

Tabla de contenido

Introducción	1
Algunos datos sobre la primera fase del proyecto	1
Versiones del cuestionario	2
Consideraciones generales	4
Uso de informante sustituto o <i>proxy</i>	4
Denominación de las variables del cónyuge	5
Análisis de parejas	5
Notas sobre el minimental	5
Algunas variables de edad	5
Ponderación de los datos	6
Valores ausentes (<i>missing values</i>)	6
Archivos de datos y documentos relacionados	7
Notas generales sobre la estructura de los <i>rosters</i>	8
El <i>roster</i> del hogar: variables disponibles	8
El <i>roster</i> de hijos: variables disponibles	9
El <i>roster</i> de hermanos: variables disponibles	10
El <i>roster</i> de padres: variables disponibles	10
El <i>roster</i> de otras personas: variables disponibles	10
Variables de posición y variables de duplicación	11
Utilización de la información contenida en los <i>rosters</i>	12
Variables calculadas	13
Errores y limitaciones identificados en los datos	14
Libro de códigos	15

Introducción

El proyecto de investigación Condiciones de salud de los adultos mayores en Puerto Rico (Puerto Rican Elderly: Health Conditions, PREHCO, por sus siglas en inglés) es un estudio sufragado con fondos federales del *National Institute on Aging* / NIA (*Grant number* RO1 AG1622090-01) mediante un subcontrato entre la Universidad de Wisconsin-Madison y la Universidad de Puerto Rico. Este proyecto cuenta también con el apoyo técnico de la Organización Panamericana de la Salud (Washington D.C.) y ha recibido ayuda económica de la Asamblea Legislativa de Puerto Rico y de la Procuradora de las Mujeres de Puerto Rico.

En su primera fase, entre mayo de 2002 y mayo de 2003, se trató de un estudio transversal de la población no institucionalizada de 60 años o más en Puerto Rico, que sirve como base de datos para el estudio de este grupo de edad y cuyo libro de códigos se explica en este documento.

Se empleó un muestreo estratificado por conglomerados, polietápico, utilizando como marco muestral el censo de Población y Viviendas del año 2000 del Bureau del Censo de Estados Unidos.

Para el proyecto se diseñó un cuestionario que incluyó secciones sobre: condiciones de salud, incapacidad física y mental, funcionalidad, uso de medicinas, necesidades de salud y servicios sociales, acceso a y uso de servicios de salud, maltrato, migración, condiciones de la vivienda, patrones de ayuda de la familia, comunidad y agencias públicas y gubernamentales, y otras.

Algunos datos sobre la primera fase del Proyecto PREHCO

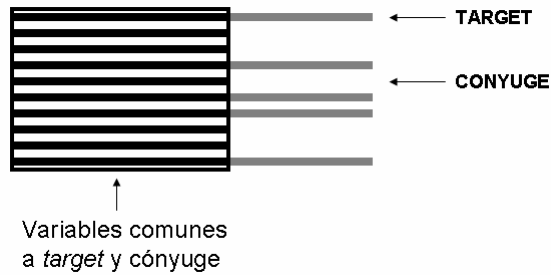
En la primera fase del proyecto PREHCO se entrevistaron 4,291 personas de 60 años o más y 1,442 cónyuges, independientemente de su edad. Con los datos obtenidos se creó una base de datos que recoge, en un mismo registro, la información de cada uno de los 4,291 participantes y la de sus cónyuges, cuando estos fueron entrevistados.

Figura 1. Estructura de la base de datos de PREHCO 1 – PREHCO1.SAV (n=4,291)



Aunque en este documento se presenta la estructura de esta base de datos (Figura 1), ya que el entrevistado y su cónyuge contestaron las mismas preguntas en muchas de las secciones del cuestionario, es posible construir, para aumentar el tamaño de la muestra, una base de datos que tenga a los entrevistados y a sus cónyuges (de 60 años o más) en registros independientes. Esta nueva base de datos tendría un total de 5,333 registros (Figura 2) y, como se mencionó incluiría sólo las variables comunes a ambos tipos de entrevista y a aquellos cónyuges de 60 años o más. Existe un procedimiento de SPSS (*syntax file*) que permite crear automáticamente esta base de datos desde la de 4,291 casos (ver sección ‘Archivos de datos y documentos relacionados’ más adelante).

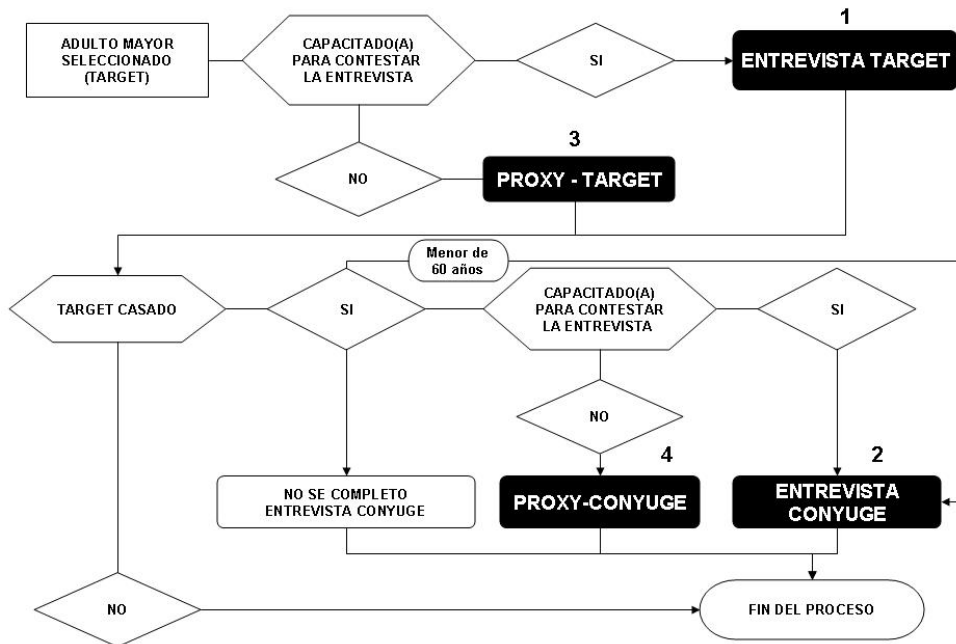
Figura 2. Estructura de la base de datos de PREHCO 1 (n=5,333)



