

Table of Contents

Introduction.....	1
Information about the second wave of the project.....	1
Versions of the Questionnaire.....	2
General Considerations.....	5
Result codes	5
Denomination of the Variables	5
Analysis of Couples	5
Notes on the Minimal.....	5
Some Age Variables	6
Weighting the Data	6
Missing Values.....	7
Data Files and Related Documents.....	8
General Notes about the Structure of the Rosters.....	9
The Household Roster: Available Variables.....	9
The Children Roster: Available Variables.....	10
The Siblings Roster: Available Variables.....	11
The Parents Roster: Available Variables	11
The Other Persons Roster: Available Variables	11
Variables of Position and Variables of Duplication	12
Utilization of Rosters' Data	13
Calculated Variables	14
Errors and Limitations Identified in Data.....	15
Codebook	17

Introduction

The research project Puerto Rican Elderly: Health Conditions (*Condiciones de salud de los adultos mayores en Puerto Rico*), known as PREHCO, is a study funded with federal funds from the National Institute on Aging / NIA (Grant number RO1 AG1622090-01) through a contract between the University of Wisconsin-Madison and the University of Puerto Rico. This project also benefits from the technical support of the Pan American Health Organization (Washington D.C.) and has received economic support from the Legislature of Puerto Rico and the Office of the Women's Advocate in Puerto Rico.

In its first wave, between May 2002 and May 2003, we carried out a cross sectional study of the non-institutionalized elderly population 60 years or older in Puerto Rico. There is ample documentation corresponding to the first wave at the PREHCO website (<http://prehco.rcm.upr.edu/docum>).

The second wave of PREHCO was completed between May 2002 and November 2007. The data collected in that second wave is presented in this codebook.

Information about the second wave of the Project

In the first phase of the PREHCO project, 4,291 elderly persons 60 years or older and 1,442 spouses were interviewed, regardless of their age. In the second wave those participants were followed. We completed 3,891 target interviews and 1,260 spouse interviews. The deceased and institutionalized participants were interviewed using a proxy. Table 1 shows the distribution of the second wave interviews by type of interview.

Table 1. Distribution of PREHCO 2 participants by type of interview.

Result	PREHCO 1 Targets (n=4,291)	PREHCO 1 Spouses (n=1,442)
Target interview	2,726	1,049
Target interview through a proxy	439	87
Deceased interview through a proxy	678	117
Institutionalized interview through a proxy	48	7
No response	400	182

Even though this codebook presents the structure of the database showed in Figure 1, since the interviewee and his/her spouse answered the same questions in many sections of the questionnaire, it is also possible to construct, in order to increase the sample size, a database that includes the targets and their spouses as independent records (Figure 2). There is a SPSS syntax file that permits us to automatically create this database (see section titled "Data Files and Related Documents" below). That procedures creates a database containing the data of the targets and that of the spouses being 60 or older at the time of the first interview.

Figure 1. Structure of the Database of PREHCO 2 – PREHCO2.SAV (n=4,291)

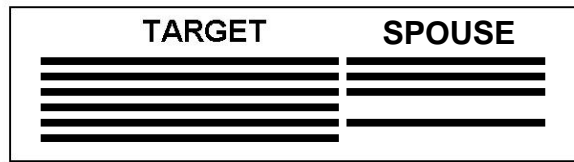
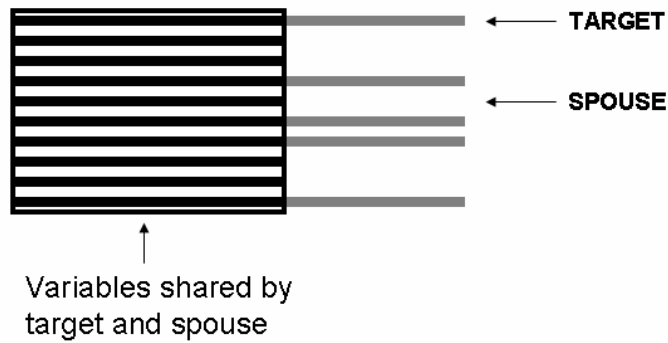


