

UNIVERSIDAD DE COSTA RICA

CRELES-RC

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

Metodología, ronda 2

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 2. 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

Panorama general de CRELES-RC

Diseño muestral y tasas de respuesta de la ronda 2 de CRELES-RC

Trabajo de campo

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

Ética de la investigación

Anexo I

Anexo II

Panorama general de CRELES-RC

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. Una Cohorte de Retiro adicional de CRELES (CRELES-RC) muestreó a personas nacidas en 1945-1955, con entrevistas de base a partir de 2010 (las entrevistas se extendieron hasta 2011, pero por simplicidad se habla de la ronda de 2010). Este documento describe los datos y métodos de la segunda ronda de CRELES-RC, un seguimiento a los dos años que comenzó en 2012 (las entrevistas se completaron a inicios de 2014, pero por simplicidad se habla de la ronda de 2012). La muestra de la primera ronda de CRELES-RC incluyó 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). La segunda ronda de CRELES-RC (que aquí se describe) intentó volver a entrevistar a todas las personas seleccionadas que respondieron el formato largo en la ronda de base, con lo que se obtuvieron 2.428 reentrevistas de personas seleccionadas en la ronda 2, más entrevistas sobre 49 personas seleccionadas fallecidas. CRELES-RC también intentó entrevistar a los cónyuges actuales o fallecidos de todas las personas seleccionadas entrevistadas en la ronda 2, con lo que se obtuvieron 1.097 reentrevistas de cónyuges, 12 entrevistas a nuevos cónyuges y entrevistas sobre 29 cónyuges fallecidos.

La primera ronda de CRELES-RC incluyó 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. A estas 491 personas no se les dio seguimiento para entrevistarlas en la segunda ronda.

Por su diseño longitudinal, 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. La ronda de 2005 de la cohorte original 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, y hay medidas de la longitud de telómeros leucocitarios para un subconjunto de participantes (para más detalles, contacte a los investigadores principales). En la segunda ronda de CRELES-RC no se midieron biomarcadores.

Otros indicadores objetivos de salud medidos en todas las rondas 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.

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 muestral y tasas de respuesta de la ronda 2 de CRELES-RC¹

Para seleccionar la muestra de base, 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 pseudosegmentos censales (cada uno formado por dos o más segmentos censales contiguos, para tener al menos 15 viviendas con personas elegibles por edad en cada pseudosegmento) a partir de un marco muestral basado en una versión corregida del Censo de 2000. En la tercera etapa, en cada pseudosegmento se seleccionaron con certeza (probabilidad uno) todas las viviendas con al menos una persona elegible por edad. 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 o “persona seleccionada”; si la persona seleccionada estaba casada (o en una unión similar al matrimonio), también se entrevistó a su cónyuge.

Las tasas de no respuesta de la segunda ronda difieren entre las personas seleccionadas y los cónyuges. La tasa total de respuesta de las personas seleccionadas fue del 89% (cuadro 1).

Cuadro 1. Tasas de respuesta de las personas seleccionadas en la ronda 2 (“ronda de 2012”) de CRELES-RC

Respuesta	Frecuencia absoluta	Frecuencia relativa
Aceptó	2.479	88,6

¹ En el anexo se incluye una descripción detallada del diseño muestral de la ronda 2. El diseño muestral y las tasas de respuesta de la ronda de base se incluyen en la documentación metodológica de la ronda 1, en creles.berkeley.edu.

Respuesta	Frecuencia absoluta	Frecuencia relativa
- Vivo	2.430	86,8
- Fallecido	49	1,8
No respuesta	319	11,4
- Rechazo	77	2,8
- No encontrado	238	8,5
- No visitado	4	0,1
Total	2.798	100,0

La no respuesta se debió principalmente a personas “no encontradas” durante el trabajo de campo (8,5%), que incluye a algunas que se sabía que aún residían en el hogar, pero que no estuvieron disponibles para la entrevista. La tasa de rechazo explícito fue cercana al 3%. Además, 4 personas no fueron visitadas porque viven en lugares que se volvieron demasiado peligrosos o lejanos para hacer la entrevista.