Versiones del cuestionario

El cuestionario principal cambia de acuerdo a varios factores. Principalmente, la capacidad del entrevistado para contestar, la existencia de un cónyuge y la capacidad de éste (ver Figura 3).

Figura 3. Flujograma para determinar el tipo de entrevista



En resumen, hay cuatro variantes principales del cuestionario (Tabla 1), estos son: (1) cuestionario del entrevistado principal o *target*, (2) cuestionario del cónyuge, (3) cuestionario del informante sustituto o *proxy*¹ del entrevistado y (4) cuestionario del informante sustituto o *proxy* del cónyuge.

1. Cuestionario del entrevistado principal o *target*: el cuestionario principal es el que se le administra al adulto mayor que está capacitado para contestarlo por sí mismo. Esta versión incluye todas las secciones del cuestionario.

¹ Criterios para cualificar al *proxy*: conocer al entrevistado por al menos un año y tener 18 años o más.

Tabla 1. Versiones del cuestionario.

1 Target	2 Cónyuge	3 Proxy-target	4 Proxy-cónyuge
Sección A: Datos personales	Completa	Completa	Completa
Sección B: Minimalista	Completa (si es mayor de 60 años)	Completa (si el cónyuge es el proxy y es mayor de 60 años) Reducida (si el proxy es otra persona y es mayor de 60 años) o preguntas	No (si el proxy es el target) Reducida (si el proxy es otra persona y es mayor de 60 años y no la ha hecho ya como proxy del target)
Sección C: Roster del hogar		Completa	
Sección D: Datos personales	D1, D15 a D20, D31 a D33	D1 a D25, D27, D28	D1, D15 a D20
Sección E: Escaleras SES	Completa		
Sección G: Estado de salud	Completa	Sólo G4-G67, G109, G114ck, G114-G131, G133-G139, G141, G143-G144, G147-G156, G161-G170	Sólo G4-G67, G109, G114ck, G114-G131, G133-G139, G141, G143-G144, G147-G156, G161-G170
Sección H: Condiciones de la niñez	Completa		
Sección I: Estado funcional	Completa	Sección Z	Sección Z
Sección J: Medicinas	Completa	Completa	Completa
Sección K: Acceso y utilización	Completa	Completa	Completa
Sección L: Datos de familiares	D11-D14	Completa	
Sección M: Transferencias			
Sección N: Historia laboral	N1-N53, N74-N76	N1-N6, N18-N26, N49-N70	N1-N53, N74-N76
Sección O: Migración	Completa		
Sección P: Vivienda		Completa	
Sección Q: Propiedades		Completa	
Sección R: Raza	Completa		
Sección S: Maltrato	Completa		
Sección T: Sexualidad			
Sección U: Medidas	Completa	Completa (si es posible)	Completa (si es posible)
Sección V: Contactos			
Sección X: Percepción del entrevistador	Completa	Completa	Completa

2. Cuestionario del cónyuge: el cuestionario del cónyuge es el que se le administra al cónyuge del adulto mayor que está capacitado para contestarlo por sí mismo. Este cuestionario no incluye algunas secciones del cuestionario principal e igualmente excluye algunas preguntas de algunas de las secciones incluidas.
3. Cuestionario del informante sustituto o proxy del entrevistado: este cuestionario es contestado por un informante sustituto, en los casos en los que el adulto mayor no está capacitado para hacerlo por sí mismo. La determinación sobre la necesidad de un informante sustituto se realiza con dos criterios: (1) la puntuación obtenida por el entrevistado en el minimal y (2) la opinión del entrevistador en algunos casos². El informante sustituto provee información relativa al adulto mayor, excluyendo algunas secciones o preguntas individuales que implican opinión o estado de ánimo. En algunas ocasiones el cónyuge actúa como informante sustituto del entrevistado y en otras, en las que no existe un cónyuge o éste no está capacitado, el informante sustituto es otra persona.
4. Cuestionario del informante sustituto o proxy del cónyuge: este cuestionario es contestado por un informante sustituto, en los casos en los que el cónyuge del adulto mayor no esté capacitado para hacerlo por sí mismo (sólo en el caso de cónyuges de 60 años o más). La determinación sobre la necesidad de un informante sustituto se realizó con los criterios ya expresados. El informante sustituto provee información relativa al cónyuge del adulto mayor, excluyendo algunas secciones o preguntas individuales que implican opinión o estado de ánimo. En algunas ocasiones el entrevistado principal actúa como informante sustituto del cónyuge y en otras el informante sustituto es otra persona.

