

Los modelos de niveles múltiples: una estrategia analítica para el estudio de los problemas de salud de la población

Multilevel models: an analysis strategy for the study of health problems in society

Irene Luppi¹

Gabriela Boggio²

¹Instituto de la Salud Juan Lazarte

²Instituto de la Salud Juan Lazarte – Consejo de Investigaciones de la Universidad Nacional de Rosario (U.N.R.) - Escuela de Estadística. Facultad de Ciencias Económicas U.N.R.

Correspondencia: Irene Luppi. Pje. 12 de Octubre 860. Rosario 2000 Argentina. E-mail: iluppi@arnet.com.ar

Resumen

Se presenta una discusión teórico metodológica sobre la aplicabilidad de modelos de niveles múltiples para el estudio de los procesos de salud/enfermedad, sus determinantes y condicionantes, en función de la estratificación de la sociedad y de las condiciones de vida de sus habitantes. Se recupera una noción de población según la perspectiva de la teoría de los sistemas complejos jerárquicos que busca no reducir la realidad, sino una construcción del problema procurando identificar distintos niveles de abstracción para su abordaje. Estos modelos constituyen una opción que supera las experiencias previas, con la aplicación de técnicas estadísticas convencionales, dado que permiten analizar simultáneamente distintos niveles de agregación conservando su estructura jerárquica. Se consideran la influencia de las variables teniendo en cuenta su pertenencia a unidades mayores y la asociación potencialmente existente entre las unidades de un mismo nivel, es decir, la correlación intracase entre variables relativas a individuos, familias, grupos, próximos entre sí, que comparten condiciones semejantes. Se evita de este modo sobredimensionar el efecto de las variables de macro nivel. Los modelos de niveles múltiples resultan particularmente adecuados para valorar desigualdades en el proceso salud/enfermedad/atención de los grupos poblacionales y analizar cómo los contextos sociales afectan los resultados y los riesgos de salud individuales. Se destaca la necesidad de desarrollar estrategias de producción de información y de análisis que permitan reconocer niveles de explicación y de intervención, para proveer insumos y desencadenar acciones adecuadas a las especificidades locales, a nivel de las micro-áreas, con miras a lograr una mayor equidad en salud.

Palabras clave: Modelos de niveles múltiples. Sistemas jerárquicos. Métodos de estudio epidemiológicos. Estudios sobre desigualdades e inequidades en salud. Espacio local y de micro-área.

Abstract

This paper presents the theoretical-methodological discussion about the applicability of multiple level models in the study of the health/sickness process, its determinants and conditioning factors, as a function of the stratification of society and the living conditions of its inhabitants. It goes back to the concept of population according to the theory of hierarchical complex systems, which seeks not to reduce reality, but rather to build the problem trying to identify different levels of abstraction in its approach. These models are options to overcome prior experiences, with the application of conventional statistical techniques, given that they make it possible to simultaneously analyze different levels of aggregation, while keeping its hierarchical structure. They consider the influence of the variables taking into account their belonging to larger units and the potential association existing between the units of a same level, that is, the intraclass correlation among variables relative to individuals, families and groups, close amongst themselves, which share similar conditions. In this manner, it tries to avoid oversizing the effect of macro level variables. The multiple level models are particularly appropriate to evaluate inequalities in the health/sickness/care process of the population groups and to analyze how social contexts affect the results and health risks of people. It highlights the need to develop information production strategies and analyses that make it possible to recognize levels of explanation and intervention to provide inputs and trigger actions suited to local specificities, at the level of micro-areas, so as to have more equity in healthcare.

Keywords: Multiple level models. Hierarchical systems. Epidemiological study methods. Studies of the inequalities and equalities in healthcare. Local space and micro area.

Interrogantes epidemiológicos y contribuciones estadísticas

Se atribuye a la Epidemiología el carácter de disciplina básica de la Salud Pública por asumir el estudio de las relaciones entre las condiciones de vida y la situación de salud en diferentes grupos de población, y contribuir al conocimiento de sus necesidades y expectativas, así como de los factores que intervienen en la efectividad de las intervenciones de los servicios de salud. Este rol se enfatiza al considerar la equidad como un propósito básico de la organización social para dar respuestas a los problemas, necesidades y demandas de individuos y grupos en relación a su salud.

Son sus funciones básicas identificar las diferencias en los patrones de morbilidad y en el perfil epidemiológico de las poblaciones en territorios particulares, así como la evaluación de las repercusiones de las distintas opciones de intervención para orientar la asignación de recursos según necesidades diferenciales y evitar el desperdicio de recursos escasos¹.

Por su parte la Estadística, que en su origen designa la medida del Estado -de allí su denominación derivada del latín *status*-, surge como respuesta a la necesidad de contabilizar los atributos de una nación. Los análisis de William Petti en la "Aritmética Política" en el siglo XVII y las estadísticas médicas de John Graunt, en su publicación *Observation upon the bills of mortality, en 1662*, sientan las bases para el tratamiento numérico de los hechos vitales, para la cuantificación de sanos y enfermos, de allí que sus obras sean reconocidas como precursoras de la Estadística, la Epidemiología y la Demografía, si bien el antecedente de los censos es de épocas pretéritas².

Los métodos empleados en la Estadística, por su capacidad para la descripción de los fenómenos biológicos y sociales, influyeron en gran medida en el desarrollo de los estudios de morbilidad y mortalidad en Inglaterra y Gales, reconociéndose a William Farr como el primer estadísti-

co médico, en el siglo XIX. Desde Farr hasta nuestros días, la Estadística, en su integración con la Clínica, contribuye a la constitución histórica de la Epidemiología, junto con un tercer elemento, el reconocimiento del carácter eminentemente colectivo de sus objetos, como campo de conocimiento comprometido con los procesos y con la transformación de la situación de salud de las poblaciones^{3,4}.

Si se pretendieran resumir los interrogantes principales que orientan las aplicaciones estadísticas en este campo, podría decirse que ellos buscan dar respuesta a las siguientes preguntas básicas: ¿en qué medida ocurren los problemas de salud/enfermedad?, ¿bajo qué circunstancias?, ¿cómo apoyar la toma de decisiones frente a distintas alternativas? Preguntas que en términos estadísticos pueden plantearse como: ¿con qué frecuencia se manifiestan los problemas?, ¿qué asociaciones es posible identificar? ¿cuáles son las diferencias significativas?

