

Original

Análisis del sesgo de selección en el piloto de un estudio longitudinal sobre envejecimiento en España

Ángel Rodríguez Laso^{a,b,*}, Elena Urdaneta Artola^{a,b}, Miguel de la Fuente Sánchez^c,
Eva Galindo Moreno^c, José Javier Yanguas Lezáun^{a,b} y Vicente Rodríguez Rodríguez^{b,d}

^a Fundación INGEMA-Instituto Gerontológico MATIA, San Sebastián/Donosti-Madrid, España

^b Unidad Asociada "El Proceso de Envejecimiento" (INGEMA-CSIC), San Sebastián/Donosti-Madrid, España

^c Demométrica, Madrid, España

^d Instituto de Economía, Geografía y Demografía, Centro de Ciencias Humanas y Sociales, Consejo Superior de Investigaciones Científicas, Madrid, España

INFORMACIÓN DEL ARTÍCULO

Historia del artículo:

Recibido el 6 de septiembre de 2012

Aceptado el 22 de noviembre de 2012

On-line el 3 de enero de 2013

Palabras clave:

Sesgo de selección
Negativa a participar
Estudios transversales
Censos
Envejecimiento

R E S U M E N

Objetivos: Demostrar que la obtención de una muestra nacional probabilística para un estudio longitudinal sobre envejecimiento en España está expuesta a sesgos de selección. Cuantificar las pérdidas que se producen al administrar el cuestionario en distintas fases.

Métodos: Estudio transversal sobre población española no institucionalizada de 50 años o más de edad realizado entre 2010 y 2011. Por medio de regresiones logísticas multivariadas se comparan características de las secciones censales de residencia de los que acceden ($n = 5813$) o no ($n = 7023$) a formar parte del marco muestral, de los que acceden ($n = 1677$) o no ($n = 2875$) a participar en el estudio, y de los individuos que contestan ($n = 1398$) o no ($n = 346$) a un cuestionario presencial posterior a uno telefónico. Además, se estudian los motivos de las negativas.

Resultados: La mala salud y la discapacidad son los motivos específicos más referidos para no formar parte del marco o no participar en el estudio (14,4% y 27,9%, respectivamente). En ambos casos, la negativa es más frecuente en las secciones censales de menor nivel socioeconómico o ubicadas en Cataluña, Guipúzcoa o Vizcaya. Los individuos mayores de 80 años responden con menos frecuencia al cuestionario presencial. En cada fase de recogida de información se pierde entre un 8,6% y un 18,4% de efectivos.

Conclusiones: Un muestreo probabilístico en puntos de muestreo elegidos por conveniencia permitiría dedicar más recursos a aumentar las tasas de respuesta en los colectivos menos participativos. Se propone la concentración de las preguntas en un solo cuestionario más breve previo a la extracción de sangre.

© 2012 SESPAS. Publicado por Elsevier España, S.L. Todos los derechos reservados.

Analysis of selection bias in the pilot study of a longitudinal study on aging in Spain

A B S T R A C T

Keywords:

Selection bias
Refusal to participate
Cross-sectional studies
Censuses
Aging

Objectives: To demonstrate that selection of a probabilistic sample at a national level for a study of aging in Spain is subject to selection bias. To quantify the losses produced after each phase of the administration of a questionnaire.

Methods: We performed a cross-sectional study of the Spanish community-dwelling population aged 50 years or older between 2010 and 2011. Through multivariate logistic regressions, the characteristics of the census tract of the patients' residence were compared between those who agreed ($n = 5,813$) or refused ($n = 7,023$) to be included in the sampling frame and between those who agreed ($n = 1,677$) or refused ($n = 2,875$) to participate in the study. The individual characteristics of persons who responded ($n = 1,398$) or refused to respond ($n = 346$) to a face-to-face questionnaire administered after a telephone interview were also compared. In addition, the reasons for refusal were studied.

Results: The most frequent specific reasons for refusing to be included in the sampling frame or to participate in the study were poor health and disability (14.4% and 27.9%, respectively). In both cases, refusal was more frequent in the census tracts of districts with a lower socioeconomic level or those located in Catalonia, Guipúzcoa or Biscay. Individuals older than 81 participated less frequently in the face-to-face questionnaire. Between 8.6% and 18.4% of participants were lost at each stage of information retrieval.

Conclusion: Probabilistic sampling in sampling points chosen by the researchers would allow more resources to be devoted to increasing response rates among the groups who are less likely to participate. Questions should be concentrated in only one shorter questionnaire, administered before blood extraction.

© 2012 SESPAS. Published by Elsevier España, S.L. All rights reserved.

* Autor para correspondencia.

Correo electrónico: angel.rodriguez@ingema.es (Á. Rodríguez Laso).

Introducción

El Estudio Longitudinal «Envejecer en España» (ELES) es un proyecto interdisciplinario que pretende el seguimiento, durante 20 años, de cohortes de población españolas nacidas antes de 1960 para analizar su proceso de envejecimiento mediante la recogida de información periódica a una muestra representativa (<http://proyectoeles.es/>). Para probar la metodología del estudio longitudinal se ha realizado un estudio piloto de grandes dimensiones.

Aparte de los objetivos longitudinales, el ELES tiene también interés en obtener estimaciones transversales nacionales, lo que llevó a elegir para el estudio piloto un muestreo probabilístico estratificado con cobertura de todo el país. Un marco muestral tan amplio dificulta el trabajo de campo, lo que puede disminuir las tasas de respuesta. Si éstas son suficientemente bajas y si hay diferencias entre los que responden o no, se producirán sesgos de selección. Para estudiar las diferencias es necesario caracterizar a los no participantes a partir de fuentes de información alternativas a los cuestionarios administrados¹⁻³.