Consideraciones generales

A continuación se presentan varias consideraciones generales que deben ser tomadas en cuenta antes de utilizar la base de datos.

Uso de informante sustituto o proxy

Un 13.5% de los *targets* fue entrevistado a través de un informante sustituto o proxy. Existen dos variables que pueden utilizarse para filtrar, si fuera necesario, esas entrevistas: CODIGO y PROXY. La variable CODIGO presenta los valores 25 (para entrevista regular) y 26 (para entrevista con *proxy*). La variable PROXY tiene los valores 0 (cuando no existe proxy), 2 (cuando el *proxy* es el cónyuge) y 3 (cuando el *proxy* es otra persona). Algunas características del proxy se recogen en las variables: PRXYAGE (edad), PRXYSEX (sexo), PRXYRL_1 (relación general con el entrevistado y PRXYRL_2 (relación específica)³.

² Se estableció como criterio que una puntuación total en el minimal de 11 o más indicaba que el entrevistado estaba capacitado para contestar por sí mismo. Una puntuación menor de 9 indicaba que era necesario utilizar un informante sustituto. Para aquellos casos en los que la puntuación fue de 9 ó 10 el entrevistador era quien determinaba si el entrevistado estaba capacitado para contestar.

³ Para más información sobre estas variables consulte la página C-1 del cuestionario.

Denominación de las variables del cónyuge

Las variables correspondientes al cónyuge del entrevistado conservan la misma nomenclatura de las de éste, pero precedida de la letra Y. Así la variable sobre el estado general de salud se denomina G1, para el entrevistado principal, e YG1 para su cónyuge.

Análisis de parejas

Con el propósito de realizar análisis de pareja pueden utilizarse dos variables de referencia. La variable SPOUSE=1 indica que existía una pareja al momento de realizar la entrevista. Sin embargo la variable CONYUGE=1 resulta más útil, ya indica que se completó la entrevista de ese cónyuge.

Notas sobre el minimal

Como se indica en la Tabla 1, la sección B del cuestionario contiene un examen mental breve o minimal. El minimal es utilizado asimismo como filtro para determinar la necesidad de un *proxy*, siguiendo los criterios expresados anteriormente. En algunos casos en los que el impedimento del entrevistado para responder por sí mismo el cuestionario era obvio, el minimal no se realizó y se seleccionó un *proxy* directamente. Estos casos pueden identificarse utilizando la variable MINRVAL (MINRVAL=1). Los participantes que presentaron un impedimento obvio no tienen puntuación (FINALSCR) en el minimal.

En algunos casos el entrevistado tuvo un impedimento físico (BRCAP=1) que le impidió realizar algunas de las pruebas del minimal que requieren destrezas como dibujar o doblar. En esos casos, la puntuación total en el minimal se obtuvo extrapolando la puntuación obtenida en las preguntas que se contestaron a la totalidad de las preguntas y pruebas incluidas en el minimal.

Algunas variables de edad

En algunas partes de la entrevista se recoge la edad en la que se produjo un suceso determinado. Inicialmente, al entrevistado se le pregunta por la edad directamente. Si desconoce la edad se le pregunta por el mes y el año. Entonces se calcula una edad final para ese suceso que indica la edad informada por el entrevistado o, en su defecto, la edad calculada con el año y el mes. Así pues para cada edad de cada suceso existe una serie de cuatro variables. Las variables de todas estas series tienen el formato: *_E (edad informada por el sujeto), *_A (año), *_M (mes) y *EDAD (edad resumen: informada o calculada). Esto quiere decir que para analizar la edad en la que se produjo un suceso determinado, debe utilizarse la última variable que es la que tiene la información más completa. En el siguiente ejemplo se presentan las variables que recogen la edad a la que terminó la primera unión del entrevistado:

d10_e (edad informada)

d10_a (año)

d10_m (mes)

d10edad (edad resumen: informada o calculada)

Ponderación de los datos

El factor de ponderación final correspondiente a cada registro, fue ubicado en la variable **FAC_T** de la base de datos (n=4,291). Para su cálculo se realizó una postestratificación por sexo y grupo de edad, considerando seis grupos de edad: 60-64, 65-69, 70-74, 75-79, 80-84 y 85 años o más. Si se construye la base de datos (n=5,333) con entrevistados y cónyuges en registros independientes (ver ‘Algunos datos sobre la primera fase del Proyecto PREHCO’ en la página 1) deberá utilizarse para la ponderación otro factor de ponderación. Este factor de ponderación está disponible en el archivo FACTORTC.SAV (ver sección ‘Archivos de datos y documentos relacionados’ más adelante).