Quizás no esté de más insistir que la Estadística basa dichas respuestas en el análisis de información relativa a un colectivo, que se obtiene a partir de alguna característica o propiedad (variable) común a todas las unidades elementales que lo integran, que pueden ser personas, unidades físicas, eventos. Los términos utilizados -población, colectivo- no son casuales, subyace la noción de grandes volúmenes de información y el instrumental que se desarrolla busca proveer elementos para su descripción e interpretación. De modo que es posible, ciertamente, definir como población cualquier conjunto de individuos agrupados según intereses particulares de investigación o de intervención, basándose en algún atributo o condición especial (portadores de una dolencia según ingreso, hábito de fumar, etc.).

En este sentido se ha desarrollado un importante arsenal conceptual y metodológico vinculado casi exclusivamente a la indagación causal y explicativa de la

Epidemiología, con una profusa aplicación de técnicas estadísticas, fundamentalmente orientadas a medir asociaciones entre “causas” y efectos en salud, en estudios de asociaciones entre riesgos y problemas de salud de individuos. El interrogante principal a responder es *¿por qué se enferman las personas?*, conocer sus características distintivas, particularmente las de aquéllas que se encuentran en los extremos de la distribución de valores. Esta inquietud, basada en las experiencias de la medicina Clínica, ha seguido moldeando el pensamiento epidemiológico, constituyendo la base de estudios de casos y controles, ocupados en descubrir cómo difieren los individuos sanos de los enfermos, asimismo de estudios de cohorte, abocados a la búsqueda de factores de riesgo que permitan identificar determinados individuos más susceptibles a contraer la enfermedad.

No se ha registrado una contribución equivalente en la conceptualización teórico metodológica de los problemas de salud abordados a nivel de los grupos sociales, cuyo atributo esencial es la interacción entre sus miembros. En este caso lo que se busca no es limitar la atención a las comparaciones intrapoblacionales, y la pregunta que interesa es *“¿por qué las poblaciones tienen un determinado perfil de salud?, o bien, ¿por qué determinados problemas de salud predominan en ciertas poblaciones?”*⁵.

Esta perspectiva supone incorporar a la preocupación predominante en la investigación epidemiológica de los problemas de salud individual, con énfasis en las enfermedades y sus riesgos, la indagación de los diferenciales de salud y enfermedad a nivel de los grupos sociales. Es decir, no limitarse a los problemas de la salud **de** las poblaciones sino abordar el estudio de la salud **en** las poblaciones^{6,*}.

La problemática en este sentido tiene implicaciones relevantes para la definición de estrategias de intervención, en sus límites y potencialidades en el espacio de la

* Resaltado del autor.

Salud Pública, que ofrece, de modo hegemónico, una direccionalidad y una racionalidad más subsidiaria de la Clínica, al apartarse progresivamente de una perspectiva poblacional hacia una individual.

La noción de población como un sistema complejo jerárquico

En términos teórico metodológicos, la cuestión central es reconocer como concepto clave que la población constituye un nivel de organización diferente al individual, cuya esencia es la interacción entre sus miembros. La comprensión de la interacción entre las partes, en una totalidad que se busca conocer en toda su amplitud, integrando las partes al interior de la propia lógica analítica, es un verdadero desafío. La noción de la complejidad surgió en la discusión científica como una vía para explicar la realidad o los sistemas vivos mediante modelos que buscan no sólo describir los elementos de los objetos sino, principalmente, las relaciones que se establecen entre ellos. Es posible identificar en los distintos niveles de organización de la realidad cualidades emergentes propias de cada nivel. Sin embargo no deja de ser cierto que esta tentativa encuentra sus límites en la "indivisibilidad" de lo real, que inevitablemente sugiere como reductora la construcción de todo modelo⁷.

En la búsqueda de explicación de lo real, la delimitación de planos de referencia se impone a la construcción teórica de los problemas como una necesidad no sólo para su comprensión, sino para encontrar modos y recursos operativos que permitan abordarla. Las conceptualizaciones proporcionadas por la teoría de los sistemas complejos y jerárquicos hacen su aporte en ese sentido, no buscando reducir la realidad sino procurando identificar distintos niveles de abstracción a los fines de su estudio*.

Estas nociones no son ajenas al campo de la salud, el ser humano se compone de

unidades menores las que son también sistemas definidos por su función particular (sistema circulatorio, digestivo, etc.) a su vez constituidos por células y éstas por otros niveles de complejidad menor. Al mismo tiempo, los individuos son miembros de un grupo social, que constituye una unidad, la que por su parte interactúa con otras unidades poblacionales generando una organización en jerarquías. La representación más clara es la de las cajas chinas o de las muñecas rusas, en las que cada vez que abrimos una caja o muñeca, aparece otra menor, que es también el mismo objeto, sólo que de menor tamaño y complejidad⁸.

Se reconoce que cada objeto de estudio que se aborda como una totalidad es en realidad un subsistema de otros mayores, los que interactúan con otros subsistemas de su mismo nivel y que asimismo está estructurado internamente por subsistemas de nivel inferior de complejidad, los que también interactúan entre sí. De este modo una población se comprende como un agrupamiento de subpoblaciones que interactúan; cada una de ellas es simultáneamente una totalidad correspondiente a un nivel inferior y una unidad integrante de una totalidad mayor. A cada nivel organizacional corresponden atributos que le son propios; son ejemplos, la denominada inmunidad de grupo que no existe en una persona aislada, del mismo modo, el poder, que sólo emerge de la interacción de sujetos en un conglomerado social. Estos atributos se expresan sólo como potencialidades propias de un nivel y a la vez están condicionadas por la forma cómo interactúan en el mismo nivel de estudio (como unidades) con otros sistemas del mismo nivel y en el contexto de las unidades mayores de las que forman parte.