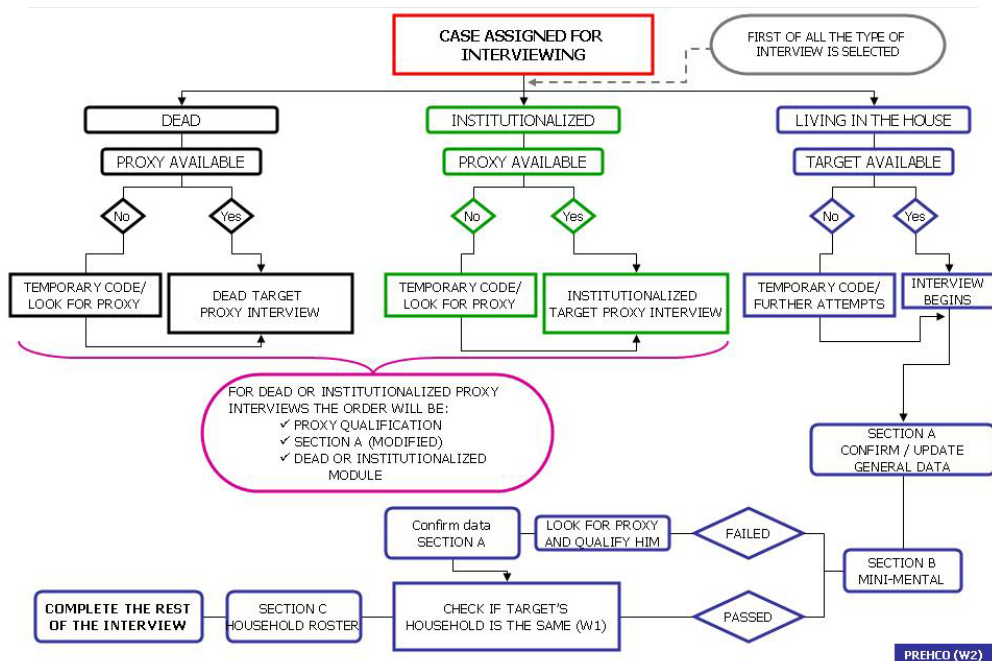
Figure 2. Structure of the PREHCO 2 database (n=3,849)



Versions of the questionnaire

The main questionnaire changes according to various factors, mainly the capacity of the interviewee to respond, and his/her actual status (living in a household, deceased or institutionalized) (see Figure 3).

Figure 3. Flowchart used to determine the type of interview



As result there are six main versions of the questionnaire (see Table 2); these are: (1) target questionnaire, (2) spouse questionnaire, (3) questionnaire of the proxy¹ for the target, and (4) questionnaire of the proxy for the spouse, (5) questionnaire of the proxy of the deceased interviewee, and (6) questionnaire of the proxy of the institutionalized interviewee.

1. Target questionnaire: the principal questionnaire is the one administered to the older adult who is capable of answering the questionnaire on his/her own. This version includes all the sections of the questionnaire.
2. Spouse questionnaire: the questionnaire of the spouse is the one administered to the spouse of the older adult capable of answering the questionnaire on his/her own. This questionnaire does not include some of the sections of the principal questionnaire and also excludes some questions in the included sections.
3. Proxy for the target questionnaire: this questionnaire is answered by a proxy, in the cases in which the older adult is not capable of responding on his/her own. The determination of whether it is necessary to have a proxy is carried out using the following criteria: (1) the score obtained by the interviewee in the minimal and (2) the opinion of the interviewer in some cases². The proxy provides information relative to the older adult, excluding some sections or individual questions that imply opinion or state of mind. In some cases the spouse acts as the proxy for the interviewee and in others, in which there is no spouse or the spouse is not capable of responding, the proxy is another person.
4. Proxy for the spouse questionnaire: this questionnaire is completed by a proxy, in the cases in which the spouse of the older adult is not capable of responding on his/her own (only for spouses 60 or older). The determination of the necessity for a proxy was realized with the criteria previously outlined. The proxy provides information relative to the spouse of the older adult, excluding sections or individual questions that imply opinion or state of mind. In some cases the target acts as a proxy for the spouse and in others the proxy is another person.
5. Questionnaire of the proxy for the deceased interviewee: if the interviewee died between the two waves of PREHCO the interview is carried out through a proxy, who answers a special reduced version of the questionnaire (see Table 2).
6. Questionnaire of the proxy for the institutionalized interviewee: if the interviewee was institutionalized between the two waves of PREHCO the interview is carried out through a proxy, who answers a special reduced version of the questionnaire (see Table 2).

¹ Criteria for qualifying as proxy: he/she must have known the interviewee for at least one year and must be at least 18 years of age.

² It was established as criteria that a total score on the minimal of 11 or more indicated that the interviewee was capable of responding to the questionnaire on his/her own. A score of less than 9 indicated that it was necessary to use a substitute informant. In those cases in which the score was 9 or 10 the interviewer was the one who determined whether the interviewee was capable of responding.