Valores ausentes (*missing values*)

En toda la base de datos se han utilizado unos códigos comunes para los valores ausentes (*missing values*). Estos valores se presentan en la siguiente tabla:

Tabla 2. Códigos y etiquetas de los valores nulos (*missing values*)

-1	No sabe	Entrevistado(a) indicó no saber la respuesta a la pregunta.
-2	No contesta	Entrevistado(a) indicó que no contestaría la pregunta.
-3	Indeterminado, dato no disponible	Datos no recopilados por algún tipo de error.
-5	No aplica a <i>proxy-target</i>	La pregunta no corresponde al cuestionario del <i>proxy</i> del <i>target</i>
-6	No aplica a <i>proxy-cónyuge</i>	La pregunta no corresponde al cuestionario del <i>proxy</i> del cónyuge
-7	No hay cónyuge	Entrevistado(a) no tiene cónyuge o pareja por lo cual no hay entrevista de cónyuge.
-8	No aplica cónyuge menor de 60 años	La pregunta no corresponde a los cónyuges menores de 60 años de edad.
-9	No aplica minimal	Cuando entrevistado(a) contesta su minimal y pasa la prueba, las variables del minimal de cónyuge y del minimal para el posible informante sustituto o <i>proxy</i> no tienen que ser contestadas.
-10	No aplica, impedimento obvio	Entrevistado(a) tiene impedimento obvio (determinado por el entrevistador antes de iniciar la prueba del minimal) y no contesta las preguntas del minimal.
-11	No aplica, impedimento físico	Entrevistado(a) tiene impedimento físico y no contesta las preguntas del minimal que requieren movimiento de la mano como escribir o dibujar.
-12	No aplica, no hay entrevista de cónyuge	Entrevistado(a) principal tiene cónyuge o pareja pero no se hizo la entrevista del cónyuge. La variable que identifica el resultado final de la entrevista de cónyuge ([ycodigo]) tiene un código que indica la razón por la cual no fue completado el cuestionario.
-13	No aplica a <i>target</i> ni a cónyuge	Este código sólo aplica a la sección Z la cual es contestada por los <i>proxy-target</i> o <i>proxy-cónyuge</i> .

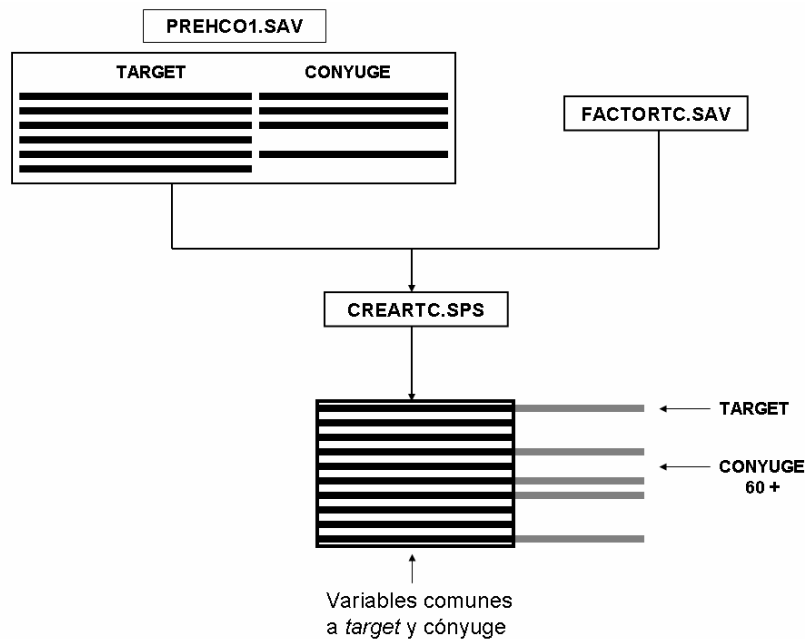
Archivos de datos y documentos relacionados

Los siguientes archivos están disponibles:

1. **PREHCO1.SAV** (aproximadamente 10Mb)
Contiene todos los datos de la primera fase de PREHCO en formato SPSS. Los datos están almacenados en 2,167 variables.
2. **creartc.sps** (19 Kb)
Este archivo de procedimiento (*syntax file*) para SPSS crea automáticamente un archivo de 5,333 registros que tiene, en registros independientes, a los entrevistados y sus cónyuges (de 60 años o más). Ya que ambas entrevistas comparten casi 700 variables, este procedimiento permite trabajar con esas variables comunes (ver la explicación sobre versiones del cuestionario más arriba) con tamaño de muestra mayor. Para ponderar el archivo resultante (PREHCOTC.SAV) debe utilizarse el factor FACTC.
3. **factortc.sav** (72 Kb)
Este archivo contiene el factor de ponderación necesario para trabajar con la base de datos de *targets* y cónyuges (n=5,333) que puede crearse usando el procedimiento CREARTC.SPS. También está en formato SPSS.

Importante: Si se desea crear el archivo con registros independientes para el *target* y el cónyuge, utilizando el procedimiento CREARTC.SPS provisto. Se deberán copiar los tres archivos de arriba en un directorio llamado C:\PREHCO1, abrir el archivo PREHCO1.SAV y ejecutar el procedimiento CREARTC.SPS. El resultado será un nuevo archivo llamado PREHCOTC.SAV.

Figura 4. Creación del archivo de datos de *targets* y cónyuges (n=5,333)



4. **intro_s.pdf** e **intro_e.pdf**

Este documento de introducción, que está disponible en español e inglés.

5. **codebook.pdf** (aproximadamente 15Mb)

Ver ‘Libro de códigos’ en la página 15. Si desea ver electrónicamente el libro de códigos de los datos puede obtener una copia del programa **ViewSav**⁴ que permite ver en tiempo real un *codebook* de una base de datos en formato SPSS.

Notas generales sobre la estructura de los *rosters*⁵

En el cuestionario se incluyen dos secciones que recogen información sobre los residentes del hogar y sobre otros familiares y personas cercanas al entrevistado. La sección C contiene los datos de los miembros del hogar y la sección L incluye cuatro *rosters*: el de hijos, el de los hermanos (hasta tres hermanos con los que el entrevistado mantiene más relación), el de los padres y otras personas cercanas.

