

UNIVERSIDAD DE COSTA RICA

CRELES-RC

Costa Rica. Estudio de Longevidad y Envejecimiento Saludable,
Cohorte de Retiro

Metodología, ronda 1

William H. Dow

Gilbert Brenes-Camacho

Luis Rosero-Bixby

Cobertura: Costa Rica, cohortes de población nacidas en 1945-1955

Financiamiento: National Institute on Aging de los Estados Unidos (grant R01AG031716)

Citación sugerida:

Dow, William H., Gilbert Brenes y Luis Rosero-Bixby (2013). CRELES: Costa Rica. Estudio de Longevidad y Envejecimiento Saludable, Cohorte de Retiro. Metodología, ronda 1. San José, Costa Rica. Centro Centroamericano de Población (CCP), Universidad de Costa Rica. Documento electrónico. Sitio Web de CRELES: <https://ccp.ucr.ac.cr/creles/>

DATOS Y MÉTODOS

Diseño de la muestra

Trabajo de campo

- Pruebas físicas, antropométricas y de movilidad y flexibilidad
 - Presión arterial
 - Mediciones antropométricas
 - Flexibilidad y movilidad

Procedimientos de laboratorio

Ética de la investigación

Anexo

Datos y métodos

El Estudio de Longevidad y Envejecimiento Saludable de Costa Rica (CRELES, Costa Rican Longevity and Healthy Aging Study) es un conjunto de encuestas longitudinales, representativas a nivel nacional, sobre la salud y las experiencias a lo largo de la vida de las personas adultas mayores de Costa Rica. CRELES forma parte del creciente conjunto de Encuestas de Salud y Jubilación (Health and Retirement Surveys) que se realizan en el mundo. Costa Rica es un caso de especial interés por su alta longevidad: la esperanza de vida es mayor que la de los Estados Unidos, a pesar de ser un país de ingreso medio, con cerca de una quinta parte del ingreso por habitante y una décima parte del gasto en salud por habitante.

CRELES está compuesto ahora por varias rondas de datos de dos cohortes: la cohorte original CRELES Pre-1945 es una muestra representativa a nivel nacional de cerca de 3.000 residentes de Costa Rica nacidos en 1945 o antes, entrevistados por primera vez en 2005. Este documento describe los datos y métodos de una nueva Cohorte de Retiro de CRELES, nacida en 1945-1955 y entrevistada a partir de 2010. Esta nueva muestra, llamada CRELES-RC 2010, incluye 2.798 entrevistas de base en formato largo con las personas elegibles por edad seleccionadas, más 1.338 entrevistas a sus cónyuges (independientemente de su edad), realizadas entre enero de 2010 y diciembre de 2011. La nueva CRELES-RC incluye además una muestra suplementaria de entrevistas en formato corto, realizadas entre enero de 2012 y enero de 2013, con 491 personas seleccionadas que inicialmente no respondieron, con el fin de estudiar los patrones de respuesta, que pueden ser especialmente sistemáticos en poblaciones en edad de trabajar.

Los datos de CRELES son muy adecuados para estudiar los determinantes de la longevidad, las relaciones entre el nivel socioeconómico y la salud, el estrés y la salud, los patrones de conductas de salud y los efectos del rápido descenso de la fecundidad que vivió Costa Rica en la década de 1960. Las encuestas CRELES se distinguen por una amplia medición de indicadores de salud y de biomarcadores. La ronda de 2005 de la cohorte CRELES Pre-1945 incluyó la toma de sangre en ayunas y la recolección de orina durante la noche, y la toma de sangre se repitió en 2007. En 2010, CRELES-RC tomó sangre venosa (sin ayuno) para medir colesterol, proteína C reactiva y HbA1c. Se ha extraído ADN en ambas cohortes. Otros indicadores objetivos de salud son la antropometría y la movilidad observada. Las encuestas CRELES se distinguen también por su vinculación con el Índice Nacional de Defunciones de Costa Rica. Esto ha permitido el monitoreo y seguimiento continuo de las defunciones, que se estudian además mediante una entrevista a un familiar sobreviviente. Los archivos de datos de uso público de CRELES contienen información sobre una amplia gama de temas, entre ellos la salud física autorreportada, la salud psicológica, las condiciones de vida, las conductas de salud, el uso de servicios de salud, el apoyo social, el trabajo y el nivel socioeconómico. En 2012 comenzó una segunda ronda de entrevistas de CRELES-RC.

CRELES-RC fue realizado por la Universidad de California, Berkeley, y el Centro Centroamericano de Población (CCP) de la Universidad de Costa Rica, con financiamiento del National Institute on Aging de los Estados Unidos (NIH grant R01 AG031716). El investigador principal es William H. Dow (Universidad de California, Berkeley), con Luis

Rosero-Bixby y Gilbert Brenes (Universidad de Costa Rica) como coinvestigadores principales.

Diseño de la muestra¹

El estudio usó un diseño muestral probabilístico en cuatro etapas. En la primera etapa se seleccionó una muestra aleatoria simple de 60 Áreas de Salud de las 102 en que se divide Costa Rica. En la segunda etapa se seleccionaron 222 seudosegmentos censales a partir de un marco muestral basado en una versión corregida del Censo de 2000; en general, cada seudosegmento estaba formado por dos segmentos censales contiguos, aunque en las zonas rurales algunos tenían 3 o 4 segmentos; los seudosegmentos se crearon para tener al menos 15 viviendas con personas nacidas entre 1945 y 1955 en cada una de estas áreas. En la tercera etapa, en cada seudosegmento se seleccionaron con certeza (probabilidad uno) todas las viviendas con al menos una persona elegible; la decisión de visitar todas las viviendas con al menos una persona elegible se tomó para tener un tamaño de muestra suficiente. Por último, en la cuarta etapa, entre todas las personas nacidas entre 1945 y 1955 en cada una de estas viviendas, se seleccionó al azar solo una como informante principal; si la persona seleccionada estaba casada, también se entrevistó a su cónyuge.