Esta propuesta de abordar la complejidad mediante su construcción teórica tiene una importancia extraordinaria para la investigación en salud, por cuanto facilita

* Estas nociones le valieron a Herber Simon el Premio Nobel de Economía en 1978 y son reconocidas como las iniciadoras de las investigaciones en inteligencia artificial.

la delimitación de los problemas de estudio y permite una mayor comprensión entre las relaciones definidas en el espacio general, particular o singular en la determinación de los fenómenos de salud⁹. Es dicha construcción teórico metodológica de la indagación la que finalmente orientará la definición de las unidades de observación y de análisis, la selección de los procedimientos y técnicas para poder dar cuenta de la integralidad de los fenómenos estudiados y en cada caso es fundamental buscar los procesos mediadores a través de los cuales las interacciones se expresan en los niveles más próximos al que se delimita como el de abordaje.

Castellanos advierte que si como estrategia de indagación se optara por abordar un único nivel de análisis, con sus características, esto no debiera oscurecer que se le aborda con el propósito de comprender en qué medida las condiciones o atributos emergentes que le son propios, son expresión de las unidades menores cuya interacción las tornan posibles o son tomadas como unidad de un sistema mayor con sus interacciones con otras unidades del mismo nivel⁶.

La diferenciación de niveles de organización de la realidad en la investigación epidemiológica es ilustrada magistralmente por Susser con los trabajos pioneros de John Snow para demostrar su hipótesis sobre la relación entre el suministro de agua contaminada y las muertes por cólera en 1853, en Londres¹⁰.

También merece destacarse la reflexión de este autor a propósito de las falacias en las que se incurre cuando se busca extraer conclusiones sin tomar recaudo de esta estructura compleja¹⁰. Esta noción se tornó polémica en la evaluación de la eficacia de los estudios ecológicos, así denominados por tomar como unidad de análisis no los individuos sino las poblaciones^{4,11}. Susser, a diferencia de la preocupación hegemónica, advierte sobre la posibilidad de error en cualquier dirección, pudiendo darse la falacia por agregación como la inversa, la falacia atomista.

En ambos casos, los autores han focalizado su interés en explicitar los errores que se cometen al establecer asociaciones, sólo que buscando identificar factores fuente de variación se ha dejado de lado una cuestión de fondo, que las falacias son consecuencia de un problema de orden metodológico debido a un desajuste entre la concepción teórica del problema a investigar, la producción de observables y los procedimientos analíticos aplicados.

La dificultad surge, tanto en las investigaciones epidemiológicas sobre condiciones de vida y salud así como las dirigidas a valorar el efecto e impacto de las intervenciones realizadas, cuando no se reconoce a la población como una estructura jerárquica, siendo que los individuos pertenecen a grupos familiares, con determinadas relaciones entre sus miembros, que a su vez estas familias están integradas a un grupo social, con particulares condiciones de vida y de salud, de acceso y de utilización de los servicios de atención; que los grupos sociales comparten un espacio social en un territorio determinado, que no es una mera delimitación geográfica sino que es producto de dinámicas sociales particulares, y que será una totalidad compleja para la identificación y desencadenamiento de acciones sobre determinantes de problemas de salud/enfermedad y sus efectos¹².

Hasta hace relativamente pocos años las opciones brindadas por la Estadística para estudiar los problemas de salud a nivel poblacional, en coherencia con un modelo interpretativo de los mismos, la teoría de la multicausalidad, consistían en analizar una multiplicidad de variables simultáneamente mediante técnicas multivariadas, desarrolladas ampliamente por numerosos especialistas¹³⁻¹⁵; muchas de ellas procuran establecer, con distinto grado de refinamiento, relaciones entre variables y medir la forma e intensidad de tal asociación. Se trata de aplicaciones donde se presentan variables "explicativas" para representar los fenómenos con influencia potencial sobre un problema de

salud o una intervención en particular.

En estas técnicas todas las variables se tratan de forma "igualitaria", en un mismo plano, lo que implica un empobrecimiento de la construcción teórica de los problemas de salud/enfermedad a un conjunto de variables sin considerar jerarquías de organización entre ellas, en una simplificación conceptual de la sociedad como estructura organizada. Abundan hasta el cansancio los ejemplos de estudios epidemiológicos en donde se mezclan en una lista heterogénea de variables, factores biológicos, físicos, químicos con otras referidas a factores demográficos, sociales, económicos, conformando un conjunto indiscriminado en que todo tiene la misma relevancia y significado^{16,17}.

Un recurso analítico para recuperar niveles jerárquicos reconocidos en la definición del problema, ha sido establecer subdominios a partir de una variable y habilitar el estudio al interior de cada uno de ellos y su comparación de modo de identificar diferencias entre grupos, según clase social, inserción laboral, niveles de pobreza, área de residencia. Se recurre a técnicas estadísticas que toman en cuenta las interrelaciones entre variables, entre ellas, las que permiten identificar y valorar el patrón de relaciones entre atributos expresados como variables categóricas, de generalizada aplicación en las ciencias sociales y de salud, donde la información queda expresada en forma de frecuencias^{15, 18,19}.

Pueden mencionarse experiencias de aplicación como la realizada al analizar las condiciones de vida y de salud en el primer año de vida de una cohorte de 652 recién nacidos en Rosario – Argentina^{20, 21}. En esa oportunidad la preocupación por restaurar de algún modo la jerarquía de niveles identificados en la concepción teórica del problema impuso un esfuerzo metodológico particular^{16, 22}.

La primera dificultad a resolver, fue maximizar la utilización de la vasta información reunida superando la limitación impuesta por el tamaño de la cohorte que no permitía análisis estadísticos con la in-

clusión de la totalidad, o bien un número importante de variables explicativas para una respuesta dada. La segunda, encontrar una estrategia analítica que posibilitara restaurar la estructura jerárquica identificada en la concepción del problema. El recurso aplicado fue encarar múltiples estudios parciales con a lo sumo cuatro variables independientes a los fines de valorar efectos, individuales y en interacción para cada respuesta. Se formularon en consecuencia, numerosas tablas de contingencia abarcando los tres niveles explicativos de ese proceso: estructural, valorado a través de la posición socioeconómica, particular, mediante las condiciones familiares de vida, e individual, expresado en las conductas maternas y condiciones biológicas del niño. Otras tablas incluyeron sólo variables expresión de factores de un nivel particular del proceso de salud-enfermedad para profundizar en el análisis de las asociaciones entre ellas.

