dentro del grupo) se debería considerar el uso de modelos de un solo nivel, y cuando los tamaños de muestra de nivel superior son relativamente pequeños, la estimación bayesiana basada en el algoritmo de cadenas de Markov con integración Monte Carlo puede tener ventaja sobre el REML si las probabilidades previas se especifican correctamente (19).

En paralelo, los modelos multinivel, al igual que los modelos de regresión convencionales, se pueden aplicar en cualquier variable que sea sujeto de ser modelada con un modelo lineal generalizado (MLG) (14). Tal como ocurre en los modelos de un solo nivel, cuando no es posible asumir que la distribución de errores se corresponde con una distribución normal, se utiliza una función de enlace para transformar el predictor. Por ejemplo, si la variable dependiente es dicotómica o se utilizan datos agregados como proporciones, se sugiere utilizar modelos de desenlace binario (*logit*, *probit* o doble *log* complementario) y se modela una transformación no lineal de la probabilidad de estar en una u otra categoría. En caso de que la variable respuesta tenga más de dos categorías, pueden considerarse otros modelos como el *logit* multinomial y el modelo de probabilidades proporcionales (*logit* acumulativo). Si se trata de conteo, cuando el tamaño de la población es grande o el evento es raro, se prefiere aplicar el modelo Poisson. Pueden considerarse modelos análogos para situaciones en las que se requiere valorar la supervivencia o el tiempo transcurrido hasta la ocurrencia de un evento determinado (4).

Como ejemplo, se cita a continuación el trabajo de Webster et al. (20), en el que a través de ecuaciones de estimación generalizadas con función de enlace *logit* se evaluó de manera conjunta la influencia del nivel socioeconómico individual y comunitario sobre el riesgo de padecer de cáncer de mama. Se construyeron tres modelos. El primero consideró el nivel socioeconómico individual, el segundo solo el nivel comunitario y el tercero correspondió a un análisis multinivel que incluyó ambos. Los resultados indican que las mujeres que vivían en comunidades con un nivel socioeconómico más alto tenían un riesgo mayor de desarrollar cáncer de

mama, independientemente de su propio nivel socioeconómico (OR = 1,30).

Concluyeron que era posible que la medida a nivel comunitario capturara un aspecto no medido del nivel socioeconómico individual.

En cuanto a la medida del efecto contextual general, que señala la proporción de la varianza individual total que es atribuible al nivel de grupo, en modelos de regresión lineal multinivel, se puede calcular a partir del coeficiente de partición de varianza (CPV) (21). Cuando los resultados son de naturaleza binaria corresponden al tiempo transcurrido hasta el evento o son recuentos enteros que denotan el número de veces que ocurrió un evento, y el ECG se puede cuantificar mediante medidas de heterogeneidad como OR mediano (MOR), Hazard ratio mediano (MHR) o razón de tasa mediana, respectivamente (22).

Estas medidas de heterogeneidad se interpretan como el cambio relativo mediano en la medida de ocurrencia del evento, cuando se compara el resultado en un individuo de un grupo seleccionado al azar con otro individuo con valores de covariables idénticos, pero seleccionado al azar de un grupo diferente, cuando los grupos están ordenados por dicha medida (22). Indican el cambio en la probabilidad individual respecto a un desenlace al cambiar de un grupo de menor riesgo a otro de mayor riesgo. Por ejemplo, un MOR igual a 1 indica que no habría diferencias entre grupos en la probabilidad de presentar el resultado. Si hubiese fuertes diferencias entre grupos, el MOR sería grande y la variable de agrupación sería relevante para comprender las variaciones de la probabilidad individual respecto al resultado (23).

Ballesteros y Moreno-Montoya (24) utilizaron un modelo *logit* de dos niveles para identificar los principales factores regionales, asociados a variaciones en la

prevalencia de limitación funcional en el adulto mayor en Colombia. Tras ajustar por variables individuales se calculó el MOR para evaluar en qué medida la probabilidad individual de limitación funcional estaba determinada por las variables a nivel de estado. El análisis mostró diferencias estadísticamente significativas en la prevalencia de limitación funcional entre los estados, particularmente en cuanto al nivel socioeconómico medido según el Índice de Desarrollo Humano (MOR = 1,22; IC 95%: 1,13-1,30).

Pasos generales para la construcción de un modelo de regresión multinivel

Para plantear un análisis de regresión multinivel primero se requiere revisar la teoría existente, las investigaciones previas y el análisis exploratorio de los datos con el fin de comprobar la existencia de una estructura jerárquica (25). Al igual que en las regresiones convencionales, es conveniente postular con anterioridad los predictores a considerar en las diferentes fases del análisis, lo mismo que sus niveles (26) y el número de sujetos a incluir en cada nivel, debido a que el poder depende tanto del número de grupos como del número de unidades en cada grupo (2). También es necesario asegurar que se midieron las características a nivel grupal, si se sospecha que los individuos pertenecen a grupos en los cuales comparten dichas características (27).

