

Muestreo por conglomerados en encuestas poblacionales

Cluster sampling in health surveys

Helmer de Jesús Zapata-Ossa, Angela M. Cubides-Munévar,
María C. López, Elisa M. Pinzón-Gómez, Paola A. Filigrana-Villegas y
Carlos A. Cassiani-Miranda

Facultad de Medicina, Fundación Universitaria San Martín. Cali Colombia.
zapatahelmer@gmail.com, angelacubides.epi@gmail.com, marlop1224@gmail.com,
elisapinzon@hotmail.com, filigranapaola@yahoo.es, Kassio30@hotmail.com

Recibido 3 Agosto 2010/Enviado para Modificación 19 Noviembre 2010/Aceptado 3 Diciembre 2010

RESUMEN

Objetivo Describir la técnica estadística del muestreo por conglomerados en encuestas de salud poblacionales.

Métodos Se describen los aspectos más importantes de cada uno de los pasos metodológicos haciendo énfasis en los fundamentos del Muestreo por Conglomerados. Éste se aplicó sobre una población de aproximadamente 100 000 habitantes de la comuna 18 de Cali. La muestra abarcó 30 manzanas y se encuestaron 3 026 personas. La información se recolectó usando el Sistema de Información de base Comunitaria para Atención Primaria en Salud.

Resultados La técnica de muestreo aplicada generó información precisa y confiable para la estimación de prevalencias poblacionales de eventos de salud en un estudio de corte transversal. En población general, el error de muestreo fue de 2,5 %; en menores de 5 años fue del 10 %; en niños de 5 a 14 años fue de 5,7 %; en población de 10 a 59 años estuvo alrededor del 3,0 % y en población de 60 y más años fue de 7,2 %.

Conclusión El método descrito permitió garantizar un error de muestreo de 2,5 % en la estimación de parámetros de prevalencia en la población general y errores de muestreo inferiores al 10 % para los grupos poblacionales establecidos.

Palabras Clave: Muestreo por conglomerados, encuestas de salud, prevalencia, estudio transversal (*fuentes: DeCS, BIREME*).

ABSTRACT

Objective Describing how the cluster sampling statistical technique can be applied to health surveys.

Methods The article describes the most important aspects for each methodological step, emphasising cluster sampling's foundations. It was applied to a population of around 100,000 inhabitants from commune 18 in Cali. The sample covered 30 city blocks and 3,026 people were interviewed. The information was gathered using the Community-based Information System for Primary Health Attention.

Results The sampling technique produced precise and reliable information for estimating population health event prevalence in a cross-sectional study. Sampling error was 2.5 % in the general population, 10% in children aged less than 5 years, 5.7% in children aged 5 to 14, around 3.0% in those aged 10 to 59 and 7.2% in the population aged more than 60.

Conclusion The method guaranteed a 2.5% sampling error in estimating general population prevalence and less than 10% sampling errors for established population groups.

Key Words: Cluster sampling, health survey, prevalence, cross-sectional (*source: MeSH, NLM*).

La aplicación de encuestas poblacionales en la salud es de gran utilidad para medir las condiciones de vida de una región en un momento dado porque facilitan el monitoreo, la vigilancia en salud pública y las intervenciones que se realicen en ellas. Este tipo de encuestas se han aplicado en los estudios mundiales sobre maltrato en el entorno familiar (1), en la medición de la violencia contra la mujer (2) y otras importantes encuestas demográficas y de salud (3,4).

En Colombia, gran parte de la información del estado de salud y sus determinantes en poblaciones humanas, se basa en datos provenientes de encuestas poblacionales, como la Encuesta de Calidad de Vida, que realizó el Departamento Nacional de Estadística (DANE) en 2008 (5), la Encuesta de Demografía y Salud, realizada por Profamilia en el año 2005 (6) y la reciente Encuesta Nacional de Salud, liderada por el Ministerio de la Protección Social y Colciencias (7). En Medellín se tuvo otra experiencia en la que se utilizó un muestreo complejo probabilístico, estratificado por estrato socioeconómico, por conglomerados de manzanas y polietápico (8).

En la ciudad de Cali, también existen iniciativas para la aplicación de este tipo de encuestas, en particular en el Sistema de Vigilancia en Salud Pública, hay experiencias en la vigilancia de los factores de riesgo de las Enfermedades Crónicas no Transmisibles (ECNT) (9), a través de la aplicación de encuestas poblacionales. Sin embargo, a pesar de su amplio uso, aspectos como la validez de las técnicas de muestreo, es un problema metodológico por resolver en la aplicación adecuada de estas encuestas. No obstante este problema metodológico, los investigadores y los tomadores de decisiones las consideran instrumentos insustituibles, porque son un medio práctico para documentar y complementar información sobre hechos vitales de uno o varios sistemas de salud.