Table 2. PREHCO 2: Versions of the Questionnaire

SECTION	TYPE OF QUESTIONNAIRE					
	TARGET	SPOUSE	PROXY TARGET	PROXY CONYUGE	PROXY DECEASED (for target and spouse)	PROXY INSTITUTIONALIZED (for target and spouse)
A Type of questionnaire	Complete	Complete	Complete	Complete	Complete	Complete
B Minimental	Complete	Complete (if 60 ≥)	Complete (if the proxy is the spouse as s/he is 60 or older) Short (if the proxy is other person 60 or older)	No (if the proxy is the target) Short (if the proxy is other person 60 or older and s/he was not proxy for the target)	Complete (if the proxy is the spouse and s/he is 60 or older) Short (if the proxy is other person 60 or older)	Complete (if the proxy is the spouse and s/he is 60 or older) Short (if the proxy is other person 60 or older)
C Roster hogar	Complete	NO	Complete	NO	NO	NO
D Marital Religión	Complete	WD15 to WD20 WD31 to WD33 WD39	WD101 to WD20	WD15 to WD20	NO	NO
G Salud	Complete	WG1 to WG3 WG4 to WG185	WG4 to WG67 WG69pre to WG60d WG109 WG114ck to WG128 WG133 to WG139 WG141 WG148 to WG156 WG161 to WG170	WG4 to WG67 WG69pre to WG60d WG109 WG114ck to WG128 WG133 to WG139 WG141 WG148 to WG156 WG161 to WG170	NO	NO
H Niñez	Complete	Complete	NO	NO	NO	NO
J Medicinas	Complete	Complete	Complete	Complete	NO	NO
K Uso servicios	Complete	Complete	Complete	Complete	NO	NO
L Roster Familiares	Complete	NO	Complete	NO	NO	NO
I Funcional	Complete	WIint to WI7ab WI9pre to WI14ab WI16 WI17	NO	NO	NO	NO
Z Proxy Funcional	WZ12 to WZ14	WZ12 to WZ14	Complete	Complete	NO	NO
M Ayudas	Complete	NO	NO	NO	NO	NO
N Laboral, Ingreso	Complete	WN5 to WN513	WN5 to WN70	WN5 to WN513	NO	NO
O Migración	Complete	Complete	NO	NO	NO	NO
P Vivienda	Complete	NO	Complete	NO	NO	NO
Q Propiedades	Complete	NO	Complete	NO	NO	NO
R Raza	Complete	Complete	NO	NO	NO	NO
T Sexualidad	Complete	NO	NO	NO	NO	NO
YY Mano dominante	Complete	Complete	NO	NO	NO	NO
U Medidas	Complete	Complete (if 60 ≥)	Complete	Complete (if 60 ≥)	NO	NO
V Contactos	Complete	Complete	Complete	Complete	NO	NO
ZZ Cuidador	Complete	NO	Complete	NO	NO	NO
F Proxy Institucionalizados	NO	NO	NO	NO	NO	Complete
Y Proxy fallecidos	NO	NO	NO	NO	Complete	NO
X Entrevistador	Complete	Complete	Complete	Complete	Complete	Complete

General Considerations

Below are some general considerations that should be taken into account before using the database.

Result Codes

The variable TCODE contains the information about the result of the interview, when it was completed, or the no response code. That way all the cases interviewed in the first wave (n=4,291) have their corresponding result code, some of them a no response code in PREHCO 2.

Denomination of the Variables

The names of the variables are similar to those in PREHCO 1, but preceded by the letter W. For instance, the variable about the health status of the respondent, named G1 in the first wave it is named WG1 in the second one. The variables corresponding to the spouse of the interviewee maintain the same nomenclature of those of the interviewee, but are preceded by the letter Y. So, for example, the variable for general health status is WG1 for the target interviewee and YWG1 for his/her spouse.

Analysis of Couples

Only the spouses already interviewed in PREHCO 1 were eligible to be interviewed in the second wave. Those of them deceased or institutionalized could be interviewed using a proxy, only if they were the last couple of the target. The variables containing the information about the status of the target and his/her present relationship with the spouse are presented in Section A: Type of questionnaire, and are showed in the following table.

Table 3. Variables about the status of the target and the spouse

WAMOR	Interviewee is alive
WVIVOA	Interviewee was married before dying
WVIVOA1	The PREHCO 1 spouse was the last couple of the deceased target
WVIVOB	Interviewee married with PREHCO 1 spouse (deceased target)
WVIVOC	PREHCO 1 spouse is alive (deceased target)
WVIVOD	Reason why the target is no longer married with the PREHCO 1 spouse (deceased target)
WVIVOE	PREHCO 1 spouse not eligible and alive (deceased target)
WAMARR1	Target married or in a union
WAMARR2	Target married with the same PREHCO 1 spouse
WTIPO	Type of interview (target)
WAINS1	PREHCO 1 spouse was the last couple of the target (target institutionalized)
WAINS2	Reason why the target is no longer married with the PREHCO 1 spouse (target institutionalized)

Notes on the Minimental

As indicated in Table 2, Section B of the questionnaire contains a brief mental examination or minimental. The minimental also works as a filter used to determine the need for a proxy, following the criteria previously explained. In the cases in which the target's impediment or incapacity to answer the questionnaire on his/her own was obvious, the minimental was not carried out and a proxy was chosen directly. These cases can be identified using the variable WBTVAL (WBTVAL=1). The participants

who demonstrated an obvious impediment do not have points (WFINSCR) in the minimal. In some cases the interviewee had a physical impediment (WBTCAP=1) that inhibited him/her from completing some of the tests on the minimal that require skills such as drawing or folding. In those cases, the total points on the minimal were obtained by extrapolating the points obtained on the questions that were answered to the entire set of questions and tests included in the minimal. Section B of the questionnaire is also used to determine the capability of the proxies, when they are 60 or older, as indicated in Table 2.