Estadísticamente, primero es necesario validar la idoneidad de un análisis multinivel. Esto es posible mediante el cálculo de la proporción de la varianza total, que es explicada por características de niveles superiores a través del CCI, que en un modelo multinivel de efectos fijos corresponde al mismo coeficiente de partición de la varianza (CPV) (28):

$$CPV = \frac{\text{varianza residual del segundo nivel}}{\text{varianza residual del primer nivel} + \text{varianza residual del segundo nivel}}$$

Para el análisis del coeficiente se debe considerar que este toma valores entre 0 y 1. Un valor de cero (0) indica la inexistencia de diferencia entre las observaciones debida a la agrupación (toda la variabilidad de la variable dependiente es explicada por características del primer nivel). Un valor de uno (1) indica que no hay diferencias dentro del grupo (toda la variabilidad de la variable dependiente es explicada por características del segundo nivel) (28). Cuanto más grande el valor del CCI, más conveniente resulta proceder con el análisis multinivel (29); habrá un mejor indicador de la existencia y necesidad de considerar la estructura anidada en los modelos analíticos (27,30).

El CCI también es importante en relación con otros aspectos del análisis multinivel, como el cálculo del tamaño de muestra efectivo. Las muestras agrupadas no son tan eficientes estadísticamente como las muestras aleatorias simples, debido a que las similitudes entre los sujetos de los conglomerados pueden reducir la variabilidad de las respuestas de un conglomerado en comparación con las esperadas de una muestra aleatoria simple, lo que puede magnificar las diferencias entre los grupos. Esta condición implica que al ajustar el tipo de muestreo se reduzca el tamaño efectivo de la muestra y por tanto la precisión en las estimaciones:

$$\text{Tamaño de muestra efectivo} = \frac{mk}{1 + \rho (m - 1)}$$

Siendo k el número de grupo, m el número de unidades por grupo, ρ el CCI. El aumento en el CCI o el número de unidades por grupo implicará menor tamaño de

muestra efectiva y menor potencia, mientras una mayor cantidad de grupos puede implicar un mayor tamaño de muestra efectiva y mayor potencia (31).

Respecto a los resultados de un modelo multinivel, como se mencionó previamente, el CCI o el CPV también pueden ser utilizados para reportar la medida del efecto contextual general; la cantidad de variabilidad que aporta el efecto de agrupación respecto de la cantidad de variabilidad del fenómeno *per se* (27,30).

A continuación, se proponen algunos pasos generales y usualmente utilizados para guiar el planteamiento de un modelo de regresión multinivel, una vez se ha confirmado teórica y estadísticamente que aplica su realización, y se ha definido el modelo estadístico, según la pregunta de investigación y la estructura de los datos.

Generar un modelo nulo

Este tiene una finalidad descriptiva, pretende servir como punto de referencia para los modelos que siguen (32). Se utiliza para analizar la variabilidad de la variable dependiente dentro y entre los niveles superiores (33). Además, sirve para valorar la magnitud de la variación que depende de los promedios formados por los grupos sin considerar los efectos de las variables explicativas.

La ecuación del modelo nulo sin variables explicativas se expresa como

$$Y_{ij} = B_{0j} + r_{ij} \quad (1)$$

Y_{ij} : valor para Y del individuo i en el grupo j.

B_{0j} : promedio de Y en el grupo j.

r_{ij} : efecto aleatorio de los individuos i en el grupo j. Diferencia entre el valor de Y para cada i en el grupo j y el promedio de Y en el grupo j (B_{0j}).

El promedio de Y en el grupo j, a su vez, se calcula como

$$B_{0j} = \gamma_{00} + U_{0j} \quad (2)$$

γ_{00} : promedio general de todos los grupos en los niveles superiores.

U_{0j} : efecto aleatorio de los grupos. Diferencia entre el promedio en cada grupo y el promedio general de todos los grupos en los niveles superiores (γ_{00}).

El modelo nulo de los dos niveles es

$$Y_{ij} = \gamma_{00} + U_{0j} + r_{ij} \quad (3)$$

Hasta este momento no se tienen variables predictoras y es posible identificar un componente fijo (γ_{00}) y el componente aleatorio ($U_{0j} + e_{ij}$). El componente fijo corresponde a la media poblacional y no cambia entre grupos, mientras que el componente aleatorio corresponde a los residuales de los dos niveles y puede variar de un grupo a otro (32).

Cada uno de los elementos de la ecuación del modelo nulo se puede identificar gráficamente en la figura 2.

Modelo multinivel general de 2 niveles

Para la estimación de un modelo de 2 niveles, se deben considerar 2 ecuaciones; la primera corresponde al nivel individual, para determinar la variación individual dentro de cada grupo (2).

Ecuación del modelo a nivel individual

$$Y_{ij} = B_{0j} + B_{1j}X_{1ij} + B_{2j}X_{2ij} \dots + r_{ij} \quad (4)$$

Y_{ij} : valor para Y del individuo i en el grupo j.

B_{0j} : promedio de Y en el grupo j.

B_{1j} : coeficiente del nivel 1.

X_{ij} : predictor del nivel 1 para el individuo i en el grupo j .

r_{ij} : efecto aleatorio del nivel 1.

La segunda ecuación considera la variación de los coeficientes entre grupos (2).

Los coeficientes B_{0j} y B_{1j} se convierten en una variable dependiente y pueden ser aleatorios, dependiendo del objetivo del análisis (32).

Ecuación del modelo a nivel grupo para B_{0j}

$$B_{0j} = \gamma_{00} + \gamma_{01}Z_j + \dots + U_{0j} \quad (5)$$

B_{0j} : intersección del primer nivel.

