

DECLARACIÓN ERRÓNEA DE LA EDAD DE ADULTOS Y ANCIANOS EN ESPAÑA: EVALUACIÓN DE LA CONSISTENCIA DE LOS DATOS PARA EL ANÁLISIS DE LA MORTALIDAD, 1900-1930

Fernando Pastor G.^a-Quismondo¹

Resumen

La carencia de información disponible acerca de la fiabilidad de los datos ha representado un impedimento para el conocimiento de los patrones de mortalidad a edades avanzadas en el pasado. El estudio de la mortalidad a edades avanzadas padece de dos problemas fundamentales, incluso en datos muy recientes. Por una parte, la inexactitud de la edad declarada, y por otra, los problemas derivados del cálculo de tasas específicas de mortalidad con contingentes de eventos muy reducidos a ciertas edades. En el presente artículo se ha empleado el método de cohortes intercensales para la detección de pautas de declaración errónea de la edad. La evaluación de la consistencia entre fuentes de información se realiza con datos censales de población de edades superiores a los 40 años, con relación al registro civil de defunciones en España y saldos migratorios a nivel nacional de 1900 a 1930. Para ello se emplearon datos censales de población por edad de los censos de 1900, 1910, 1920 y 1930; datos de defunciones anuales por edad del Movimiento Natural de la Población de los años 1900 a 1930; y las distribuciones de entradas y salidas de las Estadísticas de Pasajeros por Mar. Además, se propone un método de aproximación a las estructuras de población y defunciones para el cálculo de tasas específicas de mortalidad a edades avanzadas durante el periodo.

1 Universidad Complutense de Madrid

Palabras Clave: Mortalidad, Mortalidad a edades avanzadas, Declaración errónea de la edad.

Abstract

The lack of information available about data reliability has represented an obstacle to know about old age mortality patterns in the past. Studying old age mortality suffers from two fundamental problems, even with very recent data. On one hand, age reporting inaccuracies, and on the other hand, problems related to mortality indicators computed from reduced sample sizes at specific ages. In the present article, the intercensal cohorts method has been employed to detect age misreporting patterns. Consistency assessment between data sources is done employing census data, vital registration of deaths and net migration indicators over 40 years of age in Spain from 1900 to 1930. For this purpose have been employed data on counts by age and sex from 1900, 1910, 1920, and 1930 census. Annual data on deaths, outmigrants, and immigrants by age and sex over 40 years from 1900 to 1930 have also been gathered. Although in this paper only results for the total population are shown. Moreover, an approximation method to the true population and deaths distributions by age is proposed in order to compute more reliable age specific mortality indicators for the period.

Key Words: Mortality, Old age mortality, Age misreporting.

Résumé

La manque d'information disponible sur la fiabilité des données a représenté un obstacle pour la connaissance des tendances de la mortalité aux âges élevées pendant le passé. L'étude de la mortalité en ages élevées souffre deux problèmes fondamentaux, même pour données très récentes. D'un côté, l'inexactitude de l'âge déclarée, et de l'autre, les problèmes dérivés du calcul des taux spécifiques de mortalité avec échantillons très réduites dans certaines ages. Dans cet article on a employé le methode des cohortes intercensals pour la détection des tendances de déclaration erroné de l'age. L'évaluation de la consistance entre sources d'information a été réalisé avec des données des recensements de la population d'ages supérieures à 40 ans, par rapport au Régistre Civil de décès en Espagne et net migration au niveau national de 1900 à 1930. Pour ce but, des données de population par age des recensements de 1900, 1910, 1920 et 1930 ont été employées; aussi comme des données de décès annuels par age du Mouvement Naturel de la Population des ans 1900 à 1930; et les distributions d'entrées et

sorties des Statistiques de Passagers par Mer. Et en plus, un methode d'approximation aux structures de population et décès est proposé pour le calcul des tauxes spécifiques de mortalité en ages élevées pendant le période.

Mots clé: Mortalité, Mortalité aux âges élevées, Déclaration erroné de l'age.