La tasa de no respuesta es distinta si se toma en cuenta la muestra suplementaria de formato corto. La muestra completa tiene una tasa de no respuesta estimada del 34%, pero si se consideran solo los entrevistados en formato largo, la tasa de no respuesta sube al 43%. Las entrevistas en formato corto se usaron para calcular factores de ponderación que corrigen el posible sesgo de selección de la submuestra de formato largo. El cuestionario de cónyuges se aplicó solo a los cónyuges de quienes respondieron el formato largo; de ellos, solo se pudo entrevistar al 76% del total de cónyuges elegibles (una tasa de no respuesta del 24% entre los cónyuges contactados). Además, el equipo de campo entrevistó a 51 cónyuges cuyos participantes principales no pudieron ser entrevistados por rechazos tardíos o por problemas para contactarlos. Hay una base de datos aparte con estos 51 cónyuges.

Se hicieron mediciones antropométricas al 97% del total de la muestra de participantes principales: al 98% de quienes respondieron el formato largo y al 86% de los entrevistados en formato corto; esta última cifra es relativamente baja porque, de los 67 entrevistados en formato corto sin mediciones antropométricas, 63 fueron entrevistados por teléfono, por lo que no fue posible medirlos. Entre los cónyuges, el 99% tiene mediciones antropométricas. Los biomarcadores se recogieron solo de quienes respondieron el formato largo y de los cónyuges. En cada grupo, el 94% tiene toda la información de biomarcadores. Del total de entrevistados, solo el 1% necesitó que otra persona respondiera en su nombre (informante sustituto o “proxy”).

Se calcularon tres conjuntos de factores de ponderación. La primera variable de ponderación corresponde a la submuestra de formato largo y la submuestra suplementaria de formato corto unidas (una añadida a la otra). Esta variable debe usarse para estimar

¹ En la sección del Anexo hay una descripción detallada del diseño muestral.

medidas básicas de prevalencia, en especial de las variables de salud y socioeconómicas que forman parte de ambos cuestionarios. La segunda variable de ponderación se calculó para hacer inferencias sobre la población a partir de la submuestra de formato largo; esta variable debe usarse para analizar las variables exclusivas del formato largo. Por último, la tercera variable de ponderación corresponde al archivo de cónyuges. Debe usarse para hacer inferencias sobre la población de cónyuges y, en especial, para analizar parejas. Esta última variable de ponderación toma en cuenta que solo se pudo entrevistar al 76% de los cónyuges elegibles que efectivamente fueron contactados.

Cuando se usan los factores de ponderación, las distribuciones de frecuencias de variables seleccionadas son similares, pero no iguales, a las observadas en el censo de población de 2011 (cuadro 1). Las distribuciones por edad de la muestra completa (CRELES con entrevistas en formato corto) son ligeramente más envejecidas que las del censo y las de la submuestra de formato largo. Entre las características seleccionadas, las mayores diferencias entre las estimaciones de CRELES y las frecuencias del censo están en la variable de situación laboral. CRELES encontró una proporción mayor de personas que trabajan que el censo. Este resultado es, sin embargo, esperable: el cuestionario censal usa una sola pregunta muy general para medir la participación económica, mientras que el cuestionario de CRELES tiene varias baterías de preguntas.

Cuadro 1. Comparación de las distribuciones de frecuencias relativas de las muestras ponderadas de participantes principales de la ronda de base de la Cohorte de Retiro CRELES 1945-1955 (con y sin entrevistas en formato corto) y del Censo de Población de Costa Rica de 2011, para la población de 55 a 65 años, por sexo

Características	Hombres : CRELES con formato corto	Hombres : CRELES sin formato corto	Hombres : Censo	Mujeres : CRELES con formato corto	Mujeres : CRELES sin formato corto	Mujeres : Censo
Edad a la fecha del censo^{1/} (total)	100,0	100,0	100,0	100,0	100,0	100,0
55-58	43,6	43,5	43,6	43,6	43,7	43,7
59-62	34,2	34,3	34,2	34,3	33,8	34,4
63-65 (o más)	22,2	22,2	22,2	22,1	22,5	21,9
Estado conyugal (total)	100,0	100,0	100,0	100,0	100,0	100,0
En unión libre	14,3	14,6	15,0	7,7	7,9	8,7
Casado(a)	68,2	68,2	63,7	51,3	51,1	50,0
Separado(a)/divorciado(a)	9,3	9,0	9,4	22,4	22,8	16,5
Viudo(a)	2,6	2,6	2,4	10,9	10,6	10,3
Soltero(a)	5,6	5,5	9,6	7,7	7,6	14,4

Características	Hombres : CRELES con formato corto	Hombres : CRELES sin formato corto	Hombres : Censo	Mujeres : CRELES con formato corto	Mujeres : CRELES sin formato corto	Mujeres : Censo
Educación (total)	100,0	100,0	100,0	100,0	100,0	100,0
Sin escolaridad/primaria	56,5	56,3	56,5	58,4	58,1	58,4
Secundaria académica	19,9	19,4	21,8	20,2	21,0	21,7
Postsecundaria y técnica	23,6	24,3	21,7	21,4	20,9	19,9
Trabaja (total)	100,0	100,0	100,0	100,0	100,0	100,0
Sí	67,0	66,2	61,7	25,0	25,0	21,0
No	33,0	33,8	38,3	75,0	75,0	79,0

Nota: 1/ Fecha del Censo de 2011: 30 de mayo de 2011.

