

**Instituto Nacional de Salud Pública**  
**Centro de Investigación en Evaluación y Encuestas**

**Evaluación Externa del Programa de Desarrollo Humano Oportunidades 2007-2008**

**Nota técnica sobre el diseño muestral para la Encuesta de  
Evaluación Rural (ENCEL) 2007**

**Convenio I, 10/07/2007**

**Fecha de entrega: 31-08-07**

**Numeral: 7**

**Agosto, 2007**

## **Tabla de Contenido**

Introducción.....	1
Diseño de la muestra original de evaluación en zonas rurales.....	2
Muestra de evaluación en zonas rurales 2007 .....	3
Muestra original de 506 localidades seleccionadas en 1997 y las 151 adicionales de 2003. ....	4
B. Revisita a una submuestra de hogares de la ENSANUT 2006.....	6
D. Cuádruplas de localidades en Chiapas y Oaxaca .....	8
Asignación de cuestionario completo y corto entre los hogares de la muestra de Evaluación 2007 .....	11
Asignación de cuestionario completo y corto entre adolescentes de la muestra de Evaluación 2007 .....	12
Muestra para la evaluación operativa.....	13
Centro de Atención y Registro (CAR) .....	13
3. Mesas de Atención a Comités (MAC) y Mesas de Atención y Servicio- Mesas de entrega de apoyos (MAS-MAP).....	13
Anexo 1 Diseño Muestral de la ENSANUT 2006.....	14
Anexo 2 Distancia de Mahalanobis .....	19
Referencias .....	21

## **Introducción**

El Programa de Desarrollo Humano Oportunidades nace en 1997 como una estrategia del Ejecutivo Federal para apoyar a las familias que viven en condición de pobreza de capacidades, con el fin ampliar sus alternativas para alcanzar mejores niveles de bienestar. Para ello, el Programa se enfoca en incrementar el acceso a servicios de calidad y proporcionar apoyos económicos que permitan la continuidad de la asistencia escolar de niños y jóvenes, con el fin de que estos concluyan satisfactoriamente la educación básica y media superior y que cuenten con mayores opciones para continuar con su educación superior<sup>1</sup>. Adicionalmente, se ofrece a las mujeres embarazadas y niños complementos alimenticios y vigilancia médica para monitorear y mejorar el estado de nutrición, y proporcionarles información sobre auto-cuidado a la salud. Otro de los aspectos importantes es el fomento de la participación de los padres en la realización de las mejoras previstas en este programa.

Desde su creación, el Programa ha sido sometido a constantes evaluaciones. La intención de las mismas no sólo ha sido medir su impacto, resultados y logros, sino también servir como un mecanismo de retroalimentación que ayude a los responsables de su operación a tomar de decisiones respecto a:

- Los ajustes correctivos necesarios para hacer más eficiente al Programa
- Reorientar sus acciones
- Mejorar la funcionalidad de procesos

Desde la primera y hasta la etapa actual, la metodología muestral de estas evaluaciones ha experimentado cambios, todos ellos encaminados al fortalecimiento, actualización y consistencia con el estado y dinámica operativa del Programa y siempre buscando responder a las necesidades de información de los tomadores de decisiones.

En el presente documento se detallan los cambios metodológicos realizados a la estrategia muestral de evaluación en áreas rurales para la Encuesta de Evaluación de los Hogares Rurales 2007 así como su justificación.

## ***Diseño de la muestra original de evaluación en zonas rurales***

Para la evaluación de impacto del Programa en zonas rurales se utilizó un diseño experimental. Este estudio comprendió una muestra de localidades seleccionadas de manera aleatoria entre 6,396 que cumplían los requisitos para la incorporación al Programa, pertenecientes a siete estados del país: Guerrero, Hidalgo, Michoacán, Puebla, Querétaro, San Luis Potosí y Veracruz. La selección incluyó un total de 506 localidades de las cuales 320 sirvieron como tratamiento y 186 como control.

Por localidad tratamiento nos referimos a aquellas localidades que fueron incorporadas al Programa en 1998, mientras que el término localidades control lo aplicamos a las localidades en las que el Programa inició su operación dos años más tarde. La asignación de las localidades entre los grupos de tratamiento y control se realizó de manera aleatoria, dando una probabilidad de 0.6 a cada localidad de pertenecer al grupo intervención<sup>2</sup>.

A partir de 1998 y hasta 2000 se realizaron seis visitas a estas localidades, con diferencia temporal de seis meses entre cada visita, para recabar información que permitiera el cálculo de indicadores de desarrollo social para cada grupo (intervención y control) y la comparación de estos resultados y así poder evaluar los efectos a corto plazo del Programa<sup>3</sup>.

Seis años después, en 2003, las 506 localidades de la muestra original habían sido incorporadas al programa, lo que impedía contrastar los indicadores de localidades incorporadas al programa con los de no incorporadas. Sin embargo, con la intención de poder obtener información y cuantificar el impacto del Programa en indicadores de mediano plazo, se realizó una séptima visita a estas mismas localidades y 151 localidades adicionales similares a las originales, pero que no habían participado en el Programa, y que se encontraban ubicadas dentro de los siete estados entre los que se distribuía la muestra original con la finalidad de que sirvieran como nuevos controles de la evaluación.