Se buscó así, suplir la parcialidad de la perspectiva que ofrecía cada tabla de contingencia con un procedimiento de generación de tablas en una dinámica de pruebas sucesivas, orientada por fundamentos teóricos y no a expensas de una simple combinatoria de factores. Se construyeron numerosas estructuras parciales de análisis, que dieron lugar a una doble lectura interpretativa de la información empírica, al interior de cada recorte analítico determinado por las tablas planteadas y entre ellos. Este largo proceso exploratorio permitió visualizar la persistencia y grado de coherencia de determinados factores en las diferentes estructuras alternativas.

Similar dificultad se enfrentó en un estudio cuyo objetivo fue valorar la importancia de los atributos referidos a los antecedentes de la enfermedad de Chagas y a factores socioeconómicos y culturales en embarazadas infectadas en Rosario (Argentina). En esa oportunidad se reunieron 154 casos y un número equivalente de controles. Mediante un procedimiento próximo al mencionado, se analizaron tablas de contingencia destinadas a valorar el apor-

te relativo de los factores sociales, conductuales y biológicos de mayor relevancia que operan en la explicación de la condición de mujer infectada. Asimismo se realizaron indagaciones parciales para encontrar interrelaciones entre las condiciones biológicas del niño, las conductas de la madre y la adopción de ciertas prácticas asistenciales en la atención del parto y el recién nacido. Se logró de esta forma no sólo valorar las asociaciones más relevantes entre los factores analizados y la enfermedad de Chagas, sino también una mejor comprensión acerca de la trama de relaciones existentes entre ellos^{23,24}.

Se reconoce la contribución analítica de estos trabajos teniendo en cuenta que todo proceso de investigación es una aproximación de la realidad desde un posicionamiento y una construcción teórica que orienta su delimitación. Es justamente a partir de dicho referencial que surge la crítica al alcance de las técnicas utilizadas. Éstas resultaron oportunas en la medida que posibilitaron valorar la magnitud y el sentido de las interacciones entre los factores indagados, resultados desconocidos en las experiencias previas de indagación. Ignorar la existencia de tales interrelaciones puede explicar el fracaso de un gran número de intervenciones, de allí la potencialidad de estos análisis para interpretar correctamente los hallazgos. Sin embargo, estas experiencias encuentran su límite en la homologación de las variables, al margen de su jerarquía.

El hecho es que estas técnicas estadísticas no tienen en cuenta un modelo teórico que supone una organización de variables según planos explicativos jerárquicos que permitan avanzar en la comprensión de los determinantes de los problemas de salud/enfermedad en la población, apelando simultáneamente a características individuales, variaciones entre individuos, a nivel singular, junto con otras del orden de lo particular, reflejando las variaciones entre grupos sociales de una misma sociedad, y también aquellas destinadas a recuperar las condiciones de vida a

nivel de la sociedad a la que pertenecen, o sea a nivel general, no está contemplada⁹.

Cuando no se reconoce esta arquitectura teórica compleja, y se recurre a técnicas como las referidas, lo que se cuestiona es la validez de su aplicación para la producción de resultados, por ignorar que se trata de unidades que están agrupadas, es decir, violan el supuesto subyacente de independencia requerido entre las respuestas individuales. Dentro de cada nivel las respuestas tenderán a estar correlacionadas, por ser expresión de condiciones, atributos, de individuos, de familias, grupos sociales, que comparten condiciones semejantes.

Aún cuando esto se comprendió hace tiempo por parte de los analistas, su resolución era difícil debido a la falta de herramientas potentes de análisis de datos, en especial, programas computacionales específicos. Como se ha mencionado, una opción a la que se recurre en estudios epidemiológicos de todo tipo es definir subdominios de estudio, uno para cada nivel de complejidad identificado en la construcción teórica del problema, de modo de habilitar el análisis al interior de cada uno de ellos. La indagación de la estructura de asociaciones como los estudios ejemplificados considerando los grupos en forma simultánea es sin duda un avance al arrojar luz sobre las interacciones existentes, como expresión de condiciones mediadoras de los determinantes de los procesos de salud/enfermedad; no obstante, estos abordajes no logran dar cuenta de los desajustes entre la construcción teórica del problema y las técnicas estadísticas aplicadas para la producción de resultados empíricos.

Los modelos de niveles múltiples

Entre los numerosos modelos que provee la Estadística, los modelos de niveles múltiples (MNM), cuya designación específica es *modelos de efectos aleatorios, de coeficientes aleatorios o modelos mixtos*, ofrecen la posibilidad de integrar diferentes subdominios de estudio para analizar-

los simultáneamente conservando la estructura jerárquica reconocida en la construcción teórica del objeto de estudio.

Originalmente desarrollados en el área de la educación, la sociología, la demografía y la criminología, sólo recibieron una creciente atención en el campo de la Epidemiología y de la Salud Pública en años recientes²⁵⁻²⁷ de la mano de una reemergencia del interés por los estudios socio epidemiológicos, por conceptualizar y medir cómo los contextos sociales, los entornos residenciales afectan los resultados y los riesgos de salud individuales. Esta preocupación conlleva una tarea no siempre sencilla al definir las variables contextuales, particularmente cuando no pueden ser derivadas a partir de variables análogas medidas a nivel individual, como por ejemplo el ingreso medio familiar, sino que requieren constructos específicos.

El avance computacional y la posibilidad de desarrollar un *software* más amigable han contribuido a la expansión de su utilización en las investigaciones socio epidemiológicas como se comprueba en las diversas aplicaciones publicadas. Sin embargo no siempre han sido acompañadas de fundamentos teóricos y metodológicos que guíen su implementación²⁵.

La diferenciación en niveles jerárquicos también ha sido objeto de preocupación en otras aplicaciones estadísticas, por ejemplo, en la teoría del muestreo, en la selección de muestras aleatorias en estudios poblacionales, de carácter extensivo, mediante un diseño muestral multietápico, donde está presente el agrupamiento de unidades de selección de distinto nivel. En estas aplicaciones el objetivo no es otro que buscar la reducción de costos operativos.