La propuesta de la Organización Mundial de la Salud (OMS), de desarrollar un muestreo de 30 conglomerados para la selección de la muestra, en los monitoreos rápidos de las coberturas de vacunación en un área geográfica, es una estrategia que puede resolver el problema metodológico del diseño de muestreo en las encuestas poblacionales (10,11). En este sentido, el diseño de muestreo por conglomerados MC monoetápico puede extenderse a las encuestas poblacionales de salud en las cuales el interés sea establecer líneas de base o sistemas de vigilancia en salud pública en poblaciones relativamente homogéneas, porque es de fácil aplicación. A pesar de los esfuerzos, no se ha estandarizado la aplicación de una encuesta de salud de base poblacional, que se fundamente en el muestreo por conglomerados monoetápico (12), ni se han determinado sus implicaciones en el proceso de la inferencia estadística (13,14).

El objetivo de este estudio es documentar una técnica del MC monoetápico y la aplicación de una encuesta "Sistema de Información de Base Comunitaria Para la Atención Primaria en Salud - SICAPS" (15), en la población de la comuna 18 de la ciudad de Cali-Colombia.

MATERIALES Y MÉTODOS

Los pasos que se consideraron en la aplicación del MC de la encuesta poblacional (16), se resumen en: 1. Expresar los objetivos de la encuesta los cuales se definieron con base en la investigación "Estado de Salud y sus Determinantes en Población de la Comuna 18"; 2. Definir el universo de estudio; 3. Definir el diseño de muestreo; 4. Determinar las variables del estudio; 5. Selección y adiestramiento de los investigadores de campo; 6. Reconocimiento de la zona y levantamiento de mapas de los conglomerados geográficos; 7. Organización del manejo de los datos; y 8. Procedimientos de estimación y plan de análisis.

La población universo de estudio, correspondió a los residentes habituales en hogares y viviendas de la zona. La unidad de observación correspondió a los individuos entrevistados.

Se consideraron variables relacionadas con el estado de la vivienda, conformación de las familias y condiciones sociodemográficas de sus integrantes, eventos mórbidos y factores de riesgo para la salud.

El diseño de muestreo se clasificó en un MC monoetápico (16). Se tuvo en cuenta la variabilidad intra e interconglomerados (17). Los estimadores de medias poblacionales se obtuvieron a través de la expresión:

$$\bar{x} = \frac{\sum_{i=1}^m x_i}{\sum_{i=1}^m N_i}$$

Y su error estándar mediante la fórmula:

$$EE(\bar{x}) = \sqrt{Var(\bar{x})} = \sqrt{\left[\frac{(N-m)/(Nm\bar{M}^2)}{m-1} \right] \sum_{i=1}^m (x_i - \bar{x}N_i)^2}$$

Dado que desconoce \bar{M} este valor se reemplazó por \bar{m}

Los estimadores de las proporciones poblacionales de prevalencias se estimaron a través de la expresión:

$$\bar{p} = \frac{\sum_{i=1}^m a_i}{\sum_{i=1}^m N_i}$$

Donde N_i es el número de elementos en el i -ésimo conglomerado ($i = 1, 2, \dots, n$) y a_i el número de elementos en el conglomerado i que poseen la característica de interés.

Los errores estándar de los estimadores de las proporciones poblacionales se calcularon de la siguiente manera:

$$EE(\bar{p}) = \sqrt{Var(\bar{p})} = \sqrt{\left[\frac{(N-m)/(Nm\bar{M}^2)}{m-1} \right] \sum_{i=1}^m (a_i - \bar{p}N_i)^2}$$

El tamaño de muestra se estimó con las fórmulas del muestreo aleatorio simple MAS; La fórmula para calcular el tamaño de muestra a través de un MAS, es:

$$n_0 = \frac{P(1-P)Z^2_{1-\alpha/2}}{\varepsilon^2}$$

Donde P corresponde a la proporción esperada del evento; $(1-P)$ corresponde a la probabilidad complemento; $Z_{1-\alpha/2}$ corresponde al percentil $1-\alpha/2$ de la distribución normal estándar el cual determina el nivel de confiabilidad del estudio; ε determina la precisión absoluta en los procesos de estimación de la proporción poblacional. Dada las múltiples características a evaluar, se usó el valor de $P=0,5$ para maximizar el tamaño de muestra. El valor n_o se ajustó por un factor de corrección por población finita "fcp" (18).