1. Introducción

Las primeras décadas del siglo XX en España fueron tiempos de grandes y rápidos descensos de la mortalidad. Tablas de mortalidad realizadas para las tres primeras décadas del siglo (Dopico y Reher, 1999), indican una esperanza de vida al nacimiento de 30 años en 1900, que alcanza los 48 años en 1930, a un ritmo de crecimiento de 0.6 años de ganancia de esperanza de vida anual. Sin embargo, estas mejoras no afectan con la misma intensidad a todos los grupos de edades. Para 1900 la esperanza de vida a los 50 años es de 18 años, cifra que tan solo se incrementa en dos años hasta los 20 en 1910 y se mantiene en la misma cantidad tanto en 1920 como en 1930. Como es sabido, las primeras ganancias en esperanza de vida al nacimiento corrieron a cargo de la mortalidad infantil y juvenil. La mortalidad en el primer año de vida pasa de una tasa de 186 por mil en 1900 a 117 por mil en 1930, para seguir descendiendo posteriormente (Gómez Redondo, 1992). A continuación, durante un período de estancamiento de la mortalidad general, es la mortalidad adulta quien comienza su descenso, para dar paso, según Caselli (1991), a un pronunciado descenso de la mortalidad de ancianos a partir de la década de los años sesenta.

Por lo tanto, en el primer tercio del siglo XX, habría una falta de mejoras en la mortalidad a edades avanzadas, al menos de forma aparente. No obstante, numerosos investigadores que han estudiado la mortalidad a edades avanzadas toman con precaución estos cálculos (Coale y Caselli, 1990; Alter, 1990 y 1995; Condran Himes y Preston, 1991; Preston, Elo, Rosenwaike y Hill 1996; entre otros). Se tiene conocimiento de que durante largas épocas la información proveniente de las operaciones censales y del registro civil de acon-

tecimientos vitales acerca de individuos adultos y sobre todo de ancianos plantean dudas acerca de su fiabilidad. Esta preocupación esta fundada sobre varios fenómenos, unos concernientes al subregistro de defunciones de ancianos y subrecuento de los mismos, y otros relacionados con la declaración errónea de la edad.

El subrecuento de ancianos es típico de áreas rurales poco accesibles para los agentes censales de la época y el subregistro de sus defunciones es también común entre aquellos que viven solos y no tienen quien declare su defunción. Por otra parte están los fenómenos de declaración errónea de la edad, frecuentemente cometidos por personas de edad avanzada al ser censadas, o bien a la hora de que un tercero tenga que comunicar la edad del óbito en el registro de su defunción. Dichos fenómenos de declaración errónea de la edad reciben la denominación de preferencia de edad, o acumulación a ciertas edades. Consisten en una preferencia a declarar edades terminadas en « 0 », y en menor medida en « 5 », junto a una tendencia a exagerar su edad y están relacionados con poblaciones de bajos niveles de educación, como es el caso de los adultos y ancianos del primer tercio de siglo, así como con la pérdida de memoria asociada a las estas edades (Ewbank, 1981).

2. Breve revisión de algunos procedimientos de estimación para el análisis de la mortalidad a edades avanzadas

2.1. Método de Generaciones Extinguidas

A principios de la década de los años cincuenta Vincent (1951) desarrolla un original método para calcular la mortalidad a edades avanzadas. El nombre designado para él método fue el de generaciones extinguidas. En este se plantean dos supuestos básicos: que la generación del año n se extingue 110 años después y que la movilidad en estas edades es casi nula. Para saber, por ejemplo, los supervivientes de la generación del año n a la edad x , se generan acumulados de defunciones de la cohorte de nacidos en el año n superiores a la edad x , donde el límite superior es el año $n+110$. La

cantidad resultante es, obviamente, igual al número de supervivientes a x años de la generación del año n , suponiendo que tiene lugar una declaración exacta de la edad. De esta forma,

$$S_x^n = \sum_s^{\omega} D_x^n \quad ; \text{ y por lo tanto:}$$

$$q_x^n = D_x^n / S_x^n \quad ; \text{ donde:}$$

S_x^n son los supervivientes de la generación del año n a la edad x ,

$\sum_x^{\omega} D_x^n$ son los difuntos acumulados nacidos el año n superiores a la edad x y anteriores al año $n+110$,

q_x^n es la probabilidad de morir a la edad x habiendo nacido el año n .