γ_{00} : promedio general de todos los grupos en los niveles superiores.

γ_{01} : pendiente del segundo nivel.

Z_j : predictor del segundo nivel.

U_{0j} : efecto aleatorio del nivel 2. Mide la desviación de la intersección de cada grupo respecto a la intersección general, después de ajustar por Z_j (2).

Ecuación del modelo a nivel grupo para B_{1j}

$$B_{1j} = \gamma_{10} + \gamma_{11}Z_j + \dots + U_{1j} \quad (6)$$

B_{1j} : pendiente del primer nivel.

γ_{10} : pendiente general.

γ_{11} : pendiente del segundo nivel.

Z_j : predictor del segundo nivel.

U_{1j} : Efecto aleatorio del nivel 2. Mide la desviación de la pendiente de cada grupo respecto a la pendiente general, después de ajustar por Z_j (2).

Con base en las especificaciones anteriores, se presentan las ecuaciones de los modelos: a) solo con variable independiente individual (ecuación 7) , b) solo con

variable independiente grupal (ecuación 8), c) con variables independiente individual y grupal (ecuación 9).

$$Y_{ij} = \gamma_{00} + \gamma_{10}X_{ij} + U_{0j} + r_{ij} \quad (7)$$

$$Y_{ij} = \gamma_{00} + \gamma_{01}Z_j + U_{0j} + r_{ij} \quad (8)$$

$$Y_{ij} = \gamma_{00} + \gamma_{10}X_{ij} + \gamma_{01}Z_j + U_{0j} + r_{ij} \quad (9)$$

Finalmente, se presenta la ecuación para especificar un modelo de efectos aleatorios (intercepto y pendiente aleatoria). El modelo incluye los efectos fijos de las variables del segundo nivel (γ_{01}), las variables del primer nivel (γ_{10}) y su interacción (γ_{11}) sobre el resultado a nivel individual Y_{ij} . También incluye los componentes aleatorios de la intersección (U_{0j}), la pendiente U_{1j} y el primer nivel (r_{ij}) (2).

Ecuación del modelo de efectos aleatorios

$$Y_{ij} = \gamma_{00} + \gamma_{01}Z_j + \gamma_{10}X_{ij} + \gamma_{11}Z_jX_{ij} + U_{0j} + U_{1j}X_{ij} + r_{ij} \quad (10)$$

La figura 3 presenta los componentes fijos y aleatorios del modelo multinivel comparado con un modelo de un solo nivel.

Tanto en el modelo con 1 como con 2 niveles, es importante especificar y reportar si los predictores han sido centrados en algún nivel específico, y esto es clave en el momento de interpretar los resultados (25). Según Peugh, el centrado implica “volver a escalar una variable predictora para que un valor de cero pueda interpretarse de manera significativa”, y esto es particularmente útil cuando se utilizan variables predictoras de tipo ordinal en las que el cero no tiene un valor (27).

Evaluación del ajuste del modelo final

Para evaluar y seleccionar el modelo con mejor ajuste, se han propuesto dos técnicas principales; el análisis de la razón de verosimilitud y la evaluación mediante índices de criterio de información como el criterio de Akaike (AIC) y el criterio de información bayesiano de Schwartz (BIC) (25). Estos dos últimos procuran evitar el ajuste excesivo que puede generar la agregación de parámetros al modelo e incluyen un término de penalización para el número de parámetros (34).

La razón de verosimilitud (*likelihood ratio*) o *Deviance*, tiene una distribución χ^2 con grados de libertad iguales a la diferencia de parámetros de los modelos anidados (33), y se calcula como

$D = -2LL = -2\log\left(\frac{\lambda_0}{\lambda_1}\right)$, donde λ_0 corresponde al valor de verosimilitud (*likelihoods*) del modelo sin predictores y λ_1 del modelo final (35). La prueba sugiere que a menor valor, mejor ajuste (33).

Según el AIC, el mejor modelo es aquel que explica la mayor cantidad de variación utilizando la menor cantidad posible de variables independientes, y se calcula como $AIC = -2LL + 2q$, siendo q el número de parámetros estimados (36). El mejor modelo será aquel con los valores de AIC más bajos, pero los valores numéricos que arroja su cálculo no tienen un significado (37).

Por su parte, el BIC se calcula como $BIC = -2LL + \ln(N)q$; siendo N el número de muestras del nivel 1 (36), este criterio aumenta la penalización cuando aumenta el tamaño de muestra, y por tanto, comparado con el AIC, es más difícil lograr la significancia (34). También se considera que el modelo con mejor ajuste es aquel con BIC más bajo (Alvarado WR. Aplicación de la Teoría de Modelos Multinivel

Lineal y No-Lineal utilizando el software especializado HLM7. En: X Congreso Internacional sobre la enseñanza de la matemática asistida por computadora.

Tecnológico de Costa Rica. 2017. Disponible en:

https://www.tec.ac.cr/sites/default/files/media/doc/10_memoria_1.pdf).