Entre los cónyuges, el porcentaje de reentrevistas fue menor, pero esto se debió sobre todo a que no se intentó la entrevista del cónyuge si no se había logrado reentrevistar a la persona seleccionada. El cuadro 2 describe la no respuesta en la muestra total de cónyuges y en la submuestra con reentrevista de la persona seleccionada.

La tasa global de no respuesta de los cónyuges fue del 14%. La mayor parte de esta no respuesta corresponde a cónyuges de personas seleccionadas que no fueron encontradas (9%). Si se excluye a los cónyuges de personas seleccionadas que no fueron encontradas o que rechazaron la entrevista, la tasa de no respuesta entre los cónyuges elegibles fue del 4,3%.

Si un matrimonio se disolvió entre la primera y la segunda ronda, no se intentó volver a entrevistar al excónyuge. Hay 31 casos así, que se excluyen del denominador de la tasa de respuesta porque no son elegibles.

Cuadro 2. Tasas de respuesta de los cónyuges en la ronda 2 (“ronda de 2012”) de CRELES-RC, según la situación de reentrevista de la persona seleccionada

Respuesta	Muestra total: frec. absoluta	Muestra total: frec. relativa	Excluyendo personas seleccionadas que no respondieron: frec. absoluta	Excluyendo personas seleccionadas que no respondieron: frec. relativa
Aceptó	1.138	86,1	1.138	95,7
- Vivo	1.097	83,0	1.097	92,3
- Fallecido	29	2,2	29	2,4
- Nuevo cónyuge	12	0,9	12	1,0
No respuesta	183	13,9	51	4,3
- Rechazo	61	4,6	22	1,9

Respuesta	Muestra total: frec. absoluta	Muestra total: frec. relativa	Excluyendo personas seleccionadas que no respondieron: frec. absoluta	Excluyendo personas seleccionadas que no respondieron: frec. relativa
- No encontrado	120	9,1	29	2,4
- No visitado	2	0,2	0	0,0
Denominador de la tasa de respuesta	1.321	100,0	1.189	100,0
Unión disuelta (divorcio o separación)	31		31	
Personas seleccionadas que no respondieron			130	
Total	1.352		1.352	

Durante la primera ronda, el equipo de campo entrevistó a 51 cónyuges cuyas personas seleccionadas no pudieron ser entrevistadas por rechazos tardíos o por problemas para contactarlas. Hay una base de datos aparte de la ronda 1 con estos 51 cónyuges. Durante la segunda ronda, el equipo de campo contactó a 30 de estos cónyuges. Hay un archivo aparte con estas 30 entrevistas; no se incluyen en los cuadros de tasas de respuesta anteriores.

Durante la segunda ronda se hicieron mediciones antropométricas al 99% del total de la muestra de personas seleccionadas. Entre los cónyuges, el 92% tiene mediciones antropométricas; esta mayor no respuesta se debe a que el 6% de las entrevistas (72 casos) se hizo solo por teléfono, por lo que no fue posible tomar las medidas. Solo 3 de los 75 cónyuges entrevistados por teléfono tienen mediciones antropométricas, porque el personal de campo hizo citas especiales para tomar las medidas después de que la entrevista se había completado.

Sin contar las entrevistas sobre personas fallecidas, entre el resto de los entrevistados solo el 2,2% de las personas seleccionadas y el 1,2% de los cónyuges 1,8% necesitaron que otra persona respondiera en su nombre (informante sustituto o “proxy”).

Para analizar la primera ronda de datos se calcularon tres conjuntos de factores de ponderación. La primera variable de ponderación, llamada factor_cort, corresponde a la muestra de personas seleccionadas en formato largo y la muestra suplementaria de personas seleccionadas en formato corto analizadas juntas (una añadida a la otra). La segunda variable de ponderación, factor_princ, se calculó para hacer inferencias sobre la población a partir solo de la muestra de personas seleccionadas en formato largo. Por último, la tercera variable de ponderación, factor_cony, corresponde a la muestra de cónyuges; esta última puede usarse para hacer inferencias sobre la población de cónyuges y para analizar parejas.