Some Age Variables

In some parts of the interview the age at which a specific event took place is recorded. Initially, the interviewee is directly asked the age. If the age is unknown, the interviewee is asked for the month and the year. Then a final age is calculated for that event that indicates the age supplied by the interviewee or, in the event that it's faulty, the age calculated based on the year and the month. So that, for each age attributed to each event, there is a series of four variables. The variables for all these series have the following format: *_E (age supplied by the informant), *_A (year), *_M (month), and *AE (age summary: supplied age or calculated age). This means that in order to analyze the age in which a specific event took place, the last variable is the one that should be used, since it reflects the most complete information. In the following example, please find the variables that record the age when diabetes was diagnosed:

WG8A_E	(supplied age)
WG8A_A	(year)
WG8A_M	(month)
WG8AE	<u>(age summary: supplied or calculated)</u>

Weighting the data

The weighting factor for each record is located in the variable **FACTORT** in the database (n=4,291). Its calculation is based on a poststratification by gender and age group, considering five age groups: 64, 65-69, 70-74, 75-79, 80 years or more. If the database (n=3,849) is constructed with interviewees and spouses as separate records (see "Information about the second phase of the PREHCO Project" on page 1) a different weight factor should be used. This weighting factor is available in the file FACTORTC2.SAV (see section "Data Files and Related Documents" below).

Missing Values

The entire database uses common codes for missing values. These values are presented in the following table:

Table 4. Codes and labels for missing values.

-1	Does not know	Interviewee indicated that he/she does not know the answer to the question.
-2	Does not answer	Interviewee indicated that he/she would not answer the question.
-3	Undetermined, data unavailable	Data not compiled due to some type of error.
-5	Does not apply to the proxy-target	The question does not apply to the proxy for the target questionnaire.
-6	Does not apply to the proxy-spouse	The question does not apply to the proxy for the spouse questionnaire.
-7	There is no spouse	The interviewee does not have a spouse or partner; therefore, there is no interview of the spouse.
-8	Does not apply to the spouse younger than 60 years of age	The question does not correspond to spouses younger than 60 years of age.
-9	The target interview was not completed	The target interview was not completed. Variable TCODE contains a code indicating the reason why.
-10	Does not apply, obvious impediment	The interviewee has an obvious impediment (determined by the interviewer before starting the minimal test) and does not answer the questions of the minimal.
-11	Does not apply, physical impediment	The interviewee has a physical impediment and does not answer the questions of the minimal that require hand movements as with writing and drawing.
-12	Does not apply, there is no spouse interview	The target interviewee has a spouse or partner but the spouse/partner was not interviewed. The variable that identifies the final result of the interview of the spouse/partner ([CCODE]) has a code that indicates the reason why.
-13	Section on Caregivers does not apply	Section ZZ about caregivers does not apply to the target.
-14	Does not apply to the deceased target	That variable is not answered by the proxy of the deceased target.
-15	Does not apply to the institutionalized target	That variable is not answered by the proxy of the institutionalized target.
-16	Does not apply to the deceased spouse	That variable is not answered by the proxy of the deceased spouse.
-17	Does not apply to the institutionalized spouse	That variable is not answered by the proxy of the institutionalized spouse.
-18	Does not apply to the regular target	The question does not apply to the regular target without proxy.

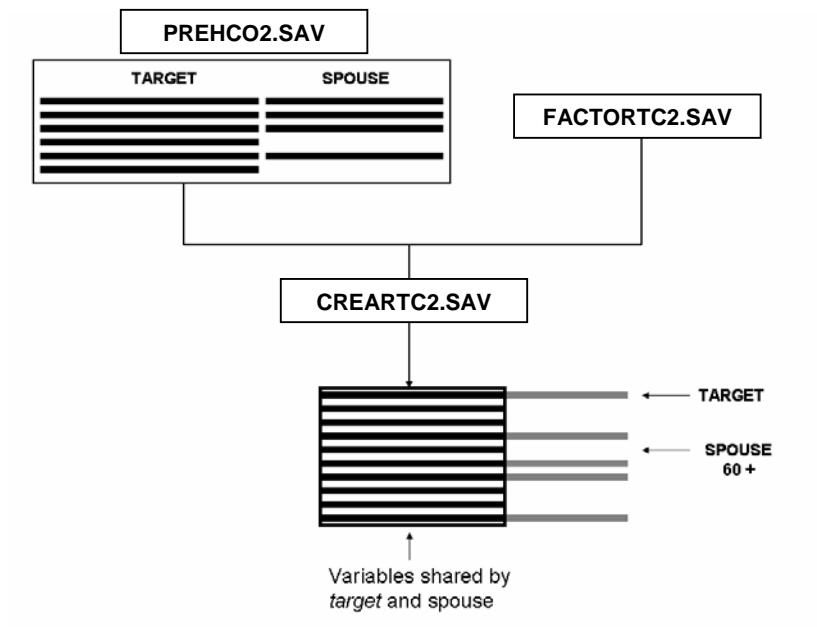
Data Files and Related Documents

The following data files are available:

1. **PREHCO2.SAV** (approximately 13Mb)
Contains all of the data of the second wave of PREHCO in SPSS format. The data is stored in 2,766 variables.
2. **CREARTC2.SPS** (19 Kb)
This SPSS syntax file automatically creates a file of 3,849 records that has the interviewees and their spouses (60 years or older in PREHCO 1) in separate records. Since both interviews share at least 665 variables, this procedure allows for working with these common variables (see above for the explanation on versions of the questionnaire) with a larger sample size. To weight the resulting file (PREHC2TC.SAV) the factor FACTORTC should be used.
3. **factortc2.sav** (53 Kb)
This file contains the weight factor necessary to work with the database of targets and spouses (n=3,849), which can be created using the procedure CREARTC2.SPS. It is also in the SPSS format.
4. **PREHCO1_EDAD CONYUGES.SAV** (43 Kb)
This file contains the ages of spouses at PREHCO 1, since it is necessary to filter those being 60 or older at PREHCO 1, as indicated above.
5. **Script para value labels al archivo de T y C.SBS** (3 Kb)
This SPSS script is executed from the CREARTC2.SPS syntax and must be located in the same directory as the other files (C:\PREHCO2) in order to run correctly.
Important: For the creation of a file with separate records for the target and the spouse using the procedure CREARTC2.SPS all five files above should be copied to a folder called C:\PREHCO2, then the file PREHCO2.SAV should be opened and the procedure CREARTC2.SPS should be executed. The result will be a new file called PREHC2TC(N=3,849).SAV (Figure 4).
6. **intro2_s.pdf e intro2_e.pdf**
This introductory document, available in English and Spanish.
7. **codbook2.pdf**
See 'Codebook' on page 17. If you would like to electronically view the codebook, it is possible to obtain a copy of the program **ViewSav**³ which allows the viewer to see a codebook of a database in SPSS format in real time.

³ ViewSav is a free Real-Time Codebook for SPSS files. It has been created by Karel Asselberghs, from the Department of Sociology of the University of Amsterdam and can be obtained at: <http://www.asselberghs.dds.nl/downloads/ViewSav.zip>

Figure 4. Creation of the data file of targets and spouses (n=3,849)



General Notes about the Structure of the Rosters

There are two sections in the questionnaire that record information about the residents of the household and about other family members and people close to the interviewee. Section C contains the data of the household members and Section L contains four rosters: the children's roster, the siblings' roster (the three with whom the interviewee maintains the most contact), the parents' roster, and other people who are close to the interviewee's roster.

The people in each roster are assigned certain positions in the database. The household roster occupies positions between 1 and 18: positions 1 to 10 are reserved for household members already in PREHCO 1. In the household roster the target always occupies the position 1. The spouse, when he/she exists, does not always necessarily occupy the same position, that is to say, he/she may theoretically be in any one of the household records between 2 and 10, since the first record always corresponds to the target. This implies that, in order to identify the record occupied by the spouse, there should be a search carried out that inquires as to the relationship of the person in each record with the target: WPAR_R(2-10)=2 or 4.

The Household Roster: Available Variables

Variable	Description
wpos(1-18)	Position in the household roster
wc4r(1-18)_1	General relationship with the interviewee
wpar_r(1-18)	Specific relationship with the interviewee
wc5r(1-18)_1	General relationship with the head of household
wc5r(1-18)_2	Specific relationship with the head of household
wsex_r(1-18)	Gender
weda_r(1-18)	Age
wc8r(1-18)	Ability to read
wc9r(1-18)	Ability to write
wesc_r(1-18)	Schooling/Education
wmar_r(1-18)	Marital status
wc12r(2-18) (a-i)	Work status (series of 9 questions)
wdd(2-18)	Duplication in another roster

The variables of the first person in the household roster will appear in the database as follows:

	wpos1	wc4r1_1	wpar_r1	wc5r1_1	wc5r1_2	wsex_r1	weda_r1	wc8r1	wc9r1	wesc_r1	wmar_r1	wdd1
1	1	0	1	1	2	2	67	1	1	13	1	0
2	1	0	1	0	1	1	91	1	1	6	4	0
3	1	0	1	2	7	2	70	1	1	13	3	0
4	1	0	1	0	1	1	66	2	2	4	6	0

The roster of children occupies positions between 21 and 53 (children already in PREHCO 1 are reserved positions 21 to 40); that of siblings occupies positions between 61 and 66 (siblings already in PREHCO 1 are reserved positions 61 to 63); that of parents between 71 and 72; and that of other persons positions between 84 and 88.

The Children Roster: Available Variables

Variable	Description
wpos(21-53)	Position in the children roster
wpar_r(21-53)	Specific relationship with the interviewee
wsex_r(21-53)	Gender
wl4r(21-53)	Indicates whether child is alive
wed1_r(21-53)	Age at the time of the interview
wed2_r(21-53)	Age at the time of death
wdon_r(21-53)	Place of residence
wesc_r(21-53)	Schooling/Education
wmar_r(21-53)	Marital Status
wl9r(21-53)	Number of Children
wl10r(21-53)	Whether he/she lived or is living in the United States
wl11r(21-53)	Received financial assistance from him/her to move to the United States
wdd(21-53)	Duplication in the household roster