En cuanto a los *softwares* para la implementación de análisis multinivel, existen programas estadísticos como MLwiN (38), diseñados específicamente para estimar estos modelos, y otros como R que son programas de análisis estadístico general, que han incluido paquetes y procesos para el análisis de datos anidados. Albright y Marinova (39) presentan una guía sobre cómo estimar modelos multinivel utilizando SPSS, Stata, SAS y R.

Experiencias en investigación en salud con análisis multinivel

El análisis multinivel se ha utilizado para dar respuesta a preguntas de investigación relacionadas con desenlaces que dependen de factores tanto individuales como ambientales y sociales, es decir, que no dependen solo del sujeto en sí mismo, sino también de su interacción con el entorno y con los demás (5). A continuación, se describen algunos estudios que a juicio de los autores pueden ilustrar la aplicación de estos modelos en la investigación sanitaria (cuadro 1).

Flores et al. (43) plantearon un análisis multinivel exploratorio para identificar asociaciones entre la mortalidad relacionada con los opioides y los factores de riesgo a nivel de vecindario en Massachusetts (Estados Unidos), en busca de información que aportara en la planeación de intervenciones poblacionales y partiendo del hecho de que, según la evidencia disponible hasta el momento, las intervenciones individuales y del lado de la oferta eran insuficientes. Plantearon

una estructura multinivel clásica de individuos anidados en bloques censales y estos, a su vez, en condados (equivalente a municipios para Colombia). Para el análisis estadístico se plantearon modelos de regresión logística multinivel de efectos fijos (interceptos aleatorios), debido a que solo estaban interesados en conocer el efecto de las variables contextuales. Encontraron que en la población de estudio las muertes relacionadas con opioides podrían estar asociadas positivamente con el porcentaje de personas que vivían en la pobreza, la tasa de inseguridad alimentaria, el número de centros de salud calificados a nivel federal y los miligramos equivalentes de hidromorfona per cápita, pero inversamente asociados con el número de asociaciones sociales por 10.000 habitantes y el número de camas de hospital por 10.000 habitantes.

Por su parte, Shrout et al. (40) evaluaron los efectos intra y entre personas de la satisfacción con las relaciones románticas sobre la inflamación (marcadores séricos) a través del estrés psicológico percibido en sobrevivientes de cáncer de mama. Midieron las variables de interés antes del tratamiento, a los 6 meses y 18 meses después de finalizar el tratamiento. Plantearon un análisis de mediación que pretende comprender cómo una variable independiente produce un efecto indirecto sobre un resultado a través de una variable interviniente o mediadora. Dado que los datos tenían una estructura jerárquica con individuos en el nivel 2 (individuos) y medidas repetidas (no independientes) de cada individuo en el nivel 1, realizaron un análisis de mediación multinivel. El estudio concluyó que tanto en el análisis intra como en el análisis entre personas una mejor satisfacción con la relación está asociada a una menor percepción de estrés, lo que a su vez se relaciona con menores niveles de inflamación.

Park et al. (41) evaluaron los factores que afectan la gravedad de los accidentes de vehículos de motor comerciales. Los accidentes fueron clasificados en cuatro grupos; mortal, mayor, menor y sin lesiones. Incluyeron 865.622 accidentes. Consideraron variables a nivel individual y variables a nivel de grupos no anidados; empresas y municipios. Dada la estructura de los datos, para el análisis de la información plantearon cuatro modelos; un modelo *logit* de un solo nivel, dos modelos *logit* multinivel convencionales y un modelo *logit* multinivel con clasificación cruzada. Identificaron una relación inversa y estadísticamente significativa entre la proporción del presupuesto de infraestructura de transporte en el municipio y la gravedad de las muertes y las lesiones de todos los tipos de accidentes.

Finalmente, si bien la presente revisión pretende abordar los elementos fundamentales del uso clásico del análisis multinivel, es importante señalar que hay evidencia emergente sobre usos específicos; uno de ellos, la aplicación en salud pública, como el análisis multinivel de heterogeneidad individual y precisión discriminatoria (MAIHDA), planteado por Evans et al. (42), el cual consiste en aplicar modelos jerárquicos multinivel en el marco de la interseccionalidad, percepción crítica que sugiere que múltiples categorías sociales no son atributos individuales, sino sistemas interrelacionados de opresión que interactúan desde lo individual hasta lo estructural. En este ámbito, el análisis multinivel permite capturar las desigualdades en salud en diferentes estratos interseccionales de la población (44).

Este trabajo resume de forma preliminar el alcance y la utilidad de los modelos estadísticos multinivel en el ámbito de la investigación en salud. Si bien esta

metodología cuenta con varias décadas desde su aparición en publicaciones científicas masivas, prevalece todavía algún rezago para su aplicación debido a las consideraciones estadísticas y técnicas necesarias. Es por eso que se presenta en este documento una breve aproximación a las consideraciones necesarias a tomar en cuenta para su planteamiento y su construcción. Así mismo, se señalan algunos ejemplos que pueden ilustrar su aplicación práctica. El alcance de esta metodología, más allá de su capacidad para ponderar el efecto de una serie de variables sobre un desenlace particular, radica probablemente en la posibilidad de permitir la integración de información con estructuras complejas, en procura de ahondar en la comprensión de la compleja trama de interacciones y relaciones que a nivel social, clínico y biológico, confluyen en la aparición de condiciones de salud, lo mismo que para la previsión de estrategias e iniciativas que, desde una mirada más amplia, incrementen los aspectos a considerar en el momento de prevenir y promover acciones sanitarias benéficas.