Para el análisis de la ronda 2 se calcularon dos nuevos conjuntos de factores de ponderación. La primera variable de ponderación, factorprinc2wave, es solo para las personas que respondieron el cuestionario de formato largo en la primera ronda y fueron entrevistadas durante la segunda ronda. La segunda variable de ponderación, factorcony2wave, es para los cónyuges entrevistados en la ronda 2, con el fin de analizar parejas. Ambos conjuntos de factores corrigen la no respuesta diferencial en la segunda ronda. Ambos deben usarse con los datos de la segunda ronda si se hace un análisis longitudinal.

El cuadro 3 compara las distribuciones de algunas variables usando los factores de ponderación de la primera ronda y los de la segunda ronda. La variable de edad se refiere a la edad en la primera ronda.

Cuadro 3. Comparación de las distribuciones de frecuencias relativas de la Cohorte de Retiro CRELES 1945-1955 para la submuestra de personas entrevistadas en la segunda ronda, ponderadas con los dos conjuntos de factores de ponderación

Características	Personas seleccionadas: ponderación de la primera ronda, formato largo	Personas seleccionadas: ponderación de la segunda ronda	Cónyuges: ponderación de la primera ronda, formato largo	Cónyuges: ponderación de la segunda ronda
Edad en la primera ronda^{1/} (total)	100,0	100,0	100,0	100,0
55-58	42,8	43,1	54,1	53,2
59-62	34,2	34,4	16,5	16,5
63-65 (o más)	23,0	22,6	29,4	30,2
Estado conyugal (total)	100,0	100,0	100,0	100,0
En unión libre	10,6	10,6	15,4	15,6
Casado(a)	60,4	60,5	84,6	84,4
Separado(a)/divorciado(a)	15,6	15,4	-	-
Viudo(a)	7,0	7,0	-	-
Soltero(a)	6,5	6,5	-	-
Educación (total)	100,0	100,0	100,0	100,0
Sin escolaridad/primaria	57,3	56,8	61,9	62,6
Secundaria académica	21,4	21,5	22,3	21,9
Postsecundaria y técnica	21,2	21,5	15,4	15,2
NS/NR	0,1	0,1	0,3	0,3
Trabaja (total)	100,0	100,0	100,0	100,0

Características	Personas seleccionadas: ponderación de la primera ronda, formato largo	Personas seleccionadas: ponderación de la segunda ronda	Cónyuges: ponderación de la primera ronda, formato largo	Cónyuges: ponderación de la segunda ronda
Sí	42,6	44,1	39,0	39,2
No	57,4	55,9	61,0	60,8

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 segunda ronda, que recogió información entre 2012 y 2014. Incluyó una entrevista estructurada, mediciones antropométricas y pruebas de funcionamiento físico; como se explicó antes, a diferencia de la primera ronda, en la segunda no se tomaron 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. Para mejorar las tasas de respuesta, 71 entrevistas de cónyuges se hicieron por teléfono en vez de en persona. Hubo un formulario de consentimiento informado especial para estas entrevistas telefónicas. El personal de campo hizo citas con algunos de los cónyuges entrevistados por teléfono para tomar las mediciones antropométricas y las pruebas de movilidad, pero la mayoría de las entrevistas telefónicas no tiene esta información.

Hay también un “cuestionario de salida” para las personas seleccionadas y los cónyuges que fueron entrevistados en la primera ronda pero fallecieron antes de ser entrevistados en la segunda. Los entrevistadores debían entrevistar a un informante sustituto que hubiera sido cercano a la persona fallecida y que tuviera suficiente información sobre sus últimos días. El “cuestionario de salida” es también una versión corta del cuestionario de formato largo. Tiene preguntas adicionales sobre limitaciones en las actividades de la vida diaria (AVD) antes de la muerte, causas generales de la muerte, estancias en hospitales y centros de cuidado, y herencias.