La selección de las 151 localidades adicionales no se realizó mediante un mecanismo de aleatorización, ya que la intención era incorporar en la muestra aquellas localidades que resultarán más similares a las de la muestra original; para ello, se implementó un procedimiento de pareamiento múltiple basado en un puntaje de propensión<sup>4</sup>.

Existen múltiples documentos que exponen con detalle diferentes aspectos sobre las variables consideradas en estos diseños y su eficiencia <sup>3 5 6</sup>.

## ***Muestra de evaluación en zonas rurales 2007***

Para la definición de la muestra de evaluación para zonas rurales 2007 se tomaron en cuenta diferentes dimensiones entre las que podemos encontrar:

- **Comparabilidad:** Se deseaba que la información derivada de esta nueva evaluación permitiera la comparación con los resultados obtenidos en las evaluaciones anteriores.
- **Longitudinalidad:** La posibilidad de que los datos obtenidos en 2007 permitieran observar la evolución de los hogares era importante tanto para los tomadores de decisiones (operadores del Programa) como para los investigadores que participaron en la definición de la muestra de evaluación.
- **Nuevas demandas de información:** Dado el crecimiento del Programa y su dispersión a lo largo del país, los responsables, nacionales y estatales, requerirán más información para apoyar sus decisiones.
- **Costos de la evaluación:** Dado que la implementación de los mecanismos de evaluación tienen un costo y las necesidades de información son diversas se tuvo que realizar un análisis detallado sobre las prioridades de información resultante de la evaluación.

Bajo la consideración de estas dimensiones se realizó un profundo análisis sobre la forma más adecuada de operar la evaluación y las localidades que conformarían la muestra de evaluación 2007. En este análisis se incorporaron los comentarios derivados de la consulta a investigadores, expertos y operadores del Programa.

Este análisis condujo al acuerdo de implementar la evaluación en tres frentes:

- A. Continuidad en localidades pertenecientes a la **muestra original de 506 localidades seleccionadas en 1997 y las 151 adicionales de 2003**.
- B. Revisita a una **submuestra de hogares de la Encuesta de Salud y Nutrición (ENSANUT)**, realizada por la SSA y el INSP en 2006 y que incorporó como variables de estratificación el tamaño de localidad y la disponibilidad del Programa Oportunidades en la localidad.
- C. Construcción de 30 **cuadruplas de localidades** pertenecientes a dos de los estados del país donde se concentra un alto porcentaje (21.1%) de los

beneficiarios de Oportunidades en zonas rurales a nivel nacional: Chiapas y Oaxaca. Como se puede apreciar en el cuadro 1, después de Veracruz, que ya había sido incluido en las evaluaciones anteriores, Chiapas y Oaxaca son los estados con las mayores fracciones de beneficiarios en zonas rurales.

A continuación se describen las consideraciones y métodos empleados en el diseño de la muestra de evaluación 2007. Esta información se presenta para cada uno de los tres frentes mencionados.

### **Muestra original de 506 localidades seleccionadas en 1997 y las 151 adicionales de 2003.**

La idea de que la nueva muestra permitiera la comparabilidad con los resultados de las evaluaciones anteriores y la posibilidad de poder realizar un análisis longitudinal sobre los hogares y sus integrantes, eran motivaciones muy importantes para regresar a las localidades anteriores. Por tanto, se decidió que la evaluación de 2007 incorporará la muestra de los hogares previamente visitados en los años de 1998, 1999, 2000 y 2003.

Sin embargo, los costos de la evaluación obligaron a tomar la decisión de incorporar solo una parte de las viviendas. A la luz del análisis de costos de la operación de la evaluación y tratando de maximizar la capacidad inferencial y de comparabilidad de la muestra, se determinó excluir de la muestra de evaluación todas aquellas localidades que tuvieran menos de veinte viviendas entrevistadas en la ENCEL 2003, ya que el costo por entrevista en las localidades pequeñas aumentaba considerablemente.

**Cuadro 1. Distribución de los beneficiarios del Programa Oportunidades por entidad federativa según tipo de localidad**

