



## **Población y Salud en Mesoamérica**

Revista electrónica publicada por el  
Centro Centroamericano de Población  
Universidad de Costa Rica, 2060 San José, Costa Rica  
<http://ccp.ucr.ac.cr>

---

## **Población y Salud en Mesoamérica**

### **Revista electrónica**

Volumen 4, número 2, informe técnico 1

Enero - junio, 2007

Publicado 1 de enero, 2007

<http://ccp.ucr.ac.cr/revista/>

## **La conveniencia del análisis multinivel para la investigación en salud: Una aplicación para Costa Rica**

*Amada Aparicio Llanos*

*Melvin Morera Salas*

© 2007 Centro Centroamericano de Población

## La conveniencia del análisis multinivel para la investigación en salud: una aplicación para Costa Rica<sup>1</sup>

*Amada Aparicio Llanos<sup>2</sup>*

*Melvin Morera Salas<sup>3</sup>*

### Resumen

Este trabajo analiza la aplicabilidad de las técnicas de análisis multinivel en la investigación en salud en el ámbito costarricense. La primera sección describe los aspectos más relevantes del análisis multinivel, las ventajas que tiene con respecto a los métodos clásicos y su aplicabilidad en investigaciones en salud. En la segunda parte se presenta un ejemplo de aplicación de regresión logística multinivel en los pacientes diabéticos de atención primaria del sistema público de salud costarricense. . El ejemplo hace énfasis en la sintaxis para su replica en uno de los programas estadísticos especializados en el análisis multinivel, el MLwiN 2.0, y la correspondiente interpretación de los resultados.

### Abstract

This work analyzes the applicability of of the multilevel analysis technique in health sciences in Costa Rica. The first section describes to the main aspects of the multilevel analysis, the advantages with respect to the classic methods and their applicability in health investigations. In the second part we present an example of application of multilevel logistic regression in the diabetic patients of the primary care. The example makes emphasis in the syntax for the application in the specialized statistical programs in the multilevel analysis, MLwiN 2,0, and the corresponding interpretation of the results.

---

<sup>1</sup> Proyecto de Investigación en Farmacoeconomía del Centro Centroamericano de Población. Este proyecto contó con el apoyo de una beca parcial sin condiciones del PPPI de la Merck Company Foundation, brazo filantrópico de Merck & Co. Inc., Whitehouse Station, New Jersey, USA.

<sup>2</sup> Médico Cirujano, Master en Salud Pública. Universidad de Costa Rica. [aaaparici@ccss.sa.cr](mailto:aaaparici@ccss.sa.cr)

<sup>3</sup> Economista de la Salud, Universidad de Costa Rica. [mmoreras2@ccss.sa.cr](mailto:mmoreras2@ccss.sa.cr)

## **Introducción**

Los modelos multinivel, son una respuesta a la necesidad de analizar la relación entre los individuos y el medio en que se desenvuelven. Con esta metodología se puede separar el papel de cada uno de los componentes de una compleja estructura interactiva, con el fin de mejorar el conocimiento de la realidad socio-sanitaria, y por ende intervenir de forma más eficiente.

Aunque, estos modelos se desarrollaron en los años ochenta del siglo pasado, no fue sino hasta que los programas computacionales fueron capaces de resolverlos, que su uso práctico se ha expandido a muchas áreas de las ciencias sociales, incluyendo su aplicabilidad en la investigación en salud (Catalán et al, 2003).

Los modelos multinivel, también denominados modelos jerárquicos, permiten solventar la limitación en el uso de modelos tradicionales de regresión que invalidan la hipótesis de independencia cuando se presenta mayor homogeneidad entre individuos de un mismo grupo respecto a individuos de distintos grupos.

Hay que reconocer, que una gran parte de las investigaciones en salud, proporcionan bases de datos que se pueden agrupar en estructuras jerárquicas y desagregar por niveles. Así, los individuos estudiados (unidades muestrales de nivel 1), pueden pertenecer a grupos o unidades mayores (unidades maestras de nivel 2, 3 ó 4), por ejemplo: el médico que los atiende, el área de salud donde reciben la atención, la zona geográfica donde viven, el hospital que los asiste; entre otros.

El presente estudio tiene como objetivo presentar a los investigadores, tanto en el campo de la salud pública como en el de los servicios sanitarios, este incipiente modelo de análisis y facilitar su aplicación mediante la aplicación de un ejemplo a datos del entorno de atención primaria de Costa Rica en el programa especializado MLwiN 2.0.

## **Conceptos básicos de la metodología multinivel**

Este primer apartado está basado en Goldstein (1999), González (2004a) y Rasbash et al (2004). En él se describen los aspectos más relevantes del análisis multinivel, las ventajas que tiene con respecto a los métodos clásicos y su aplicabilidad en investigaciones en salud.

### **Características del objeto de estudio del modelo multinivel**

Los individuos pertenecientes a un mismo contexto tenderán a ser más similares en su comportamiento entre sí, que respecto a los que pertenezcan a distintos contextos. Así, las personas que pertenezcan a una misma área de salud podrían tener hábitos de vida más parecidos entre sí que respecto a personas de otra área, por el hecho de tener distintas culturas y políticas de promoción de salud en las respectivas zonas. Esta similitud entre los individuos dentro de los grupos establece una estructura de correlación

intracontextual que impide el cumplimiento de la hipótesis de independencia sobre la que están basados los modelos de regresión tradicionales e invalida, por tanto, sus métodos de estimación, lo que se traduce en estimaciones incorrectas de los errores estándar (Davis et al, 2001).

En el uso del modelo multinivel, las unidades de análisis deben estar bien definidas, esto requiere de claridad en el concepto de estructura jerárquica; en el campo de la salud se reconoce la existencia de estas estructuras en los datos, como en la evaluación de indicadores de salud que se distribuyen por áreas, en la satisfacción de los usuarios en relación con los especialistas que los asisten y el centro de salud donde se atienden, en pruebas clínicas donde se aplican diferentes tratamientos, y en la evolución de la salud de un paciente en el tiempo.

Las jerarquías a su vez están expresadas en niveles. Para tal efecto, se aclara que se deben tener datos para un mínimo de 2 niveles, donde el nivel 1 debe ser la mínima unidad estudiada y los siguientes niveles (2, 3, 4...), deben contener al anterior. En educación, por ejemplo, una estructura jerárquica en 3 niveles ésta dada por los alumnos (unidad de nivel 1) agrupados dentro de cursos (unidad de nivel 2), y esto a su vez dentro de colegios (unidad de nivel 3). En el campo de las ciencias sociales, los votantes agrupados dentro de distritos o áreas de votación conforman una estructura jerárquica de 2 niveles. En el área de la salud un ejemplo es los pacientes en el nivel 1, dentro de hospitales (nivel 2) y éstos a su vez agrupados dentro de áreas geográficas (nivel 3).

### **Problemas que resuelve el modelo multinivel**

Los modelos multinivel resuelven dos problemas que se presentan cuando se utilizan los modelos de un único nivel a datos que son jerárquicos, (Sánchez y Ocaña, 1999).

- ◆ Problemas estadísticos de correlación entre los individuos en la estimación de los mínimos cuadrados ordinarios ineficiente y significaciones espurias.
- ◆ Problemas conceptuales ya que se emplea el nivel equivocado (analizar los datos a un nivel y extraer conclusiones a otro).
  - Falacia ecológica (interpretar datos agregados a nivel individual).
  - Falacia atomística (interpretación agregada a partir de datos individuales, agregar datos de distintas subpoblaciones como si fuera una).

Además resuelven otros problemas tales como: determinar el efecto directo de las variables explicativas individuales y de grupo, determinar si las variables explicativas del grupo “moderan” las relaciones a nivel individual (interacciones entre niveles) y determinar qué porcentaje de la variabilidad de la variable explicada o dependiente, una vez controlada por las variables explicativas, es imputable al individuo y que porcentaje es imputable al grupo.

### **Características de los datos y objetivos de los modelos multinivel**

Las características que deben tener los datos para aplicar el modelo multinivel son:

- ◆ Los datos de la población estudiada deben tener estructura jerárquica, anidada.

- ◆ Los datos pueden provenir de una muestra multiéptica de la misma estructura de datos jerárquica.
- ◆ Los individuos pertenecen a grupos.
- ◆ Los individuos de un mismo grupo comparten ciertas características, experiencias o influencias.
- ◆ Las variables están medidas en los diferentes niveles.

Entre los principales objetivos de los modelos multinivel destacan:

- ◆ La capacidad de combinar y analizar información de diferentes niveles.
- ◆ Analizar las relaciones entre niveles “micro” y “macro”.
- ◆ Estimar la magnitud de las varianzas que operan a los diferentes niveles y cómo se relacionan con las variables explicativas.