D_x^n son los difuntos de la generación del año n a la edad x , y

Este método es apropiado para poblaciones recientes con un alto grado de fiabilidad en sus datos estadísticos, y donde no tenga lugar el fenómeno de declaración errónea de la edad. Además necesita de registros de defunciones por edad y generación de nacimiento. Este método ha sido empleado por Kannisto (1990) para la Finlandia contemporánea y parcialmente por Preston, Elo, Rosenwaike y Hill (1996), para la mortalidad anciana de la población negra actual en los Estados Unidos.

2.2. Técnicas con variables de crecimiento

Con la creación de un método para la estimación de la bondad de los registros de defunciones con relación a variables de crecimiento « r » de la población, Preston y Coale (1982), abren la puerta a toda una serie de derivaciones del mismo llevadas a cabo por ellos y otros autores. El método necesita información disponible sobre la tasa de crecimiento de la población y sobre defunciones y población por edades quinquenales. Se supone que la población es estable, es decir, que la mortalidad y la fecundidad no ha cambiado en el pasado, y que hay ausencia de migraciones. En una población estable

con una tasa de crecimiento positiva, la población de edad x es mayor que la suma de las defunciones anuales de x y más años. Es decir, como la población crece, las defunciones anuales sobre la edad x pertenecen a una cohorte menor de nacimientos que la cohorte de nacimientos de donde ha sobrevivido la población de edad x . De esta forma, el ajuste de las defunciones surge de la comparación de la suma de defunciones de x y más años con la población de edad x . Este ajuste se hace multiplicando las mismas por un factor exponencial resultante de la tasa de crecimiento de la población y la edad media del grupo. Posteriormente, las defunciones se acumulan como estimación de la población a una cierta edad. Estimación que se compara con la población observada en esa edad en forma de razón, dando como resultado la bondad de las defunciones registradas.

Los supuestos implícitos en este método como la mortalidad y fecundidad constantes, así como la ausencia de migraciones, son condiciones con frecuencia muy distintas a las que se dan en las poblaciones reales. Además, los resultados podrían estar sesgados si la declaración errónea de la edad es diferente en las defunciones que en la población.

Derivado de este método se desarrolla el método de Bennett y Horiuchi (1984), para estimar el subregistro de defunciones sin asumir que la población es estable. Para éste son necesarias las distribuciones de población por grupos quinquenales en dos censos y las defunciones registradas entre los mismos. Supone que la bondad de las fuentes y la declaración errónea de la edad son iguales en ambos censos, y la migración es nula. Esta técnica estima la población en cada grupo de edad y lo compara con la población observada en cada grupo. La población estimada se basa en las defunciones registradas, poblaciones en edad exacta derivadas de la información censal y de tasas específicas de crecimiento calculadas con las poblaciones de ambos censos. Cualquier diferencia de cobertura o declaración errónea de la edad entre los dos censos, generaría unos resultados sesgados. Aplicaciones de métodos con variables de crecimiento pueden observarse en Cachinero Sánchez (1985) para el siglo XVIII español o Alter (1990) para principios del siglo XX en Estados Unidos.

2.3. *Matrices correctoras de la declaración errónea de la edad*

Como hemos mencionado anteriormente, aparte de la cobertura de las fuentes, el principal problema que presentan las fuentes de información estadística de adultos y ancianos es la declaración errónea de la edad. Si bien estas matrices no son un método de estimación de la mortalidad por si mismas, son filtros correctores de las estructuras de población y defunciones a las que posteriormente aplicar otros procedimientos. Por esta razón se incluyen en este apartado.

Estas matrices están compuestas por las edades declaradas en uno de sus ejes, y por la edad real en el otro. De esta forma se obtiene la distribución real de las edades de los individuos o eventos en cada edad declarada, permitiendo añadir a cada edad o grupo de edades los que están erróneamente ubicados fuera del mismo, y sustraer los que están equivocadamente dentro. La dificultad en este caso es como llegar a construir tales matrices. Una forma en que se ha realizado es mediante el cruce de dos fuentes de información con datos individuales. El procedimiento consiste en vincular a los individuos en ambas fuentes y comparar las edades que figuran en las mismas. Obviamente, se tiene que haber establecido, *a priori*, una jerarquía entre las fuentes. Concediéndole a una mayor fiabilidad que a la otra en lo concerniente a la edad declarada. La más fiable será la ubicada en el eje de las edades reales. Otra de las formas en que se puede generar la matriz es mediante la simulación de pautas de declaración errónea de la edad (de las que se habla más adelante), tomando las estructuras corregidas como edades reales.