Durante la primera ronda, el equipo de campo registró con dispositivos GPS las coordenadas geográficas del lugar de residencia de cada participante. Esta tarea no se hizo en la segunda ronda, porque la mayoría de las personas no había cambiado de residencia. Todos los datos se recogieron en los hogares de los participantes, por lo general en una sola

visita. Los participantes dieron su consentimiento informado mediante su firma en ambas entrevistas. La entrevista duraba unos 90 minutos; durante ese tiempo, la presión arterial se midió dos veces. Al final de la entrevista se tomaron las mediciones antropométricas y se realizaron las pruebas de funcionalidad física.

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 66 entrevistas (1,85% del total de entrevistas que no son de salida) se realizaron con la ayuda de un informante sustituto.

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. En la cohorte CRELES-RC, el modelo de PDA usado en el trabajo de campo fue el HP-IPAQ-111 Pocketbook.

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. Como la entrevista del cónyuge era más corta, a veces el intervalo entre las dos mediciones también lo fue.

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.

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.

La talla y la altura de rodilla se midieron solo en la primera ronda, porque los investigadores suponen que, a estas edades, cambian poco en un intervalo de dos años. Por lo tanto, no hay información de talla ni de altura de rodilla en las bases de datos de la segunda ronda. El procedimiento para tomar estas medidas en la primera ronda fue el siguiente:

Talla: se usó un estadiómetro marca Seca para medir la estatura de las personas adultas mayores durante la primera ronda. 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).

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.

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, el 23 de febrero de 2011 y el 18 de abril de 2012 (referencias: VI-2878-2009, VI-5308-2009, VI-1313-2012 y VI-2403-2012), 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. Durante la primera ronda, algunas personas de la muestra suplementaria (las “entrevistas cortas”) fueron entrevistadas por teléfono; durante la segunda ronda, también algunos cónyuges fueron entrevistados 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 I

Diseño muestral de la ronda de base y de la segunda ronda

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

Población de estudio:

Primera ronda: 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).

Segunda ronda: personas nacidas entre 1945 y 1955 que vivían en Costa Rica entre enero de 2012 y febrero de 2014, y sus cónyuges corresidentes (independientemente de su año de nacimiento). A las personas que se mudaron a otros países no se les dio seguimiento en la segunda ronda.

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

Tamaño final de la muestra:

Primera ronda: 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.

Segunda ronda: se completaron entrevistas con 2.472 personas seleccionadas que estaban vivas y con 1.142 cónyuges (“entrevistas de cónyuges”). De las 2.472 entrevistas de personas seleccionadas, 50 corresponden a entrevistas de salida sobre personas fallecidas. Se lograron 30 entrevistas de salida sobre cónyuges fallecidos. Además, en la segunda ronda se entrevistó por primera vez a 13 personas que habían iniciado una unión con 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 pseudosegmentos censales descritos antes.

Procedimiento de trabajo de campo:

Primera ronda: los entrevistadores iban a cada pseudosegmento censal divididos en dos grupos de 3 entrevistadores y un chofer (uno de ellos era el supervisor de campo). Visitaban cada vivienda del pseudosegmento 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 50 cónyuges sin información sobre la persona seleccionada.

Segunda ronda: el jefe de campo organizó el trabajo para tener el mayor número posible de entrevistas dentro de un área determinada de antemano. También lo organizó para que hubiera al menos 18 meses entre la entrevista de base y la de seguimiento. El equipo se redujo a un solo grupo de solo 3 entrevistadores y 1 supervisor/chofer. La información de los hogares que no respondían se obtuvo de los vecinos o de informantes de contacto cuyos datos se anotaron en la bitácora física de cada entrevista de hogar durante la primera ronda. Los entrevistadores buscaban el nombre y el número de identificación de cada entrevistado en la PDA para mantener la misma identificación y vincular los registros de la segunda ronda con los de la primera. Si la persona seleccionada se había casado recientemente, el registro del nuevo cónyuge se introducía en la base de datos electrónica de la PDA. Algunos cónyuges entrevistados en la primera ronda sin la información de la persona seleccionada se volvieron a entrevistar en la segunda ronda. Hay 29 cónyuges en la base de datos de la segunda ronda sin información sobre la persona seleccionada. No se calcularon factores de ponderación para estos cónyuges sin información sobre la persona seleccionada.