Por otro lado, el carácter interdisciplinario del ELES requiere numerosas preguntas para cubrir todas las áreas de interés. Para no sobrecargar a los participantes con largas entrevistas, se decidió administrar el cuestionario en distintos momentos y con diferentes técnicas. La pérdida de individuos en cada fase también puede introducir sesgos.

El objetivo principal de este artículo es explorar los sesgos de selección existentes en el estudio piloto del ELES. Para ello se caracteriza a los no respondedores según: 1) sus motivos para no participar, 2) las características sociodemográficas de las secciones censales donde viven, y 3) las variables del cuestionario telefónico de los que no participan posteriormente en un cuestionario administrado por encuestador con ayuda de ordenador (CAPI). Como objetivo secundario se plantea evaluar cuántas pérdidas en efectivos supone dividir un cuestionario general en distintas fases y tipos.

Métodos

Diseño de muestreo

La metodología del estudio piloto del ELES se encuentra detallada en otra publicación⁴. El diseño ha sido polietápico con estratificación por comunidad autónoma y tamaño del hábitat de residencia de las unidades muestrales de primera etapa (municipios-secciones censales). La estratificación ha sido proporcional a la población mayor de 50 años de los estratos, aunque la del País Vasco está sobrerrepresentada. Las secciones censales se han elegido por muestreo aleatorio proporcional a sus tamaños poblacionales. Las unidades de segunda etapa (hogares) y las que proporcionan la información (un individuo por hogar) se han seleccionado aleatoriamente hasta alcanzar una representación proporcional en los estratos en cuanto a sexo y grupos de edad decenales. El marco muestral lo ha constituido el directorio de hogares con teléfono, en el que consta la sección censal.

Contacto y recogida de información

Puesto que no existe un marco muestral nacional de personas por encima de una determinada edad fácilmente accesible, ha sido necesario construirlo. Para ello se realizaba una primera llamada en la que se averiguaba si vivía en el domicilio seleccionado alguna persona de 50 o más años de edad. De ser así, se preguntaba si querían recibir información adicional del estudio por correo. En caso de negativa, se recogían dos motivos. Si accedían, se anotaba el número

de personas por sexo y grupos de edad decenales mayores de esa edad que vivían en el hogar, y se enviaba la información.

Posteriormente se realizaba una segunda llamada y se pedía ya la participación de una persona de sexo y edad concretos, seleccionada de forma aleatoria. En caso de negativa del hogar, se sustituía por otro de la misma sección censal. Si la negativa era del individuo, podía intentarse la sustitución por otra persona del domicilio del mismo grupo de edad y sexo. Por la estructura de los hogares, esta situación se ha podido dar en contadas ocasiones. Se requerían también dos motivos para esta negativa. Si el individuo accedía a participar, se le realizaba la entrevista telefónica y se concertaba una visita de la enfermera, que realizaba una extracción de sangre y administraba el cuestionario cognitivo *Mini-Mental*⁵. Si las puntuaciones de unas partes específicas del *Mini-Mental* alcanzaban unos umbrales preestablecidos, dejaba un cuestionario autocumplimentado. A esta visita seguía otra, también concertada, de un encuestador para realizar el CAPI, que podía ser contestado por el cuidador si la persona seleccionada estaba incapacitada para hacerlo. El encuestador recogía el cuestionario autocumplimentado, ayudando al entrevistado a completarlo si no estaba relleno y el individuo quería contestarlo.

Se empezó a llamar a los hogares en diciembre de 2010 y el trabajo de campo finalizó en septiembre de 2011.

El estudio piloto del ELES fue aprobado por el Subcomité de Ética del Consejo Superior de Investigaciones Científicas. Todos los participantes firmaron consentimientos informados para cada fase.

Variables explicativas para el análisis de la negativa a formar parte del marco muestral y a participar en el cuestionario telefónico (variables dependientes)

De las que ha calculado el Instituto Nacional de Estadística a partir del censo de 2001 para las secciones censales⁶, se han empleado los porcentajes de habitantes con 80 y más años de edad respecto a la población mayor de 65 años, de hogares unipersonales, de hogares que mencionan problemas de delincuencia o vandalismo en la zona, y de personas de referencia con estudios secundarios o universitarios; la condición socioeconómica (media para los habitantes de la sección de una puntuación que combina su ocupación, actividad y situación profesional; las puntuaciones originales se han multiplicado por 10 y en el modelo multivariado se han introducido centradas en la media, para evitar la colinealidad con su término cuadrático); y el estado medio de los edificios destinados principalmente a viviendas (promedio de los valores asignados a los estados: ruinoso 0, malo 2, deficiente 4, bueno 10; por su marcada asimetría negativa se ha utilizado el logaritmo de la puntuación reflejada, por lo que valores mayores de la variable transformada significan «peor estado»).

Además, se emplean el tamaño del hábitat de residencia (menos de 10.000 habitantes, entre 10.000 y 100.000, entre 100.000 y 500.000, más de 500.000), la comunidad autónoma (con las tres provincias vascas por separado dada su sobrerrepresentación) y la banda horaria en que se contactó con el hogar (de 10 a 13:30, de 13:31 a 15:30, de 15:31 a 20:00 y de 20:01 a 22 horas).

Variables explicativas para el análisis de la negativa a participar en el CAPI (variable dependiente)

De las del cuestionario telefónico previo se seleccionaron la edad, el sexo, el nivel de estudios, si la persona vive sola, el tamaño del hábitat, la comunidad autónoma de residencia y el ítem de satisfacción con el estado de salud del *Personal Wellbeing Index*⁷.