The variables of the first person in the Children Roster will appear in the data base in as follows:

	wpos21	wlvr21	wpar_r21	wsex_r21	wl4r21	wed1_r21	wed2_r21	wdon_r21	wesc_r21	wmar_r21	wl9r21	wl10r21	wl11r21	wdd21	wpos22
1	21	1	5	2	1	46	-	4	16	1	4	3	-	0	22
2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3	21	2	5	1	2	-	54	-	-	-	2	2	1	0	22
4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

The Siblings Roster: Available Variables

Variable	Description
wpos(61-66)	Position in the Siblings Roster
wpar_r(61-66)	Specific relationship with the interviewee
wsex_r(61-66)	Gender
weda_r(61-66)	Age
wdon_r(61-66)	Place of Residence
wmar_r(61-66)	Marital Status
wdd(61-66)	Duplication in the Household Roster

The Parents Roster: Available Variables

Variable	Description
wpos(71-72)	Position in the Parents Roster
wdon_r(71-72)	Place of residence
wmar_r(71-72)	Marital status
wesc_r(71-72)	Schooling/Education
wdd(71-72)	Duplication in the Household Roster

The Other Persons Roster: Available Variables

Variable	Description
wpos(84-88)	Position in the Other Persons Roster
wsex_r(84-88)	Gender
weda_r(84-88)	Age
wpar_r(84-88)	Specific Relationship with the Interviewee
wdon_r(84-88)	Place of residence
wmar_r(84-88)	Marital status
wdd(84-88)	Duplication in the Household Roster

As can be seen in the previous tables, the variables common to all rosters (see Table 3 above) have a common nomenclature. This means that the marital status of the child that occupies position 22 will be stored in the variable WMAR_R22 and the marital status of the brother that occupies position 61 will be stored in the variable WMAR_R61. This occurs the same way with gender (WSEX_R##), age (WEDA_R##), relationship (WPAR_R##), and place of residence (WDON_R##).

Though each roster's variables maintain their name, which corresponds to the question on the questionnaire, they also include the position of that person and, consequently, a reference to his/her primary roster. Therefore, the number of children of the child who occupies position 35 will be stored in WL9R35, and the number of children of the child who occupies position 22 will be stored in WL9R22.

The structure of these rosters is not homogeneous; that is, not all rosters compile the same information.

Variables of Position and Variables of Duplication

The sections of the rosters include the so-called variables of position, whose names follow the sequence WPOS1 to WPOS18 for the household, WPOS21 to WPOS53 for the children, WPOS61 to WPOS66 for the siblings, WPOS71 to WPOS72 for the parents, and WPOS84 to WPOS88 for other persons. These variables have a value corresponding to their position (for example: WPOS23=23) if a person exists in that position, or they are empty if no one in that position exists.

Since the household roster is not exclusive of other rosters, a person could belong to the household roster and to any other of the rosters. This duplication can only occur between the household roster and any other of the rosters. A child who resides in the home could occupy position 5 in the household roster (WPOS5=5) and the position 22 in the children roster (WPOS22=22). This is an example of a duplicated record.

Each roster, even if it has variables in common with other rosters, as was presented in Table 3, contains some information not included in the others (see sections C and L in the questionnaire.) For this reason, it is convenient to understand whether a person is duplicated and, as such, has additional information in another roster. The variables of duplication are used to identify if a record is duplicated.

The duplication variables follow the same nomenclature as those of position (WDD1 to WDD18, WDD21 to WDD53, WDD61 to WDD66, WDD71 to WDD72 and WDD84 to WDD88). These variables can have a value of zero (0) when the person is not duplicated, a value corresponding to the position in which the duplicated person is found, or a missing value if there is no one in this record.

To follow the previous example of the son who is found in position 5 of the household and is duplicated in position 22 of the children roster, the variable WDD5 would have a value of 22 (WDD5=22) and the variable WDD22 would have a value of 5 (WDD22=5).

Utilization of Rosters' Data

The information in these rosters can be used for various purposes. One of them is the analysis of the family arrangements and family structure of the interviewee, and his/her network of support. Also, it can be used for the purpose of identifying characteristics of certain people with whom the interviewee has contact. This task implies a certain complexity and requires explanation.

There are three sections in the questionnaire: section I (Functional Status), section M (Transfers), and section N (Labor History and Income), that use the information recorded in the rosters. Basically, in certain parts of these sections, the interviewee is asked about the assistance he/she receives from or provides to others. The interviewee who declared that he/she received or provided assistance for a given activity, indicates who is the person who assists him/her or whom among those in the rosters, is the recipient of his/her help. This information is codified using the number of position of that person in the rosters. If the selected person is duplicated, for example in the household roster and in the children roster, in other words, a child who lives in the home, the number indicating his/her position in his/her primary roster, the household roster, will appear. The following example illustrates this situation:

If an interviewee indicates having difficulty preparing meals for him/herself (WI4=1 or WI4=3), he/she will be asked about the person who helps in the preparation of meals (WI4A). The value of the variable WI4A indicates the position of that person in the rosters. In this way, if WI4A=4, the person who helps the interviewee would also be the one in the fourth position of the household roster, his/her age would be WEDA_R4, his/her gender would be WSEX_R4, his/her marital status would be WMAR_R4, his relation to the interviewee WPAR_R4, and his/her place of residence, in this example would be the home. But, in addition to these common variables, there are other particular variables of the household roster in which there is additional information about the person. In this case all of these variable names will have the ending R4.