Las personas de cada *roster* tienen asignadas ciertas posiciones en la base de datos. El *roster* del hogar ocupa las posiciones entre la 1 y la 10. En el *roster* del hogar el *target* siempre ocupa la posición 1. El cónyuge, cuando existe, no siempre ocupa necesariamente la misma posición, es decir puede encontrarse teóricamente en cualquiera de los registros del hogar entre el 2 y el 10, ya que el primer registro siempre corresponde al entrevistado principal. Esto implica que, para identificar el registro ocupado por el cónyuge se deberá realizar una búsqueda por el parentesco de la persona de cada registro con el entrevistado: PARE_R(2-10)=2 ó 4.

El *roster* del hogar: variables disponibles

Variable	Descripción
pos(1-10)	Posición en el <i>roster</i> del hogar
c4r(1-10)_1	Parentesco general con el entrevistado
pare_r(1-10)	Parentesco específico con el entrevistado
c5r(1-10)_1	Parentesco general con el jefe del hogar
c5r(1-10)_2	Parentesco específico con el jefe del hogar
sexo_r(1-10)	Sexo
edad_r(1-10)	Edad
c8r(1-10)	Habilidad para leer
c9r(1-10)	Habilidad para escribir
c10r(1-10)	Escolaridad
mari_r(1-10)	Estado marital
c12r(1-10) (a-i)	Estado laboral (serie de 9 preguntas)
dd(1-10)	Duplicación en otro <i>roster</i>

⁴ ViewSav es un *Real-Time Codebook* gratuito para archivos de SPSS. Ha sido creado por Karel Asselberghs, del Departamento de Sociología de la Universidad de Amsterdam y puede obtenerse en la dirección: <http://www.asselberghs.dds.nl/downloads/ViewSav.zip>

⁵ Una presentación detallada sobre la estructura de los *rosters* (ROSTERS.PDF) está disponible en nuestro web site <http://prehco.rcm.upr.edu>.

Así pues las variables de la primera persona en el *roster* del hogar aparecerían en la base de datos de esta manera:

	caseid	c1num	c3	pos1	c4r1_1	pare_r1	c5r1_1	c5r1_2	sexo_r1	edad_r1	c8r1	c9r1	c10r1	mar_r1	dd1
1	11050110	2	2	1	0	1	1	2	2	64	1	1	13	1	0
2	11050210	1	1	1	0	1	0	0	1	87	1	1	6	3	0
3	11050310	5	2	1	0	1	2	7	2	66	1	1	13	3	0
4	11050410	1	1	1	0	1	0	0	1	62	2	2	4	6	0
5	11050610	1	1	1	0	1	0	0	2	70	1	1	10	4	0
6	11050810	2	1	1	0	1	0	0	1	75	1	1	13	1	0
7	11050910	1	1	1	0	1	0	0	1	82	1	1	15	6	0
8	11060510	1	1	1	0	1	0	0	2	78	2	2	2	4	0

El *roster* de hijos ocupa las posiciones entre la 21 y la 40, el de los hermanos las posiciones entre la 61 y la 63, el *roster* de padres las posiciones entre la 71 y la 72 y el *roster* de otras personas las posiciones entre la 81 y la 83.

El *roster* de hijos: variables disponibles

Variable	Descripción
pos(21-40)	Posición en el <i>roster</i> de hijos
pare_r(21-40)	Parentesco específico con el entrevistado
sexo_r(21-40)	Sexo
l4r(21-40)	Indica si el hijo está vivo
edad_r(21-40)	Edad a la entrevista o edad al momento de morir
dond_r(21-40)	Dónde vive
l7r(21-40)	Escolaridad
mari_r(21-40)	Estado marital
l9r(21-40)	Número de hijos
l10r(21-40)	Vive o vivió en Estados Unidos
l11r(21-40)	Lo ayudó económicamente a mudarse a Estados Unidos
dd(21-40)	Duplicación en el <i>roster</i> del hogar

Así pues las variables de la primera persona en el *roster* de hijos aparecerían en la base de datos de esta manera:

	caseid	l1num	pos21	pare_r21	sexo_r21	l4r21	edad_r21	dond_r21	l7r21	mar_r21	l9r21	l10r21	l11r21	dd21
39	11100820	2	21	5	2	1	28	1	15	6	0	3	.	3
40	11101310	5	21	5	2	1	47	5	15	4	3	1	2	0
41	11110110	1	21	5	2	1	47	4	16	4	3	3	.	0
42	11110120	2	21	5	2	1	47	4	16	1	1	3	.	0
43	11110210	5	21	5	1	1	52	6	16	6	0	3	.	0
44	11110310	1	21	5	1	2	0	2	1	0
45	11110410	2	21	5	2	1	41	1	13	6	0	3	.	4
46	11110510	3	21	5	2	1	47	4	13	2	3	3	.	0
47	11110610	2	21	5	2	1	44	1	15	6	0	3	.	3
48	11110710	3	21	5	1	1	40	4	12	1	4	3	.	0

El roster de hermanos: variables disponibles

Variable	Descripción
pos(61-63)	Posición en el <i>roster</i> de hermanos
pare_r(61-63)	Parentesco específico con el entrevistado
sexo_r(61-63)	Sexo
edad_r(61-63)	Edad
dond_r(61-63)	Dónde vive
mari_r(61-63)	Estado marital
dd(61-63)	Duplicación en el <i>roster</i> del hogar