Análisis estadístico

Se realizaron comparaciones de medias o pruebas de ji al cuadrado con corrección de Rao entre los que respondieron afirmativamente o negativamente a la primera llamada, los que aceptaron o no contestar al cuestionario telefónico, y los que participaron o no en el CAPI. Se realizaron regresiones logísticas binarias multivariadas de la aceptación de la participación en el marco, de la participación en el cuestionario telefónico y de la negativa a participar en el CAPI, introduciendo inicialmente todas las variables y eliminando manualmente una a una todas aquellas que no fueran significativas ($p \geq 0,05$) y cuya exclusión no produjera un cambio mayor del 10% ni un aumento de los intervalos de confianza de las *odds ratios* (OR) de las restantes (modelo final). Se comprobó el supuesto de linealidad en el *logit*. En el modelo para la negativa al CAPI se probaron las interacciones del sexo con la edad, el nivel de estudios y el hecho de vivir solo, y de esto último con la edad.

Todos los análisis se han realizado con Stata 9.2, definiendo las muestras como complejas.

Resultados

Flujo de contacto y recogida de información

Se representan en la **figura 1**. El esfuerzo realizado en el trabajo de campo ha sido importante, ya que sólo un 14,7% de los hogares hallados con mayores de 50 años han participado en la entrevista

telefónica, y únicamente el 8,6% han pasado por todas las fases de recogida de información.

Motivos para negarse a formar parte del marco muestral

Los proporcionaron 7078 individuos (98,1% de los que se negaron). Los más frecuentes fueron inespecíficos, como falta de interés o tiempo (67,7%), seguidos por enfermedad o discapacidad para oír, entender o contestar (14,4%), y reticencias frente a la metodología de encuestas (6,1%). El resto de las causas no superó el 5%.

Motivos para no participar en el cuestionario telefónico

Proporcionaron información 2648 individuos (89,7% de los que se negaron). Nuevamente los más declarados fueron motivos inespecíficos (55,1%) y problemas de salud o discapacidad (27,9%); en tercer lugar, ausencias (6,7%).

Variables de los hogares asociadas a aceptar formar parte del marco muestral

Se presentan en las cuatro primeras columnas de la **tabla 1**. La muestra total queda reducida a 12.836 hogares (frente a los 13.242 iniciales) porque se ha excluido a los de secciones nuevas inexistentes en 2001 (**fig. 2**).

Aceptar ser parte del marco muestral se asocia a un mayor nivel educativo, menor proporción de hogares unipersonales, mayor

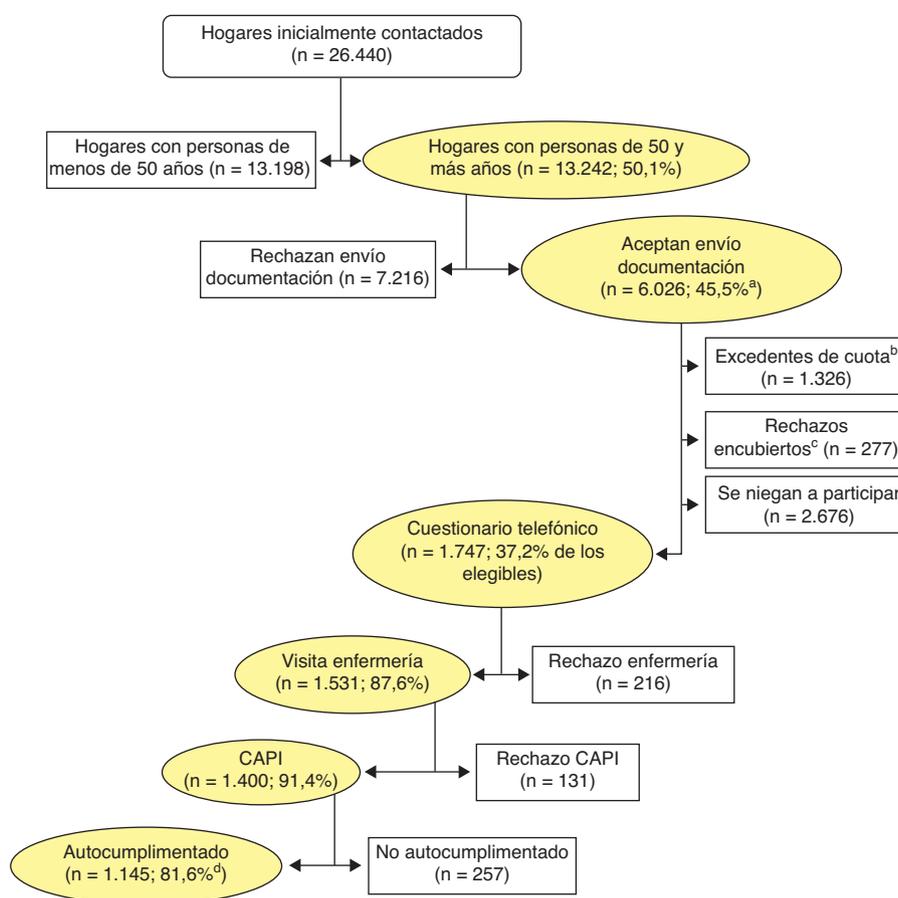


Figura 1. Flujo de contacto y participación en el estudio piloto del ELES.

^a Los porcentajes indican la participación respecto a la fase previa salvo que se especifique otra cosa.

^b Excedentes de cuota son los hogares que a pesar de estar interesados en participar no han sido llamados porque eran sobrantes una vez cubiertas las necesidades de efectivos por sexo y grupos de edad decenales. Para el cálculo del porcentaje de participación en el cuestionario telefónico se les ha excluido del denominador.

^c Rechazos encubiertos son los hogares que no vuelven a contestar al teléfono o posponen reiteradamente la participación.

^d Dos personas contestaron el autocumplimentado pero no el CAPI, luego el porcentaje se ha calculado para n = 1.143.