$$n_1 = \frac{n_o}{1 + \frac{n_o}{N}}$$

Este valor se incrementó considerando un ajuste por no respuesta % NR del 30 %. La forma n_1 de ajustar fue,

$$n_2 = \frac{n_1}{1 - \% NR}$$

A su vez el valor ajustado por con ajuste por "% NR", se incrementó por efecto de diseño multiplicando n_2 por el efecto de diseño *deff* con el propósito de corregir la varianza de los estimadores que generalmente fluctúa entre 1,5 y 3,0 (12). El *deff* corresponde a la medida de homogeneidad entre los elementos al interior de cada conglomerado.

$$n = n_2 * deff$$

Dado que los tamaños de los conglomerados son variables, el *deff* y la varianza del estimador de la proporción se puede obtener a través de la fórmula propuesta por Cochran (17),

$$Deff = (N - n) \frac{\sum_{i=1}^{i=M} M_i (p_i - P)^2}{N - 1}$$

Donde M corresponde al total de conglomerados; N es el número de elementos en la población; n es el tamaño de muestra de elementos; p_i corresponde a la prevalencia esperada del evento en el conglomerado i .

Cuando " n " es grande como en esta situación, el efecto de diseño se puede aproximar con la siguiente fórmula Cochran (17),

$$Deff = M \frac{\sum_{i=1}^{i=M} (p_i - P)^2}{NPQ}$$

También se calculó el coeficiente de correlación intraclase rho usando la expresión $1+(b-1)rho$; donde "b" es el número de elementos por conglomerados que se calculó utilizando el valor medio del tamaño de los conglomerados "n/m", donde "n" es el número de elementos de la muestra y "m" es el número de conglomerados seleccionados. Para eventos de salud, los valores rho generalmente fluctúan entre 0,1 y 0,3 (19).

La selección aleatoria sistemática de los "m" conglomerados se realizó a partir del marco muestral de los "M" conglomerados suministrados por Planeación del Municipio. La fracción de muestreo se determinó por la expresión $f=m/M$.

Se controló el sesgo en las estimaciones a través del factor de expansión, el cual se obtuvo a través del inverso de la fracción de muestreo "1/f". Este valor a su vez, se ajustó por un factor de cobertura de la muestra, que permitió corregir problemas de marcos muestrales incompletos y la no participación de algunos o todos los miembros del hogar (20).

El operativo de campo se realizó con personal previamente entrenado. En el reconocimiento, se utilizaron instructivos y mapas con codificación de barrios y manzanas del municipio de Santiago de Cali (21). Se estableció un organigrama con definición de responsables y actividades de coordinación, de trabajo de campo, y de articulación con los representantes de la comunidad. La recolección de los eventos de salud se realizó a partir de las tarjetas SICAPS, las cuales fueron sometidas a revisión crítica previa a su digitación en el aplicativo SICAPS.

El proyecto fue aprobado por el Comité de Ética de la Fundación Universitaria San Martín FUSM.

RESULTADOS

Los pasos de la aplicación del MC usado en la medición del estado de salud y sus determinantes en la población de la comuna 18 de la ciudad de Cali, se detallan a continuación.

Población objetivo. El estudio se realizó en la comuna 18, ubicada al suroccidente de la ciudad de Cali en una zona de alta vulnerabilidad social. Para el 2008 la comuna tenía 97 707 habitantes según proyecciones de población DANE (22), distribuidas según sexo, así: 48,1 % hombres y 51,9 % mujeres.

Muestra. Los parámetros usados en la aplicación del MC fueron: tamaño promedio del conglomerado de 35 viviendas (9); relación de hogares por vivienda 1,1 hogar por vivienda (22); y 3,8 personas por hogar (5). De las M=697 manzanas identificadas en los mapas suministrados por la oficina de planeación del municipio (21), se planeó una muestra de 30 y se reconocieron 36 considerando un ajuste por cobertura del 15 % sobre el número de conglomerados; este procedimiento permitió prever que algunas manzanas coincidieran con parques, iglesias, clubes o canchas.

Cobertura de la muestra. Los encuestadores reconocieron en las 30 manzanas: 1 038 viviendas y 1 142 hogares. Al momento de la aplicación de la encuesta, los entrevistadores visitaron 808 hogares de los 1 142 planeados que corresponden a una cobertura del 71 %. En los hogares visitados se logró recoger información de 3 026 personas, la distribución de la población total de la comuna y la muestra obtenida según grupos poblacionales se presentan en la Tabla 1.