Probabilidad de selección:

*Primera ronda*²: 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:

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

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

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

² Describimos la fórmula de la probabilidad de selección aplicada solo a quienes respondieron el formato largo en la ronda 1 porque, durante la segunda ronda, se entrevistó a personas que en la primera ronda habían respondido solo el cuestionario de formato largo.

- (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:
- 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).
- (d) **factor4:** la probabilidad de responder el cuestionario de formato largo en vez del de formato corto. Del total de entrevistas a personas seleccionadas, 2.798 (85%) respondieron el cuestionario de formato largo y 491 respondieron la entrevista de formato corto. Usamos dos modelos de regresión logística (uno para hombres y otro para mujeres) para calcular las probabilidades diferenciales de responder el cuestionario de formato largo, controlando por varias covariables. Los modelos se simplificaron 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). Entre los hombres, los más educados y quienes vivían fuera de Cartago tuvieron más probabilidad de responder el formato largo. Entre las mujeres, las de

más edad, las que reportaron peor salud, las no casadas y las que trabajaban tuvieron más probabilidad de responder el formato largo; además, las mujeres que vivían en Cartago, Puntarenas o Alajuela tuvieron menos probabilidad de responder este tipo de cuestionario³.

La fórmula final de la probabilidad de selección es:

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

$$* factor1 * factor2 * factor3 * factor4$$

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

$$* factor1 * factor2 * factor3 * factor4$$

Segunda ronda: el factor de ponderación de la segunda ronda corrige los factores de la primera ronda por la no respuesta diferencial en la nueva etapa de entrevistas. Por lo tanto, el factor de ponderación puede resumirse como el inverso del producto de la probabilidad de selección en la primera ronda por la probabilidad de responder en la segunda ronda.

$$P(X, 2 \text{ ronda}) = P(UPM) * P(USM) * P(UTM) * P(U4M)$$

$$* factor1 * factor2 * factor3 * factor4 * P(\text{segunda ronda})$$

$$P(X, 2 \text{ ronda}) = \frac{60}{102} * \frac{Mt_{\alpha}}{984.89} * 1 * \frac{1}{\text{Número de personas elegibles por edad en el hogar}}$$

$$* factor1 * factor2 * factor3 * factor4 * P(\text{segunda ronda})$$

La probabilidad de responder en la segunda ronda se estima con un modelo logístico para cada sexo (hombres y mujeres), en el que la variable dependiente binaria es igual a uno si la persona no respondió el cuestionario de la segunda ronda, y las covariables provienen de la base de datos de la primera ronda. El cuadro 4 describe el modelo estimado. En ambos grupos, residir en el Área Metropolitana y ser inmigrante aumentan la probabilidad de no responder en la segunda ronda. Estos factores se observaron durante el trabajo de campo. Fue más difícil completar entrevistas en el Área Metropolitana, donde está la capital, San José; se trató de un problema de horarios y de otras características de la vida urbana, más que de problemas de accesibilidad geográfica. Las personas nacidas en el extranjero fueron más difíciles de encontrar en la segunda ronda porque algunas se mudaron de nuevo a otro

³ Para ver el modelo, consulte el documento de muestreo y metodología correspondiente a la primera ronda.

país, temporal o permanentemente; el 2% de las personas seleccionadas cambió de dirección, algunas al extranjero, por lo que el personal de campo no pudo entrevistarlas. Los hombres jubilados tuvieron más probabilidad de responder la entrevista, mientras que las mujeres que trabajaban fueron más difíciles de encontrar en la segunda ronda. Además, los hombres que reportaron al menos una limitación funcional tuvieron menos probabilidad de ser entrevistados.