<b>Entidad</b>	<b>Rurales</b>		<b>Urbanos</b>		<b>Total</b>	
	<b>Frecuencia</b>	<b>Porcentaje</b>	<b>Frecuencia</b>	<b>Porcentaje</b>	<b>Frecuencia</b>	<b>Porcentaje</b>
Aguascalientes	14444	0.4	11517	0.7	25961	0.5
Baja California	6553	0.2	9860	0.6	16413	0.3
Baja California Sur	5960	0.2	4861	0.3	10821	0.2
Campeche	33940	1.0	23237	1.5	57177	1.1
Coahuila de Zaragoza	23831	0.7	14672	0.9	38503	0.8
Colima	7378	0.2	9678	0.6	17056	0.3
Chiapas	403066	11.7	151080	9.6	554146	11.1
Chihuahua	45834	1.3	13760	0.9	59594	1.2
Distrito Federal	1343	0.0	16957	1.1	18300	0.4
Durango	58650	1.7	16761	1.1	75411	1.5
Guanajuato	173020	5.0	60410	3.8	233430	4.7
Guerrero	219253	6.4	100898	6.4	320151	6.4
Hidalgo	162813	4.7	33054	2.1	195867	3.9
Jalisco	87962	2.6	58497	3.7	146459	2.9
México	225791	6.6	152419	9.7	378210	7.6
Michoacán	183764	5.4	88696	5.7	272460	5.4
Morelos	30626	0.9	37757	2.4	68383	1.4
Nayarit	37834	1.1	9199	0.6	47033	0.9
Nuevo León	28288	0.8	17522	1.1	45810	0.9
Oaxaca	323825	9.4	84964	5.4	408789	8.2
Puebla	242270	7.1	148808	9.5	391078	7.8
Querétaro Arteaga	56648	1.7	18487	1.2	75135	1.5
Quintana Roo	27654	0.8	21145	1.3	48799	1.0
San Luis Potosí	146661	4.3	42368	2.7	189029	3.8
Sinaloa	96195	2.8	30663	2.0	126858	2.5
Sonora	36340	1.1	25156	1.6	61496	1.2
Tabasco	114069	3.3	36979	2.4	151048	3.0
Tamaulipas	52628	1.5	31660	2.0	84288	1.7
Tlaxcala	22473	0.7	29815	1.9	52288	1.0
Veracruz	418924	12.2	165078	10.5	584002	11.7
Yucatán	51425	1.5	82158	5.2	133583	2.7
Zacatecas	91092	2.7	21330	1.4	112422	2.2
<b>Total</b>	<b>3430554</b>	<b>100.0</b>	<b>1569446</b>	<b>100.0</b>	<b>5000000</b>	<b>100.0</b>

Como resultado de esta decisión se excluyeron un total de 103 localidades de las 657 (506 y 151) que participaron en la ENCEL 2003. De estas localidades excluidas:

- 37 pertenecían al grupo de las 320 que recibieron el tratamiento en 1998.
- 10 pertenecían al grupo de las 186 que conformaron el grupo control en 1998 y que recibieron el tratamiento en 2000.

- 55 al grupo de las 151 localidades que se incorporaron como controles en 2003

Así, la muestra de evaluación 2007 incorporó 554 localidades de la muestra original de evaluación. La distribución del número de localidades en muestra, de acuerdo a su papel en las evaluaciones previas, aparece en el cuadro 2.

**Cuadro 2. Distribución de las localidades en muestra para la ENCEL 2007 según tipo.**

<i><b>Tipo de Localidad</b></i>	<i><b>Frecuencia</b></i>	<i><b>Porcentaje</b></i>
Tratamiento 1998	283	51.1
Tratamiento 2000	176	31.8
Control 2003	95	17.1
Total	554	100.0

Para considerar estas cifras en el contexto de número de hogares entrevistados como parte de la ENCEL 2003, debemos mencionar que la ENCEL involucró un total de 34,202 hogares, lo que implica que con esta decisión, sólo el 3.3% de los hogares fue excluido de la nueva muestra (ver cuadro 3).

**Cuadro 3. Número de hogares por tipo de localidad según número de hogares en muestra para la ENCEL 2003**

<i><b>Tamaño de muestra ENCEL 2003</b></i>	<i><b>Tipo de localidad</b></i>	<i><b>Hogares</b></i>	<i><b>Porcentaje</b></i>
<i>Localidades con menos de 20 hogares en la ENCEL 2003</i>	Tratamiento 1998	511	45.6
	Tratamiento 2000	149	13.3
	Control 2003	460	41.1
	Total	1120	100.0
<i>Localidades con 20 o más hogares en la ENCEL 2003</i>	Tratamiento 1998	16139	48.8
	Tratamiento 2000	10685	32.3
	Control 2003	6258	18.9
	Total	33082	100.0

## **B. Revisita a una submuestra de hogares de la ENSANUT 2006**

Contemplando las nuevas necesidades de información, derivadas del crecimiento del Programa y de su dispersión, se incorporó en la nueva muestra de evaluación una fracción de los hogares entrevistados en la ENSANUT 2006. Con esto se buscó obtener datos que permitieran conocer el estatus del Programa en la región Norte y Occidente de México.



La ENSANUT es una encuesta probabilística, basada en un diseño muestral polietápico, estratificado y de conglomerados, que permite estimar proporciones mayores o iguales al 8.1% a nivel estatal, con precisión y confianza propias<sup>7</sup>. La metodología muestral de la ENSANUT 2006 se presenta en el Anexo 1.