Trabajo de campo

Por ser longitudinal, el estudio consta de una recolección de datos de base en los hogares y de una encuesta de seguimiento en los hogares dos años después. Este informe presenta los resultados de la ronda de base (primera ronda), que recogió información durante 2010 y 2011. Incluyó una entrevista estructurada, mediciones antropométricas, pruebas de funcionamiento físico y la toma de muestras de sangre. El cuestionario incluía preguntas sobre historia conyugal, características de los hijos (vivos y fallecidos), salud y conductas de salud, ingresos por salarios y actividades empresariales, nivel socioeconómico percibido, características de la vivienda, transferencias intergeneracionales y apoyo social. Había baterías especiales de preguntas sobre estrés percibido, actividad física (la escala IPAQ), estado cognitivo, síntomas de depresión (una versión corta de la Escala de Depresión Geriátrica) y amistad. Para entrevistar al cónyuge del informante principal, cuando lo hay, se usa una versión más corta del cuestionario, centrada sobre todo en preguntas de salud y apoyo social.

En el campo también se registraron, con dispositivos GPS, las coordenadas geográficas del lugar de residencia de cada participante. Todos los datos y especímenes se recogieron en los hogares de los participantes, por lo general en una sola visita. Los participantes dieron su consentimiento informado mediante su firma, respondieron un cuestionario principal de unos 90 minutos y la presión arterial se midió dos veces durante la entrevista. Al final de la entrevista se tomaron las muestras de sangre, se hicieron las mediciones antropométricas y se realizaron las pruebas de funcionalidad física. El proyecto no requirió que los entrevistados estuvieran en ayunas, porque los biomarcadores elegidos no son sensibles a si la persona estaba en ayunas o no.

Al inicio de la entrevista principal se incluyó una evaluación cognitiva que, junto con el criterio del entrevistador, determinaba si se necesitaba o no un informante sustituto

("proxy") que ayudara a responder la encuesta. Solo 36 entrevistas (1,28% del total) se realizaron con la ayuda de un informante sustituto.

El trabajo de campo de la muestra suplementaria de entrevistas en formato corto, con personas difíciles de contactar que habían rechazado la entrevista en formato largo, se realizó entre enero de 2012 y enero de 2013. El cuestionario estructurado de esta muestra es una versión corta del cuestionario principal; incluyó preguntas sobre salud, historia conyugal, transferencias intergeneracionales y apoyo social. También se hicieron mediciones antropométricas y pruebas de funcionamiento físico, excepto en el 12% de las entrevistas en formato corto, que se hicieron por teléfono después de repetidos intentos fallidos de hacerlas en persona. Todas las entrevistas en los hogares las completaron dos equipos de 3 entrevistadores cada uno, que estuvieron continuamente en el campo durante un período de dos años, por lo que los entrevistadores tenían mucha experiencia. Las tomas de sangre las hizo un flebotomista capacitado, en su mayoría después de terminada la entrevista.

Todos los datos del trabajo de campo se registraron en computadoras de mano Pocketbook, un tipo de "asistente digital personal" (PDA), con una aplicación desarrollada en el Centro Centroamericano de Población para este estudio. Esto incluyó el cuestionario principal, con patrones complejos de saltos y vínculos entre preguntas. Se usaron la misma tecnología y el mismo procedimiento que en las cohortes nacidas antes de 1945 y entrevistadas entre 2004 y 2009. Durante la prueba piloto de ese estudio, dos entrevistadores registraron las respuestas al cuestionario simultáneamente en la PDA y en papel, con un nivel de concordancia sumamente alto (Hidalgo-Céspedes, Rosero-Bixby et al. 2007). La PDA muestra en pantalla el texto de cada pregunta que el entrevistador debe leer y, cuando hace falta, también da instrucciones. Las respuestas se registran por lo general tocando en la pantalla la opción seleccionada de una lista, pero también pueden ingresarse números o texto directamente en "graffiti" o con un teclado virtual, si el entrevistador lo prefiere. La PDA controla el flujo de la entrevista: hace los saltos de preguntas y aplica filtros según las respuestas anteriores; también ejecuta verificaciones de consistencia programadas de antemano y genera automáticamente ciertas variables, como la fecha y la hora. La PDA no permite registrar datos inconsistentes o fechas fuera de rango, ni saltarse preguntas de la secuencia. Los datos se respaldaban diariamente en el campo y se cargaban con regularidad para monitorear su calidad en tiempo real durante el trabajo de campo.

Pruebas físicas, antropométricas y de movilidad y flexibilidad

A continuación se describen los materiales, equipos y métodos usados en las mediciones físicas: presión arterial, mediciones antropométricas, pruebas de flexibilidad y movilidad, fuerza de la mano y flujo espiratorio máximo. Hay más detalles de las pruebas en los manuales del entrevistador, en el sitio web del proyecto (<http://ccp.ucr.ac.cr/creles/index.htm>).