If we observe the variable WDD4, the variable of duplication for that member of the household, and it has a value of 22, this would mean that the person is duplicated in the children roster in position 22. In this record of the children roster there is also additional information about that person. In this case all these variable names will end in R22.

Even though the norm is that the interviewee would choose from among the members in the rosters, in the previously mentioned questions in sections I, M, and N, the interviewee was also given the opportunity to identify other people who might not be included in any of the rosters. In those cases some information is collected for those new persons (gender, age, relationship with the interviewee, education, marital status and place of residence). If that new person added to the rosters occupies the position 85, that information will be stored in the following variables: wsex_r85, weda_r85, wpar_r85, wesc_r85, wmar_r85 and wdon_r85.

Calculated Variables

In the data base there is a group of calculated variables⁴ that present important data from some of the sections of the questionnaire. These variables offer the user more clear or summarized information on some of the specific data regarding the interviewees. Some of the variables are the product of the combination of two or more variables, while others simply the re-codification of one variable of the questionnaire. Tables 6 and 7 list the constructed variables for both the target and the spouse, indicating the section of the questionnaire to which they refer, its position in the database, its name, and its description. As with the original variables of the questionnaire of the spouse, his/her calculated variables begin with the letter “yw”.

Table 6. List of calculated variables of the target in the database

Variable	Description
tipo_e	TYPE OF INTERVIEW TARGET (CALCULADA)
factort	WEIGHTING FACTOR (CALCULADA)
wyage	AGE AT DEATH TARGET (calculada)
wb9t	RESPONSE TO THE PROVERB IS CORRECT (CALCULADA)
wfinscr	FINAL MINIMENTAL SCORE TARGET (CALCULADA)
wpasomin	MINIMENTAL PASS/FAIL TARGET (CALCULADA)
wbpx1	PROXY CANDIDATE LIVES WITH THE TARGET (CALCULADA)
wbpx2	PROXY CANDIDATE LIVED WITH THE TARGET (INST. OR DECEASED) (CALCULADA)
wprxyage	PROXY-TARGET AGE (CALCULADA)
wprxysex	PROXY-TARGET GENDER (CALCULADA)
wprxyr_1	GENERAL RELATIONSHIP WITH PROXY-TARGET (CALCULADA)
wprxyr_2	SPECIFIC RELATIONSHIP WITH PROXY-TARGET (CALCULADA)
wqjefe	IDENTIFIES THE HOUSEHOLD HEAD (CALCULADA)
wpos1- wpos18	POSITION IN THE HOUSEHOLD ROSTER (C1)
wdd1- wdd18	INDICATES DUPLICATED HOUSEHOLD MEMBER IN OTHER ROSTER
wdepre_y	DEPRESSION SCALE SCORE (TARGET) (wg171-wg185)
wl1numt	TOTAL NUMBER OF CHILDREN (CALCULADA)
wl1num	TOTAL NUMBER OF CHILDREN ALIVE (CALCULADA)
wpos21- wpos53	POSITION IN THE CHILDREN ROSTER (H21)
wdd21- wdd53	INDICATES POSITION OF SON/DAUGHTER DUPLICATED IN HOUSEHOLD ROSTER
wpos61- wpos66	POSITION IN SIBBLINGS ROSTER (N61)
wdd61- wdd66	INDICATES POSITION OF BROTHER/SISTER DUPLICATED IN HOUSEHOLD ROSTER
wpos71- wpos72	POSITION IN PARENTS ROSTER (P71)
wdd71- wdd72	INDICATES POSITION OF PARENT DUPLICATED IN HOUSEHOLD ROSTER
wpos84- wpos88	POSITION IN OTHERS ROSTER (O84)
wurbmi	MORBID OBESITY (CALCULADA)
wget_up	RESULT OF THE GET UP & GO (CALCULADA)

⁴ All the calculated variables include in their label the word CALCULADA.

Table 7. List of calculated variables of the spouse in the database

Variable	Description
ytipo_e	TYPE OF INTERVIEW SPOUSE (CALCULADA)
ywyage	AGE AT DEATH SPOUSE (calculada)
ywb9c	RESPONSE TO THE PROVERB IS CORRECT (CALCULADA)
ywfinscr	FINAL MINIMENTAL SCORE SPOUSE (CALCULADA)
ywpasomi	MINIMENTAL PASS/FAIL SPOUSE (CALCULADA)
ywbpx1	PROXY CANDIDATE LIVES WITH THE SPOUSE (CALCULADA)
ywbpx2	PROXY CANDIDATE LIVED WITH THE SPOUSE (INST. OR DECEASED) (CALCULADA)
ywpxyage	PROXY-SPOUSE AGE (CALCULADA)
ywpxysex	PROXY- SPOUSE GENDER (CALCULADA)
ywpxyr_1	GENERAL RELATIONSHIP WITH PROXY-SPOUSE (CALCULADA)
ywpxyr_2	SPECIFIC RELATIONSHIP WITH PROXY- SPOUSE (CALCULADA)
ywdeprey	DEPRESSION SCALE SCORE (SPOUSE) (ywg171-ywg185)
ywurbmi	MORBID OBESITY (CALCULADA)
ywget_up	RESULT OF THE GET UP & GO (CALCULADA)