### **Ventajas y desventaja del modelo multinivel**

Ventajas:

- ◆ Dan una versión más realista.
- ◆ Modelan cada nivel de jerarquía.
- ◆ No requieren la hipótesis de independencia entre las medidas de la variable resultado.
- ◆ Dan estimaciones más precisas.

Desventajas:

- ◆ Presentan mayor complejidad tanto en el marco teórico como del modelo propuesto para analizar los datos.
- ◆ Mayor dificultad en la comunicación de los resultados.

### **Aplicaciones del análisis multinivel al campo de la salud**

En la década de los ochenta, se publicaron los primeros trabajos donde se proponen y desarrollan las técnicas de multinivel como una metodología estadística para el análisis de datos con estructura jerárquica, técnica que dio respuesta a la problemática en que las rectas de regresión difieren tanto en el intercepto como en la pendiente. Sus primeras aplicaciones se desarrollan en el campo de las ciencias sociales, específicamente en el área de la educación, (Davis et al, 2001 y Tilling et al, 2001).

Una revisión bibliográfica realizada por Catalán et al (2003) en Medline entre 1995 y 2001, encontró 222 artículos con aplicación de modelos multinivel al campo de la salud. En el cuadro 1 se presentan los temas con mayores aplicaciones.

Los modelos multinivel, desarrollados para variables de respuesta de tipo continua y discreta, también se aplican en datos de medidas repetidas, en donde un individuo es medido en más de una ocasión o en datos longitudinales (Conesa, et al, 2004). Otra aplicación son los estudio de supervivencia y los meta análisis (Hox y de Leeuw, 2003).

### **Categorías y niveles de los modelos multinivel**

En la revisión bibliográfica Catalán et al (2003), identificaron 12 categorías referente al tipo de modelo multinivel aplicado a los datos en la investigación o a la explicación de la metodología multinivel en el campo de la salud (cuadro 2).

De la revisión anterior, la mayoría de estudios publicados utilizan dos niveles de análisis (65%), un 9% usan tres niveles, un 3% cuatro niveles y el resto no especifican los niveles (Catalán et al, 2003).

### **Software disponibles**

En la actualidad existen 19 diferentes software para estimar modelos multinivel, en el cuadro 3 se presenta una breve descripción de los mismos.

### **Aplicación Práctica: Determinantes del adecuado abordaje del paciente diabético: aplicación del análisis multinivel para Costa Rica.**

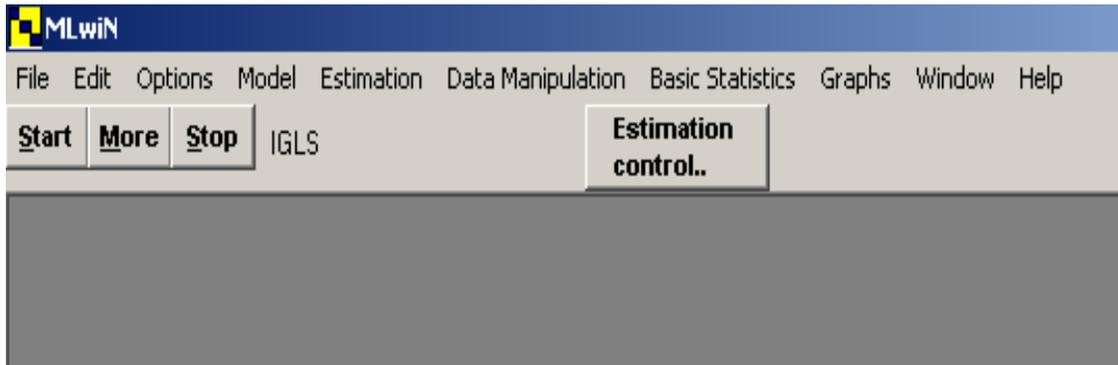
En este apartado, se especifica la metodología multinivel para una variable dependiente categórica (binaria). Los datos fueron tomados de la investigación realizada por Aparicio y Morera (2006), donde la variable dependiente se refiere al adecuado control (medido como un nivel de HbA1c menor o igual a 7%) de una muestra de pacientes diabéticos que fueron atendidos desde los servicios de las diferentes Áreas de Salud (AS) de la Caja Costarricense del Seguro Social (CCSS).

El objetivo de dicha investigación fue identificar los factores asociados al adecuado control del diabético y examinar el grado de variación entre áreas de salud. Los datos poseen una estructura jerárquica de dos niveles con 3.595 pacientes diabéticos (nivel 1) atendidos en 82 AS (nivel 2). El cuadro 4 presenta las variables utilizadas.

### **Cómo introducir los datos en el programa MLwiN 2.0.**

El programa utilizado para ejemplificar el análisis multinivel es el MLwiN 2.0. Este paquete es específico para modelos multinivel y presenta las características básicas de los paquetes de computación en ambiente Windows, al cual se puede acceder sin costo en el sitio Web [www.mlwin.com](http://www.mlwin.com) .

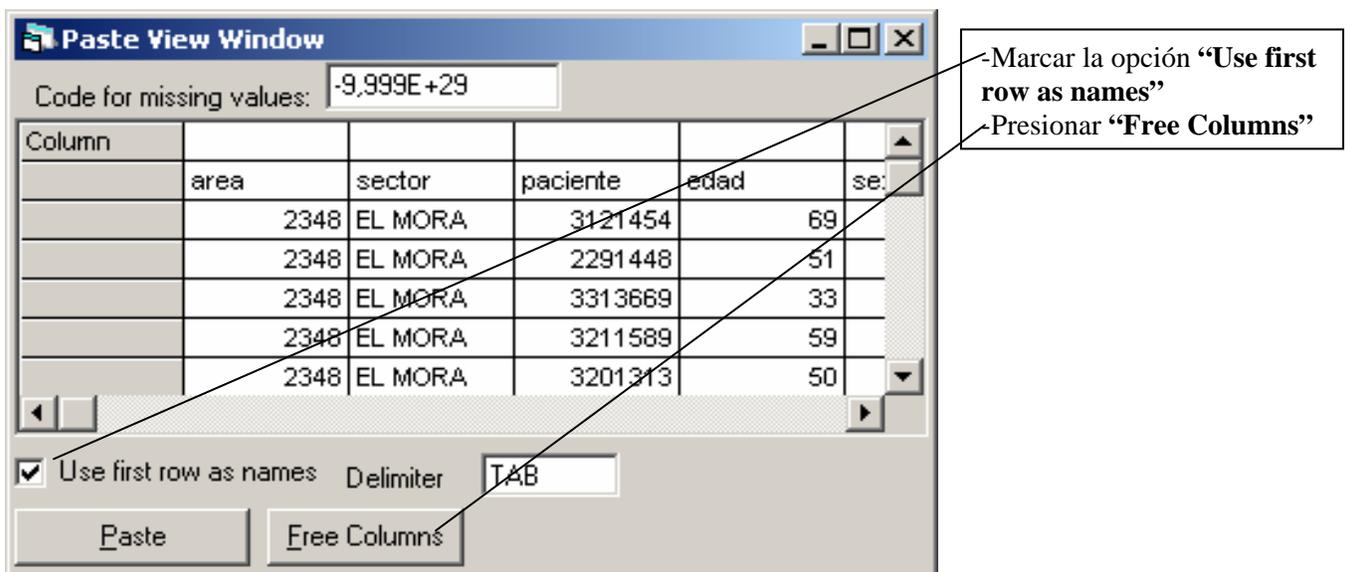
**Figura 1. Ventana de entrada en el programa con el menú de opciones.**



Para empezar a trabajar con el modelo, el primer paso es tener un archivo con extensión ws (archivo de lectura de MLwiN). El proceso más sencillo es trabajar la base de datos en una hoja electrónica tipo Excel y luego trasladar los datos al ambiente de MLwiN<sup>4</sup>:

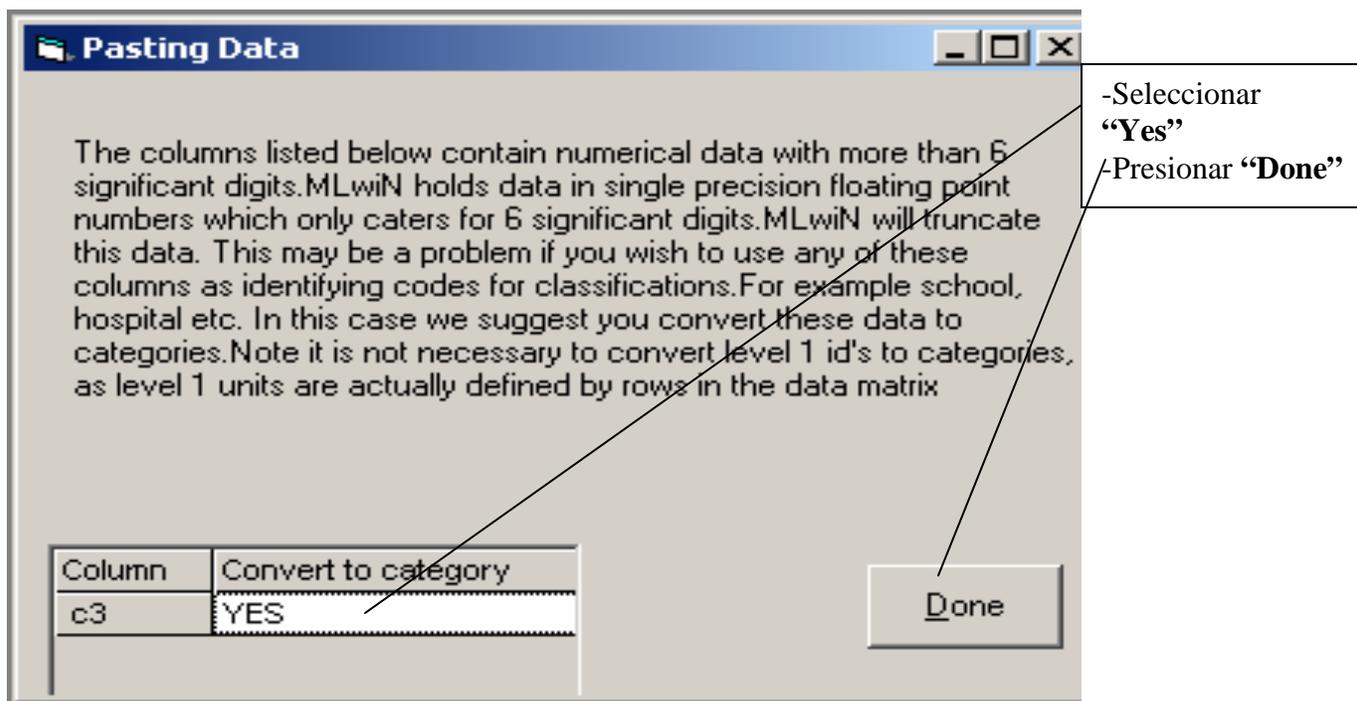
- ◆ Marcar el rango con los datos en la hoja de Excel (incluyendo el nombre de las variables) y darle la opción de copiar.
- ◆ Abrir el programa MLwiN y en el menú “**Edit**” elegir “**Paste**”.

**Figura 2. Subcuadro de dialogo del menú.**



Si la variable tiene más de seis caracteres le aparecerá una ventana de dialogo para que usted recodifique. Si la variable es continua debe reducir los caracteres en la hoja electrónica que usted está utilizando como base (Excel por ejemplo) y si es una variable nominal ejemplo “paciente” se puede convertir la variable en categórica, seleccionando “**Yes**” en la siguiente ventana.

<sup>4</sup> Para un mayor detalle del proceso de traslado de datos ver Rasbash et al (2004)

**Figura 3. Ventana de dialogo para recodificar variable.**

Para guardar los datos se debe utilizar la opción "Save Worksheet" del menú "File". Luego especificar el nombre (ejemplo, diabetes), MLwiN guarda el archivo con la extensión ws.

### Estimación de modelos de regresión logística multinivel en MLwiN

#### *Modelo vacío*

El modelo<sup>5</sup> más simple de descomposición de la varianza es el modelo sin variables explicativas, que admite una descomposición de la varianza en variación entre AS y variación entre diabéticos dentro de la misma área.

Si llamamos  $y_{ij}$  al valor de la variable dependiente para el diabético  $i$  que pertenece al AS  $j$  y a  $\pi_{ij}$  la probabilidad de que un diabético  $i$  que pertenece al AS  $j$  esté bien compensado, la formulación del modelo sería:

<sup>5</sup> Esta sección y las dos siguientes están basadas en Goldstein (1999), González (2004) y Rasbash et al (2004)

$$\pi_{ij} = P(y_{ij} = 1)$$

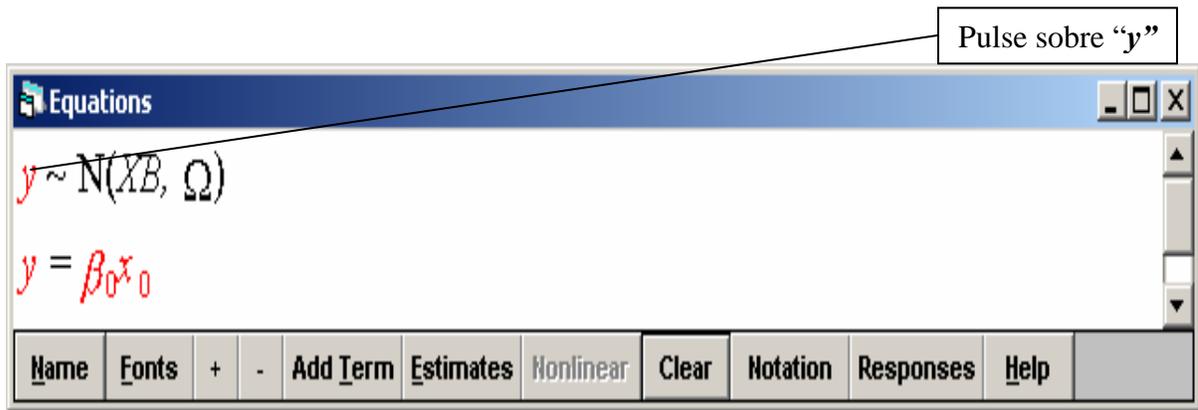
$$\log \text{it}(\pi_{ij}) = \log\left(\frac{\pi_{ij}}{1 - \pi_{ij}}\right) = \beta_{0j}$$

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

En la especificación anterior, se observa que el intercepto tiene dos términos:  $\beta_0$  y un componente para un área específica, el efecto aleatorio  $\mu_{0j}$ . Se asume la hipótesis de que  $\mu_{0j}$  se distribuye normalmente con media cero y varianza  $\sigma_{\mu_0}^2$ .

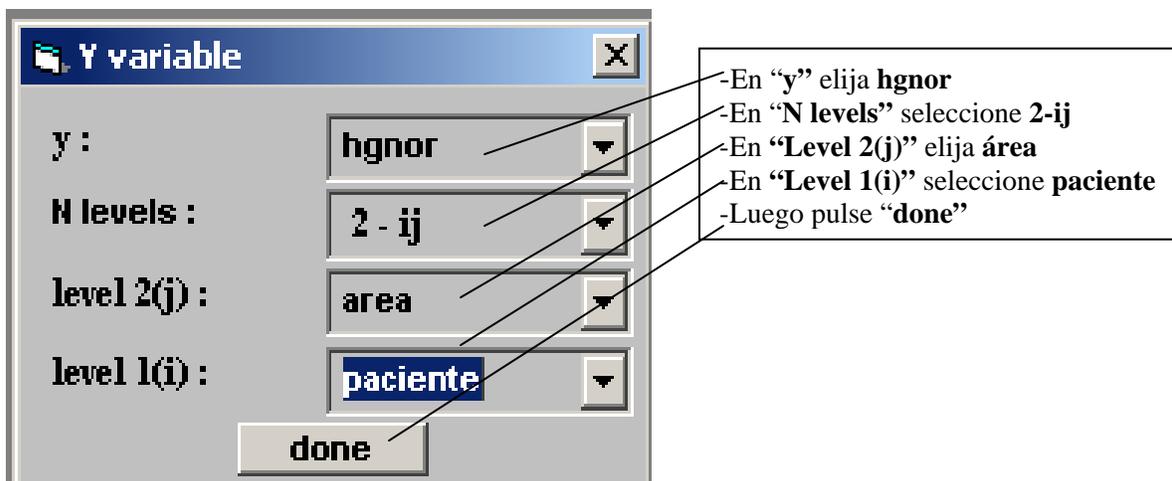
En el menú “**Model**”, seleccione “**Equations**”, para estimar el modelo vacío.

**Figura 4. Ventana de dialogo para determinar la ecuación del modelo.**



Una vez que a pulsado sobre “y” aparecerá la ventana denominada “**Y variable**”.

**Figura 5. Subcuadro de dialogo “Y variable”.**



En la ventana “Equations Pulse sobre “N”, y aparecerá la ventana “Response type”.