Los hogares que se visitarán en la muestra de evaluación 2007 y previamente visitados en la ENSANUT 2006, pertenecen a los estados de Durango y Sinaloa, en la región norte y de Aguascalientes y Nayarit en la región Occidente del país. La selección de estos estados se realizó buscando maximizar el número de hogares en el estrato Oportunidades-Rural de la ENSANUT. Esta muestra en total incluye 93 localidades entre las que se distribuyen 1,212 viviendas. El número de viviendas por entidad se muestra en el siguiente cuadro 4.

**Cuadro 4** *Viviendas Oportunidades en la ENSANUT 2007 en las entidades seleccionadas*

<i>Región</i>	<i>Entidad</i>	<i>Viviendas en muestra</i>
Occidente	Aguascalientes	252
	Nayarit	252
Norte	Sinaloa	371
	Durango	337

Con estos tamaños de muestra y considerando un error relativo de 25%, una confianza del 95%, un efecto de diseño de 1.7 y una tasa de no respuesta de 10%, se pueden estimar proporciones mayores o iguales al 18.7% para la región occidente y mayores al 14.1% para la región norte.

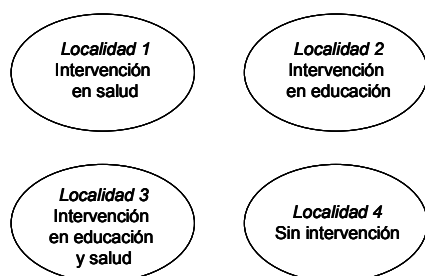
Dado que la Muestra de Evaluación Rural 2007 incorpora al estrato Oportunidades de la ENSANUT en su totalidad para los estados mencionados y que el cálculo de las probabilidades de selección (y por lo tanto de los factores de expansión) de la ENSANUT se realizan por estrato, las probabilidades de selección y los factores de expansión para los casos que se revisitarán en la Muestra de Evaluación Rural 2007 se calcularán usando el mismo procedimiento, con la salvedad de que serán actualizados de acuerdo al número de viviendas reportado por los conteos que se realizan en campo al momento de la entrevista y en su caso por la no respuesta para el levantamiento de la Encuesta de evaluación 2007.

## **D. Cuádruplas de localidades en Chiapas y Oaxaca**

Se decidió extraer la muestra en los estados de Chiapas y Oaxaca, dado que en estos estados se concentra una importante fracción del total de los beneficiarios del Programa Oportunidades.

Considerando que al tener cuatro localidades similares se pueden evaluar los efectos marginales y conjuntos de intervenciones en las dimensiones de salud y educación, ver figura 1, se decidió agrupar la muestra de estos estados en cuádruplas de localidades.

*Figura 1. Escenarios de intervención*



La selección de 30 cuádruplas implicó la selección de 120 localidades. En cada estado se seleccionaron 60 localidades distribuidas en 15 cuádruplas para cada estado.

El marco de localidades que podían participar en la selección, incluyó todas las localidades rurales de Chiapas y Oaxaca incorporadas al Programa de Desarrollo Humano Oportunidades y que contara con al menos 20 hogares.

Para cada estado se construyeron 6 estratos correspondientes a las categorías cruzadas de las variables:

- Tamaño de localidad: Según su distribución en terciles de población como: pequeñas, medianas y grandes.
- Tipo de proveedor de salud donde los beneficiarios de localidad cumplen con sus responsabilidades: IMSS-OPORTUNIDADES y SSA.

En el cuadro 5 se presenta la distribución de las localidades y su población, por estrato y entidad federativa.

**Cuadro 5. Distribución de las localidades y su población, por estrato y entidad federativa**

<b>Estrato</b>	<b>Chiapas</b>		<b>Oaxaca</b>		<b>Total</b>	
	<b>Localidades</b>	<b>Población</b>	<b>Localidades</b>	<b>Población</b>	<b>Localidades</b>	<b>Población</b>
Pequeña-SSA	869	122752	768	96662	768	96662
Pequeña-IO	537	76618	546	69493	546	69493
Mediana-SSA	779	228157	708	189237	708	189237
Mediana-IO	636	187544	612	169112	612	169112
Grande-SSA	644	531597	720	617686	720	617686
Grande-IO	767	683308	598	508844	598	508844
Total	4232	1829976	3952	1651034	3952	1651034

El número de cuádruplas asignadas a cada estrato, se realizó proporcionalmente al tamaño del estrato dentro de la entidad. Esto dio como resultado la siguiente asignación:

**Cuadro 6 Asignación del número de cuartetos por estrato**

Estado	Estrato	Cuartetos
Chiapas	Pequeña-SSA	3
	Pequeña-IO	2
	Mediana-SSA	3
	Mediana-IO	2
	Grande-SSA	2
	Grande-IO	3
	Total	15
Oaxaca	Pequeña-SSA	3
	Pequeña-IO	2
	Mediana-SSA	3
	Mediana-IO	2
	Grande-SSA	3
	Grande-IO	2
	Total	15

Para la construcción de cada cuádrupla se buscó que la distancia promedio entre las 4 localidades (en total 6 distancias) se minimizara. La definición de distancia entre localidades empleada fue la distancia de Mahalanobis, donde la matriz de Varianza-Covarianza que se utilizó en los cálculos de la distancia corresponde a la de todas las localidades con posibilidades de ser seleccionadas en el estado. La decisión de usar la distancia de Mahalanobis resulta de la necesidad de incorporar variables que se

encuentran en diferentes unidades de medición ya que esta es invariante bajo cambios de escala. En el anexo 2 se describe el cálculo de la distancia de Mahalanobis.