El roster de padres: variables disponibles

Variable	Descripción
pos(71-72)	Posición en el <i>roster</i> de padres
pare_r(71-72)	Parentesco específico con el entrevistado
sexo_r(71-72)	Sexo
edad_r(71-72)	Edad
dond_r(71-72)	Dónde vive
dd(71-72)	Duplicación en el <i>roster</i> del hogar

El roster de otras personas: variables disponibles

Variable	Descripción
pos(81-83)	Posición en el <i>roster</i> de otras personas
pare_r(81-83)	Parentesco específico con el entrevistado
l28r(81-83)	De qué hijo o hermano es hijo
sexo_r(81-83)	Sexo
dond_r(81-83)	Dónde vive
dd(81-83)	Duplicación en el <i>roster</i> del hogar

Como se aprecia en las tablas anteriores, las variables comunes a todos los *rosters* (ver Tabla 3 más arriba) tienen una nomenclatura común. Esto quiere decir que el estado marital del hijo que ocupe la posición 22 estará almacenado en la variable MARI_R22 y el del hermano que ocupe la posición 61 en la variable MARI_R61. Esto ocurre igualmente con el sexo (SEXO_R##), la edad (EDAD_R##), el parentesco (PARE_R##) y el lugar en el que reside (DOND_R##).

Igualmente las variables que son particulares de cada roster, aunque mantienen su nombre, que corresponde a la pregunta del cuestionario, incluyen la posición de esa persona y por consiguiente su *roster* primario de pertenencia. Así el número de hijos del hijo que ocupa la posición 35 se almacena en L9R35, y el número de hijos del hijo que ocupa la posición 22 en L9R22.

La estructura de estos *rosters* no es homogénea, es decir en todos ellos no se recoge la misma información. En la Tabla 3 puede observarse la presencia o ausencia de una serie de variables básicas en cada uno de los *rosters*.

Tabla 3. Presencia de algunas variables básicas en cada uno de los cinco *rosters*.

Variable	<i>Roster</i>				
	Hogar	Hijos	Hermanos	Padres	Otros
Relación con el entrevistado	X	X	X	X	X
Género	X	X	X	X	X
Edad	X	X	X	X	no
Estado marital	X	X	X	no	no
Donde vive	⁶	X	X	X	X

Variables de posición y variables de duplicación

Las secciones de los *rosters* incluyen las llamadas variables de posición, cuyo nombre sigue la secuencia POS1 a POS10 para el hogar, POS21 a POS40 para los hijos, POS61 a POS63 para los hermanos, POS71 a POS72 para los padres y POS81 a POS83 para otras personas. Estas variables tienen el valor correspondiente a la posición (por ejemplo: POS23=23) si existe una persona en esa posición o están vacías si no existe.

Ya que el *roster* del hogar no es excluyente de los otros *rosters*, una persona podría pertenecer al *roster* del hogar y a otro cualquiera de los *rosters*. Esta duplicación sólo puede ocurrir entre el *roster* del hogar y cualquiera de los otros *rosters*. Un hijo que resida en el hogar podría ocupar la posición 5 en el *roster* del hogar (POS5=5) y la posición 22 en el *roster* de hijos (POS22=22). Este es un ejemplo de un registro duplicado.

Cada *roster*, aunque comparta variables comunes con los otros, como se presentó en la tabla 3, contiene alguna información no incluida en los demás (ver en el cuestionario las secciones C y L). Por esta razón es conveniente conocer si una persona está duplicado y, por lo tanto, tiene información adicional en otro *roster*. Para identificar si un registro está duplicado se utilizan las variables de duplicación.

Las variables de duplicación siguen la misma nomenclatura de las de posición (DD1 a DD10, DD21 a DD40, DD61 a DD63, DD71 a DD72 y DD81 a DD83). Estas variables pueden tener el valor cero (0) cuando la persona no está duplicada, un valor correspondiente a la posición en la que se encuentra el duplicado o un valor missing si no hay ninguna persona en ese registro.

Así pues, para el hijo del ejemplo anterior, que se encuentra en la posición 5 del hogar y se encuentra duplicado en la posición 22 del roster de hijos, la variable DD5 tendría el valor 22 (DD5=22) y la variable DD22 tendría el valor 5 (DD22=5).

⁶ Aunque no existe una variable que indique el lugar en el que viven los miembros del hogar, se entiende que, siguiendo la lógica de las variables sobre el lugar de residencia de los otros *rosters*, todas las personas del hogar tendrían el valor 1 (Vive en el hogar) para poder ser comparadas con otras personas.

Utilización de la información contenida en los *rosters*

La información de los *rosters* puede utilizarse para varios propósitos. Uno de ellos sería analizar los arreglos familiares y la estructura familiar del entrevistado y su red de apoyo. También podría utilizarse con el propósito de identificar las características de ciertas personas con las que el entrevistado se relaciona. Esta tarea implica una cierta complejidad y requiere explicación.