Figura 6. Subcuadro de dialogo Tipo de respuesta

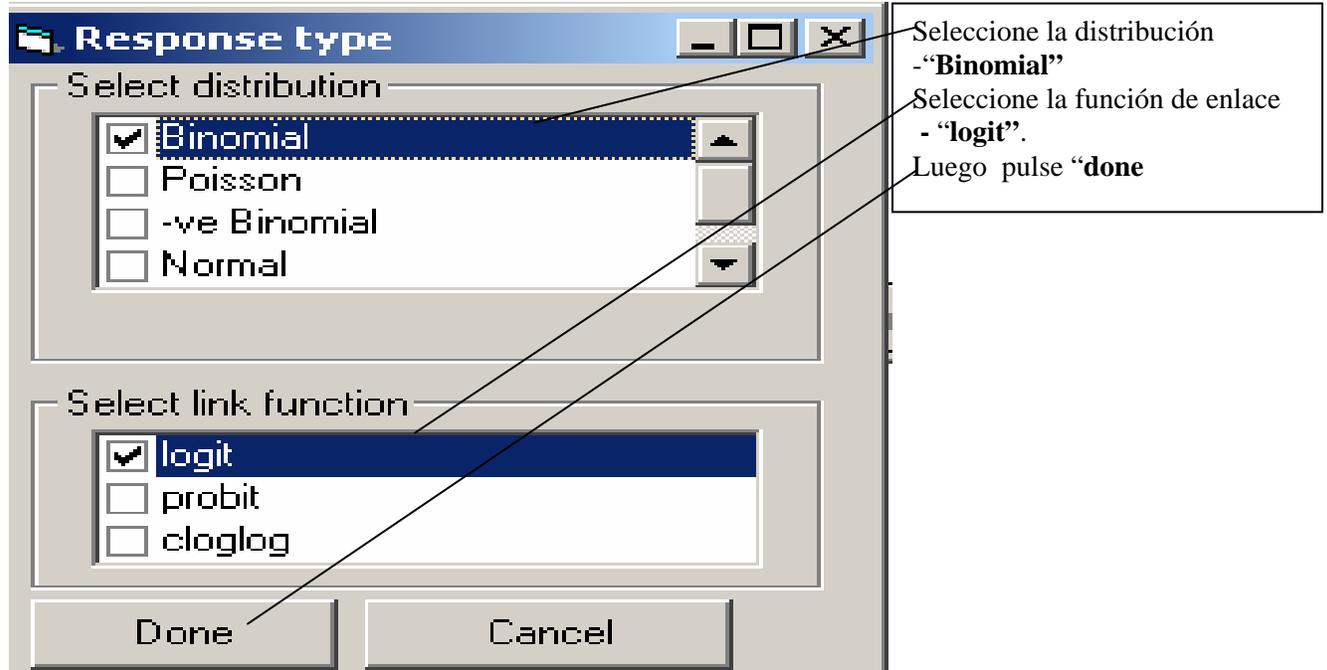


Figura 7. Ventana “Equations” con la ecuación del modelo completa.



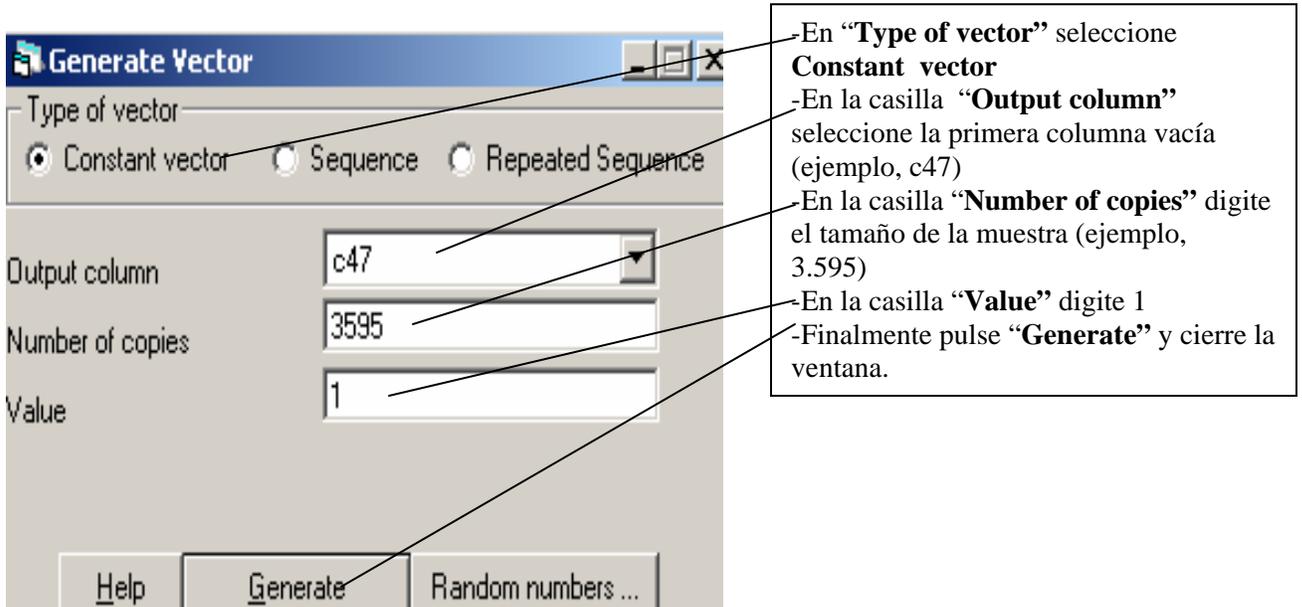
La primera línea de esta ventana indica que la variable sigue una distribución con parámetros  $n_{ij}$  y  $\pi_{ij}$ . El primer parámetro es conocido como el denominador y en el caso de datos binarios,  $n_{ij}$  es igual a 1 para todas las unidades.

Ahora usted debe crear dos variables; la primera que denominaremos *cons* será la constante en el modelo y a la segunda variable  $n_{ij}$  que llamaremos *denom*. Las dos variables son vectores con valor de uno ya que son constantes del modelo.

El procedimiento consiste en crear un vector con valor de 1, con tamaño igual al número de datos que posee la muestra (3.595).

En el menú “**Data Manipulation**”, seleccione “**Generate vector**”.

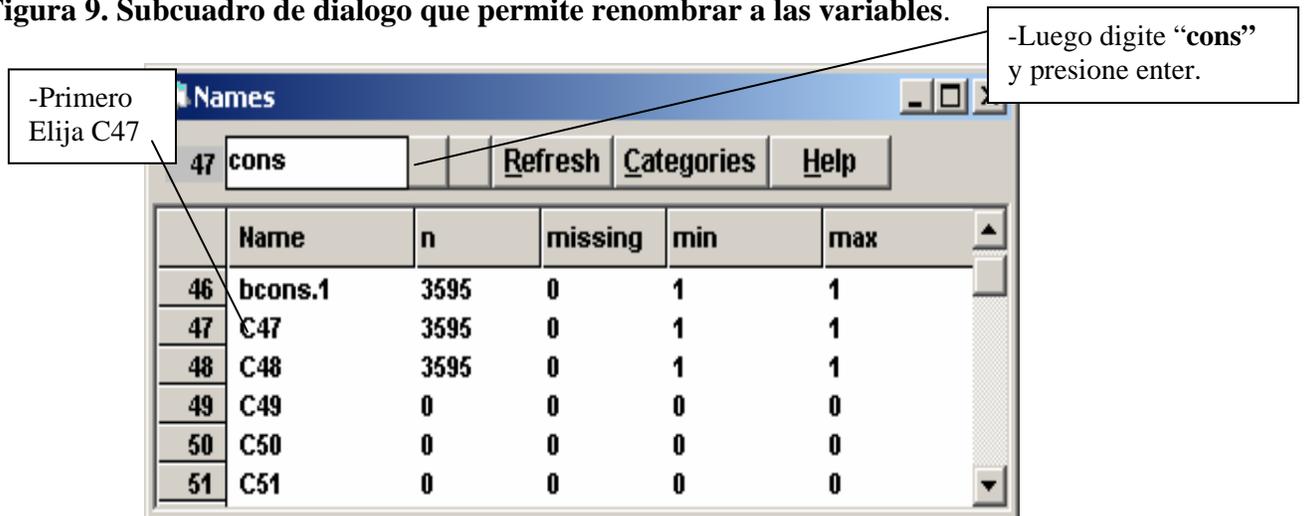
**Figura 8.** Subcuadro de dialogo para generar un vector.



Repita el paso anterior y elija la columna c48, luego de un nombre nuevo a las variables creadas.

Para dar un nombre nuevo a las variables creadas, vaya al menú y elija “**Data Manipulation**”, seleccione “**Names**” y se abrirá una ventana del mismo nombre. Primero debe elegir la columna y dar “click”, entonces se habilitará la casilla a la que le antecede el número de columna elegida, anote en esta el nuevo nombre y dar click en “enter”, de esta manera la variable quedara renombrada.