Presión arterial. Se midió en dos ocasiones durante la entrevista principal, con un intervalo promedio de 20 minutos entre una y otra; la medición se hizo con monitores digitales marca OMRON de inflado automático, modelo HEM-711 (precisión: ± 3 mmHg),

que se calibraban periódicamente. El brazalete se ajustaba al grosor del brazo de la persona.

Mediciones antropométricas. Las tomaron los entrevistadores, que fueron capacitados y certificados para ello, con una capacitación de actualización después de un año de trabajo de campo. Las mediciones tomadas y el equipo utilizado son los siguientes:

Peso corporal: se usó una balanza marca Life Source, A&D Medical, modelo UC-321; se colocaba sobre un piso plano y sin alfombras, la medición se hacía sin zapatos y, a los participantes vestidos, se les pedía sacar de los bolsillos los objetos pesados.

Talla: se usó un estadiómetro marca Seca para medir la estatura de las personas adultas mayores. La medición no se hacía si la persona tenía deformaciones importantes de la columna.

Altura de rodilla: la medición se hacía en la pierna derecha, siempre que la persona entrevistada no tuviera una lesión en ella. Se usaba un inclinómetro para indicar el ángulo de 90 grados y luego se medía la altura con un estadiómetro fabricado por Shorr Productions (USES Knee-Height Caliper).

Circunferencias abdominal y de cadera: estas mediciones se hacían con los participantes de pie, en posición semianatómica (con los pies separados y la palma de las manos apoyada en la cara lateral del muslo). Se usaron cintas métricas marca Quick Medical.

Circunferencia de pantorrilla: la persona estaba sentada, con la pierna derecha descubierta.

Circunferencia de brazo: con la persona sentada o de pie, la circunferencia se medía en el punto medio entre el acromion (hueso posterior del hombro) y el olécranon (hueso saliente del codo).

Pliegues cutáneos tricipital y subescapular: el entrevistador hacía las mediciones con los pulgares y los índices, para asegurarse de tomar solo el tejido adiposo y no músculos ni nervios. Para ello se usó un calibrador de pliegues Lange, de Beta Technology Incorporated.

Fuerza de la mano. Se tomaron dos mediciones de la fuerza de la mano (en el análisis se usa el valor más alto), con la persona de pie y el brazo dominante extendido junto al cuerpo. Se usó un dinamómetro de Creative Health Products Inc., modelo T-18.

Flexibilidad y movilidad. Las pruebas de flexibilidad y movilidad tenían el propósito de medir (1) el equilibrio y el balance, (2) la agilidad y (3) la velocidad de la marcha. Los ejercicios realizados fueron los siguientes:

Equilibrio y balance: para medirlos se hicieron dos pruebas: (1) permanecer de pie con los pies juntos durante 10 segundos y (2) levantarse cinco veces desde la posición de sentado, con los brazos cruzados sobre el pecho.

Agilidad: la agilidad se midió con la capacidad de la persona de agacharse, recoger un lápiz y enderezarse. Si la persona no podía hacerlo en menos de 30 segundos, la prueba no continuaba. La prueba tampoco se hacía si la persona había tenido una operación de cataratas u otro procedimiento de retina en las seis semanas anteriores.

Velocidad de la marcha: para medir la capacidad de levantarse de una silla y caminar, se pedía a la persona que se levantara de una silla y caminara una distancia de 3 metros como lo hace normalmente, ni más lento ni más rápido. La prueba se registraba con un cronómetro, anotando el tiempo en segundos que tomaba completarla.

Procedimientos de laboratorio

La muestra de sangre se obtenía por venopunción, en su mayoría poco después de realizada la entrevista. Se recogían dos tubos de sangre: uno con anticoagulante (VACUTAINER / EDTA) de 3-4 ml, que luego se centrifugaba para separar el plasma de las células, y otro tubo sin anticoagulante con activador de coágulo (VACUTAINER SST, 5 ml) para obtener suero. En el laboratorio se separaba una fracción de suero en un tubo cónico tipo Eppendorf para las pruebas de colesterol total, HDL (lipoproteína de alta densidad) y PCR (proteína C reactiva), y 1 ml de sangre completa en el tubo con EDTA para el análisis de HbA1c (hemoglobina glicada).

Los biomarcadores de las muestras de sangre del proyecto CRELES-RC se analizaron en el laboratorio clínico de la Oficina de Bienestar y Salud de la Universidad de Costa Rica (UCR). Los tubos con los especímenes se enviaban al final del día al laboratorio para su análisis. Las muestras se mantenían a la temperatura adecuada en una hielera con hielo. Las fracciones restantes de suero y plasma se dividían en alícuotas en criotubos de tapa roja y se almacenaban en ultracongelación (-140 °C). Cuando el equipo de campo entrevistaba en regiones fuera de la capital, San José, las muestras de sangre se guardaban en refrigeradores de clínicas y hospitales públicos de la Caja Costarricense de Seguro Social (el principal proveedor público de servicios de salud del país).