Las variables consideradas para el cálculo de las distancias se enlistan en el cuadro 7.

***Cuadro 7. Variables consideradas para la construcción de las cuádruplas en los estados de Chiapas y Oaxaca***

---

Distancia a la localidad en donde se ubica el centro de salud (IMSS-Oportunidades) más cercano

Distancia a la localidad en donde se ubica el centro de salud (SSA) más cercano

Distancia a la localidad en donde se ubica el centro de educación primaria más cercano

Distancia a la localidad en donde se ubica el centro de educación secundaria más cercano

Grado promedio de escolaridad

% Población de 15 a 24 años que asiste a la escuela

% Población de 15 años y más sin escolaridad

% Población de 15 años y más con educación básica incompleta

% Población de 15 años y más con educación posbásica

% Población de 5 años y más que habla alguna lengua indígena

% Población de 5 años y más que habla alguna lengua indígena y no habla español

Total de hogares

% Hogares con jefatura femenina

Promedio de ocupantes por cuarto en viviendas particulares habitadas

% Viviendas particulares habitadas con piso de tierra

% Viviendas particulares habitadas que disponen de excusado o sanitario

% Viviendas particulares habitadas que no disponen de agua entubada de la red pública

% Viviendas particulares habitadas que no disponen de drenaje

% Viviendas particulares habitadas que no disponen de agua entubada de la red pública

% Viviendas particulares habitadas sin ningún bien

% Viviendas particulares habitadas que disponen de televisión

% Viviendas particulares habitadas que disponen de refrigerador

% Viviendas particulares habitadas que disponen de lavadora

% Viviendas particulares habitadas que disponen de computadora

% Población de 0 a 4 años

% Población de 6 a 11 años

% Población de 12 a 14 años

% Población de 15 a 59 años

% Población de 65 años y más

Promedio de hijos nacidos vivos

% Población sin derechohabencia a servicios de salud

% Población derechohabiente del IMSS

% Población derechohabiente del ISSSTE

% Población derechohabiente por el Seguro Popular

% Población de 15 años y más analfabeta

% Población de 6 a 11 años que no asiste a la escuela

% Población de 12 a 14 años que no asiste a la escuela

Como criterios de restricción para la conformación de las cuádruplas se utilizaron los siguientes:

1. Una clínica no puede ser el proveedor de dos localidades de la cuádrupla

2. Las cuádruplas pertenecientes a diferentes estratos no pueden compartir clínica

### ***Asignación de cuestionario completo y corto entre los hogares de la muestra de Evaluación 2007***

Adicionalmente, por razones operativas y presupuestales, se diseñó un instrumento de hogar en dos versiones: una completa y una corta. De acuerdo a los tiempos disponibles para el levantamiento, se estimó el número total de entrevistas de hogar completas que podrían realizarse en alrededor de 28 mil hogares.

Para asignar a los hogares la versión completa, se consideraron dos criterios principales: garantizar que los hogares con seguimiento desde 1998 (localidades tratamiento 1998 y 2000) tuvieran un cuestionario completo y tener un número mínimo de cuestionarios completos por localidad.

Estos criterios de decisión para seleccionar los hogares en que se aplicaría la versión completa del cuestionario de hogar se ajustaron de acuerdo al tipo de localidad: para las localidades que participaron como tratamiento o control en la muestra original de evaluación y para las localidades que conformaron los nuevos controles para la evaluación incorporadas en 2003.

En las localidades de la muestra original los criterios de selección fueron los siguientes:

- A todos los hogares identificados como de la muestra original (es decir, que participaron en la evaluación desde 1998) y que permanecían en la muestra en el año 2003, se les aplicó el cuestionario completo con probabilidad uno.
- De los hogares pertenecientes a estas localidades que no participaron en la muestra de 1998, pero que fueron incorporados a las encuestas de evaluación antes de 2007, se seleccionaron de manera aleatoria:

$$20 * \frac{t_{2003}}{e_{2007}}$$

Donde:

$t_{2003}$  = Número de hogares nuevos encontrados en el momento de la visita a la localidad entre 1998 y 2003

$e_{2007}$  = Total esperado de hogares en la localidad en 2007

- De los hogares identificados como nuevos al momento de la visita en 2007

$$20 * \frac{e_{2007}}{t_{2007}}$$

donde:

$e_{2007}$  = Número esperado de hogares nuevos encontrados en el momento de la visita a la localidad en 2007

$t_{2007}$  = Total esperado de hogares en la localidad en 2007

En las localidades que conformaron el grupo de nuevos controles el criterio de selección donde se aplicó el cuestionario completo fue:

- Selección aleatoria de hasta 26 hogares en la localidad.