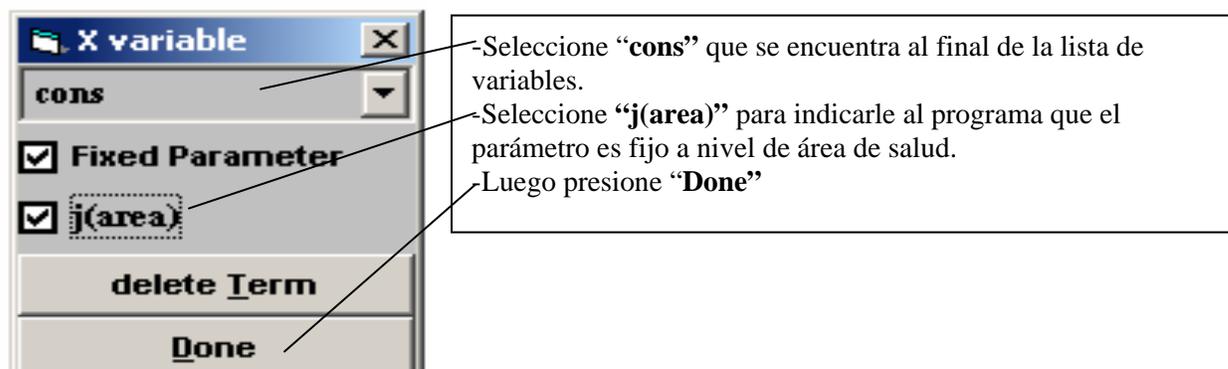
**Figura 9.** Subcuadro de dialogo que permite renombrar a las variables.



Repita el paso anterior para nombrar C48 como “denom” y así sucesivamente según sea el caso.

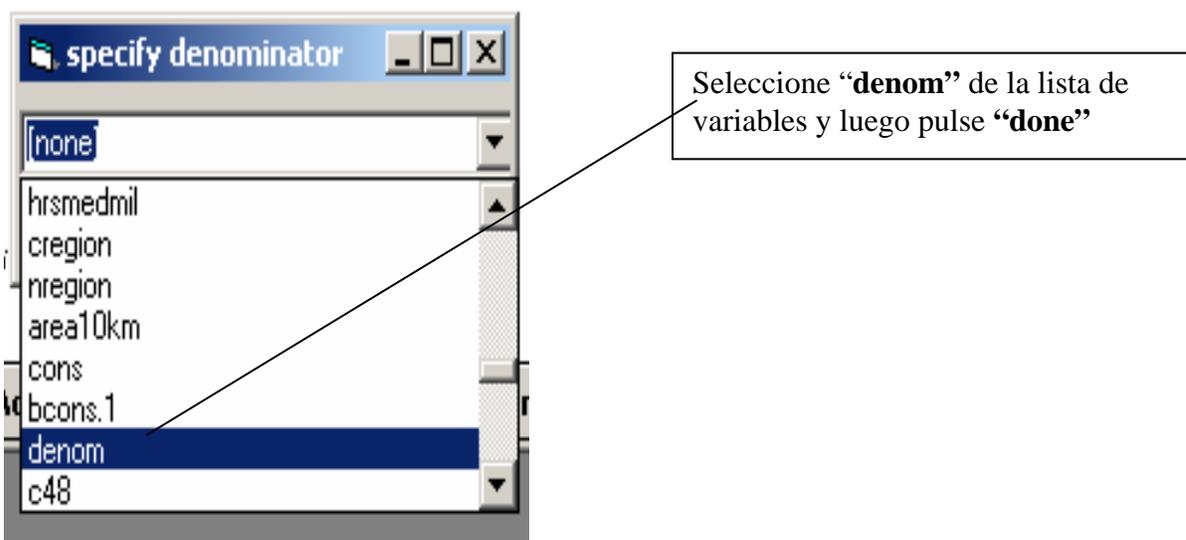
Para especificar las constantes del modelo vaya a la ventana “Equations” haga click sobre “ $B_0 x_0$ ” y de inmediato aparece un subcuadro de dialogo.

**Figura 10. Subcuadro de dialogo para especificar la primera constante.**



Para especificar la segunda constante, vaya a la ventana “Equations” precione *ni* y de inmediato aparece un subcuadro de dialogo.

**Figura 11. Subcuadro de dialogo para especificar la segunda constante.**



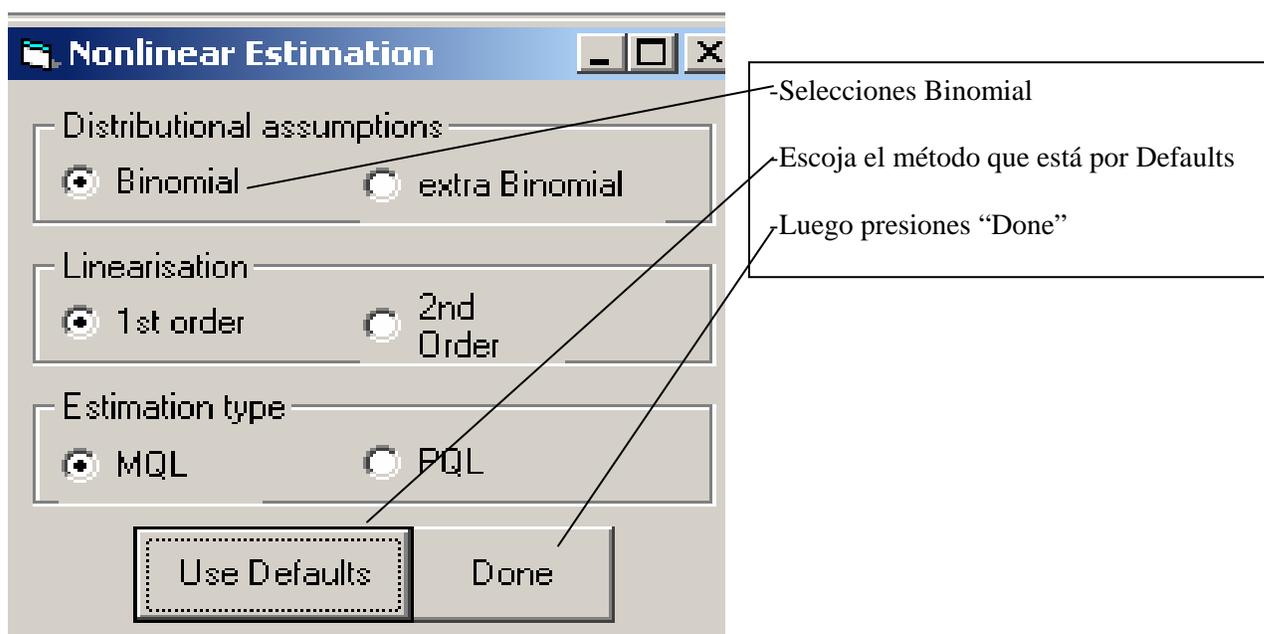
El paso siguiente es analizar el proceso de estimación que se usará. Para este ejemplo, en ambos modelos (vacío y de intercepto aleatorio) se utilizó el método de estimación Penalised Quasilikelihood (PQL), dado que se recomienda para eliminar el sesgo que se presenta cuando el número de unidades del primer nivel agrupadas en cada unidad de

segundo nivel es reducido o bien la variancia de segundo nivel es grande<sup>6</sup> (Rasbash et al, 2004).

Rasbash et al (2004), recomiendan que primero se realice la estimación con el método que está dado en MLwiN y luego se aplique el método seleccionado por el investigador, de acuerdo a las características que presenten los datos analizados.

Para estimar el modelo primero pulse sobre “**Nonlinear**” en la ventana “**Equations**” y luego en “**Nonlinear Estimation**”.

**Figura 12. Subcuadro de dialogo de estimación del modelo.**

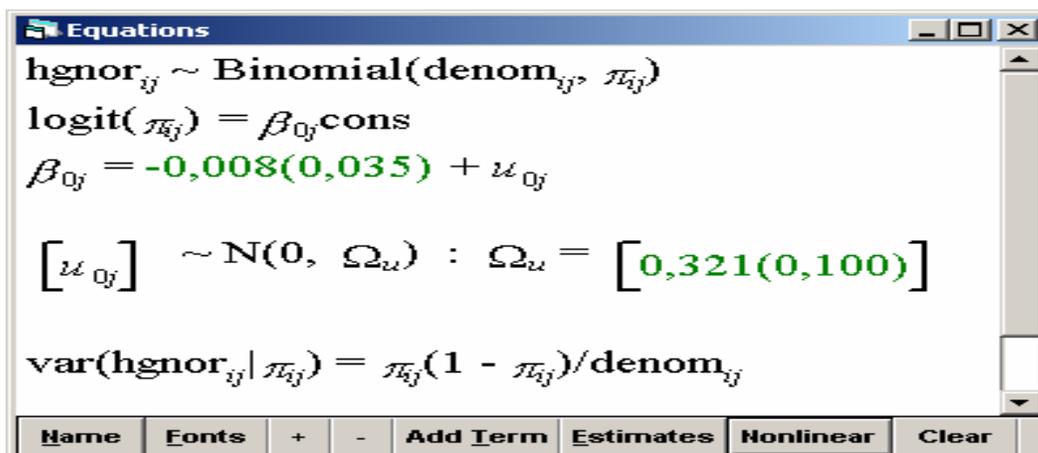


Para estimar el modelo debe seguir los siguientes pasos en el orden expuesto:

- ◆ Presionar “**Start**” en la parte superior izquierda de la ventana del menú principal.
- ◆ Pulse dos veces sobre “**Estimates**” en la ventana “**Equations**”.
- ◆ Seleccione el método
- ◆ Pulse “**Nonlinear**” en la ventana “**Equations**”
- ◆ Pulse las casillas “**2nd Order**”, “**PQL**”
- ◆ Pulse “**done**”.
- ◆ Pulse “**More**” en la parte superior izquierda de la ventana principal y aparecen los resultados de la estimación actualizados.