Tabla 1
Características (porcentajes o medias de variables) de los hogares que rechazan y aceptan formar parte del marco muestral, y que rechazan y aceptan participar en el cuestionario telefónico del estudio piloto del ELES

	Rechazan formar parte del marco (IC95%) (n = 7023)	Aceptan formar parte del marco (IC95%) (n = 5813)	p	Rechazan participar en el cuestionario telefónico (IC95%) (n = 2875)	Aceptan participar en el cuestionario telefónico (IC95%) (n = 1677)	p
Personas de 80 y más años en sus secciones (%)	20,2	20,3	0,820	20,2	20,4	0,145
Hogares unipersonales en sus secciones (%)	19,0	18,6	0,024	18,4	18,6	0,462
Personas con estudios secundarios o universitarios en sus secciones (%)	60,0	61,1	0,028	60,4	62,4	0,001
Condición socioeconómica media de sus secciones	10,1 (10,0-10,2)	10,1 (10,0-10,2)	0,815	10,0 (9,9-10,2)	10,2 (10,0-10,3)	<0,001
Hogares que declaran problemas de vandalismo o delincuencia en la zona en sus secciones (%)	20,1	21,1	0,047	20,4	21,1	0,223
Logaritmo de la puntuación media de los edificios reflejada en sus secciones (rango 0-1,78)	0,46 (0,41-0,51)	0,45 (0,40-0,50)	0,400	0,46 (0,41-0,51)	0,45 (0,40-0,50)	0,492
Municipios <10.000 hab.	22,0	22,0	0,375	21,6	21,6	<0,001
Municipios 10.000-100.000 hab.	41,5	39,5		41,6	36,7	
Municipios 100.000-500.000 hab.	26,1	27,3		27,9	28,5	
Municipios >500.000 hab.	10,4	11,3		8,9	13,1	
Andalucía	7,6	11,3	<0,001	9,7	12,1	0,001
Aragón	1,0	0,6				
Asturias	1,7	1,9				
Baleares	2,0	1,2				
Canarias	2,4	3,1		2,1	2,9	
Cantabria	0,9	1,2				
Castilla-La Mancha	3,4	3,1		3,5	3,2	
Castilla y León	4,3	4,7		4,0	5,7	
Cataluña	14,4	12,4		13,3	12,0	
Extremadura	1,8	3,1		3,0	2,7	
Galicia	6,5	7,3		6,5	5,7	
Madrid	10,0	11,3		10,1	11,4	
Murcia	1,1	1,9				
Navarra	0,9	1,3				
Rioja	1,2	0,6				
Valencia	6,2	7,2		5,6	7,8	
Álava	3,3	3,6		3,9	4,3	
Guipúzcoa	11,8	8,4		10,2	8,5	
Vizcaya	19,6	15,9		20,3	14,9	
Resto de España ^a				7,9	8,7	
Llamada entre las 10 y las 13:30 h	3,4	2,7	<0,001	3,8	2,2	0,024
Llamada entre las 13:31 y las 15:30 h	30,5	37,3		37,3	37,3	
Llamada entre las 15:31 y las 20 h	62,5	57,7		56,3	57,8	
Llamada entre las 20:01 y las 22 h	3,6	2,4		2,7	2,6	

IC95%: intervalo de confianza del 95%.

^a En el análisis de la participación se han agrupado el resto de las comunidades autónomas por la escasa muestra en cada una de ellas.

percepción de vandalismo en la sección censal y a recibir la llamada en la hora de la comida. Hay diferencias por comunidades autónomas. En el modelo multivariado (cuatro primeras columnas de la tabla 2), la asociación de una mayor participación con una mejor condición socioeconómica desplaza a la del nivel educativo y además tiene un componente cuadrático que la incrementa. Disminuye la aceptación, también de forma cuadrática, en las secciones más envejecidas. Se mantienen el resto de las asociaciones. Es más frecuente la aceptación en Extremadura, Murcia, Andalucía, Cantabria y Castilla y León, y menos en La Rioja, Baleares, Guipúzcoa, Cataluña, Aragón y Vizcaya.

Variables asociadas a participar en el cuestionario telefónico

Aparecen en las tres últimas columnas de la tabla 1. La muestra queda reducida a 4552 individuos después de excluir, de entre los que aceptan participar en el marco muestral, a los excedentes de cuota con variables en cuanto a sección censal (fig. 2).

Responder al cuestionario telefónico se asocia con el nivel educativo y socioeconómico de la sección, vivir en ciudades grandes y en determinadas comunidades autónomas, y ser llamado después de comer. En el análisis multivariado (tres últimas columnas de la tabla 2) persiste la asociación con la condición socioeconómica. La participación es más frecuente en Andalucía y Castilla y León, y menos en Cataluña y País Vasco.

Variables individuales asociadas a la negativa a participar en el CAPI

Tres individuos no evaluaron su satisfacción con su estado de salud y fueron excluidos del análisis (n = 1744). El bloque de interacciones resultó no significativo (p = 0,210). Hay una gran desproporción de personas de 80 y más años de edad entre los no participantes, que además tienen menos estudios, viven solos con más frecuencia y están menos satisfechos con su salud (tabla 3). En la regresión múltiple (tabla 4) sólo es significativa la asociación con la edad y no tener estudios secundarios.