Son muchos los estudios que ofrecen este tipo de matrices correctoras: Dechter y Preston (1991) comparando la edad declarada de los individuos con su fecha de nacimiento en Latinoamérica; Preston, Elo, Rosenwaike y Hill (1996) con certificados de defunción y registros de la seguridad social; Alter (1995) cruzando registros de defunciones con registros de las pensiones del ferrocarril a principios del siglo XX en Estados Unidos; o Mari Bhat (1995) con registros de defunciones y simulaciones para la India, etc.

2.5. *Simulación de pautas de declaración errónea de la edad*

Otra de las técnicas de estimación de estructuras de población y defunciones para el análisis de la mortalidad a edades avanzadas, es la simulación de pautas de error en la declaración de la edad para corregir esas estructuras. Inicialmente se establece el ritmo con que la declaración errónea aumenta con la edad y obtener la proporción total de edades mal declaradas en cada edad. Posteriormente, decidir que proporción de estos se adjudica a edades superiores y que proporción se adjudica a edades inferiores a cada edad. Finalmente se establece que fracción de los que se adjudican por encima y de los que se adjudican por debajo en cada edad, pertenecen a una, dos, tres o más edades por encima o por debajo de cada edad. Además, pueden emplearse diferentes pautas en los censos y en las defunciones. Evidentemente, el establecimiento de estas pautas siempre tiene un componente de arbitrariedad. Sin embargo, unos sólidos fundamentos teóricos a la hora de ser establecidas pueden hacer que resulten esclarecedoras. Trabajos como los de Condran, Himes y Preston (1991), Alter (1990) o Mari Bhat (1995) han empleado las simulaciones para definir las pautas de declaración errónea de la edad en diversos contextos.

Sin embargo, muchos de los trabajos mencionados no utilizan únicamente uno de estos métodos, sino varios de ellos en distintas fases del proceso, dada la complejidad del fenómeno. En nuestro caso, tras haber contrastado los diferentes métodos con relación a los datos de que disponemos para las primeras tres décadas españolas, no se ha empleado íntegramente ninguno de los métodos expuestos por los siguientes motivos. A principios de siglo en España no se dispone de datos de defunciones por edad y año de nacimiento. El grado de declaración errónea de la edad es elevado y diferente en los censos y en las defunciones. No sigue una regularidad aparente en sus pautas, y no ha respondido favorablemente a las pautas encontradas en otros ámbitos. Por lo tanto se ha optado por aplicar el método de cohortes intercensales (que describimos a continuación) para la detección de las pautas de declaración errónea de la edad sin imponer a los datos estructura alguna, en pri-

mera instancia. Posteriormente, se ha desarrollado un procedimiento de aproximación a las estructuras reales de población y defunciones para el posterior cálculo de indicadores de mortalidad.

2.5. El método de Cohortes Intercensales

Este método ha sido empleado principalmente para el cálculo de saldos migratorios a partir de dos estructuras censales de población y las cohortes de defunciones que tienen lugar entre las mismas. Sin embargo, si se dispone de los saldos migratorios de las cohortes durante el período intercensal o hay ausencia de migración, permite la medición de la consistencia que existe entre la diferencia del tamaño notificado de las cohortes del primer al segundo censo y el número de defunciones ocurridas en las cohortes entre los dos censos. La ecuación para estimar la población en el momento del segundo censo, a partir de la población observada en el primero y de las defunciones intercensales por cohortes sería:

$$N'_{i+j}(t+j) = N_i(t) - D_i + M_i$$

donde:

$N_i(t)$ = Población observada de edad i en el primer censo.

D_i = Defunciones intercensales registradas en la cohorte de edad i en el primer censo a lo largo del período intercensal, hasta alcanzar la edad $i+j$ en el segundo censo, efectuado j años después del primer censo.

M_i = Migración neta en la cohorte de edad i en el primer censo a lo largo del período intercensal, hasta alcanzar la edad $i+j$ en el segundo censo, efectuado j años después del primer censo.