Errors and Limitations Identified in Data

There are some errors or known limitations in the database that the users should consider:

1. In Section G, where personal information from household members is collected (positions 1 to 20) some variables were skipped for those household members already in the first interview. Those members are located in positions 1 to 10, being the target the number one. Missing questions for those members were those about the ability to read and write (WC8R## and WC9R## respectively), educational level (WESC_R##) and age (WEDA_R##). The last one stores the calculated age at the time of the second interview, considering the date of the first one. Because a programming error the age considered to skip the question about education was that of the first interview. As consequence those being three years old or younger were not asked about their educational level. Although for some cases the information was obtained by phone, for others is unknown. In the database these cases are properly identified with the value -3.
2. In Section G on Health there is a series of gender related questions (WG132 to WG142 for female respondents about, breast self-exam, hysterectomy and hormones consumption, and WG143 to WG147 for male respondents). Because a programming error, some respondents did not answer the appropriate questions for their gender. In many cases the information was gathered by the interviewer using a copy of the questionnaire, recorded on paper and corrected in the electronic version of the interview. Those cases with missing information are properly identified in the database with the value -3.
3. In Section K the respondent is asked about health insurance. The beginning of the fieldwork in 2006 matched the release of the new Medicare Part D (prescription drug coverage) in Puerto Rico. That fact caused some confusion about how to code the type of health insurance, specially Medicare, of the respondent at the time of the interview. Because of that the interviewers were instructed to gather detailed information for question WK1INS (variables WK1INS_1, WK1INS_2, WK1INS_3

and WK1INS_4). Afterwards that information was reviewed and the data coded accordingly. This situation applies to both target and spouse interviews.

4. In Section L, where data about the children is collected (positions 21 to 53) two different errors were identified. The first one is a programming error and is related with age and the skip patterns in the questions about education (WESC_R##), marital status (WMAR_R##) and number of children (WL9R##). This error affected only to children already listed in the first interview (positions 21 to 40). The second one was interviewer related and occurred while identifying if the child lived with the target. After gathering the child's data the interviewer confirmed if that child was a household member, that way a link was established with the household roster data and unnecessary duplication of data was avoided. If by error the child was identified as being a household member some information was lost: age (WEDA_R##), place of residence (WDON_R##), educational level (WESC_R##) and marital status (WMAR_R##). We call this an entity duplication error in the rosters. For many cases the information was obtained by phone, for others is unknown. In the database, these cases are properly identified with the value -3.
5. In Section L, where data about the siblings is collected (positions 61 to 66) three different errors were identified. One of them was the same error described for the children in point 4. Because of this error the interviewer wrongly identified a sibling as a household member and those questions common to both rosters were not asked. This problem caused the lost of data about place of residence (WDON_R##), marital status (WMAR_R##) and educational level (WESC_R##). The other two errors affected the skip patterns for the total number of siblings (WL20A) and the reason why the target did not maintain a relationship with his/her siblings as closest as in the first interview (WL20CR61, WL20CR62, WL20CR63). For many cases the information was obtained by phone, for others is unknown. In the database, these cases are properly identified with the value -3.
6. In Section I on Functional Status the target's interviews in slightly different from that of the spouse since the spouse's has not rosters. As a consequence the spouse cannot identify the exact person he/she gives help to or receives help from, but if he/she receives/gives any help. Because an initial programming error some of the spouses did not answer the questions about the frequency of the help received (YWI1AB, YWI2AB, YWI3AB, YWI4AB, YWI5AB, YWI6AB, YWI7AB, YWI9AB, YWI10AB, YWI11AB, YWI12AB, YWI13AB y YWI14AB). In the database, these cases are properly identified with the value -3.
7. In Section M on Transfers an interviewer related error was identified. Questions WM10M2 and WM10O2 ask about the son/daughter being helped by the target taking care of the grandsons/granddaughters. In some cases the questions was confuse and the target identified the grandson/granddaughter instead of the son/daughter. The interviewer used the grandson/granddaughter code while the son/daughter was the right one. Cases where this error cannot be corrected were properly identified in the database with the value -3.
8. In Section U about Anthropometric Measurements was a programming error in question WU5C. As a result of that error 18 cases have missing information in variable WU5D. Those cases were participants in wheelchairs at the moment of the

interview and skipped directly to question WU5 (knee height). The same applies to 7 spouse cases in question YWU5D. In the database, these cases are properly identified with the value -3.

Codebook

The codebook book is available in the document CODBOOK2.PDF. For each one of the variables in the database there is the following information: name of the variable, label of the variable, position in the database, text of the corresponding question in the questionnaire, type of variable, size of variable, missing values, codes and labels used, and distribution of frequencies and commentaries, the latter for some complex variables. The distribution of frequencies is not presented for those variables that have more than 20 categories, except in the case of important variables such as the age of the interviewee.