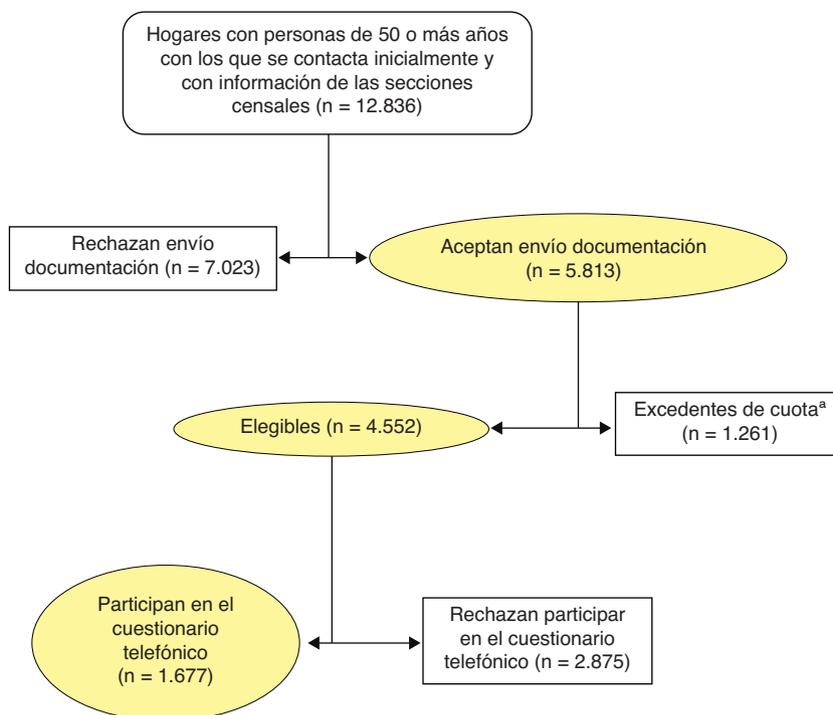


Figura 2. Tamaños de muestra para los análisis de aceptación de formar parte del marco muestral y de participar en el cuestionario telefónico del estudio piloto del ELES. ^aExcedentes de cuota son los hogares que a pesar de estar interesados en participar no han sido llamados porque eran sobrantes una vez cubiertas las necesidades de efectivos por sexo y grupos de edad decenales.

Discusión

Durante el proceso de selección de la muestra del estudio piloto del ELES se ha dado un elevado porcentaje de negativas que se han asociado de forma consistente con un menor nivel socioeconómico, vivir en Cataluña, Guipúzcoa y Vizcaya, y la edad. Menos de la mitad de los hogares hallados con mayores de 50 años de edad aceptaron formar parte del marco muestral. Sólo han completado todas las fases de recogida de información el 24% de los individuos elegibles.

Aunque son distintos los procesos que subyacen a no aceptar la inclusión en el marco muestral tras una llamada “fría” por teléfono (al no contar con ningún marco que permitiera el envío de información previa) y a rechazar participar tras recibir información detallada, ambos coinciden en su asociación con un menor nivel socioeconómico del hogar. La literatura confirma este hallazgo⁸⁻¹², salvo un artículo reciente con algún fallo metodológico¹³.

Las negativas a participar en los dos contactos iniciales han sido mayores en Cataluña y País Vasco. Podría deberse a que el contacto inicial se hizo desde Madrid en castellano, pero la información postal para el segundo contacto sí se envió en catalán y euskera (con cartas de apoyo de la Generalitat y el Gobierno Vasco), y las asociaciones son muy parecidas. Otras encuestas nacionales, como la Encuesta de Población Activa¹⁴ o la Encuesta Nacional de Salud de 2006 (ENS-2006)¹⁵, confirman nuestros resultados. Puede haber sucedido que los habitantes de estas comunidades se sientan menos identificados con los objetivos de un estudio de ámbito nacional.

Vivir en una sección muy envejecida se asocia con el rechazo a ser incluido en el marco (que puede interpretarse como un rechazo a las encuestas en general), y tener más de 81 años de edad se asocia a ser encuestado en el domicilio. Nuestros resultados concuerdan con los de Chou et al.¹⁶, pero no con los de otros autores, lo que puede deberse a quién promueve la encuesta¹⁷, al método de recogida de la información¹⁷ o a cuándo se realizó el estudio¹⁸. En otro artículo los resultados no están desagregados por grupos de edad⁹. Las secciones con más hogares unipersonales han sido menos propensas a formar parte del marco. Está descrito que es difícil captar a

este tipo de hogares¹⁹. El método empleado de llamar por teléfono antes de enviar al encuestador puede haber evitado que las secciones con mayor percepción de vandalismo hayan participado menos, resultado tradicionalmente asociado a la entrevista personal²⁰. De hecho, han participado más, aunque las diferencias son de escasa relevancia.

Desde la perspectiva de la salud, el sesgo más preocupante es el condicionado por la menor participación de las personas con mal estado de salud o discapacidad. Entre un 14% y un 28%, dependiendo de la fase de contacto, han alegado esta circunstancia como motivo para no participar. Como comparación, sólo el 4,6% de los participantes en el estudio piloto ha señalado que su salud es mala o muy mala, frente al 17% en prácticamente el mismo segmento de edad de la ENS-2006¹⁵. Sin embargo, en el análisis multivariado de participación en el CAPI no hay asociación con el estado de salud subjetivo. Ayudará a esclarecer la existencia o no de un sesgo el estudio de la mortalidad de participantes y no participantes, que la literatura refiere mayor entre estos últimos^{16,21}.

En cuanto a las pérdidas en las fases de recogida de información, el mayor porcentaje se ha producido en la respuesta al cuestionario autoadministrado, pero esto se ha debido principalmente a sus estrictos criterios de cumplimentación. Si se excluye a los que no los reunían, la participación ha sido del 91,9%, comparable a la participación en el CAPI. No se trata de pérdidas muy altas, pero acumuladas suponen que sólo el 65,5% de los inicialmente entrevistados tienen información de todas las variables del estudio (si se excluye el cuestionario autoadministrado, el 80%). La fase en que se han producido más pérdidas por negativa ha sido en la visita de enfermería, probablemente porque es la primera que requería dejar entrar en el domicilio a un desconocido que además iba a realizar un procedimiento molesto. Con todo, es de destacar positivamente que sólo se pierda un 12,4% en esta transición.