Cuadro 4. CRELES-RC. Modelos de regresión logística para estimar la probabilidad de no responder en la segunda ronda, dado que la persona respondió el cuestionario de formato largo en la primera ronda, por sexo

Covariables	Coefficiente	Error estándar	Valor p
Hombres			
Reside en el Área Metropolitana	0,926	0,184	0,000
Jubilado	-0,783	0,242	0,001
Con limitaciones funcionales	0,630	0,211	0,003
Nacido en el extranjero	0,596	0,318	0,061
Intercepto	-2,313	0,159	0,000
Mujeres			
Reside en el Área Metropolitana	0,602	0,175	0,001
Trabaja	0,476	0,176	0,007
Nacida en el extranjero	0,659	0,271	0,015
Intercepto	-2,787	0,153	0,000

Probabilidad de selección de los cónyuges (o para análisis de parejas):

Primera ronda: 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. 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. Las probabilidades diferenciales de responder el cuestionario de cónyuges, controlando por varias covariables, se calcularon con otro modelo logístico⁴, simplificado 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,

⁴ Para ver el modelo, consulte el documento de muestreo y metodología correspondiente a la primera ronda.

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:

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

$$P(X. sp) = P(X \text{ con formato largo})$$

$$* P(\text{cónyuge entrevistado} \mid \text{persona seleccionada entrevistada})$$

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

$$* factor1 * factor2 * factor3 * factor4$$

$$* P(\text{cónyuge entrevistado} \mid \text{persona seleccionada entrevistada})$$

Segunda ronda: en la segunda ronda debían entrevistarse todos los cónyuges, salvo cuando la unión había terminado. Como se mencionó antes, la tasa de no respuesta de los cónyuges llega al 14%. Por lo tanto, el factor de ponderación de los cónyuges es el inverso del producto de la probabilidad de selección del cónyuge en la ronda 1 por la probabilidad de ser entrevistado en la ronda 2. La fórmula de la probabilidad de selección de los cónyuges puede expresarse así:

$$P(X. sp. 2 \text{ ronda}) = P(\text{cónyuge} \mid X \text{ casado y con formato largo})$$

$$* P(\text{cónyuge responde en la 2.ª ronda})$$

$$P(X. sp. 2 \text{ ronda}) = P(X. sp) * P(\text{cónyuge responde en la 2.ª ronda})$$

La probabilidad de que un cónyuge responda en la segunda ronda varía entre grupos. Para identificar las fuentes más importantes de variación en esa probabilidad, usamos una regresión logística para cada sexo para estimar el conjunto de probabilidades. Usamos un procedimiento por pasos (stepwise) para estimar un modelo parsimonioso, solo con las variables más relevantes. Luego estimamos las probabilidades de respuesta con el modelo logístico y multiplicamos el inverso de estas probabilidades por los factores de ponderación de la primera ronda para calcular los factores de ponderación de los cónyuges para la ronda 2. Para tener más información, las covariables elegidas para el modelo son variables de la base de datos de las personas seleccionadas, que tiene más variables. Esta decisión toma en cuenta, además, que la mayor parte de la no respuesta entre los cónyuges se debe a la no respuesta de las personas seleccionadas.

En los hombres, residir en áreas urbanas y en el Área Metropolitana aumenta la probabilidad de no respuesta; además, si la persona seleccionada tiene mala salud, es más probable que su cónyuge no responda la encuesta. En las mujeres, residir en el Área Metropolitana aumenta la probabilidad de no respuesta. Si la persona seleccionada trabaja, es más probable que el cónyuge no responda. También es más probable que el cónyuge no responda si la persona seleccionada tiene limitaciones funcionales, pero es más probable que responda si la persona seleccionada reporta mala salud.

Cuadro 5. CRELES-RC. Modelo de regresión logística para estimar la probabilidad de que un cónyuge entrevistado en la primera ronda no respondiera en la segunda ronda

Características de las personas seleccionadas	Coeficiente	Error estándar	Valor p
Hombres			
Reside en áreas urbanas	0,634	0,306	0,038
Reside en el Área Metropolitana	0,582	0,296	0,049
La persona seleccionada tiene mala salud	0,707	0,267	0,008
Intercepto	-3,137	0,320	0,000
Mujeres			
Reside en el Área Metropolitana	0,677	0,219	0,002
La persona seleccionada tiene limitaciones funcionales	0,828	0,272	0,002
La persona seleccionada trabaja	0,584	0,261	0,025
La persona seleccionada tiene mala salud	-0,451	0,236	0,055
Intercepto	-2,461	0,297	0,000

Anexo II

Descripción del archivo de seguimiento Track_RC_V1.dta

El archivo de seguimiento Track_RC_V1.dta contiene información básica sobre cada una de las personas entrevistadas durante el trabajo de campo de CRELES-RC. El archivo puede usarse para saber cuáles entrevistados tienen datos longitudinales.