Si bien los modelos siguen requiriendo sendos conocimientos estadísticos, la existencia de *softwares* como MLwiN facilitan la tarea de su construcción, a la vez que llaman la atención sobre su uso riguroso. Bancos de información con estructuras más complejas de anidamiento, frecuencias de ocurrencia bajas en el desenlace de interés, desbalanceo en la cantidad de información disponible, entre otras situaciones, constituyen múltiples desafíos para el uso de esta metodología; sin embargo, dado su aporte en la extensión de los modelos de regresión convencionales, el modelamiento multinivel se convierte en uno de los pilares fundamentales a fortalecer en lo concerniente a la indagación cuantitativa en salud. Su uso no debería depender de los gustos particulares del investigador,

sino de la necesidad de explicar efectos de manera independiente en grupos que naturalmente se aglomeran, lo cual resalta su importancia en el estudio de problemas de salud pública y colectiva.

Conflictos de intereses

Los autores declararon no tener ningún conflicto de interés.

Financiación

No se recibió financiación para el desarrollo ni la publicación de esta revisión.

Referencias

1. De la Cruz F. Modelos multinivel. *Rev Per Epidemiol.* 2008;12:1-8.
2. Diez-Roux AV. Multilevel analysis in public health research. *Annu Rev Public Health.* 2000;21:171-92. <https://doi.org/10.1146/annurev.publhealth.21.1.171>
3. Usami S. Generalized sample size determination formulas for experimental research with hierarchical data. *Behav Res Methods.* 2014;46:346-56. <https://doi.org/10.3758/s13428-013-0387-1>
4. University of Bristol. Centre for Multilevel Modelling. Sample sizes for multilevel models. Fecha de consulta: 26 de diciembre de 2021. Disponible en: <http://www.bristol.ac.uk/cmm/learning/multilevel-models/samples.html>
5. Diez Roux AV. La necesidad de un enfoque multinivel en epidemiología. *Región Soc.* 2008;20:77-91.
6. Fisher AJ, Medaglia JD, Jeronimus BF. Lack of group-to-individual generalizability is a threat to human subjects research. *Proc Natl Acad Sci USA.* 2018;115:E6106-15. <https://doi.org/10.1073/pnas.1711978115>
7. University of Bristol. Centre for Multilevel Modelling. What are multilevel models and why should I use them? Fecha de consulta: 16 de noviembre de 2022.

Disponibile en: <http://www.bristol.ac.uk/cmm/learning/multilevel-models/what-why.html>

8. Damtie Y, Kefale B, Yalew M, Arefaynie M, Adane B. Multilevel analysis of determinants of polygyny among married men in Ethiopia. *BMC Public Health*. 2021;21:1677. <https://doi.org/10.1186/s12889-021-11701-z>
9. Dessie ZG, Zewotir T, Mwambi H, North D. Multivariate multilevel modeling of quality of life dynamics of HIV infected patients. *Health Qual Life Outcomes*. 2020;18:80. <https://doi.org/10.1186/s12955-020-01330-2>
10. Hagadorn JI, Shaffer ML. Hierarchical data structures and multilevel modeling. *J Pediatr*. 2019;212:250-1. <https://doi.org/10.1016/j.jpeds.2019.05.042>
11. Huang F. Multilevel Modeling Myths. *Sch Psychol Q*. 2018;33:492-9. <https://doi.org/10.1037/spq0000272>
12. Ntani G, Inskip H, Osmond C, Coggon D. Consequences of ignoring clustering in linear regression. *BMC Med Res Methodol*. 2021;21:139. <https://doi.org/10.1186/s12874-021-01333-7>
13. Mumper M. American Psychological Association. 2017. Multilevel modelling. Fecha de consulta: 16 de noviembre de 2022. Disponible en: <https://www.apa.org/science/about/psa/2017/01/multilevel-modelling>
14. University of Bristol. Centre for Multilevel Modelling. Multilevel models: An introduction and FAQs. Fecha de consulta: 15 de noviembre de 2022. Disponible en: <http://www.bristol.ac.uk/cmm/learning/multilevel-models/>
15. Barker KM, Dunn EC, Richmond TK, Ahmed S, Hawrilenko M, Evans CR. Cross-classified multilevel models (CCMM) in health research: A systematic