La principal limitación de este trabajo es que no se diseñó específicamente para el estudio de la no respuesta. De todas formas, obtener información más detallada de los que se niegan a responder habría generado también negativas, lo que ya se aprecia

Tabla 2
Regresión logística de la aceptación de formar parte del marco muestral y de participar en el cuestionario telefónico del estudio piloto del ELES

	Modelo final ^a de la aceptación de formar parte del marco (n = 12.836)			Modelo final de la participación en el cuestionario telefónico (n = 4552)		
	OR	(IC95%)	p	OR	(IC95%)	p
Personas de 80 y más años en la sección (%)	1,00	(0,99-1,02)	0,534			
Cuadrado del porcentaje de personas de 80 y más años en la sección	1,00	(1,00-1,00)	0,015			
Hogares unipersonales en la sección (%)	0,99	(0,98-1,00)	0,005			
Condición socioeconómica media de la sección (centrada; rango: -4,2-2,8)	1,17	(1,11-1,23)	<0,001	1,25	(1,17-1,33)	<0,001
Cuadrado de la condición socioeconómica centrada	1,04	(1,02-1,06)	0,001			
Hogares que declaran problemas de vandalismo o delincuencia en la zona en la sección (%)	1,00	(1,00-1,01)	0,039			
Andalucía ^b	1,48	(1,23-1,78)	<0,001	1,59	(1,32-1,92)	<0,001
Aragón	0,70	(0,57-0,85)				
Asturias	1,09	(0,85-1,39)				
Baleares	0,56	(0,32-0,97)				
Canarias	1,46	(0,83-2,54)		1,50	(0,94-2,40)	
Cantabria	1,39	(1,04-1,85)				
Castilla-La Mancha	1,03	(0,89-1,19)		0,97	(0,74-1,27)	
Castilla y León	1,15	(1,01-1,30)		1,36	(1,07-1,73)	
Cataluña	0,68	(0,58-0,79)		0,67	(0,54-0,84)	
Extremadura	1,85	(1,55-2,20)		1,41	(0,94-2,13)	
Galicia	1,16	(0,99-1,35)		0,83	(0,64-1,07)	
Madrid	0,86	(0,73-1,01)		0,85	(0,71-1,02)	
Murcia	1,73	(1,33-2,25)				
Navarra	1,22	(0,59-2,51)				
Rioja	0,52	(0,34-0,81)				
Valencia	1,02	(0,75-1,38)		1,15	(0,92-1,45)	
Álava	0,91	(0,57-1,44)		0,83	(0,71-0,96)	
Guipúzcoa	0,66	(0,55-0,78)		0,69	(0,52-0,91)	
Vizcaya	0,81	(0,70-0,93)		0,71	(0,59-0,84)	
Resto de España ^c				1,02	(0,86-1,22)	
Llamada entre las 10 y las 13:30 h (ref. 15:31-20)	1,14	(0,88-1,46)	<0,001			
Llamada entre las 13:31 y las 15:30 h (ref. 15:31-20)	1,33	(1,22-1,44)				
Llamada entre las 20:01 y las 22 h (ref. 15:31-20)	0,77	(0,59-1,00)				
Valor p del modelo	<0,001			<0,001		
Valor p de la prueba de bondad de ajuste de Hosmer y Lemeshow	0,530		0,187			

OR: *odds ratio*; IC95%: intervalo de confianza del 95%.

^a Modelo final: obtenido tras eliminar manualmente una a una todas las variables inicialmente consideradas que no fueran significativas ($p \geq 0,05$) y cuya exclusión no produjera un cambio mayor del 10% ni un aumento de los intervalos de confianza de las OR de las restantes.

^b Las OR han sido calculadas respecto a la media de participación de las comunidades autónomas.

^c En el análisis de la participación se han agrupado el resto de las comunidades autónomas por la escasa muestra en cada una de ellas.

en una pregunta tan simple como el motivo de la no participación. Lo que hemos hecho para conseguir más información de los que se niegan, asignando a los hogares y a los individuos las características de las secciones censales donde se ubican, es un ejemplo de falacia ecológica²², pero para las variables socioeconómicas empleadas no parece descabellado pensar que hay una estrecha relación entre secciones, hogares e individuos, por lo menos en poblaciones de cierto tamaño. Por otro lado, la información del censo corresponde a 2001, así que hay una diferencia de 10 años con el piloto del ELES. No creemos que la posición relativa de la mayoría de las secciones censales haya cambiado mucho en una década. Otra limitación es que el marco muestral del estudio piloto del ELES, el listado de teléfonos, no cubre a toda la población. Sin embargo, de la Encuesta sobre Equipamiento y Uso de Tecnologías de la Información y Comunicación en los Hogares²³ puede estimarse que la cobertura por teléfono fijo en los mayores de 50 años puede superar el 92%. Por último, este trabajo se basa en el análisis de las negativas y no analiza la falta de contacto, que presenta peculiaridades distintas².

En definitiva, nuestros resultados constatan la dificultad de hacer participar en una encuesta sobre envejecimiento en España a los muy mayores, los de menor nivel socioeconómico y los que viven en Cataluña y País Vasco. Aunque se sustituya a los no respondedores, no puede asegurarse que los que han participado finalmente sean equiparables a los que rechazaron inicialmente. Algunos trabajos encuentran que los sesgos de no respuesta pueden ser pequeños aunque haya una elevada tasa de no respuesta y diferencias entre respondedores y no respondedores^{8,11,21}, pero utilizan como "población" de referencia la que ha contestado a encuestas previas con su propio sesgo de selección, por lo que sus conclusiones no pueden extenderse a la población general.