Variable	Descripción
idquest	Tipo de entrevista: personas seleccionadas con el cuestionario de formato largo, cónyuges y personas seleccionadas con el cuestionario de formato corto.
idsujeto	Número de la persona adulta mayor: número de identificación de cada entrevistado
hhid	Identificador del hogar
idlote	Unidad Primaria de Muestreo
nwaves	Número de rondas con datos. Las personas que respondieron el formato corto fueron entrevistadas en una sola ronda.
sex	Sexo
track_w1	Situación en la ronda 1: 1 “Interviewed” 6 “Not yet spouse” 8 “Interviewed-Spouse wo/target” Los clasificados como “Not yet spouse” (aún no cónyuge) son cónyuges de personas seleccionadas que iniciaron su unión después de la primera ronda, pero antes de la segunda. Los clasificados como “Interviewed-Spouse wo/target” (cónyuge entrevistado sin persona seleccionada) son cónyuges que fueron entrevistados, pero cuya persona seleccionada correspondiente no fue entrevistada. El personal de campo no tenía la intención de entrevistarlos, pero algunos se entrevistaron por inconsistencias durante el trabajo de campo.
track_w2	Situación en la ronda 2: “Interviewed” 2 “Death by wave date” 3 “Lost between waves 1 & 2” 4 “No follow-up-Short form” 7 “No longer a spouse” 8 “Interviewed-Spouse wo/target” 10 “Lost-Sp.wo/target” (entrevistado; fallecido a la fecha de la ronda; perdido entre las rondas 1 y 2; sin seguimiento, formato corto; ya no es cónyuge; cónyuge entrevistado sin persona seleccionada; cónyuge sin persona seleccionada perdido) El trabajo de campo no dio seguimiento a los entrevistados con el

Variable	Descripción
	formato corto. Algunos cónyuges sin persona seleccionada se perdieron para el seguimiento en la ronda 2.
fentr1	Fecha de la entrevista, ronda 1
fentr2	Fecha de la entrevista, ronda 2
dwave1	Fecha mediana en que se visitó la Unidad Primaria de Muestreo en la ronda 1
dwave2	Fecha mediana en que se visitó la Unidad Primaria de Muestreo en la ronda 2
agew1	Edad en la ronda 1
agew2	Edad en la ronda 2
proxyw1	Variable indicadora de si la entrevista requirió un informante sustituto (proxy) en nombre del informante principal en la ronda 1
proxyw2	Variable indicadora de si la entrevista requirió un informante sustituto (proxy) en nombre del informante principal en la ronda 2
rantrow1	Variable indicadora de si el entrevistado tiene datos de mediciones antropométricas en la ronda 1
rantrow2	Variable indicadora de si el entrevistado tiene datos de mediciones antropométricas en la ronda 1
phonew1	Variable indicadora de si la entrevista se hizo por teléfono en la ronda 1. En la ronda 1, solo algunas entrevistas de formato corto y algunas de cónyuges se hicieron por teléfono.
phonew2	Variable indicadora de si la entrevista se hizo por teléfono en la ronda 2. En la ronda 2, solo algunas entrevistas de cónyuges se hicieron por teléfono.
biomarkers	Variable indicadora de si el entrevistado tiene información de biomarcadores.
sp_wo_target	Variable indicadora de si el entrevistado es un cónyuge cuya persona seleccionada correspondiente no fue entrevistada.
ponderador	Factor de ponderación de la ronda 1
ponde_r2	Factor de ponderación de la ronda 2
ponde_shortlong	Factor de ponderación de las personas seleccionadas si se juntan los entrevistados de formato corto y largo.
gam	Variable indicadora de si el entrevistado vive en la Gran Área Metropolitana.