Tabla 1. Población y muestra obtenida según grupo poblacional, resultado de la aplicación de la encuesta, Comuna 18 de Santiago de Cali, año 2008

Grupo Poblacional	Población	Tamaño de muestra obtenido
No. Hogares	24 049	808
Población total	97 707	3 022
Niños(as) menores de cinco años	9 253	184
Niños(as) de cinco a catorce años	19 559	579
Hombres y Mujeres de 10 a 52 años	75 379	2 019
Hombres y Mujeres de 10 a 59 años	80 229	2 195
Hombres y Mujeres de 40 a 59 años	18 835	639
Hombres y Mujeres de 60 y más años	7 376	358

Fuente: Población DANE, Cálculos de los investigadores

Errores de muestreo. Se recalcularon los errores de muestreo para cada uno de los grupos poblacionales. Los cálculos se realizaron teniendo en cuenta lo siguiente: prevalencia del evento de salud del 50 %, confiabilidad del 95 %; ajuste por población finita de acuerdo al tamaño de la subpoblación y un ajuste por efecto de diseño (13). Se observa que el máximo error de muestreo se presenta para las estimaciones en el dominio de estudio conformado por los menores de cinco años siendo del 10,2 %, en el resto de grupos poblacionales los errores fueron inferiores al 10 % (Tabla 2).

Tabla 2. Tamaños de muestra para una variabilidad máxima del evento en salud y un error máximo, por grupo poblacional de los módulos de la encuesta SICAPS*

Características	Número	Muestra estimada				Tamaño de muestra n ajustado por DEFF=2	Tamaño de muestra obtenido
		Error de muestreo máximo	Muestra planeada mediante un MAS	Muestra planeada ajuste por población finita	Muestra planeada ajuste por población finita		
No. Hogares	24 049	4,8	410	403	806	808	
Población total	97 707	2,5	1 537	1 513	3 026	3 022	
Niños(as) menores de cinco años	9 253	10,2	93	92	185	184	
Niños(as) de cinco a catorce años	19 559	5,7	293	288	576	579	
Hombres y Mujeres de 10 a 52 años	75 379	3,1	1 026	1 012	2 024	2 019	
Hombres y Mujeres de 10 a 59 años	80 229	2,9	1 111	1 096	2 192	2 195	
Hombres y Mujeres de 40 a 59 años	18 835	5,4	326	320	640	639	
Hombres y Mujeres de 60 y más años	7 376	7,2	184	180	359	358	

*Prevalencia del evento=50; Coeficiente de confiabilidad al 95 %=1,96; Fuente: Población DANE, Cálculos; de los investigadores; Nota; MAS: Muestreo Aleatorio Simple; DEFF: Efecto de Diseño.

Factores de expansión. En la Tabla 3, se presenta la fracción de muestreo de selección de la manzana ($f=m/M = 30/687=0,0437$); también se calculó los factores de expansión o "pesos" usando el inverso de f . Estos "pesos" fueron ajustados por la cobertura de la muestra. Los resultados indican que cada persona está representando a 32 individuos de la población de la comuna, lo cual se debe considerar en el plan de muestreo (22). Se observa la representatividad de la muestra para los grupos poblacionales obtenidas a partir de las ponderaciones. En ésta, también se dispone del indicador relacionado con el efecto de la conglomeración el cual estuvo por debajo del 0,1 lo que indica que las variables están distribuidas al azar al interior de los conglomerados.

Tabla 3. Población representada en la muestra según los pesos PSI y factores de cobertura para cada una de las unidades de análisis por grupo poblacional

Características	Número	Muestra estimada				Tamaño de muestra n ajustado por DEFF=2	Tamaño de muestra obtenido
		Error de muestreo máximo	Muestra planeada mediante un MAS	Muestra planeada ajuste por población finita	Muestra planeada ajuste por población finita		
No. Hogares	24 049	4,8	410	403	806	808	
Población total	97 707	2,5	1 537	1 513	3 026	3 022	
Niños(as) menores de cinco años	9 253	10,2	93	92	185	184	
Niños(as) de cinco a catorce años	19 559	5,7	293	288	576	579	
Hombres y Mujeres de 10 a 52 años	75 379	3,1	1 026	1 012	2 024	2 019	
Hombres y Mujeres de 10 a 59 años	80 229	2,9	1 111	1 096	2 192	2 195	
Hombres y Mujeres de 40 a 59 años	18 835	5,4	326	320	640	639	
Hombres y Mujeres de 60 y más años	7 376	7,2	184	180	359	358	