En el Cuadro 8 se muestra la distribución del número de cuestionarios largos y cortos según el tipo de localidad.

**Cuadro 8. Número de cuestionarios por longitud y tipo de localidad**

<i>Longitud</i>	<i>Tipo de localidad</i>	<i>Esperados</i>	<i>%</i>
<i>Corto</i>	Tratamiento 1998	2972	16.4
	Tratamiento 2000	2255	12.4
	Control 2003	4022	22.1
	Chiapas-Oaxaca	8910	49.1
	Total	18159	100.0
<i>Largo</i>	Tratamiento 1998	13167	48.9
	Tratamiento 2000	8430	31.3
	Control 2003	2236	8.3
	Chiapas-Oaxaca	3120	11.6
	Total	26953	100.0
<i>Total</i>	Tratamiento 1998	16139	35.8
	Tratamiento 2000	10685	23.7
	Control 2003	6258	13.9
	Chiapas-Oaxaca	12030	26.7
	Total	45112	100.0

### **Asignación de cuestionario completo y corto entre adolescentes de la muestra de Evaluación 2007**

En la misma lógica de los aspectos logísticos y de costos, .y considerando que el tamaño de muestra esperado excedía el necesario para los análisis propuestos en variables de comportamientos de riesgo, se implementó una versión corta de la encuesta de jóvenes, que se aplicó sistemáticamente a 1 de cada dos jóvenes.

## ***Muestra para la evaluación operativa***

### **Centro de Atención y Registro (CAR)**

Selección con probabilidad proporcional a la cantidad de localidades de la ENCEL atendidas por los CAR (cabezas de las zonas de atención). Se seleccionó un CAR por estado. En el marco muestral se incluyeron sólo aquellos CAR que tuvieran las siguientes características: atender a más de tres localidades de la muestra ENCEL, disponer de la dirección para localizarlos y tener tanto MAC como MAS en operación durante octubre (en los estados que no tendrían MAS no se usó este último criterio).

### **3. Mesas de Atención a Comités (MAC) y Mesas de Atención y Servicio-Mesas de entrega de apoyos (MAS-MAP)**

Estas unidades se seleccionaron en dos etapas:

1. Selección de dos zonas de atención por estado (una de las cuales fue la misma que la seleccionada para aplicar los instrumentos del CAR) con probabilidad proporcional a la cantidad de localidades de la ENCEL atendidas por cada zona de atención. Se seleccionó una zona de atención para sustitución en caso de que una de las primeras no tuviera MAS o MAC operando durante octubre.
2. Se visitó el CAR de cada zona de atención seleccionada y se solicitó listas de las MAS y MAC que operarían durante los cinco días subsiguientes. Con base en esas listas se seleccionaron dos MAS y una MAC de manera aleatoria.

## **Anexo 1 Diseño Muestral de la ENSANUT 2006**

La Encuesta Nacional de Salud y Nutrición 2006 (ENSANUT 2006) fue diseñada para recabar información relacionada con el estado de salud y nutrición de la población mexicana, la calidad y respuesta de los servicios de salud y el gasto en salud que realizan los hogares mexicanos. Entre otros aspectos que se pudieron explorar utilizando la información derivada de la ENSANUT 2005, se pueden mencionar:

- Estado nutricional de los niños y mujeres en México.
- Estado de Salud de la población mexicana y la prevalencia de algunos padecimientos crónicos e infecciosos.
- Percepción de la población sobre la calidad del sistema de salud en el estado.
- Percepción de la población sobre la respuesta del sistema de salud en el estado.
- Características sociodemográficas de los hogares que incurren en un gasto catastrófico como consecuencia de la afectación de la salud de sus miembros.
- Estudiar el impacto en salud del Programa Oportunidades.

### **Cobertura Geográfica y de Unidades**

La información obtenida a través de la ENSANUT 2006 permitió el cálculo de indicadores estatales con representatividad de las áreas urbanas y rurales de cada entidad federativa.

Por razones de tipo operativo no se incluyeron las viviendas colectivas, como son: instalaciones militares, cárceles, conventos, hoteles, asilos, etc.

### **Tamaño de muestra**

Para cumplir con los objetivos de la ENSANUT 2006, se consideró que el indicador relevante a estimar de menor magnitud era un resultado que se esperaba se presentaría en una proporción de aproximadamente 8.1%. Además, contemplando que la encuesta debería permitir obtener estimaciones estatales con un error relativo máximo de 25%, una confianza del 95%, una tasa de no respuesta del 20% y un efecto de diseño de 1.7; se determinó un tamaño de muestra de aproximadamente 1,476 hogares usando la siguiente fórmula:



$$n = \frac{Z_{\alpha/2}^2 (1-P) deff}{r^2 P TR}$$

donde:

$n$  Tamaño de muestra

$P$  Proporción a estimar

$Z_{\alpha/2}^2$  Cuantíl de una distribución normal asociado a un nivel de confianza deseado  $\alpha$

$r$  Error relativo máximo que se está dispuesto a aceptar

$deff$  Efecto de diseño, que es la pérdida o ganancia en la eficiencia de diseño, por tratarse de un diseño complejo

$TR$  Tasa de respuesta esperada

El tamaño de muestra a nivel nacional fue de 47,232 viviendas y permitió estimar proporciones mayores o iguales a 0.3% con la misma precisión y confianza que las estimaciones estatales.