<sup>6</sup> Es último problema es el que se presenta en la investigación analizada.

Figura 13. Ventana de resultados de la estimación.



Este resultado permite construir el coeficiente de correlación intraclase, que mide la proporción de la variancia total que es explicada por las diferencias entre grupos. Algebraicamente se puede expresar de la siguiente forma:

$$\rho_1 = \frac{\sigma_{\mu 0}^2}{\sigma_{\mu 0}^2 + \sigma_e^2}$$

Según Rasbash et al (2004), en una distribución logística los residuos del nivel 1 poseen una variancia igual a  $\pi^2/3$  y por lo tanto, para un modelo logístico de dos niveles de intercepto aleatorio, con variancia del intercepto igual a  $\sigma_o^2$ , la correlación intraclase<sup>7</sup> es:

$$\rho_1 = \frac{\sigma_o^2}{\sigma_{\mu 0}^2 + \frac{\pi^2}{3}}$$

Con los resultados del modelo vacío que se estimó anteriormente, el valor de coeficiente es:

$$\rho_1 = \frac{0,322}{0,322 + 3,29} = 0,089$$

Estos resultados implican que aproximadamente el 9% de la variancia en el control del diabético, es atribuible a las diferencias entre AS.

<sup>7</sup> Goldstein (1999) propone un método alternativo para estimar la proporción de la variancia para datos discreto, la cual es una extensión directa de la definición dada para el caso lineal, aplicada a la variable binaria  $y_{ij}$  de resultado. Este proceso y la macro (vpc.txt) correspondiente para calcular la proporción de la variancia atribuible al nivel 2 se detallan en el Manual de Usuario de MLwiN 2.0 (Rasbash et al, 2004).

*Regresión logística multinivel de intercepto aleatorio*

Si al modelo vacío añadimos variables explicativas a nivel individual (nivel 1) y grupal (nivel 2) las probabilidades de ser un diabético compensado variarán entonces entre individuos dentro de las AS. El modelo resultante es:

$$\text{logit}(y_{ij}) = \beta_{00} + \sum_{k=1}^K \beta_{k0} X_{kij} + \sum_{h=1}^H \beta_{h0} Z_{hj} + \mu_{0j} + e_{ij}$$

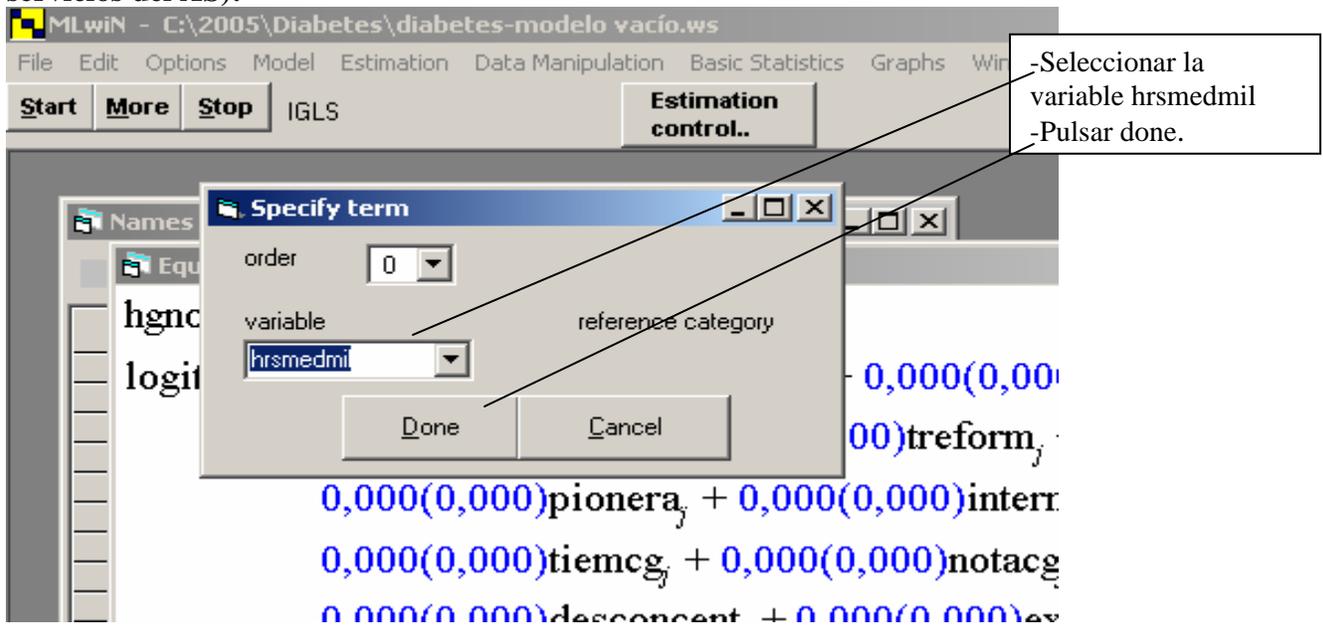
Donde  $y_{ij}$  es la adecuación del diabético, que toma valor de uno si el individuo  $i$  del AS  $j$  esta bien controlado (un nivel de HbA1c menor o igual al 7%) y cero en caso contrario, las variables explicativas del nivel 1 están denotadas como “x” y las variables explicativas del nivel 2 se denotan por “z”. El término de error que es compuesto, divide la parte no explicada de la variable dependiente en dos partes, una propia del AS ( $\mu$ ) y la otra propia del diabético ( $e$ ). Se supone que ambos componentes de la varianza, tienen media cero y varianza constante.

Para estimar este modelo en MLwiN, se agregaron al modelo vacío que se estimó en la sección anterior, tres variables explicativas del nivel 1 y diecisiete del nivel 2.

Pasos para agregar una a una las variables:

- ◆ Pulse sobre el botón “Add term” de la ventana “Equations”
- ◆ En la ventana “Specify term” seleccione de la lista, una por una, las siguientes variables: edad, sexo, imc treform, pionera, intermed, tiemcg, notacg, desconcent, externa, nivel1refor, nivel2, vespert, urgencias, ninos, adultomay, secund, univer, kmhosp y hrsmedmil.
- ◆ Pulse “done” para que se incluya la variable en la ecuación.

**Figura 14. Ventana demostrativa para la variable hrsmedmil (mide la oferta de servicios del AS):**



Una vez realizada la operación anterior, pulse el botón “**Estimates**” en la ventana “**Equations**”.

**Figura 15. Subcuadro de dialogo de estimación**

The screenshot shows a software window titled "Equations" with the following content:

$$hgnor_{ij} \sim \text{Binomial}(\text{denom}_{ij}, \pi_{ij})$$

$$\text{logit}(\pi_{ij}) = \beta_{0j} \text{cons} + 0,015(0,003)\text{edad}_{ij} + -0,183(0,077)\text{sexo}_{ij} + 0,002(0,007)\text{imc}_{ij} + 0,052(0,063)\text{treform}_j + -1,106(0,403)\text{pionera}_j + -0,654(0,233)\text{intermed}_j + -0,075(0,049)\text{tiemcg}_j + 1,123(0,654)\text{notacg}_j + 0,546(0,131)\text{desconcent}_j + -0,312(0,145)\text{externa}_j + 0,565(0,153)\text{nivellrefor}_j + 0,751(0,173)\text{nivel2}_j + -0,162(0,115)\text{vespert}_j + -0,209(0,111)\text{urgencias}_j + 0,054(4,847)\text{ninoss}_j + 16,580(6,555)\text{adultomay}_j + -3,314(1,092)\text{secund}_j + -1,961(1,005)\text{univer}_j + -0,012(0,003)\text{kmhosp}_j + -0,001(0,000)\text{hrsmedmil}_j$$

$$\beta_{0j} = -0,395(1,270) + u_{0j}$$

$$[u_{0j}] \sim N(0, \Omega_u) : \Omega_u = [0,357(0,107)]$$

$$\text{var}(hgnor_{ij} | \pi_{ij}) = \pi_{ij}(1 - \pi_{ij}) / \text{denom}_{ij}$$

At the bottom of the window is a toolbar with buttons: Name, Fonts, +, -, Add Term, **Estimates**, Nonlinear, Clear, Notation, Responses, Help.