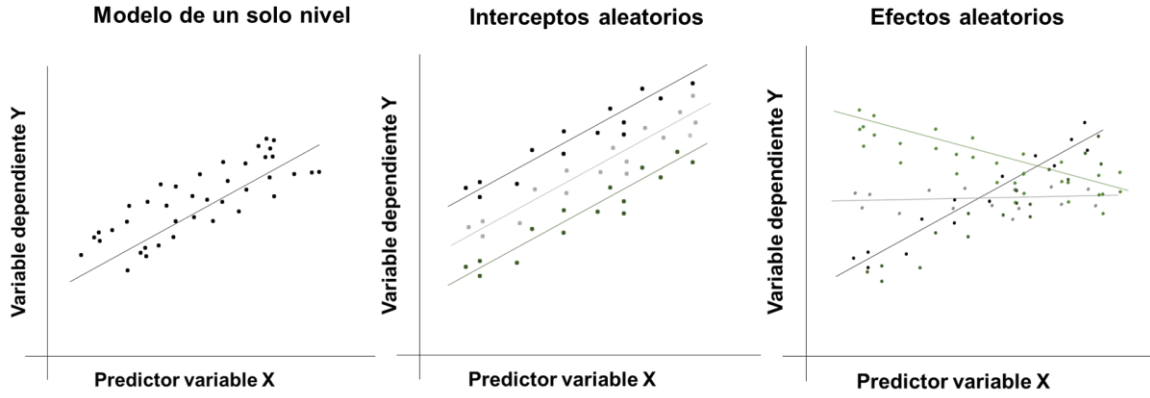
- review of published empirical studies and recommendations for best practices. *SSM Popul Health*. 2020;12. <https://doi.org/10.1016/j.ssmph.2020.100661>
16. Finch H, Bolin JE, Kelley K. *Multilevel modeling using R*. 2.^a ed. Boca Raton: Chapman and Hall/CRC; 2019. p. 252.
17. University of Bristol. Centre for Multilevel Modelling. Random intercept models. Fecha de consulta: 27 de diciembre de 2021. Disponible en: <https://www.bristol.ac.uk/cmm/learning/videos/random-intercepts.html>
18. Speed TP. Restricted Maximum Likelihood: Overview. En: Balakrishnan N, Colton T, Everitt B, Piegorisch W, Ruggeri F, Teugels JL, editores. *Wiley StatsRef: Statistics Reference Online*. Wiley; 2014. <https://doi.org/10.1002/9781118445112.stat01451>
19. Bolin JH, Finch WH, Stenger R. Estimation of random coefficient multilevel models in the context of small numbers of level 2 clusters. *Educ Psychol Meas*. 2019;7:217-48. <https://doi.org/10.1177/0013164418773494>
20. Webster TF, Hoffman K, Weinberg J, Vieira V, Aschengrau A. Community- and Individual-Level Socioeconomic Status and Breast Cancer Risk: Multilevel Modeling on Cape Cod, Massachusetts. *Environ Health Perspect*. 2008;116:1125-9. <https://doi.org/10.1289/ehp.10818>
21. Austin PC, Wagner P, Merlo J. The median hazard ratio: a useful measure of variance and general contextual effects in multilevel survival analysis. *Stat Med*. 2017;36:928-38. <https://doi.org/10.1002/sim.7188>
22. Austin PC, Stryhn H, Leckie G, Merlo J. Measures of clustering and heterogeneity in multilevel Poisson regression analyses of rates/count data. *Stat Med*. 2018;37:572-89. <https://doi.org/10.1002/sim.7532>

23. Merlo J, Chaix B, Ohlsson H, Beckman A, Johnell K, Hjerpe P, et al. A brief conceptual tutorial of multilevel analysis in social epidemiology: using measures of clustering in multilevel logistic regression to investigate contextual phenomena. *J Epidemiol Community Health*. 2006;60:290-7.
<https://doi.org/10.1136/jech.2004.029454>
24. Ballesteros SM, Moreno-Montoya J. Factores individuales y departamentales asociados con la prevalencia de limitación funcional entre ancianos colombianos: Un análisis multinivel. *Cad Saude Publica*. 2018;34:12.
<https://doi.org/10.1590/0102-311X00163717>
25. Dedrick RF, Ferron JM, Hess MR, Hogarty KY, Kromrey JD, Lang TR, et al. Multilevel Modeling: A Review of Methodological Issues and Applications. *Rev Educ Res*. 2009;79:69-102. <https://doi.org/10.3102/0034654308325581>
26. Catalán-Reyes MJ, Galindo-Villardón MP. Utilización de los modelos multinivel en investigación sanitaria. *Gac Sanit*. 2003;17(Supl. 3):35-52.
27. Peugh JL. A practical guide to multilevel modeling. *J Sch Psychol*. 2010;48:85-112. <https://doi.org/10.1016/j.jsp.2009.09.002>
28. Osorio AM, Romero GA, Bonilla H, Aguado LF. Socioeconomic context of the community and chronic child malnutrition in Colombia. *Rev Saúde Pública*. 2018;52:1-12. <https://doi.org/10.11606/S1518-8787.2018052000394>
29. Gabriëlle I, Jongmans M. Intra-class correlation testing to examine Intra-group differences [Bachelor thesis]. Enschede: University of twente; 2021.
30. Yamana H. Introduction to Multilevel Analysis. *Ann Clin Epidemiol*. 2021;3:5-9.
https://doi.org/10.37737/ace.3.1_5