Existen tres secciones del cuestionario: la sección I (Estado Funcional), la sección M (Transferencias) y la sección N (Historia Laboral e Ingreso) que utilizan la información recogida en los *rosters*. Básicamente, en ciertas partes de estas secciones se pregunta al entrevistado sobre la ayuda que recibe de o que presta a otras personas. El entrevistado que declaró recibir o prestar ayuda para una determinada actividad, indica quién es la persona que le presta ayuda o que es objeto de su ayuda entre las personas de los *rosters*. Esta información se codifica con el número de posición de esa persona en los *rosters*. Si la persona seleccionada está duplicada, por ejemplo en el *roster* del hogar y el *roster* de hijos, es decir, es un hijo que vive en el hogar, aparecerá el número que indique su posición en su *roster* primario, que sería el del hogar. El siguiente ejemplo ilustra esta situación:

Si una persona indica tener dificultad preparando sus comidas (I4=1 o I4=3), se le pregunta por la persona que le ayuda preparando las comidas (I4A). El valor de la variable I4A indica la posición de esa persona en los rosters. Así, si I4A=4, la persona que ayuda al entrevistado sería aquella en la cuarta posición del roster del hogar, su edad sería EDAD_R4, su sexo SEXO_R4, su estado marital MARI_R4, su parentesco con el entrevistado PARE_R4 y el lugar donde reside, en este ejemplo el hogar. Pero, además de estas variables comunes, existen otras variables particulares del roster del hogar en las que igualmente existe información adicional sobre esa persona. En este caso todas esas variables tendrán la terminación R4.

Si observamos la variable DD4, la variable de duplicación para ese miembro del hogar, y tiene el valor 22, esto querrá decir que esa persona está duplicada en el roster de hijos en la posición 22. En ese registro del roster de hijos existe igualmente información adicional sobre esa persona. En este caso todas esas variables tendrán la terminación R22.

Aunque la norma es que el entrevistado seleccione la persona de entre los miembros de los *rosters*, en las preguntas ya mencionadas de las secciones I, M y N, se le dio la posibilidad de identificar otras personas que no estuvieran incluidas en ninguno de los *rosters*, ni siquiera el de otras personas, cuando se le preguntaba por las ayudas dadas o recibidas. Como resultado, en algunas de las variables se observa con cierta frecuencia el valor -77, indicando que la persona mencionada no está en ninguno de los *rosters*. El resultado neto es que no pueden conocerse las características de esas personas, salvo que en principio no son miembros del hogar, ni hijos, ni hermanos, ni padres. Igualmente en las preguntas sobre la persona que ayuda de la sección N pueden aparecer valores como -99 (el entrevistado no tiene ese gasto) o -98 (pagan todos en el hogar).

VARIABLES CALCULADAS

En el banco de datos aparece un grupo de variables calculadas⁷ que presentan datos importantes de algunas de las secciones del cuestionario. Estas variables le ofrecen al usuario información más clara o resumida de algunos datos específicos de los entrevistados. Algunas de las variables son el producto de la combinación de dos o más variables, otras simplemente la recodificación de una variable del cuestionario. En las Tablas 4 y 5 se presentan las listas de las variables construidas para el *target* y el cónyuge, indicando la sección del cuestionario a la que se refieren, su posición en la base de datos, su nombre y su descripción. Al igual que las variables originales del cuestionario del cónyuge, sus variables calculadas comienzan con la letra “y”.

Tabla 4. Lista de variables calculadas del *target* en la base de datos

SECCION	POSICION	VARIABLE	DESCRIPCION
Inicial	8	fac_t	Factor de ponderación
	9	targvi	Targets en la vivienda
	10	totarget	Total de targets en vivienda
B	64	finalscr	Puntuación final del minimental
	65	pasomin	¿Entrevistado(a) pasó el minimental?
C	73	qjefe	Quién es el jefe del hogar
G	642	depre_y	Puntuación Escala de Depresión
	476	g77_code (R)*	¿Qué le dijo el médico que era ese dolor o malestar en el pecho?
	615	g168fisi (R)	¿Qué tipo de pasatiempo (actividades físicas) realizó regularmente?
	616	g168soci (R)	¿Qué tipo de pasatiempo (actividades sociales) realizó regularmente?
	617	g168ocio (R)	¿Qué tipo de pasatiempo (actividades de ocio) realizó regularmente?
H	648	code_h4 (R)	Ocupación del padre
	649	rcode_h4 (R)	Ocupación del padre recodificada
	656	code_h6d (R)	Ocupación de la madre
	657	rcodeh6d (R)	Ocupación de la madre recodificada
K	743	k60a_cod (R)	Servicio de salud que no pudo conseguir
N	1188	code_n7 (R)	Ocupación actual (semana pasada)
	1189	code_n9 (R)	Tipo de establecimiento o negocio donde trabaja
	1215	code_n18 (R)	Ultimo empleo
	1216	code_n20 (R)	Tipo de establecimiento o negocio donde trabajaba
	1228	code_n25 (R)	Ocupación de la mayor parte de su vida
	1229	cod_n25b (R)	Tipo de establecimiento o negocio donde trabajó la mayor parte de su vida
	1241	n50_code (R)	Tipo de problema de salud que padece
O	1339	o10_code (R)	Razón principal para regresar de EEUU a PR
U	1460	bmi	Índice de Masa Corporal
	1461	rbmi	Índice de Masa Corporal recodificado
	1469	legstand	Resultado ejercicio pararse en una sola pierna
	1473	get_up	Resultado ejercicio levantarse y caminar

*(R)=Recodificada de una variable alfanumérica original

⁷ Todas las variables calculadas incluyen en su etiqueta la palabra CALCULADA.