Fuente: Población DANE. Cálculos de los investigadores; Nota: Total de manzanas en la comuna = 687; DEFF=1 + (b-1) rho; Fracción de muestreo 0,0437; Ponderación o factor de expansión 22,9; Ponderación ajustada por cobertura del estudio (71 %) de 32,2

DISCUSIÓN

El muestreo por conglomerados ha tenido diferentes usos en salud; el método utilizado en el Plan Ampliado de Inmunización PAI (10, 23) usa 30 conglomerados artificiales constituidos por 7 niños, donde se estiman coberturas de vacunación con error máximo de 10 %, considerando una prevalencia de 50 %, una confiabilidad de 95 % y un efecto de diseño de 2. El método que se propone en este artículo, con los mismos parámetros, permite garantizar un error inferior a 5 % en la estimación de parámetros de prevalencia en la población general y errores de muestreo entre 5 % y 10 % para la estimación de prevalencias en los grupos poblacionales menores de cinco y mayores de sesenta años.

Con relación a la selección de los conglomerados, en el método PAI se exige la selección aleatoria de un individuo dentro de un conglomerado o punto inicial; la selección del individuo comienza en el punto inicial hasta que se consiguen siete, cuyas edades interesan en función de su estado de inmunización; de esta manera se establecen los conglomerados artificiales. De manera similar, en el método modificado de muestreo (11), utilizado para evaluación rápida de necesidades después de un desastre, los encuestadores llegan a un punto cerca del centro de cada uno de los 30 conglomerados, se encaminan en una dirección elegida al azar por lanzamiento de una moneda, hasta llegar a la unidad de observación habitada más cercana en la que entrevistan a un habitante adulto. Después siguen caminando a la próxima unidad hasta que completan siete entrevistas. En el caso de viviendas compartidas por varias familias, solamente se entrevistan a las personas de la primera vivienda.

El método propuesto es de fácil aplicación pues solo requiere de la selección aleatoria de forma sistemática de los conglomerados geográficos, los que generalmente están disponibles en los mapas de las alcaldías de los municipios. La muestra constituida tiene las características de "Muestra Maestra", lo que permite disponer de información para la caracterización de toda una población o una subpoblación específica, manteniendo errores de muestreo aceptables.

El "SICAPS" tiene como propósito el monitoreo de la salud y sus condicionantes, por lo que se ha aplicado de manera censal. En este estudio se dio un uso alternativo a este sistema de vigilancia, considerándolo como encuesta poblacional por muestreo. Para su aplicación hubo necesidad de establecer ajustes frente al objetivo del estudio y documentar el proceso para su aplicación, lo que permite poner de manifiesto una alternativa de su uso para el sector de la salud.

Con relación a las limitaciones del estudio, se puede mencionar el alto porcentaje de no respuesta, que afectó la composición y tamaño de los grupos poblacionales seleccionados en la muestra, influenciando el error de muestreo. También se puede considerar como otra limitación el contar con muestras relativamente pequeñas en algunos grupos poblacionales lo que limita los procesos de estimación en eventos de salud de baja frecuencia, generando como consecuencia, que el error de muestreo pueda ser incluso superior a la prevalencia del evento, por lo que el resultado de la estimación puede llegar a ser no válido. Se resalta como una desventaja el hecho de disponer de información cartográfica que al momento de seleccionar las manzanas, pudieron estar con algún grado de desactualización ♦

Agradecimientos: La investigación en su diseño, ejecución, recolección de información, procesamiento, análisis y redacción del manuscrito fue financiada por la Fundación Universitaria San Martín, Sede Cali. Los autores agradecen al Centro de Investigaciones Multidisciplinarias para el Desarrollo CIMDER, la participación activa de sus funcionarios en el proceso de capacitación, ajuste del instrumento y apoyo logístico en el desarrollo de esta investigación.