### **Unidades de análisis**

Las unidades de análisis definidas para la encuesta son las siguientes:

- Hogar, es el conjunto de personas, relacionadas por algún parentesco o no, que habitualmente duermen en una vivienda bajo el mismo techo, beneficiándose de un ingreso común, aportado por uno o más de los miembros del hogar.
- Utilizadores de servicios de salud, son las personas del hogar que buscaron o recibieron atención dentro de los seis meses anteriores a la fecha de la encuesta por enfermedad, lesión, accidente, vacunación, revisión del niño sano, planificación familiar, cuidado del embarazo, prueba del Papanicolau, rehabilitación, servicio dental, obtención de certificado médico, atención de las brigadas de salud, pláticas de educación para la salud, Programa de control de diabetes o hipertensión arterial.
- Niños, son las personas del hogar entre cero y nueve años de edad.
- Adolescentes, son las personas del hogar en el grupo de edad de 10 a 19 años de edad inclusive.
- Adultos, son las personas del hogar mayores de 19 años.

## **Estratificación**

El diseño muestral de la ENSANUT 2006 es probabilístico, polietápico, estratificado y por conglomerados. La estratificación de las unidades muestrales se realizará considerando las variables: tamaño de localidad y estatus de incorporación al Programa Oportunidades. Los estratos considerados por la ENSANUT 2006 se muestran en las siguientes tablas.

<b>Estrato</b>	<b>Descripción</b>
I. Ciudad o Área Metropolitana	<ul style="list-style-type: none"><li>• Ciudades y áreas metropolitanas</li><li>• Resto de las ciudades de 100,000 y más habitantes y/o capitales de estado.</li></ul>
II. Complemento Urbano	<ul style="list-style-type: none"><li>• Localidades de 2,500 a 99,999 habitantes.</li></ul>
III. Rural	<ul style="list-style-type: none"><li>• Localidades con menos de 2,500 habitantes.</li></ul>

<b>Estrato</b>	<b>Descripción</b>
I. Oportunidades	<ul style="list-style-type: none"><li>• Localidades incorporadas al Programa Oportunidades</li></ul>
II. No Oportunidades	<ul style="list-style-type: none"><li>• Localidades incorporadas al Programa Oportunidades</li></ul>

El tamaño de muestra para cada estrato se asignará de manera proporcional al número de habitantes dentro del estrato.

## **Esquema de Selección**

Por cuestiones de tipo operativo, el esquema de selección depende del estrato que se esté observando. En los estratos I y II el esquema de selección es idéntico, mientras que para el estrato III existe una variación.

### *Estratos I y II*

- En la primera etapa se seleccionarán  $k$  AGEB's con probabilidad proporcional al tamaño (PPT) definido por el número de viviendas particulares habitadas en el AGEB.
- Para cada uno de los AGEB's seleccionados se seleccionaron 6 manzanas con probabilidad proporcional al número de viviendas en la manzana.
- Posteriormente en cada una de las manzanas seleccionadas se realizará una selección de 6 viviendas usando muestreo sistemático con arranque aleatorio.
- En cada una de las viviendas seleccionadas y siempre que la composición del hogar lo permita se seleccionará mediante el uso de muestreo aleatorio simple a: un adulto, un adolescente, un niño y un utilizador de servicios, según las definiciones dadas anteriormente. La encuesta se aplicará en todos los hogares de la vivienda.
- A cada uno de los individuos seleccionados se les aplicará la entrevista siguiendo el cuestionario correspondiente.

La constante  $k$  estará en función del tamaño de muestra asignado al estrato.

### *Estrato III*

En el caso de las localidades que conformaron el estrato III el esquema de selección como sigue:

- Las unidades primarias de muestreo eran AGEB's que contenían localidades que conformaban el estrato III. La probabilidad de seleccionar un AGEB era proporcional al número de viviendas particulares habitadas en el AGEB. Se seleccionaron tantos AGEB's como correspondía al tamaño del estrato en la entidad.
- La unidad secundaria de muestreo (USM) estaba constituida por las localidades o grupos de localidades agrupadas en un mínimo de 120 viviendas. Estas USM se seleccionaron con probabilidad proporcional al número de viviendas particulares habitadas que contengan.
- Para cada localidad o grupo de localidades seleccionadas se extrajo una muestra aleatoria sistemática de 3 segmentos de doce viviendas cada uno.