Los métodos de análisis usados para los biomarcadores son los siguientes:

Biomarcador	Método
Colesterol total	Prueba colorimétrica enzimática; longitud de onda: 505 nm; temperatura: 37 °C; reactivos Roche; equipo: analizador Cobas c501
HDL	Prueba colorimétrica enzimática; longitud de onda: 600 nm; temperatura: 37 °C; reactivos Roche; equipo: analizador Cobas c501
PCR	Inmunoensayo turbidimétrico potenciado por partículas, con anticuerpos monoclonales anti-PCR; equipo: analizador Cobas c501
HbA1c	Inmunoensayo de inhibición turbidimétrica (TINIA) para sangre total hemolizada; reactivos Roche; equipo: analizador Cobas c501

Ética de la investigación

El estudio fue aprobado en abril de 2009 por el Comité para la Protección de Sujetos Humanos (CPHS II) de la Universidad de California en Berkeley. El estudio y los procedimientos de trabajo de campo y de consentimiento informado también fueron aprobados por el Comité Ético Científico de la Universidad de Costa Rica en las sesiones del 24 de abril de 2009, el 10 de agosto de 2009 y el 23 de febrero de 2011 (referencias: VI-2878-2009, VI-5308-2009 y VI-1313-2012), como parte del proyecto de investigación

número 828-A2-825. Todas las bases de datos del estudio se anonimizaron (se eliminaron el nombre y otros identificadores) para evitar riesgos a la privacidad de los participantes. El consentimiento informado por escrito se firmó durante la primera ronda de entrevistas, en la que se explicó que habría visitas de seguimiento después de dos años. Algunas personas de la muestra suplementaria (las “entrevistas cortas”) fueron entrevistadas por teléfono; el consentimiento de estos participantes quedó registrado en un “Formulario de consentimiento informado para entrevistas telefónicas” especial, aprobado por el Comité Ético Científico de la Universidad de Costa Rica en la sesión del 18 de abril de 2012 (referencia: VI-2403-2012).

Referencias

Hidalgo-Céspedes, J., L. Rosero-Bixby, et al. (2007). “Mejora de la calidad y disminución de costos de censos y encuestas utilizando computadores de mano (PDA). Una aplicación en Costa Rica.” *Población y Salud en Mesoamérica* 5(1).

Anexo

Diseño muestral de la ronda de base

Cohorte de Retiro CRELES 1945-1955 (CRELES-RC 2010)

Población de estudio: personas nacidas entre 1945 y 1955 que vivían en Costa Rica entre 2010 y 2011, y sus cónyuges corresidentes (independientemente de su año de nacimiento).

Tamaño de muestra planeado: 3.330 personas seleccionadas. El estudio supuso una tasa de no respuesta del 10%.

Tamaño final de la muestra: se completaron entrevistas en formato largo con 2.798 personas seleccionadas (“entrevistas principales”) y 1.338 cónyuges (“entrevistas de cónyuges”). Para investigar los patrones de no respuesta en las entrevistas en formato largo, se hicieron 491 entrevistas adicionales en formato corto, para un total de 3.289 entrevistas a personas seleccionadas.

Diseño muestral: el estudio usó un diseño muestral probabilístico en cuatro etapas:

- (a) Unidades primarias de muestreo (UPM): Áreas de Salud. Se seleccionó una muestra aleatoria simple de 60 Áreas de Salud de las 102 que cubren el territorio costarricense. Estas UPM se habían seleccionado originalmente para la cohorte CRELES-2004 (la cohorte nacida antes de 1946). En estas Áreas de Salud vivía el 55% de la población adulta mayor en 2002-2004 y el 62% de la población nacida en 1945-1955 (de 54 a 65 años en la ronda de base). El estudio decidió mantener las mismas UPM para la nueva cohorte, con el fin de facilitar la planificación del trabajo de campo y de aprovechar la información administrativa sobre los servicios de salud de esas áreas recogida en 2009.
- (b) Unidades secundarias de muestreo (USM): pseudosegmentos censales, contruidos a partir de conglomerados de segmentos del Censo de 2000. Los segmentos del Censo de Población de 2000 se agruparon para tener al menos 15 viviendas con personas nacidas entre 1945 y 1955 en un territorio geográficamente continuo. Estos conglomerados se llaman pseudosegmentos censales. La mayoría está formada por 2 segmentos censales; sin embargo, en algunos casos (sobre todo en zonas muy rurales), los conglomerados tenían 3 o 4 segmentos. El estudio seleccionó 222 USM con probabilidad proporcional al tamaño (PPT), con un procedimiento sistemático para cubrir toda la región definida por las 60 Áreas de Salud.
- (c) Unidades terciarias de muestreo (UTM): viviendas. El estudio seleccionó todas las viviendas dentro del pseudosegmento censal (las USM).
- (d) Unidades de muestreo de cuarta etapa (U4M): personas. Entre todas las personas nacidas entre 1945 y 1955 que vivían en cada vivienda, se selecciona una al azar como informante principal.

Marco muestral: el marco muestral es la base de datos principal del Censo de Población de Costa Rica de 2000. Para corregir la información desactualizada, estimamos razones de

sobrevivencia con el padrón electoral de 1998 y 2006. El padrón electoral tiene un registro por cada ciudadano costarricense; las estimaciones calculadas en el Centro Centroamericano de Población sugieren que, en el rango de 55 a 65 años, la población del padrón electoral representa el 97% de la población total. Dividimos la población de 51 a 61 años en 2006 entre la población de 43 a 53 años en 1998, para cada distrito electoral. Dado que se trata de un período de ocho años, elevamos la razón a la potencia 1,25 para estimar una razón de 10 años (que correspondía a 2010, año en que empezó el trabajo de campo). La población de cada segmento censal se multiplicó por esta razón. Este procedimiento tiene dos supuestos:

- a) La razón de cada distrito electoral representa correctamente los patrones de mortalidad y de migración interna de las cohortes en los segmentos censales ubicados en ese distrito electoral.
- b) Los patrones de mortalidad y de migración interna de los ciudadanos costarricenses (que figuran en el padrón electoral) son similares a los de los no ciudadanos.