Una manera de disminuir el sesgo es aumentar la tasa de respuesta de los grupos que participan menos, lo que podría conseguirse, a igualdad de presupuesto, optando por un muestreo probabilístico en localidades elegidas por conveniencia donde sea más fácil conseguir un gran apoyo institucional y mediático. Es el modelo empleado por el estudio de envejecimiento canadiense²⁴. Además, esta opción disminuiría los costes y los sesgos no

Tabla 3
Características (medias o porcentajes de variables) de los individuos participantes y de los que se niegan a participar en el CAPI del estudio piloto del ELES

	Participantes (n = 1398)	Negativas (n = 346)	p
Edad de 50 a 60 años	34,8	24,0	
Edad de 61 a 70 años	29,5	27,2	<0,001
Edad de 71 a 80 años	23,5	23,4	
Edad de 81 y más años	12,2	25,4	
Mujeres (%)	55,4	57,8	0,372
Menos estudios que primarios	34,3	45,7	
Estudios primarios	22,4	20,2	<0,001
Estudios secundarios	21,3	12,7	
Estudios universitarios	22,0	21,4	
Viven solos (%)	15,1	22,8	0,001
Satisfacción con el estado de salud (rango 0-10)	7,1	6,8	0,013
Municipios <10.000 hab.	22,5	24,6	
Municipios 10.000-100.000 hab.	35,8	36,4	
Municipios 100.000-500.000 hab.	27,5	26,9	0,703
Municipios >500.000 hab.	14,1	12,1	
Andalucía	12,5	15,0	
Aragón	2,4	2,0	
Asturias	1,8	1,2	
Baleares	1,4	1,2	
Canarias	3,0	2,0	
Cantabria	1,0	1,4	
Castilla-La Mancha	3,1	2,6	
Castilla y León	5,6	4,9	
Cataluña	11,4	12,1	
Extremadura	2,7	2,3	0,842
Galicia	5,4	5,8	
Madrid	11,4	9,0	
Murcia	1,5	2,0	
Navarra	1,4	1,4	
Rioja	0,5	1,2	
Valencia	8,2	9,5	
Álava	4,2	3,8	
Guipúzcoa	8,5	6,4	
Vizcaya	13,9	16,2	

Tabla 4
Regresión logística de la no participación en el CAPI del estudio piloto del ELES

	Modelo final ^a (n = 1744)		p
	OR	(IC95%)	
Edad de 61 a 70 años (ref. de 50 a 60)	1,22	(0,83-1,79)	
Edad de 71 a 80 años (ref. de 50 a 60)	1,24	(0,84-1,83)	<0,001
Edad de 81 años o más (ref. de 50 a 60)	2,58	(1,77-3,76)	
Estudios primarios (ref. menos que primarios)	0,82	(0,59-1,14)	
Estudios secundarios (ref. menos que primarios)	0,56	(0,39-0,81)	0,022
Estudios universitarios (ref. menos que primarios)	0,91	(0,67-1,25)	
Valor p del modelo	<0,001		
Valor p de la prueba de bondad de ajuste de Hosmer y Lemeshow	0,962		

OR: *odds ratio*; IC95%: Intervalo de confianza al 95%.

^a Modelo final: obtenido tras eliminar manualmente una a una todas las variables inicialmente consideradas que no fueran significativas ($p \geq 0,05$) y cuya exclusión no produjera un cambio mayor del 10% ni un aumento de los intervalos de confianza de las OR de las restantes.

relacionados con el muestreo, facilitaría el seguimiento de los individuos y permitiría obtener información detallada de los no respondedores, que es en realidad la única forma de saber si se ha producido un sesgo de no respuesta. Sin embargo, supondría renunciar a proporcionar datos representativos de ámbito nacional. En cuanto a la forma de administrar las preguntas, parece razonable tratar de concentrarlas en una sola entrevista anterior a la extracción de sangre, reduciendo la longitud del cuestionario para no afectar a la fiabilidad de la información obtenida en las últimas preguntas.

¿Qué se sabe sobre el tema?

La existencia de sesgos de no respuesta pone en peligro la representatividad de cualquier estudio poblacional, aunque se realice un muestreo probabilístico sobre un marco completo. Repartir las preguntas de un cuestionario general en distintas fases supone ir perdiendo muestra en cada una.

¿Qué añade el estudio realizado a la literatura?

Realizar un estudio piloto de grandes dimensiones como base para un estudio longitudinal ha permitido detectar indicios de sesgos de no respuesta relacionados con el estado de salud, el nivel socioeconómico y la comunidad autónoma de residencia de los participantes. Factores como la posibilidad de aumentar las tasas de respuesta, la logística de la recogida de información (incluyendo la de los no respondedores) y el seguimiento de los sujetos, hacen más recomendable un muestreo probabilístico en áreas geográficas seleccionadas por conveniencia. Es conveniente concentrar las preguntas en un solo cuestionario más reducido previo a la extracción de sangre.

Contribuciones de autoría

A. Rodríguez Laso concibió y diseñó el trabajo, consiguió parte de los datos, los analizó e interpretó, y escribió una primera versión del artículo. M. De la Fuente Sánchez colaboró en el diseño del trabajo, consiguió parte de los datos, interpretó los resultados e hizo una revisión crítica del artículo. E. Urdaneta Artola, E. Galindo Moreno, J. Yanguas Lezáun y V. Rodríguez Rodríguez participaron en la interpretación de los resultados y la revisión crítica del artículo. Todos los autores aprobaron la versión final.

Financiación

El estudio piloto del ELES se ha financiado con ayudas del Gobierno Vasco, Programa SAIOTEK 2009; del Ministerio de Ciencia e Innovación, dentro del Programa de Acciones Complementarias 2009; y de la Fundación Obra Social de Caja Madrid.