De igual modo, se reconoce la existencia de una estructura jerárquica en los estudios longitudinales. Es el caso de contar con varias mediciones de una característica bajo estudio, para un mismo sujeto en diferentes puntos del tiempo. En este contexto de datos longitudinales, donde se espera encontrar mucha más variación en las respuestas de diferentes individuos, que

entre las provenientes del mismo individuo en diferentes momentos, se desarrollaron procedimientos estadísticos apropiados, los cuales han sido puestos a prueba incluso en problemáticas regionales^{28,29}. Estas técnicas, denominadas modelos marginales y modelos de transición, si bien reconocen la correlación entre las respuestas, no consideran la estructura jerárquica de los datos de manera explícita en la formulación del modelo estadístico^{30, 31}.

En los modelos clásicos el componente sistemático permite medir el efecto de las variables explicativas sobre la respuesta y el componente aleatorio contiene la variación de la respuesta no explicada por las variables o factores incluidos en el modelo; se supone que ese componente es independiente de una observación a otra. Cuando esta pretensión de independencia no puede ser alcanzada, en razón del agrupamiento entre las unidades de un mismo nivel analítico, es viable la aplicación de los MNM. Estos modelos, que pueden utilizarse en un amplio rango de situaciones donde se diferencian unidades de nivel menor anidadas dentro de unidades de nivel mayor, valoran la influencia de las variables teniendo en cuenta el nivel de pertenencia de las mismas y la asociación potencialmente existente entre las unidades de un mismo nivel, es decir la correlación intraclase provocada por la proximidad de respuestas de individuos que comparten condiciones semejantes. Con ello se evita sobredimensionar el efecto de las variables de nivel superior.

Cuando el problema comprende dos niveles, se diferencian dos tipos de covariables: las “entre-grupo” o de nivel macro, cuyos valores varían de grupo a grupo (siendo el mismo para todas las unidades de un mismo grupo), y las “intra-grupo” o de nivel micro (cuyos valores varían entre las unidades del mismo grupo). Estos modelos jerárquicos permiten examinar simultáneamente el efecto de variables de nivel macro y micro sobre respuestas a nivel micro, considerando la falta de independencia de las respuestas dentro de los

grupos. Esto es posible mediante la descomposición de la variabilidad total en una variación intra-grupo y otra variación entre-grupos.

La incorporación de los denominados efectos aleatorios en el MNM, los cuales al no ser fijos sino variables –estocásticas– con una distribución de probabilidad asociada, permite especificar indirectamente la forma de asociación más pertinente a las relaciones que se buscan. Existe abundante literatura específica que amplía estos conceptos³²⁻³⁵.

Puede agregarse que el uso de efectos aleatorios en modelos lineales para variables respuesta con distribución normal ha sido extensamente estudiado; su utilización en cambio para el caso de variables respuesta categóricas es de desarrollo reciente, con la formulación de una clase de modelos denominada *modelos lineales generalizados mixtos*. Estos modelos abarcan variables respuesta no normales (muy frecuentes en los estudios poblacionales de salud/enfermedad) y efectos fijos y aleatorios en el predictor, permitiendo la inclusión de covariables de distinto nivel^{15, 36}.

Si bien estos modelos logran los dos objetivos principales perseguidos por los modelos lineales generalizados mixtos, incorporar correlación y permitir inferencias más amplias, su naturaleza no lineal conlleva complicaciones que no se presentan en el caso de respuesta con distribución normal. En lo que se refiere a los procedimientos de estimación, se utilizan frecuentemente métodos de aproximación a la función de verosimilitud por medio de la integración numérica. La convergencia de ellos depende en gran medida de los valores iniciales asignados y surgen problemas cuando los datos están esparcidos y las componentes de variancia son grandes. En definitiva todavía resta mucho por

hacer al desarrollo estadístico para el ajuste de este tipo de modelos.

Un ejemplo en particular, para datos binarios, es el modelo denominado “modelo logístico-normal”, que contempla un enlace logit y un predictor lineal con efectos aleatorios distribuidos normalmente. Este resulta formalmente definido por las siguientes especificaciones:

- La esperanza condicional de la variable respuesta correspondiente al i-ésimo individuo del j-ésimo grupo, Y_{ij} , depende de efectos fijos y aleatorios a través del siguiente predictor lineal:

$$\log \left\{ \frac{P(Y_{ij} = 1 / b_j)}{P(Y_{ij} = 0 / b_j)} \right\} = \mathbf{X}_{ij}\beta + b_j.$$

- Condicional a un único efecto aleatorio b_j , las Y_{ij} son independientes y tienen distribución de probabilidad Bernoulli.
- El efecto aleatorio único b_j se asume tiene una distribución normal univariada.

Este modelo también puede expresarse por medio de dos ecuaciones* lo que permite visualizar cómo la variación individual dentro de cada grupo se explica por la ecuación de nivel individual y la variación entre los grupos se explica por una ecuación a nivel grupal. La ventaja al especificar un MNN en términos de una serie de modelos es mostrar cuáles son las covariables que están operando en cada nivel del modelo.

Este caso sencillo del modelo logístico-normal (de intercepto aleatorio) se hace más complejo, al considerarse aleatorios los coeficientes asociados a algunas variables del nivel individual (pendientes aleatorias); esto significa que el efecto de dichas variables sobre la respuesta puede diferir, según cuál sea la unidad de nivel superior. Asimismo es factible incorporar covariables del segundo nivel, las que pueden incluso interactuar con las del primer nivel. Ello conduce al modelo más general de dos niveles**. Valga la aclaración que la

* 1) nivel individuo: $\eta_{ij} = b_j + \mathbf{X}_{ij}\beta$; 2) nivel grupo: $b_j = \beta_0 + v_{0j}$.