Tabla 5. Lista de variables calculadas del cónyuge en la base de datos

SECCION	POSICION	VARIABLE	DESCRIPCION
Inicial	1483	conyuge	Identifica entrevista de cónyuge completada
	1488	y marital	Estado marital
	1489	y leer	Sabe leer
	1490	y escribe	Sabe escribir
	1491	y escolar	Escolaridad
B	1543	y finascr	Puntuación final del minimental
	1544	y pasomin	Pasó el minimental
G	1690	yg77code (R)*	¿Qué le dijo el médico que era ese dolor o malestar en el pecho?
	1829	yg168fis (R)	¿Q ué tipo de actividad física realizó regularmente?
	1830	yg168soc (R)	¿Q ué tipo de actividad social realizó regularmente?
	1831	yg168soci (R)	¿Q ué tipo de actividad de ocio realizó regularmente?
	1856	y depre_y	Puntuación de Escala de Depresión
H	1862	y code_h4 (R)	Ocupación del padre
	1863	y rcodeh4 (R)	Ocupación del padre recodificada
	1870	y codeh6d (R)	Ocupación de la madre
	1871	y rcodh6d (R)	Ocupación de la madre recodificada
K	1957	y k60a_co (R)	Servicio de salud que no pudo conseguir
N	1997	y code_n7 (R)	Ocupación actual (semana pasada)
	1998	y code_n9 (R)	Tipo de establecimiento o negocio donde trabaja
	2024	y coden18 (R)	Ultimo empleo
	2025	y coden20 (R)	Tipo de establecimiento o negocio donde trabajaba
	2037	y coden25 (R)	Ocupación de la mayor parte de su vida
	2038	y codn25b (R)	Tipo de establecimiento o negocio donde trabajó la mayor parte de su vida
	2050	y n50code (R)	Tipo de problema de salud que padece
O	2106	y o10code (R)	Razón principal para regresar de EEUU a PR
U	2146	y bmi	Índice de Masa Corporal
	2147	y rbmi	Índice de Masa Corporal recodificado
	2155	y legstand	Resultado ejercicio pararse en una sola pierna
	2159	y get_up	Resultado ejercicio levantarse y caminar

*(R)=Recodificada de una variable alfanumérica original

Errores y limitaciones identificados en los datos

Existen algunos errores o limitaciones conocidos en la base de datos que los usuarios deben considerar:

1. Las variables G55 (se ha caído después de los 60 años) y G56 (número de caídas en el último año) presentan algunos casos sin información (22 y 53, respectivamente) debido a un error de programación, corregido durante el levantamiento de los datos. Aunque para algunos casos pudo obtenerse la información por teléfono, para otros casos los datos son desconocidos. En la base de datos aparecen debidamente identificados con el valor -3.
2. Ver el último párrafo de la sección ‘Utilización de la información contenida en los rosters’ sobre la limitación en la identificación de las características de algunas personas que prestan o reciben ayuda (valor -77).

3. La sección S sobre maltrato fue contestada por un número reducido de participantes. Esto es debido a las instrucciones dadas a los entrevistadores. Por lo delicado de las preguntas, se les instruyó a no realizar la sección si alguna persona se encontraba presente en el momento de la entrevista. Aquellos usuarios de la base de datos interesados en estudiar el maltrato hacia los adultos mayores deben tener en cuenta esta limitación.
4. En la sección G sobre Salud se incluyó experimentalmente una sub-sección de síntomas tomados del ámbito clínico con el doble propósito de validar algunas de las enfermedades informadas por el entrevistado y poder analizar el poder predictivo de los síntomas. Por un error inicial de programación, posteriormente subsanado, algunos de los entrevistados no contestó las preguntas relacionadas a los síntomas respiratorios (G92-G92I3).
5. En la sección O sobre Migración, las preguntas desde la O10a a la O10d se introdujeron una vez comenzado el trabajo de campo por lo que algunos de los casos no tienen información válida. Este problema afecta solamente a algunos de los entrevistados que hubieran residido en Estados Unidos y están debidamente identificados con el valor -3.
6. Debido a un error de programación, el estado marital de algunos de los hermanos en la sección L no tiene valor.
7. La pregunta sobre el diagnóstico de bilharzia (G107) contienen un gran cantidad de datos *missing* por un error en la programación del cuestionario.
8. Las variables M16 (hijo o hija que más le ayuda), M17 (hermano o hermana que más le ayuda) y M18 (persona que más le ayuda) no tienen información para algunos casos.
9. Alguna información sobre activos e ingresos se recogió usando variables categóricas en vez de variables continuas. Por tal razón, estas variables sólo permiten realizar análisis limitados sobre los efectos del ingreso y los activos. Después de terminado el campo se realizó una encuesta telefónica a una submuestra aleatoria de los participantes de PREHCO para obtener información sobre estas variables. Próximamente se hará disponible una nueva versión del archivo completo de los datos que incluirá variables continuas sobre activos e ingresos.

Libro de códigos

El libro de códigos se encuentra en el documento CODEBOOK.PDF. Para cada una de las variables en la base de datos se presenta la siguiente información: nombre de la variable, etiqueta de la variable, posición en la base de datos, texto de la pregunta correspondiente del cuestionario, tipo de variable, tamaño de la variable, valores ausentes o *missing values*, códigos y etiquetas utilizados, distribución de frecuencias y comentarios, esto último para algunas variables complejas. La distribución de frecuencias no se presenta para aquellas variables que tienen más de 20 categorías, excepto en el caso de algunas variables de importancia como la edad del entrevistado.