### *Nota técnica sobre el diseño muestral*

- En cada una de las viviendas de los segmentos seleccionados se aplicará la encuesta en todos los hogares que contenga. Siempre que la composición del hogar lo permitió, se seleccionó: un adulto, un adolescente, un niño y un utilizador de servicios, según las definiciones dadas anteriormente. Esta selección se realizó usando muestreo aleatorio simple.

## Anexo 2 Distancia de Mahalanobis

La **distancia euclidiana** entre dos vectores  $x = (x_1, x_2 \dots x_p)^T$  y  $y = (y_1, y_2 \dots y_p)^T$  se define como:

$$D_E(x) = \sqrt{\sum_{i=1}^p (x_i - y_i)^2}$$

Como puede verse, esta distancia toma en cuenta sólo la diferencia entre coordenadas, lo cual no produce problemas cuando los vectores son determinísticos, pero cuando se trabaja con vectores aleatorios no sólo se debe tomar en cuenta la diferencia escalar, si no también el efecto que produce un vector sobre otro (la relación de dependencia lineal, que se mide por medio de la matriz de varianza-covarianza).

Debido a esta pérdida de información, se han desarrollado técnicas que permiten incorporar dicho efecto.

La **distancia de Mahalanobis** fue introducida por P. C. Mahalanobis en 1936. Dicha distancia se basa en las correlaciones entre variables con diferentes patrones que pueden ser identificados y analizados. Difiere de la distancia Euclidiana, pues toma en cuenta las correlaciones entre los elementos del conjunto de datos, y es invariante de la escala que se esté utilizando (no depende de las unidades de medición).

Formalmente la distancia de Mahalanobis de un grupo de valores con media  $\mu = (\mu_1, \mu_2 \dots \mu_p)^T$  y matriz de varianza-covarianza  $\Sigma$  para un vector multivariado  $x = (x_1, x_2 \dots x_p)^T$  se define como:

$$D_M(x) = \sqrt{(x - \mu)^T \Sigma^{-1} (x - \mu)}$$

es decir, permite tener una medida de discrepancia entre la observación y la media del vector.

Esta definición puede extenderse para medir la distancia que existe entre dos vectores aleatorios  $\bar{x}$  y  $\bar{y}$  e idénticamente distribuidos con matriz de varianza-covarianza  $\Sigma$  :

$$d(\bar{x}, \bar{y}) = \sqrt{(\bar{x} - \bar{y})^T \Sigma^{-1} (\bar{x} - \bar{y})}$$

Si la matriz de varianza-covarianza es la matriz identidad, la distancia de Mahalanobis se reduce a la distancia euclidiana, esto ocurre en el caso en el que las variables son independientes y con varianza igual a uno. Si la matriz de varianza-covarianza es diagonal (esto implica independencia entre las variables), la medida resultante es llamada **distancia Euclidiana normalizada** y tiene la siguiente forma:

$$d(\bar{x}, \bar{y}) = \sqrt{\sum_{i=1}^p \frac{(x_i - y_i)^2}{\sigma_i^2}}$$

En donde  $\sigma_i$  es la desviación estándar de los  $x_i$  sobre el conjunto de muestra.

## **Referencias**

---

<sup>1</sup> Acuerdo por el que se emiten las Reglas de Operación del Programa de Desarrollo Humano Oportunidades, para el ejercicio fiscal 2007. DIARIO OFICIAL (Cuarta Sección-Vespertina) 21, Miércoles 28 de febrero de 2007

<sup>2</sup> Behrman, J., and P. E. Todd. 1999a. Randomness in the Experimental Samples of PROGRESA (Education, Health, and Nutrition Program). March. International Food Policy Research Institute, Washington, D.C.

<sup>3</sup> Instituto Nacional de Salud Pública, Coordinación Nacional del Programa de Desarrollo Humano Oportunidades 2006. Nota metodológica general rural. Enero. Disponible en: <http://evaluacion.oportunidades.gob.mx>.

<sup>4</sup> Díaz JJ, Sudhanshu Handa S. An assessment of propensity score matching as a non experimental impact estimator: evidence from Mexico's PROGRESA program. J Hum Res, Volume 41, Number 2, Spring 2006, pp. 319-345(27).

<sup>5</sup> Behrman, J., and P. E. Todd. 1999b. A Preliminary Evaluation of the Sample Sizes used for the Evaluation of the Education, Health, and Nutrition Program (PROGRESA) of Mexico. January. International Food Policy Research Institute, Washington, D.C.

<sup>6</sup> Todd P. Technical note on using matching estimators to evaluate the Oportunidades program for six year follow-up Evaluation of Oportunidades in rural areas. Informe Técnico; 2004.

<sup>7</sup> Olaiz-Fernández G, Rivera-Dommarco J, Shamah-Levy T, Rojas R, Villalpando-Hernández S, Hernández-Avila M, Sepúlveda-Amor J. Encuesta Nacional de Salud y Nutrición 2006. Cuernavaca, México: Instituto Nacional de Salud Pública, 2006.