** En la formulación de los modelos a dos niveles se considera Y_{ij} la respuesta para la i-ésima unidad del nivel 1 en el j-ésimo grupo del nivel 2 (continua, binaria o de conteo) y un vector (fila) \mathbf{X}_{ij} de covariables de dimensión $1 \times p$, que pueden estar definidas en cada uno de los dos niveles. Se contemplan los siguientes supuestos: 1) La distribución condicional de cada Y_{ij} dado un vector de efectos aleatorios \mathbf{b}_j pertenece a la familia de distribuciones exponenciales con $\text{var}(Y_{ij} / \mathbf{b}_j) = v[E(Y_{ij} / \mathbf{b}_j)]\phi$, donde $v(\cdot)$ es una función de variancia conocida que depende de la media condicional $E(Y_{ij} / \mathbf{b}_j)$ y ϕ es el parámetro de escala o parámetro de dispersión. Dados los efectos aleatorios \mathbf{b}_j , las Y_{ij} son independientes entre sí. 2) La media condicional de Y_{ij} depende de efectos fijos y aleatorios a través de un predictor lineal: $\eta_{ij} = \mathbf{X}_{ij}\beta + \mathbf{Z}_j\mathbf{b}_j$, $g[E(Y_{ij} / \mathbf{b}_j)] = \eta_{ij} = \mathbf{X}_{ij}\beta + \mathbf{Z}_j\mathbf{b}_j$ para alguna función enlace conocida $g(\cdot)$. 3) Los efectos aleatorios tienen una distribución de probabilidad, en general la normal multivariada con media cero y matriz de covariancia \mathbf{G} .

persistencia de variación significativa en interceptos o pendientes, después de la inclusión de variables del nivel grupal, sugiere que otros factores de nivel grupal –posiblemente responsables de esa variación– necesiten ser explorados.

Está claro que la decisión de incluir u omitir determinadas covariables e interacciones (entre variables del mismo o de diferente nivel) y de considerar determinados efectos como fijos o aleatorios debe basarse en la comprensión teórica que orienta la construcción del objeto de investigación.

Otra ventaja al considerar efectos aleatorios para representar los grupos definidos por una unidad de mayor nivel, es que se requiere estimar un único parámetro: la varianza de la distribución de los grupos, teniendo presente que la cantidad de grupos debe ser lo suficientemente grande como para obtener una buena estimación de dicha variancia. Por el contrario, si se representan los grupos mediante efectos fijos, la cantidad de parámetros a estimar es del orden de la cantidad de grupos, lo que puede conducir a un modelo demasiado parametrizado, con estimaciones ineficientes cuando el número de grupos es muy grande y el número de individuos por grupo, pequeño.

Sin embargo, esta ventaja del modelo de efectos aleatorios sobre el modelo de efectos fijos conlleva un costo adicional: una mayor dificultad y complejidad en el cálculo del tamaño muestral y la potencia, la que depende tanto de la cantidad de grupos como de unidades dentro de cada grupo.

En Rosario se llevó a cabo una experiencia a los fines de probar el ajuste de modelos logístico-normales, tomando como punto de referencia los partos acontecidos durante el año 2003 en una maternidad pública y planteando como variable respuesta el peso al nacer en el recién nacido³⁷. Se buscó organizar las variables según determinantes definidos teóricamente como de distinto nivel jerárquico: individual –que abarca antecedentes sociales, obstétricos y de cuidado materno- y grupal

–relacionado con la micro-área de residencia de la madre, el contexto socio espacial de pertenencia-. Se realizaron distintos ajustes de modelos logístico-normales, con efectos fijos para las variables explicativas de nivel individual y un efecto aleatorio asociado a una variable de nivel espacial en cada caso. La utilización de información de fuente secundaria constituyó un obstáculo, por cuanto el nivel de agregación disponible sólo posibilitó recurrir a indicadores proxy, aún reconociendo la potencial distorsión que podían incorporar a los fines de delimitar áreas según condiciones de vida próximas (distrito municipal, seccional policial, área de residencia - un indicador elaborado ad-hoc basado en un agrupamiento de seccionales policiales según la ubicación geográfica y el nivel de prevalencia de talla, baja o muy baja, en niños en edad escolar, según fuera revelado en un trabajo sobre Situación Nutricional de Escolares realizado en el ámbito de la Secretaría de Salud Pública de la Municipalidad de Rosario⁻³⁸).

A pesar de esta limitación, la exploración realizada en esa primera prueba, incluyendo el contexto de pertenencia de la madre como fuente adicional de variabilidad en el peso del recién nacido, posibilitó alcanzar resultados más confiables que los que se obtendrían con técnicas estadísticas convencionales (modelos ingenuos), en razón de tener en cuenta la correlación debida a la pertenencia de los individuos a grupos. En otras palabras, al considerar que las respuestas individuales tienden a estar correlacionadas por ser expresión de atributos de individuos próximos entre sí, que comparten condiciones de vida y salud semejantes, al estar insertos en un mismo espacio social, un mismo barrio.

Recuperar la noción de población como una estructura compleja es fundamental cuando se plantea como objetivo valorar desigualdades en el proceso salud/enfermedad/atención de los grupos poblacionales y analizar cómo los contextos sociales afectan los resultados y los riesgos de salud individuales. Se reconoce como

una tarea compleja pero necesaria, la reflexión sobre los modos de análisis e interpretación de observables en la investigación epidemiológica con la expectativa de proveer insumos útiles para la caracterización socio espacial, según condiciones de vida y salud, y la definición de intervenciones adecuadas a las realidades locales, en sus especificidades a nivel de las microáreas, con miras a lograr una mayor equidad en salud³⁹.

La posibilidad de profundizar el estudio en este sentido se estima que supera las experiencias previas, por cuanto estos modelos estadísticos, que incluyen la discriminación por niveles - con la inclusión de efectos aleatorios asociados a los agrupamientos definidos- aportan al logro de una mayor validez en los hallazgos y una mayor coherencia con la concepción teórica del problema.

Estudios de desigualdades e inequidades

La preocupación por proporcionar modelos estadísticos útiles a los estudios epidemiológicos data de tiempos históricos. La dificultad, también histórica, reside en el examen de los supuestos de dichos modelos y su coherencia con la comprensión de los problemas de salud/enfermedad a nivel de la población, de sus determinantes y condicionantes en función de la estratificación de la sociedad, según las condiciones de vida de sus habitantes, apelando a una definición compleja de la realidad.

Los MNM ofrecen una alternativa analítica para la investigación en epidemiología que posibilita restaurar la dimensión poblacional o de sociedad, frente al arsenal de estudios de enfermedades de poblaciones, reducidos a factores de riesgo individuales. Con estos modelos es posible contribuir a la descripción y explicación de los determinantes y condicionantes organizando las variables según niveles teóricos de distinta jerarquía: individual, particular y general, sin distorsionar esta construcción conceptual.

Se ha señalado la importancia de producir información epidemiológica para examinar condiciones locales y regionales e identificar las brechas de salud en el seno de la población, a los fines de sustentar las intervenciones que se necesitan para poder distribuir adecuadamente los servicios a los grupos más desprotegidos. Con lo cual se plantea el problema de la falta de equidad y la búsqueda de una solución; también la necesidad de potenciar la utilización de información a nivel de lo local, que permita expresar desigualdades⁴⁰.