31. Killip S, Mahfoud Z, Pearce K. What is an intracluster correlation coefficient? Crucial concepts for primary care researchers. *Ann Fam Med.* 2004;2:204-8. <https://doi.org/10.1370/afm.141>
32. Ramos Rodríguez FJ, Lara Porras AM, Molina Muñoz D. Competencia matemática de los estudiantes andaluces: un análisis multinivel de la encuesta PISA 2015. *Pi-InnovaMath.* 2019;2. <https://doi.org/10.5944/pim.2.2019.24130>
33. Alarcón R, Blanca MJ, Arnau J, Bono R. Modelado jerárquico por pasos: análisis multinivel del estrés cotidiano en adolescentes. *Rev Mex Psicol.* 2015;32:124-33.
34. Vrieze SI. Model selection and psychological theory: A discussion of the differences between the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). *Psychol Methods.* 2012;17:228-43. <https://doi.org/10.1037/a0027127>
35. Oliver JC, Rosel J, Jara P. Modelos de regresión multinivel: aplicación en psicología escolar. *Psicothema.* 2000;12:487-94.
36. Kim S, Jeong Y, Hong S. The impact of ignoring a crossed factor in cross-classified multilevel modeling. *Front Psychol.* 2021;12:637645. <https://doi.org/10.3389/fpsyg.2021.637645>
37. Portet S. A primer on model selection using the Akaike Information Criterion. *Infect Dis Model.* 2020;5:111-28. <https://doi.org/10.1016/j.idm.2019.12.010>
38. University of Bristol. Centre for Multilevel Modelling. 2022. MLwiN. Fecha de consulta: 28 de febrero de 2023. Disponible en: <http://www.bristol.ac.uk/cmm/software/mlwin/>

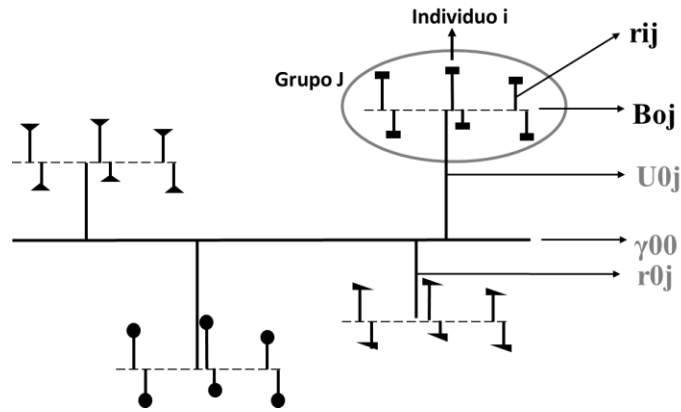
39. Albright JJ, Marinova DM. Estimating Multilevel Models using SPSS. Stata. 2010; Fecha de consulta: 28 de febrero de 2023. Disponible en:
<https://scholarworks.iu.edu/dspace/handle/2022/19737>
40. Flores MW, Cook BL, Mullin B, Halperin-Goldstein G, Nathan A, Tenso K, et al. Associations between neighborhood-level factors and opioid-related mortality: a multilevel analysis using death certificate data. *Addict Abingdon Engl*. 2020;115:1878-89. <https://doi.org/10.1111/add.15009>
41. Shrout MR, Renna ME, Madison AA, Alfano CM, Povoski SP, Lipari AM, et al. Relationship satisfaction predicts lower stress and inflammation in breast cancer survivors: a longitudinal study of within-person and between-person effects. *Psychoneuroendocrinology*. 2020;118:104708.
<https://doi.org/10.1016/j.psyneuen.2020.104708>
42. Park HC, Kim DK, Kho SY, Park PY. Cross-classified multilevel models for severity of commercial motor vehicle crashes considering heterogeneity among companies and regions. *Accid Anal Prev*. 2017;106:305-14.
<https://doi.org/10.1016/j.aap.2017.06.009>
43. Evans CR, Williams DR, Onnela JP, Subramanian SV. A multilevel approach to modeling health inequalities at the intersection of multiple social identities. *Soc Sci Med*. 2018;203:64-73. <https://doi.org/10.1016/j.socscimed.2017.11.011>
44. Arias-Uriona AM, Losantos M, Bedoya P. La interseccionalidad como herramienta teórico-analítica para estudiar las desigualdades en salud en las Américas. *Rev Panam Salud Pública*. 2023;47:1.

Figura 1. Comparación entre modelo de un solo nivel, modelo de interceptos aleatorios y modelo de efectos aleatorios



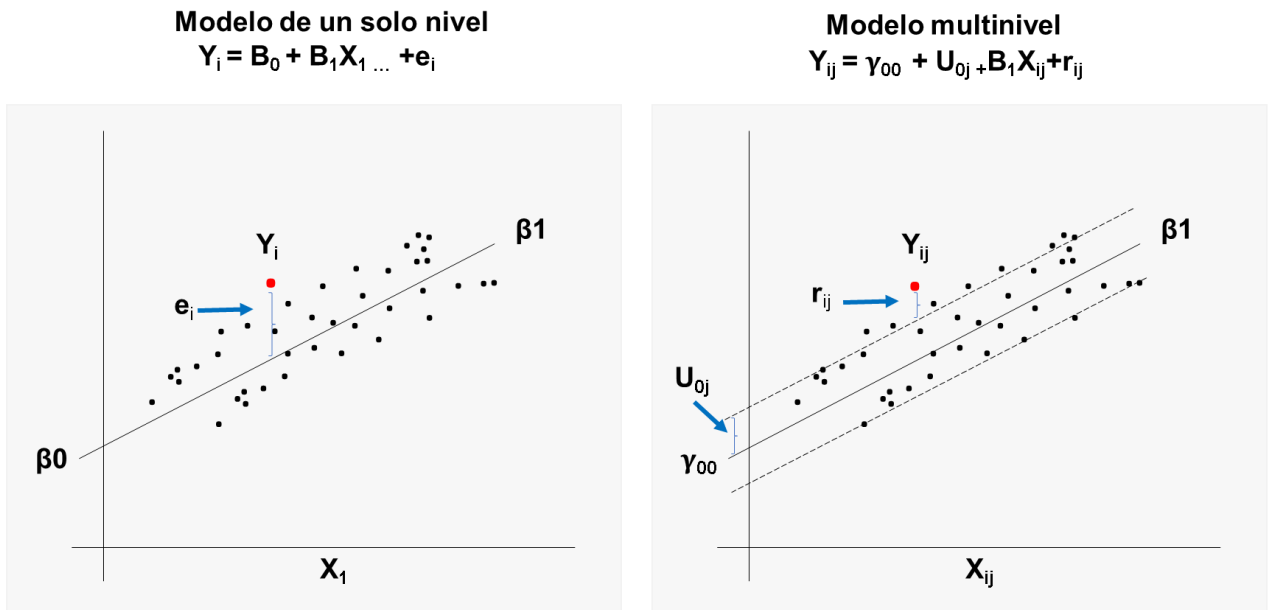
Fuente: Construcción propia.