Después de corregir la población del marco muestral, los datos se agruparon en los seudosegmentos censales descritos antes.

Procedimiento de trabajo de campo: los entrevistadores iban a cada seudosegmento censal divididos en dos grupos de 3 entrevistadores y un chofer (uno de ellos era el supervisor de campo). Visitaban cada vivienda del seudosegmento para determinar cuáles tenían residentes nacidos entre 1945 y 1955. Cuando era posible, la información sobre los hogares que no respondían se obtenía de los vecinos; a los hogares que seguían sin responder se les asignó una probabilidad estimada de tener un residente elegible, para efectos de ponderación. En cada vivienda se seleccionaba al azar a una persona como entrevistada principal, mediante un algoritmo generador de números aleatorios programado en la computadora de mano usada para la captura de datos. Si la persona seleccionada estaba casada o vivía en unión libre, su pareja corresidente se seleccionaba con certeza (condicionado a que se completara la entrevista principal) para una entrevista de cónyuge. Algunos cónyuges se entrevistaron sin tener la información de la persona seleccionada, porque esta rechazó la entrevista o no pudo ser contactada después de entrevistar al cónyuge; hay 51 cónyuges sin información sobre la persona seleccionada.

Probabilidad de selección: el factor de ponderación es igual al inverso de la probabilidad de selección. Cada entrevistado tiene una probabilidad de selección distinta. Por lo tanto, la muestra no es autoponderada. Hay tres conjuntos de probabilidades de selección y, por ende, tres conjuntos de factores de ponderación: (1) para todos los entrevistados principales, incluida la muestra suplementaria de entrevistas en formato corto; (2) solo para los entrevistados principales del cuestionario de formato largo, y (3) para los cónyuges (que también pueden usarse al analizar parejas).

La probabilidad de selección planeada, sin correcciones, sería:

$$P(X) = P(UPM) * P(USM) * P(UTM) * P(U4M)$$

$$P(X) = \frac{60}{102} * \frac{Mt_{\alpha}}{984,89} * 1 * \frac{1}{\text{Número de personas elegibles por edad en el hogar}}$$

donde Mt_{α} es el número de viviendas con al menos una persona nacida entre 1945 y 1955. En un diseño con probabilidad proporcional al tamaño (PPT), Mt_{α} es el tamaño de los conglomerados en el marco muestral².

La probabilidad de selección se refiere a seleccionar 60 Áreas de Salud de 102, por la probabilidad de seleccionar un pseudosegmento censal, por la probabilidad de seleccionar a una persona dentro de la vivienda. Como los pseudosegmentos se seleccionaron con probabilidad proporcional al tamaño (PPT), la probabilidad de seleccionar una USM varía de una a otra, porque cada pseudosegmento tiene un tamaño distinto. La probabilidad de seleccionar una USM varía según el tamaño del pseudosegmento, Mt_{α} , que equivale al número de viviendas con al menos una persona nacida entre 1945 y 1955 en cada área. La probabilidad de seleccionar una USM es $Mt_{\alpha}/984,89$, donde 984,89 es igual al número de hogares con al menos un residente de 55 a 65 años (218.645) en el marco muestral, dividido entre el número de segmentos censales que se planeaba seleccionar (222). La probabilidad de seleccionar una UTM es uno, porque se visitan todas las viviendas.

Probabilidad de selección de todos los entrevistados principales (entrevistas en formato largo y corto)

La probabilidad de selección efectiva incluye tres factores de corrección, debidos a limitaciones del trabajo de campo.

- (a) factor1 = 0,99. Se excluyeron por completo ocho pseudosegmentos censales (equivalentes a 16 segmentos censales). Cuatro de estas USM están en barrios sumamente peligrosos, por lo que se excluyeron para evitar cualquier daño al personal de campo. En tres USM no había ninguna persona elegible, y otro pseudosegmento no pudo visitarse porque no tenía acceso fácil en vehículo. Además, 11 USM se cubrieron solo parcialmente: en 4 de ellas, uno de los segmentos censales se visitó y el otro estaba en un barrio peligroso, por lo que solo se visitó la mitad del pseudosegmento; en cada una de otras 7 USM, uno de los segmentos tenía informantes y en el otro no se encontraron informantes. El número total de viviendas con informantes esperados en estas áreas sin información equivalía al 1% del total de viviendas esperadas en toda la muestra.
- (b) factor2 = (viviendas efectivamente visitadas en la USM X / número total de viviendas esperadas con información en la USM X). Hubo diferencias importantes entre pseudosegmentos en el número de viviendas en las que se encontraron personas elegibles, en comparación con el número esperado según el marco muestral. La razón promedio por USM fue de 0,42, pero el mínimo fue 0,03 y el máximo 1,6 (en 6 USM, el equipo de campo encontró más viviendas con personas elegibles de las esperadas). Esta inconsistencia podría deberse a:

² Kish, Leslie (1965). Survey Sampling. Nueva York: Wiley.

- i) Un marco muestral desactualizado.
 - ii) Un alto porcentaje de viviendas con personas elegibles que no pudieron encontrarse (o identificarse como tales).
 - iii) Rechazos.
- (c) factor3: una tasa de respuesta diferencial según sexo, edad y educación. Con base en el Censo de Población de 2011, encontramos que la muestra ponderada por factor1 y factor2 sobreestima a las mujeres respecto a los hombres, a las personas de más edad respecto a las más jóvenes y a las menos educadas respecto a las más educadas. Calculamos factores de corrección para 30 grupos distintos, definidos por sexo (hombres, mujeres), grupos de edad (54-56, 57-58, 59-60, 61-62, 63-65) y años de escolaridad (solo primaria, al menos un año de secundaria académica, al menos un año de educación técnica o postsecundaria).