La investigación epidemiológica, desde una perspectiva basada en un principio de equidad, puede contribuir a que se tengan en cuenta las necesidades de los grupos sociales y de los individuos y a partir de esa consideración, habilitar la elaboración de estrategias compensatorias⁴¹.

Cuando el propósito fundamental es la comprensión de las desigualdades e inequidades en salud, la diferenciación según niveles jerárquicos de explicación y de intervención en salud, es crucial. Si bien el nivel socioeconómico continúa siendo uno de los predictores más consistentes de bienestar, es necesario profundizar en el análisis de los procesos políticos, culturales e institucionales en el nivel local, además de los niveles de ingreso y pobreza, éstos son importantes contribuyentes a los resultados en salud. Específicamente, son las características de las microunidades sociales, las condiciones sociales y materiales de vida, la organización social y política de los espacios donde se realizan las intervenciones, las que aportan significativamente a la comprensión de los procesos sociales que contribuyen positiva y negativamente en la producción de salud.

No abundan los estudios de carácter extensivo que contemplen la definición de variables y su análisis simultáneo desde una concepción compleja de la realidad, recuperando diferentes niveles de abstracción. Un campo de promisorio desarrollo es el de la investigación sobre inequidades y desigualdades en salud, que adquirió mayor énfasis a partir de la publicación de *The*

Black Report and the Health Divide, por Townsend y Davidson en 1982, evidenciando las enormes disparidades en salud de la población británica en los años ochenta⁴² y que ha dado lugar a numerosos interrogantes y debates sobre definiciones, conceptos, así como de estrategias analíticas y de interpretación de los hallazgos. Se ha manifestado el interés por desarrollar estrategias analíticas para pequeñas áreas y de múltiples niveles jerárquicos, junto con las de geo referenciamiento y otras técnicas aptas para cuantificar y monitorear la disparidad de los grupos poblacionales según las distintas concentraciones de riqueza, incluidos análisis temporales, a los fines de valorar la extensión de las desigualdades entre los individuos en una población y contemplar al mismo tiempo las inequidades, que serían las desigualdades a través de los grupos sociales⁴³.

En este terreno la producción de referentes empíricos mediante los MNM, desde una construcción teórica basada en una idea de realidad organizada jerárquicamente en diferentes dimensiones, puede abrir un camino de integración de los dos abordajes que se individualizan en la evaluación de las inequidades en salud.

Medir la distribución del status de salud mediante los individuos en una población, o bien, describiendo las diferencias entre grupos sociales, de qué manera las mismas están moldeadas por los factores sociales, económicos y culturales, por la pertenencia a un determinado espacio, que también está social e históricamente determinado.

Se trata de un campo en construcción, tanto en sus dimensiones teóricas, metodológicas como técnicas, que requiere de no pocos esfuerzos para desarrollar formas de abordaje que contribuyan a profundizar en la comprensión de la articulación de la problemática individual y colectiva en el análisis de los procesos de salud/enfermedad de la población.

En el mismo sentido, es necesario abordar el problema de la producción de mediciones teniendo en cuenta requerimientos analíticos esenciales y con mayor énfasis aún, la utilización de la voluminosa información de rutina, relativa a distintos niveles de agregación geográfica, para valorar desigualdades y sus determinantes, de modo de potenciar el diseño de intervenciones orientadas a lograr una mayor equidad en la provisión y acceso a los servicios de salud.

Referencias

1. Dussault G. La epidemiología y la gestión de los servicios de salud. *Boletín epidemiológico* 1995; 16(2): 1-5.
2. Szwarcwald CL, de Castillo EA. Os caminhos da Estatística e suas incursões pela Epimiologia. *Cad Saúde Pública* 1992; 8(1): 5- 21.
3. Almeida Filho N. Uma breve história da Epidemiologia, in Roquayrol MZ, Almeida Filho N (org.). *Epidemiologia e Saúde*, 5ª ed. Rio de Janeiro: MEDSI; 1999, p. 1-14.
4. Almeida Filho N. *La ciencia tímida. Ensayos de deconstrucción de la Epidemiología* Buenos Aires: Lugar Editorial; 2000.
5. Rose G. Individuos enfermos y poblaciones enfermas. *Boletín Epidemiológico OPS* 1985; 6(3): 1 - 8.
6. Castellanos PL. Epidemiologia, saúde pública, situação de saúde e condições de vida. Considerações conceituais, in Barata R. (org.). *Condições de vida e situação de saúde Saúde e Movimento 4*. Rio de Janeiro: ABRASCO, 1997, p. 31-75.
7. Czeresnia D. O conceito de saúde e a diferença entre prevenção e promoção, in Czeresnia D, Machado de Freitas C (org.). *Promoção da saúde. Conceitos, reflexões, tendências*. Rio de Janeiro: Editora Fiocruz, 2003, p. 39-53.
8. Susser M, Susser E. Choosing a Future for Epidemiology: I. Eras and Paradigms. *Epidemiology: II. From Black Box to Chinese Boxes and Eco-Epidemiology. Am J Public Health* 1996; 86(5): 668-677.
9. Castellanos PL. Sobre el concepto de salud-enfermedad. Un punto de vista epidemiológico. *Cuadernos Médicos Sociales* 1987; 42: 15-23.
10. Susser M. *Conceptos y Estrategias en epidemiología. El pensamiento causal en las ciencias de la salud* Biblioteca de la salud. México: Fondo de Cultura Económica; 1991.
11. Almeida Filho N, Rouquayrol MZ. *Disenhos de pesquisa em epidemiologia*, in Roquayrol MZ, Almeida Filho N (org.). *Epidemiologia e Saúde*. 5ª ed. Rio de Janeiro: MEDSI; 1999. p. 149-170.