REFERENCIAS

1. Organización Mundial de la Salud [Internet]. Estudio multipaís de la OMS sobre salud de la mujer y violencia doméstica contra la mujer (WorldSafe).2005. Disponible en: http://www.who.int/gender/violence/who_multicountry_study/summary_report/chapter1/es/index6.html. Consultado 12 de enero de 2010.
2. Sagot M. [Internet]. Medir la Violencia. Un desafío para la elaboración de políticas efectivas: El caso de la Encuesta Internacional de Violencia contra las Mujeres (IVAWS). Disponible en: <http://www.google.com.co/search?hl=es&q=IVAWS+violencia+contra+la+mujer&meta> Consultado 30 de enero de 2010.
3. Macro International Inc [Internet]. Demographic and Health Surveys: MEASURE DHS. Disponible en: <http://www.measuredhs.com> Consultado 30 de enero de 2010.
4. CDC. [Internet]. International Reproductive Health Surveys and Comparative. Disponible en: <http://www.cdc.gov/reproductivehealth/Surveys/> Consultado 20 de enero de 2010.
5. DANE. [Internet]. Encuesta sobre calidad de vida, año 2008. Disponible en: http://www.dane.gov.co/index.php?option=com_content&task=category§ionid=35&id=32&Itemid=144. Consultado 10 de enero 2010.
6. Profamilia. [Internet]. Encuesta de demografía y salud, año 2005. Disponible en: <http://www.profamilia.org.co/encuestas/index.htm>. Consultado 20 de enero de 2010
7. Ministerio de la Protección Social, Colombia. [Internet]. Encuesta Nacional de Salud 2007. Disponible en: <http://www.minproteccionsocial.gov.co/salud/PaginasEncuestaNacionaldeSaludPublica.aspx>. Consultado 30 de enero de 2010.
8. Cardona D, Agudelo HB, Segura AM. Un diseño de muestreo complejo en el análisis de la calidad de vida de la población adulta. Medellín, 2005. Colomb Med. 2008; 39 (2): 161-74.
9. Centro para el Desarrollo y Evaluación de Políticas y Tecnología en Salud Pública (CEDETES). Estudio de Prevalencia de Factores de Riesgo Asociados a las ECNT. Boletín informático del CEDETES. Santiago de Cali: CEDETES; 2005.

10. Henderson RH, Sundaresan T. Cluster sampling to assess immunization coverage: a review of experience with a simplified sampling method. *Bull World Health Organ.* 1982;60:253-260.
11. Malilay J, Flander WD, Brogan D. Método modificado de muestreo por conglomerados para la evaluación rápida de necesidades después de un desastre. *Rev Panam Salud Pública.* 1997; 2(1): 7-1.
12. Kish L. *Muestreo de Encuesta.* 1a edición. México. Editorial Trillas; 1982.
13. Silva LC. *Muestreo para la Investigación en Ciencias de la Salud.* Madrid: Díaz de Santos; 1993.
14. Mandujano F. *Teoría del Muestreo: Particularidades del Diseño Muestral En Estudios de la Conducta Social.* Universidad de Playa Ancha Valparaíso, Chile. *Revista Electrónica de Metodología Aplicada.* 1998; 3 (1):1-15. Disponible en: <http://www.psico.uniovi.es/rema/v3n1/indice.html>. Consultado 10 de enero de 2010.
15. Centro de Investigaciones Multidisciplinarias para el desarrollo CIMDER. *Sistema de Información de Base Comunitaria para Atención Primaria en Salud* año 2004.
16. Mendenhall W, Scheaffer R, Ott L. *Elementos de Muestreo.* México. Grupo Editorial Iberoamérica; 1987.
17. Cochran W. *Técnicas de Muestreo.* 3ª Impresión. México: Editorial Continental S.A.; 1969.
18. Daniel WD. *Bioestadística. Base para el análisis de las ciencias de la salud.* Cuarta edición en español. México: Limusa Wiley; 2008.
19. Bennett S, Woods T, Liyanage WM, Smith DL. A simplified general method for cluster-sample survey of health in developing countries. *World Health Stat Q.* 1991;44(3):98-106.
20. Dean AG, Shah SP, Churchill JE, DoEpi. Computer-assisted instruction in epidemiology and computing and a framework for creating new exercises. *Am J Prev Med.* 1998; 14(4):367-71.
21. Escobar G. *Calí en Cifras 2005.* Santiago de Cali. Departamento Administrativo de Planeación (DAP)- Alcaldía de Santiago de Cali. Feriva; 2005.
22. DANE. [Internet]. *Resultados del Censo, año 2005.* Disponible en: www.dane.gov.co. Consultado 11 de enero de 2010.
23. Brogan D, Flagg EW, Deming M, Waldman R. Increasing the Accuracy of the Expanded Programme on Immunization's Cluster Survey Design. *Ann Epidemiol.* 1994;4(4):302-11.