La fórmula final de la probabilidad de selección es:

$$P(X) = P(UPM) * P(USM) * P(UTM) * P(U4M) * factor1 * factor2 * factor3$$

$$P(X) = \frac{60}{102} * \frac{Mt_a}{984.89} * 1 * \frac{1}{\text{Número de personas elegibles por edad en el hogar}} * factor1 * factor2 * factor3$$

Probabilidad de selección solo para los cuestionarios de formato largo: del total de entrevistas a personas seleccionadas, 2.798 (85%) respondieron el cuestionario de formato largo y 491 respondieron la entrevista de formato corto. Para analizar la información exclusiva de la entrevista en formato largo, calculamos un conjunto de factores de ponderación basado en las probabilidades de selección específicas de contactar a la persona y de lograr que aceptara responder el cuestionario de formato largo (en vez de rechazar la entrevista).

Como el cuestionario de formato corto se aplicó para obtener información básica de personas difíciles de contactar, estos entrevistados son relativamente distintos de quienes aceptaron responder el cuestionario de formato largo. Por lo tanto, estimamos dos modelos de regresión logística (uno para hombres y otro para mujeres), en los que la variable dependiente es si la persona seleccionada respondió el cuestionario de formato largo o el de formato corto. El modelo se simplificó con un procedimiento de selección por pasos hacia atrás (stepwise backward), con una probabilidad de salida de 0,20 (decidimos usar una significancia relativamente alta para dar prioridad a la clasificación sobre la explicación del evento). Estimamos la probabilidad de responder el cuestionario de formato largo, y la probabilidad de selección de esta submuestra es igual a la probabilidad de selección de haber sido entrevistado, por la probabilidad de responder el cuestionario de formato largo:

$$P(X \text{ con formato largo}) = P(X) * P(\text{formato largo} | X)$$

El cuadro siguiente resume los modelos finales usados para calcular las probabilidades de responder el formato largo, dado que se aceptó la entrevista.

Cuadro 1. CRELES-RC. Modelos de regresión logística para estimar la probabilidad de responder el cuestionario de formato largo en vez del corto, dado que la persona aceptó la entrevista, por sexo

Covariables	Coeficiente	Error estándar	Valor p
Hombres			
Primaria vs. más que primaria	0,362	0,140	0,010
Reside en Cartago vs. en otras provincias	-0,226	0,171	0,187
Intercepto	1,722	0,114	0,00
Mujeres			
Edad	0,077	0,025	0,002
Mala salud vs. buena salud	0,407	0,152	0,007
No casada vs. casada	0,519	0,200	0,009
Trabaja vs. no trabaja	0,416	0,160	0,009
Provincia de residencia (ref.: resto del país)			
- Cartago	-0,554	0,187	0,003
- Puntarenas	-0,467	0,221	0,035
- Alajuela	-0,351	0,202	0,083
Intercepto	-3,352	1,485	0,024

Probabilidad de selección de los cónyuges (o para análisis de parejas): los cónyuges solo se entrevistaron si el entrevistado principal aceptó la entrevista de formato largo (excepto 51 cónyuges que se entrevistaron antes de que fallara la entrevista de la persona seleccionada). Entre quienes respondieron el cuestionario de formato largo, el 24% de los cónyuges elegibles no fue entrevistado por rechazo o por dificultades del trabajo de campo. Por lo tanto, se calcularon por separado las probabilidades de selección de los cónyuges (excepto para los 51 cónyuges sin entrevista principal, ya que estos casos no tienen información suficiente para analizar el patrón de no respuesta). Estos factores de ponderación de cónyuges también pueden usarse al analizar la submuestra formada solo por parejas en las que ambos completaron la entrevista. La probabilidad de selección del cónyuge es igual a la probabilidad de selección del informante principal multiplicada por la tasa de respuesta de los cónyuges:

$$P(X. sp) = P(\text{cónyuge} \mid X \text{ casado y con formato largo})$$

Como la tasa de respuesta varía según varias características, estimamos otro modelo de regresión logística, en el que la variable dependiente es si el cónyuge fue entrevistado o no. El modelo se simplificó con un procedimiento por pasos (stepwise), con una probabilidad de entrada de 0,05. Los cónyuges con tasas de respuesta más bajas son los que trabajan, los de mayor educación, los que viven en unión libre, los que viven en la provincia de Cartago y las cónyuges mujeres. Con esta ecuación se calcularon las probabilidades de ser entrevistado. La probabilidad de selección de los cónyuges se calcula entonces como la

probabilidad de selección del informante principal multiplicada por esta probabilidad de entrevista del cónyuge:

$$P(X, sp) = P(X \text{ con formato largo}) * P(\text{cónyuge entrevistado} | \text{persona seleccionada entrevistada})$$

Las probabilidades estimadas con el modelo logístico varían entre 0,43 y 0,91, con una media de 0,79. El cuadro siguiente describe el modelo de regresión logística.