En la ventana anterior se presentan los parámetros estimados para cada variable y para la variancia de los errores (con sus respectivos errores estándar entre paréntesis). Con estos datos se pueden calcular los valores “Z” e intervalos de confianza, para realizar pruebas de hipótesis individuales y los odds ratio para facilitar la interpretación de los resultados. A continuación se presenta un ejemplo para una variable del nivel 1 y otra del nivel 2 (cuadro 5).

Estos resultados muestran que por cada 100 hombres controlados (hemoglobina glicosilada menor al 7%), existen 83 mujeres controladas. Lo cual se puede aproximar a que las mujeres tienen un 16,7% menos de posibilidad de estar controladas que los hombres, dadas las demás variables del modelo. Este resultado es estadísticamente significativo al 5% de confianza.

Las AS que iniciaron tempranamente la reforma (entre 1995-1997) tienen un 67% menos de posibilidad de presentar pacientes controlados que las AS que iniciaron en un período intermedio (1998-2000) o tardío (después del 2000), manteniendo fijas las demás variables del modelo, y este resultado es estadísticamente significativo al 1% de confianza.

El coeficiente de correlación intraclase para el modelo regresión logística multinivel de intercepto aleatorio es el siguiente:

$$\rho_1 = \frac{0,357}{0,357 + 3,29} = 0,0979$$

Resultado que muestra que aproximadamente el 10% de la variancia en el control del diabético, es atribuible a las diferencias entre AS, después de controlar por la variables explicativas.

## **Discusión**

Dado que la mayor parte de investigaciones en salud proporciona bases de datos en las que se puede establecer estructuras jerárquicas, es indispensable emplear modelos multinivel para eliminar el sesgo que producen las estimaciones de los modelos tradicionales.

Los modelos multinivel contribuyen a aislar y medir los efectos contextuales y poder separar el papel de cada uno de los componentes para mejor conocimiento de la realidad socio-sanitaria y poder intervenir de forma más eficiente.

En esta investigación se demostró la utilidad de la metodología multinivel para analizar los determinantes del adecuado abordaje del paciente diabético en el ámbito costarricense. Entre los temas donde la aplicación de modelos multinivel tendría gran utilidad, están los estudios de utilización de medicamentos, dado que la metodología permite integrar la información de los pacientes que se encuentran en los expedientes clínicos, con información de los médicos prescriptores, AS y zonas de referencia a Hospitales de mayor complejidad. La principal limitación es que en Costa Rica solo 12 unidades (cobertura de aproximadamente el 20% de la población) poseen expediente electrónico con lo cual se facilitaría y agilizaría el proceso de recolección de la información.

Otros estudios que han aplicado modelos multinivel y que aprovechan bases de datos secundarias como la base de egresos de la CCSS, están relacionados con el análisis de pequeñas áreas y estudian los determinantes de la variabilidad en procedimientos quirúrgicos o diagnósticos.

La investigación que sirvió de base para este estudio presenta la limitante de no incorporar variables a nivel 1 (paciente diabético) que se ha demostrado son explicativas del control del diabético y que de incluirse se considera mejorarían la bondad del ajuste del modelo. Por lo que se recomienda mejorar el registro de datos y sistemas de información que posee la CCSS.

## Bibliografía

1. Aparicio, Amada y Morera, Melvin (2006). Determinantes del adecuado abordaje del paciente diabético: aplicación del análisis multinivel para Costa Rica. En revisión. In press 2006.
2. Catalán-Reyes y Galindo Billardón (2003). *Utilización de modelo multinivel en investigación sanitaria*. Gaceta Sanitaria, 2003; 17(Supl 3): 35-52.
3. Conesa, M.D., Rodríguez Laso, A. y Otero, A. (2004). *Aplicaciones de los Modelos Multinivel al análisis de medidas repetidas en estudios longitudinales*. Revista Española de Salud Pública, 78, pp. 177-188.
4. Davis K, Taylor B, Furniss D (2001). *Narrative account of tracking the rural domestic violence survivors' journey: a feminist approach*. Health care Women Int 2001;22:333-47. Citada en Catalán y Galindo, 2003.
5. Flack RS, Wang J, Siegal HA, Carlson RG (2000). Longitudinal application of the medical outcomes study 36-item short-form health survey with not-in-treatment crack-cocaine user. Med Care 2000;38:902-10 Citada en Catalán y Galindo, 2003.
6. Goldstein H (1999). *Multilevel Statistical Models*. Institute of Education, Multilevel Models Project. London, april 1999. Disponible en: <http://www.arnordpublishers.com/support/goldstein.htm>.
7. González, Beatriz (2004). *Nuevos instrumentos de análisis de los determinantes de la prescripción y prestación farmacéutica*. Barcelona: Masson, 2004; 147-172.
8. González, B (2004a). *Curso Regresión Multinivel*. Departamento de Métodos Cuantitativos, Universidad Las Palmas, Gran Canaria. 2004
9. Hox, J.J. y de Leeuw, E.D. (2003). *Multilevel models for meta-analysis*. En S.P. Reise & N. Duan (Eds.),. *Multilevel Modeling. Methodological Advances, Issues, and Applications* (pp. 90-111). Mahwah, NJ: Lawrence Erlbaum Associates.
10. Rasbash Jon, Steele Fiona, Browne William y Prosser Bob (2004). *A User's Guide to MLwiN Version 2.0*. Centre for Multilevel Modelling Institute of Education University of London, Printed November 2004. Disponible en: <http://www.mlwin.com/download/userman20.pdf>.
11. Sánchez-Cantalejo E y Ocaña-Riola R (1999). *Los modelos multinivel o la importancia de la jerarquía*. Gaceta Sanitaria 1999; 13(5):391-398.
12. Tilling K, Sterne JA, Wolfe CD (2001). *Multilevel growth curve models with covariate effects. Application to recovery after stroke*. Stat Med 2001;20:685-704. Citada en Catalán y Galindo, 2003.
13. Web Site Centre of Multilevel Modelling, University of Bristol. [www.mlwin.com](http://www.mlwin.com). Consulta realizada el 10-11-2005.

**Cuadro 1. Temas en salud con mayores aplicaciones, según la revisión bibliográfica**

<b>Tema en salud</b>	<b>Artículos publicados</b>
Psiquiatría-Psicología	19
Tabaco alcohol y drogas	16
Salud de la mujer y embarazo	15
Metodología estadística en salud	13
Administración y economía en salud	11
Salud infantil y adolescente	11
Mortalidad de la población	9
Calidad de cuidado y práctica médica	8
Comunicación profesional de salud/paciente	8
Salud dental	5
Otros temas en salud	107

**Fuente:** Adaptado de Catalán et al (2003).

**Cuadro 2. Tipos de modelos según la característica de la variable de respuesta**

<b>Tipo de Modelo</b>	<b>Variable de respuesta</b>
<i>Lineal</i>	Donde hay una variable de respuesta continua
<i>Logístico</i>	Para variable de respuesta dicotómica
<i>Longitudinal</i>	Para una respuesta media en diferentes tiempos u ocasiones, donde lo fundamental es la evolución en el tiempo
<i>Medidas repetidas</i>	Donde la respuesta se mide en diferentes ocasiones, pero lo fundamental no es el tiempo
<i>Multinomial</i>	La respuesta tiene más de dos categorías
<i>Poisson</i>	Para recuento de eventos
<i>Multivariante</i>	Cuando se estudian varias respuestas
<i>Multivariante-longitudinal</i>	Para varias respuestas medidas en diferentes tiempos
<i>Supervivencia</i>	Donde la respuesta es la aparición de un evento
<i>Clasificación cruzada</i>	Para unidades que pertenecen a diferentes grupos
<i>Metanálisis</i>	Para el análisis del tamaño del efecto de varios estudios
<i>Metodología</i>	Cuando el interés principal de la investigación es dar a conocer la metodología multinivel
<i>Categoría no específica</i>	Se utiliza cuando el resumen y las referencias no proporcionan información sobre el tipo de modelo