Figura 2. Elementos incluidos en el modelo nulo



Fuente: Construcción propia.

Figura 3. Componentes fijos y aleatorios del modelo multinivel comparado con un modelo de un solo nivel



Fuente: Construcción propia.

Cuadro 1. Ejemplos de aplicación del análisis multinivel en investigación en salud

Estudio	Población y estructura	Modelos	Resultados	Análisis complementario
Asociaciones entre los factores a nivel de vecindario y la mortalidad relacionada con los opioides: un análisis multinivel utilizando datos de certificados de defunción. Flores et al. (2020) (40).	Estructura multinivel clásica: Nivel 1: individuos; 3.809 muertes prematuras relacionadas con opioides y 8.729 no relacionadas con opioides. Nivel 2: 2.517 grupos de bloques censales. Nivel 3: 14 condados.	Modelo 1: De intersecciones aleatorias sin covariables. Modelo 2: Con datos a nivel individual y a nivel de área.	Algunos resultados: Mayor mortalidad a mayor porcentaje de personas en la pobreza (OR = 1,01, IC 95 % = 1,00 – 1,01) y a mayor tasa de inseguridad alimentaria (OR = 1,21, IC 95 % = 1,07 – 1,37). Menor mortalidad a mayor número de camas de hospital por 10.000 personas (OR = 0,78, IC 95% = 0,68–0,88).	Tres análisis de sensibilidad. En los dos primeros se identificaron hallazgos similares a los del modelo principal, en el tercero (se agregaron secuencialmente categorías de factores de riesgo a nivel de área) se observaron varias asociaciones que no eran frecuentes en el modelo principal.
La satisfacción con la relación predice un menor estrés e inflamación en sobrevivientes de cáncer de mama: un estudio longitudinal de los efectos dentro de la persona y entre personas. ShROUT et al. (2020) (41).	Medidas repetidas Nivel 1: 3 mediciones; antes, durante y después del tratamiento. Nivel 2: 139 participantes recibieron tres visitas.	Modelos multinivel para proteína C reactiva (PCR) sérica y compuesto de citoquinas. 1. Dentro de las personas. 2. Entre personas.	Respecto a los resultados relacionados con la PCR a nivel intrapersona: A mayor satisfacción en la relación, menor estrés percibido ($\beta = 0,19^*$). Mayor estrés se relacionó con niveles más altos de PCR ($\beta = 0,55^{**}$). La satisfacción con la relación no fue asociada directamente con la PCR, pero sí indirectamente a través del estrés percibido ($\beta = -0,10^*$).	Modelos hipotéticos para explorar si el estrés estaba indirectamente relacionado con la inflamación a través de la satisfacción con la relación tuvieron peor ajuste que los modelos planteados inicialmente: AIC = 1825,72 vs. 1954,36 para PCR. AIC = 2064,40 vs. 2192,36 para el compuesto de citoquinas.
Modelos multinivel de	Clasificación cruzada	Modelo 1: un solo nivel.	Algunos hallazgos incluyen:	No se reporta

<p>clasificación cruzada para la gravedad de los accidentes automovilísticos comerciales teniendo en cuenta la heterogeneidad entre empresas y regiones. Park et al. (2017) (42).</p>	<p>Nivel 1: 86.622 choques de vehículos.</p> <p>Nivel 2: Dos grupos no anidados; 1.875 empresas y 230 municipios.</p>	<p>Modelo 2: dos niveles; uno para empresa y otro para región.</p> <p>Modelo 3: combinó empresa y región.</p> <p>Cada modelo se planteó para 4 tipos de transporte. Total: 12 modelos.</p> <p>La mayoría con CCI > 12,1 %.</p>	<p>Consecuencias menos graves cuando el conductor era el dueño del taxi ($\beta = -0,490^*$).</p> <p>En municipios con presupuestos de infraestructura de transporte relativamente grandes hubo menos accidentes graves para camiones grandes ($-0,0139^{**}$), autobuses ($-0,008^{**}$), y taxis ($-0,005^{**}$).</p>	
---	---	---	---	--

* $p < 0,05$, ** $p < 0,01$