Conflictos de intereses

Ninguno.

Agradecimientos

A Alfredo Heras Flor por su dedicación en la obtención de los datos para este trabajo.

Bibliografía

- Groves RM. Nonresponse rates and nonresponse bias in household surveys. *Public Opin Q.* 2006;70:646–75.
- Stoop IAL. What is nonresponse and why to be concerned. En: *The hunt for the last respondent.* The Hague: Social and Cultural Planning Office of the Netherlands; 2005. p. 23–63.
- Lynn P. The problem of nonresponse. En: de Leeuw ED, Hox JJ, Dillman DA, editores. *International handbook of survey methodology.* New York: Lawrence Erlbaum Associates; 2008. p. 35–55.
- Teófilo Rodríguez J, González Cabezas AN, Díaz Veiga P. Estudio longitudinal Envejecer en España: el proyecto ELES. *Boletín sobre el envejecimiento "Perfiles y tendencias" n° 50.* Madrid: Instituto de Mayores y Servicios Sociales; 2011, 44 p. (Actualizado el 24/10/2012, consultado el 12/4/2012.) Disponible en: www.imsersomayores.csic.es/documentos/documentos/boletinsobreenvejec50.pdf
- Folstein MF, Folstein SE, McHugh PR, et al. MMSE. Examen Cognoscitivo Mini-Mental. (Traducción y adaptación al español: Lobo A, Saz P, Marcos G y Grupo ZARADEMP). Barcelona: TEA Ediciones; 2002.
- INÉbase. Censos de población y viviendas 2001. Resultados definitivos [Internet]. Instituto Nacional de Estadística de España. (Actualizado el 24/10/2012, consultado el 12/4/2012.) Disponible en: http://www.ine.es/censo/es/seleccion_inframunicipal.jsp
- International Wellbeing Group. *Personal Wellbeing Index – Adult (PWI-A). Manual.* Melbourne: Australian Centre on Quality of Life, Deakin University; 2006, 36 p. (Actualizado el 24/10/2012, consultado el 12/4/2012.) Disponible en: http://www.deakin.edu.au/research/acqol/instruments/wellbeing_index.htm
- Vives A, Ferreccio C, Marshall G. Comparación de dos métodos para corregir el sesgo de no respuesta a una encuesta: sustitución muestral y ajuste según propensión a responder. *Gac Sanit.* 2009;23:266–71.
- Hoeymans N, Feskens EJ, Van Den Bos GA, et al. Non-response bias in a study of cardiovascular diseases, functional status and self-rated health among elderly men. *Age Ageing.* 1998;27:35–40.
- Jackson R, Chambless LE, Yang K, et al. Differences between respondents and nonrespondents in a multicenter community-based study vary by gender ethnicity. *The Atherosclerosis Risk in Communities (ARIC) Study Investigators.* *J Clin Epidemiol.* 1996;49:1441–6.
- Boshuizen HC, Viet AL, Picavet HS, et al. Non-response in a survey of cardiovascular risk factors in the Dutch population: determinants and resulting biases. *Public Health.* 2006;120:297–308.
- Eckholm O, Gundgaard J, Rasmussen NK, et al. The effect of health, socio-economic position, and mode of data collection on non-response in health interview surveys. *Scand J Public Health.* 2010;38:699–706.
- Lee S, Brown ER, Grant D, et al. Exploring nonresponse bias in a health survey using neighborhood characteristics. *Am J Public Health.* 2009;99:1811–7.
- Ballano Fernández C, Martínez Vidal MA. Características de la no respuesta en la Encuesta de Población Activa. *Estadística Española.* 2000;42:263–77.
- Encuesta Nacional de Salud 2006 [Internet]. Ministerio de Sanidad, Política Social e Igualdad. (Actualizado el 24/10/2012; consultado el 22/3/2011.) Disponible en: <http://pestadistico.msc.es/PEMSC25/ArbolNodos.aspx>
- Chou P, Kuo HS, Chen CH, et al. Characteristics of non-participants and reasons for non-participation in a population survey in Kin-Hu, Kinmen. *Eur J Epidemiol.* 1997;13:195–200.
- Van den Akker M, Buntinx F, Metsemakers JF, et al. Morbidity in responders and non-responders in a register-based population survey. *Fam Pract.* 1998;15:261–3.
- Adams MM, Scherr PA, Branch LG, et al. A comparison of elderly participants in a community survey with nonparticipants. *Public Health Rep.* 1990;105:617–22.
- Stoop IAL. The usual suspects: hard to reach and reluctant to cooperate. En: *The hunt for the last respondent.* The Hague: Social and Cultural Planning Office of the Netherlands; 2005. p. 64–80.
- Stoop IAL. Studying nonrespondents. En: *The hunt for the last respondent.* The Hague: Social and Cultural Planning Office of the Netherlands; 2005. p. 105–32.
- Cohen G, Duffy JC. Are non-respondents to health surveys less healthy than respondents? *J Off Stat.* 2002;18:13–23.
- Liberatos P, Link BG, Kelsey JL. The measurement of social class in epidemiology. *Epidemiol Rev.* 1988;10:87–121.
- INÉbase Encuesta sobre Equipamiento y Uso de Tecnologías de la Información y Comunicación en los Hogares [Internet]. Instituto Nacional de Estadística de España. (Actualizado el 24/10/2012; consultado el 25/4/2012.) Disponible en: <http://www.ine.es/jaxi/menu.do?type=pcaxis&path=/t25/p450&file=inebase>
- Parminder SR, Wolfson C, Kirkland SA, et al. The Canadian Longitudinal Study on Aging (CLSA). *Can J Aging.* 2009;28:221–9.