**Fuente:** Adaptado de Catalán et al (2003)

**Cuadro 3**  
**Software disponible para el análisis multinivel**

Nombre	Especificaciones	Sitios en la Red
AML	Concentración de eventos históricos y modelos multiprocesos	<a href="http://www.applied-ml.com">www.applied-ml.com</a>
ASREML	Estimación de Máxima verosimilitud. Fracciones similares a GENSTAT.	<a href="http://www.vsn-intl.com/asreml/">www.vsn-intl.com/asreml/</a>
BAYESX	Generalización del cuadrado medio del error estimado Respuesta continuas y discretas con datos anidados y de estructura de clasificación cruzada Regresión semiparamétrica sobre concentraciones .	<a href="http://www.stat.uni-meunchen.de/lang/bayesx/bayesx.html">www.stat.uni-meunchen.de/lang/bayesx/bayesx.html</a>
BMDP	Modelo de componentes de la varianza y correlación seriada para estructuras de datos anidados. Máxima verosimilitud y GEE.	<a href="http://www.statsol.ie/bmdp/bmdp.htm">www.statsol.ie/bmdp/bmdp.htm</a>
EGRET	Variables discretas para datos anidados en dos niveles. Estimación de máxima verosimilitud.	<a href="http://www.Cytel.com/products/egret">www. Cytel.com/products/egret</a>
GENSTAT	Variables discretas y continuas. Datos anidados y tabla cruzada. Estimación de máxima verosimilitud.	<a href="http://www.nag.co.uk/stat/tt_soft.asp">www.nag.co.uk/stat/tt_soft.asp</a>
HLM	Variables discretas en tres niveles. Estructura de series correlacionadas; medidas del error. Estimación de máxima y "cuasiverosimilitud".	<a href="http://www.ssicentral.com/hlm">www.ssicentral.com/hlm</a>
LIMDEP	Programa general de respuesta econométrica. Respuesta de variables continuas y discretas. Datos anidados. Estimación de máxima verosimilitud.	<a href="http://www.limdep.com">www.limdep.com</a>
LISREL	Ecuaciones de estructura multinivel. Estructura de datos anidados. Estimación de máxima verosimilitud	<a href="http://www.ssicentral.com/lisrel.html">www.ssicentral.com/lisrel.html</a>
MIXOR, MIXREG	Programa para respuestas continuas y discretas multicatóricas de tres niveles. Estimación de máxima verosimilitud	<a href="http://tigger.uic.edu/~hedeker/mix.html">http://tigger.uic.edu/~hedeker/mix.html</a>
MLWIN	Paquete de respuesta general. Respuestas continuas y discretas para datos anidados, tablas cruzadas y grupos de estructura múltiple para cualquiera de los niveles. Estructura de series correlacionadas; modelo de eventos históricos; análisis de factores; medida de error. Estimación de máxima verosimilitud y estimación de cuadrado medio del error estimado	<a href="http://www.multilevel.ioe.ac.uk">www.multilevel.ioe.ac.uk</a>
MPLUS	Variables de respuesta continua y discreta. Datos anidados. Ecuaciones de estructura multinivel.	<a href="http://www.atatmodel.com/mplus">http://www.atatmodel.com/mplus</a>

Estimación de máxima verosimilitud

OSWALD	Trabajar con <i>s plus</i> para análisis de series correlacionadas y datos de eventos históricos. Máxima verosimilitud y estimación GEE	<a href="http://www.maths.lancs.ac.uk/software/oswald">http://www.maths.lancs.ac.uk/software/oswald</a>
SAS (versión 8)	Paquete de respuesta general. Respuestas continuas y discretas para datos anidados y de estructura de clasificación cruzada de más de dos niveles. Estructura de serie correlacionadas. Máximos y estimación de "cuasiverosimilitud".	<a href="http://www.sas.com/products/sassystem/release82/">http://www.sas.com/products/sassystem/release82/</a>
S_PLUS 2000	Paquete de respuesta general. Estructura continua y discreta para datos anidados	<a href="http://www.insightful.com">http://www.insightful.com</a>
SPSS (versión 10)	Paquete de respuesta general. Estructura continua y discreta para datos anidados	<a href="http://www.spss.com/spss10">http://www.spss.com/spss10</a>
STATA	Paquete de respuesta general. Respuestas continuas y discretas. Datos anidados y de estructura de clasificación cruzada de más de dos niveles. Modelo de ecuación estructural (programa para Modelos Lineales Generalizados y Mixtos con variables Latentes, GLLAMM), serie correlacionadas. Estimación de máxima verosimilitud.	<a href="http://www.spssscience.com/systat">http://www.spssscience.com/systat</a> <u>para modelos gllamm ver:</u> <a href="http://www.gllamm.org">www.gllamm.org</a>
SYSTAT	Paquete de respuesta general. Respuestas continuas. Datos anidados de más de dos niveles. Serie correlacionadas. Estimación de máxima verosimilitud.	<a href="http://www.spssscience.com/systat">http://www.spssscience.com/systat</a>
WINBUGS	Paquete de respuesta general. Usos de cuadrado medio del error estimado para ajustar el mejor rango vía lenguaje estadístico controlado. Respuestas continuas y discretas, medidas de error, análisis de factores, serie correlacionada y más.	<a href="http://www.mrc-bsu.cam.ac.uk">http://www.mrc-bsu.cam.ac.uk</a>

---

**Fuente:** Web Site de los programas.

**Cuadro 4. Descripción de las variables**

	<b>Variable</b>	<b>Indicador</b>	<b>Descripción</b>
<b>N i v e l  I</b>	Hgnor (Hemoglobina glicosilada normal )	Porcentaje de hemoglobina glicosilada	1= cuando la hemoglobina glicosilada es menor o igual a 7%. 0= cuando la hemoglobina glicosilada tiene otro valor.
	Sexo	Número de hombres	1= Hombre
	Edad	Número de mujeres	0= Mujer
	Estado nutricional (Índice de masa corporal)	Edad del paciente IMC	Edad en años cumplidos del diabético. Índice de masa corporal del diabético
<b>N i v e l  II</b>	Reforma (Momento de adopción de la reforma)	Pionera	1= si el AS inició la reforma entre 1995-1997 0= si la opción de la reforma fue en cualquier otro año.
		Intermed	1= si el AS inició la reforma entre 1998-2000. 0= si la opción de la reforma fue en cualquier otro año.
		Tardía	1= si el AS inició la reforma posterior al año 2000. 0= si la opción de la reforma fue en cualquier otro año.
	Tiempo de reforma (1995-2004)	treform	Variable que mide los años desde que se el AS inició la reforma
	Disponibilidad de recursos (horas médico por mil habitantes)	hrsmedmil	Horas médico disponible por cada 1.000 personas adscritas al en el año 2004.
	Tiempo en Compromiso de gestión	Tiemcgr	Años de participar en el Compromiso de Gestión
	Nota en Compromiso de Gestión	Notacg	Nota promedio (de cero a uno) en la evaluación en Compromisos de Gestión en el quinquenio 2000-2004.
		Concent	1= si el AS es de gestión concentrada 0= cualquier otra forma de gestión
	tipogest (tipo de gestión)	Desconcent	1= si el AS es de gestión desconcentrada 0= cualquier otra forma de gestión
		Externa	1= si el AS es de contratación externa. 0= cualquier otra forma de gestión

**Cuadro 4. Descripción de las variables (continuación)**

Variable	Indicador	Descripción
tipocentro (tipo de centro)	Nivel1	1= si el AS es de primer nivel 0= si el AS tiene cualquier otro nivel de complejidad
	Nivel1refor	1= si el AS tiene médico internista en el primer nivel. 0= si el AS tiene cualquier otro nivel de complejidad
	Nivel2	1= que tiene conjuntamente atención de primer nivel y atención ambulatoria especializada 0= si el AS tiene cualquier otro nivel de complejidad
horario (horario atención)	Diurno	1= áreas con horario normal de 8 horas de atención. 0= si el tiene cualquier otro horario de atención
	Vespert	1= que poseen, además del horario normal de 8 horas, una extensión de horario no menor a 24 horas. 0= si el tiene cualquier otro horario de atención
	Urgencias	1= que poseen, además del horario normal de 8 horas, un servicio de urgencias de 24 horas. 0= si el tiene cualquier otro horario de atención
Presión asistencial	Ninos	Proporción de niños con menos de 6 años dentro de la población total adscrita al AS.
	Adultomay	Proporción de población con 65 años y más dentro de la población total adscrita al AS.
	Primaria	Proporción de la población adscrita a cada AS con un nivel educativo de primaria o menos.
Nivel educativo	Secund	Proporción de la población adscrita a cada AS con un nivel educativo de secundaria completa (académica o técnica).
	Univer	Proporción de la población adscrita a cada AS con un nivel educativo parauniversitaria o universitaria
Acceso al hospital (Distancia)	kmhosp	Distancia en kilómetros de la Sede del AS al Hospital de referencia.

**Fuente:** Aparicio y Morera (2006)N  
i  
v  
e  
l  
  
II

**Cuadro 5. Cálculo de Odd Ratio**

<b>Variable</b>	<b>Parámetro estimado (<math>\beta</math>)</b>	<b>Error Estándar (EE)</b>	$Z = \frac{\beta}{EE}$	<b>Odd Ratio (<math>e^\beta</math>)</b>
Sexo	-0,183	0,077	$Z = \frac{-0,183}{0,077} = 2,37$	$\exp^{-0,183} = 0,833$
Pionera	-1,106	0,40	2,74	0,33