12. Vilaça Mendes E, Fontes Teixeira C, Cardoso Araújo E, Logiodice Cardoso MR. Distritos Sanitários: conceitos-chave, in Vilaça Mendes E (org.). *Distrito Sanitário. O processo social de mudança das práticas sanitárias do Sistema Único de Saúde*. São Paulo - Rio de Janeiro: HUCITEC-ABRASCO; 1993. p. 159-185.
13. Johnson RA, Wichern DW. *Applied Multivariate Statistical Analysis*. 4th. Edition. New York: Prentice Hall Inc.; 1998.
14. Neter J, Wasserman W, Kutner M. *Applied Linear Regression Models*. Homewood, IL: Richard D. Irwin; 1989.
15. Agresti A. *Categorical Data Analysis*. 2nd edition. New York: Wiley & Sons; 2002.
16. Luppi I. Condicionantes sociobiológicos de la salud del niño en el primer año de vida: una aproximación mediante modelos logit, in *Ciencias Sociales y Medicina*. Santiago de Chile: Edit. Universitaria; 1992. p. 30-37.
17. Goldberg M. Este obscuro objeto da epidemiologia, in *Epidemiologia. Teoría e objeto* 2a. ed. São Paulo – Rio de Janeiro: HUCITEC –ABRASCO; 1994. p. 87-136.
18. Collett D. *Modelling binary data*. 2nd edition. Florida: Chapman & Hall/CRC Press; 2002.
19. Christensen R. *Log-linear models and logistic regression*. 2nd edition. New York: Springer-Verlag; 1997.
20. Bloch C, Quinteros Z, Troncoso MC, Belmartino S, Torrado S. El proceso de salud-enfermedad en el primera año de vida (primera parte). *Cuadernos Médico Sociales* 1985; 32: 5-19.
21. Bloch C, Quinteros Z, Troncoso, MC, Belmartino S, Torrado S. El proceso de salud-enfermedad en el primer año de vida (segunda parte). *Cuadernos Médico Sociales* 1985; 33: 3-34.
22. Boggio G, Luppi, I, Torres de Quinteros Z. Condicionantes socio-biológicos de la salud del niño en el primer año de vida: una aproximación mediante modelos logit. Actas del XVII Coloquio Argentino de Estadística. Sociedad Argentina de Estadística. Santa Fe; 1989, p. 77-83.
23. Boggio G, Luppi I, Sánchez SM, Troncoso MC. Enfermedad de Chagas connatal en el Hospital Roque Saénz Peña de la ciudad de Rosario. *Cuadernos Médico Sociales* 1995; 72: 25-46.
24. Boggio G. Modelo de regresión logística aplicado a un estudio sobre enfermedad de Chagas. *Cad Saúde Pública* 1997; 13(3): 389-396.
25. O'Campo P. Invited Commentary: Advancing theory and methods for multilevel models of residential neighborhoods and health. *Am J Epidemiol* 2003; 157: 9-13.
26. Diez-Roux AV. Bringing context back into epidemiology: variables and fallacies in multilevel analysis. *Am J Public Health* 1998; 88(2): 216-222.
27. Diez-Roux AV. A glossary for multilevel analysis. *J Epidemiol Community* 2002; 56: 588-594.
28. Hachuel L, Boggio G, Mendez F. Modelos marginales para el estudio de la desocupación en Rosario 2001-2002. [CD-ROM] XXXI Coloquio Argentino de Estadística. Sociedad Argentina de Estadística. San Juan; 2003.
29. Hachuel L, Boggio G, Mendez F, Servy E. Dinámica de la desocupación en el Gran Rosario en el período 2001-2002. Uso de modelos de transición. *Octavas Jornadas de Investigaciones en la Facultad de Ciencias Económicas y Estadística*. U.N.R. Rosario; 2003. www.fcecon.unr.edu.ar, 07/2005.
30. Diggle PJ, Liang KY, Zeger SL. *Analysis of longitudinal data*. Oxford Statistical Science. Oxford: Oxford University Press; 1994.
31. Fitzmaurice G, Laird N, Ware J. *Applied longitudinal analysis*. New York: John Wiley & Sons Inc.; 2004.
32. McCulloch C, Searle S. *Generalized, linear and mixed models*. New York: John Wiley & Sons; 2002.
33. Goldstein H, Browne W, Rasbash J. Tutorial in Biostatistics. Multilevel modelling of medical data. *Statist Med* 2002; 21: 3291-3315.
34. Goldstein H. *Multilevel Statistical Models*, 3rd edition. London: Edward Arnold; 2003.
35. Demidenko, E. *Mixed Models. Theory and Applications*. New Jersey: John Wiley & Sons; 2004.
36. Agresti A, Booth J, Hobert JP, Caffo B. Random effects modeling of categorical response data. *Sociol Methodol* 2000; 30: 27-81.
37. Hachuel L, Boggio G, Wojdyla D. *Modelos logit mixtos: una aplicación en el área de la salud*. Novenas Jornadas de Investigaciones en la Facultad de Ciencias Económicas y Estadística U.N.R, Rosario, marzo 2004. www.fcecon.unr.edu.ar, 07/2005.
38. Aronna A, Balparda L, Bisio S, Gómez M, Liborio M, Moyano C, Salarano D. *Situación Nutricional de Escolares* Publicación de la Municipalidad de Rosario. Argentina; 2003.
39. Luppi I, Boggio G, Hachuel L. *Estudio de desigualdades en salud desde una perspectiva socioespacial mediante análisis de niveles múltiples*. Actas VI. Jornadas de Debate Interdisciplinario en Salud y Población. Instituto Gino Germani. Facultad de Ciencias Sociales. U.B.A. Buenos Aires, julio 2005.
40. Castillo SC. Mediciones de las desigualdades en salud. *Rev Panam Salud Pública* 2002; 12(6): 371-372.
41. Barata RB. Ineqüidade e saúde: a determinação social do processo saúde-doença. *Revista USP* 2001; 51: 138-145.

42. Szwarcwald CL, Bastos FI, Tavares CLA. Medidas de desigualdad en salud: la discusión de algunos aspectos metodológicos con una aplicación para la mortalidad neonatal en el Municipio de Río de Janeiro, 2000. *Cad Saúde Pública* 2002; 18(4): 959-970.

43. Barbosa JS, Barros MBA. Epidemiologia e desigualdade: notas sobre a teoria e a historia. *Rev Panam Salud Pública* 2002; 2(6): 375-382.

Recebido em: 09/08/05

Versão reformulada reapresentada em: 17/11/05

Aprovado em: 14/12/05