Cuadro 2. CRELES-RC. Modelo de regresión logística para estimar la probabilidad de entrevistar a un cónyuge entre quienes respondieron el formato largo

Covariables	Coeficiente	Error estándar	Valor p
El cónyuge trabaja	-1,082	0,122	0,000
Educación (ref.: primaria o menos)			
- Educación secundaria	-0,533	0,148	0,000
- Educación postsecundaria	-0,378	0,168	0,024
Pareja en unión libre	-0,413	0,152	0,007
Reside en Cartago vs. en otras provincias	-0,315	0,153	0,040
Cónyuge hombre	0,722	0,128	0,000
Intercepto	1,646	0,122	0,000

Tasa de no respuesta: la muestra suplementaria de entrevistas en formato corto se seleccionó para investigar las propiedades sistemáticas de la no respuesta entre los entrevistados principales. Los equipos de campo indicaron que gran parte de la no respuesta se debía a personas que habían sido invitadas a participar pero se habían negado porque la encuesta de 90 minutos era demasiado larga. Las 2.798 entrevistas de formato largo (90 minutos) se completaron hasta enero de 2012, y las entrevistas de formato corto se hicieron de enero de 2012 a enero de 2013. La mayoría de las entrevistas cortas completadas fueron en viviendas ya visitadas, en las que se conocía el número de personas elegibles, pero en las que la persona seleccionada se había negado antes a completar una entrevista de formato largo después de repetidos intentos. El cuadro 3 muestra cómo se calculó la tasa de respuesta con y sin las entrevistas cortas.

La no respuesta puede clasificarse en cuatro categorías: (1) Rechazo explícito de personas elegibles (las personas dicen abiertamente que no quieren ser entrevistadas). (2) Rechazo explícito de hogares que no dan ninguna información, por lo que se desconoce el número de personas elegibles. (3) Rechazo implícito (personas que aceptan ser entrevistadas más adelante, pero no están disponibles cuando se las visita de nuevo). Según el cuadro 3, en la muestra principal la tasa de respuesta es del 57%. El rechazo explícito fue relativamente bajo (3,1%), pero el rechazo implícito fue mayor: 26%. No todos los 1.269 casos de rechazo implícito pueden considerarse como tales; el personal de campo tuvo que descartar visitar a algunas personas contactadas por primera vez que no pudieron ser visitadas de nuevo porque sus viviendas estaban demasiado lejos de la ruta programada. (4) Además, los entrevistadores encontraron 6.280 viviendas en las que nunca se pudo contactar a ningún habitante, aun después de varias visitas al barrio. Si aproximadamente el 10% de todas las

viviendas de Costa Rica tiene una persona nacida entre 1945 y 1955, y si suponemos que las viviendas cuyos habitantes no pudieron identificarse tienen esta misma proporción de personas elegibles, entonces se perdieron aproximadamente 661 viviendas con personas elegibles durante el trabajo de campo: 628 en viviendas sin contacto y 33 en viviendas que se negaron a dar información sobre sus habitantes. Entre los hogares con los que se logró algún tipo de contacto, la tasa de respuesta en formato largo fue del 85% (2.798+154+330, donde 330 es el número total de viviendas con rechazo explícito a revelar el número de personas elegibles).

Cuando se incluye la submuestra de entrevistas en formato corto, la tasa de respuesta sube al 66%. El rechazo explícito aumentó levemente (de 3,8% a 4,1%), pero el rechazo implícito disminuyó considerablemente: de 26% a 16%. La tasa de no respuesta por viviendas sin información de contacto no varía en absoluto, porque el personal de campo no regresó a esas viviendas. Entre los hogares con los que se logró algún tipo de contacto, la tasa de respuesta total (formato corto + formato largo) fue del 86%.

Cuadro 3. Estimación de la tasa de respuesta con base en las viviendas visitadas con información sobre sus habitantes, clasificadas según el tipo de respuesta a la encuesta. CRELES-2010, ronda 1

Tipo de respuesta	Formato largo: frec. abs.	Formato largo: frec. rel. (%)	Total corto + largo: frec. abs.	Total corto + largo: frec. rel. (%)
Respuesta	2.798	57,3	3.289	66,1
No respuesta total	2.084	42,7	1.684	33,9
Rechazo explícito con personas elegibles en la vivienda	154	3,1	205	4,1
No respuesta estimada de personas seleccionadas en viviendas con rechazo explícito a revelar el número de personas elegibles (10% de las viviendas visitadas)	33	0,7	33	0,7
Rechazo implícito (viviendas con número conocido de personas elegibles a las que fue imposible entrevistar)	1.269	26,0	819	16,5
Viviendas estimadas sin contacto y con número desconocido de personas elegibles (10% de las viviendas visitadas)	628	12,9	627	12,6
Total	4.882	100,0	4.973	100,0

Nota: los totales de los dos grupos no son iguales porque algunas cifras son estimadas.

Diferencias entre las muestras de formato largo y corto: el análisis de la muestra suplementaria de entrevistas en formato corto permite hacer inferencias sobre un posible sesgo de selección por tasas de respuesta diferenciales. El cuadro 4 compara las muestras según 5 características que se sabe producen respuesta diferencial en Costa Rica: sexo, edad, educación, estado conyugal y situación laboral. En comparación con las entrevistas de formato largo, las de formato corto tienen relativamente más hombres, más personas de 59 a 62 años, más hombres casados, más mujeres viudas, más personas con educación posprimaria, menos hombres que trabajan y más mujeres que trabajan. En general, la submuestra de quienes respondieron el formato largo tiene cierto grado de sesgo de selección, en comparación con quienes respondieron la entrevista de formato corto.