$N'_{i+j}(t+j)$ = La población estimada de edad $i+j$ en el segundo censo, efectuado j años después del primer censo.

La consistencia entre las fuentes viene dada por la razón entre la población observada en el segundo censo y la población prevista:

$$N_{i+j}(t+j) / N'_{i+j}(t+j)$$

Siendo la razón 1 la consistencia perfecta entre las fuentes.

Es pertinente subrayar aquí que nos referimos a consistencia, lo que no implica necesariamente la exactitud. Es decir, puesto que ambas fuentes podrían cometer pautas idénticas de error, una razón igual a la unidad podría estar escondiendo errores que se manifiestan de igual forma en todas las fuentes. Es decir, si el subregistro y las pautas de declaración errónea de la edad fueran iguales en cada cohorte en las tres fuentes la razón sería igual a 1 aunque los datos serían inexactos. Además, la desviación de la unidad no es directamente imputable a una de las fuentes sino a una combinación del conjunto de las desviaciones de cada una de ellas. A pesar de todo, ésta razón no solamente constituye la mejor prueba, sino en opinión de algunos autores la prueba disponible más fiable de la calidad de datos para edades avanzadas, (Condran, Himes, y Preston 1991).

La aplicación del método requiere de datos censales de población por grupos de edad o edades simples en dos censos, de defunciones por edad, y de emigrantes e inmigrantes por grupos de edad. Los datos de defunciones y migraciones para cada uno de los años del período intercensal, siempre tomando el 31 de Diciembre de cada año como fecha de referencia.

Para este estudio se han recopilado las defunciones por edad para el total de la nación, a partir de 40 años de edad en todos los Movimientos Naturales de la Población (MNP) desde 1900 a 1930, excluidas las posesiones africanas. Siempre que existía la distribución por años simples, la hemos preferido a la distribución quinquenal. La conversión de grupos quinquenales a defunciones por edad simple en los años donde el MNP no las aporta (1900-1907, 1909 y 1918), se calcula aplicando al contingente de cada grupo la distribución media proporcional por edades simples de los dos años colindantes más próximos. Los datos de población se toman de los censos de 1900, 1910, 1920 y 1930 por grupos quinquenales. Los datos de migraciones se toman de las Estadísticas de Pasajeros por Mar: entradas y salidas en grandes grupos de edad. Para la aplicación del método con los datos españoles, se homogeneizan los grupos de edad de los Censos, los MNP y las Estadísticas de Pasajeros por Mar a grupos quinquenales que comienzan con edad terminada

en los dígitos « 0» o « 5». En todas las fuentes el contingente de la categoría « Sin edad» se ha distribuido de forma proporcional a la distribución por edades de dicha fuente. Puesto que los periodos intercensales son de diez años y los grupos de edad son de cinco años, todos los supervivientes al segundo censo formarán parte del contingente de población dos grupos de edad superiores al que pertenecieron en el primer censo teniendo en cuenta el saldo migratorio de la cohorte entre los censos.

Una ventaja de éste método es que permite que los resultados estén basados en datos empíricos más que en estimaciones indirectas. Lo que por otra parte conlleva un mayor esfuerzo de recolección y procesamiento de datos, ya que el seguimiento de las defunciones de las cohortes implica un gran volumen de información. La aplicación del método de cohortes intercensales lleva implícitos algunos supuestos, cuyo conocimiento es imprescindible para la correcta interpretación de los resultados. En primer lugar se considera que los nacimientos y las defunciones se encuentran equidistribuidos a lo largo del año. En segundo lugar se supone que las distintas fuentes cometen pautas idénticas de error, si no carecen de las mismas, y por tanto las diferencias se atribuyen al subregistro de defunciones, al subrecuento de población, a la mala declaración de la edad, o a una combinación de las tres en una o varias fuentes

3. Resultados

3.1. *Primer Acercamiento: Observación a partir de datos brutos*

En una fase preliminar se ha aplicado el método de cohortes intercensales a los datos originales de las fuentes con los procedimientos de desagregación e interpolación menos sofisticados con el propósito de observar las pautas de error por grupos de edad de la forma más bruta posible. Facilitando la comprensión de los errores posibles.

La desagregación en edades simples de la población y las defunciones se realiza de forma lineal. En el caso de las migra-

ciones las entradas y salidas de los mayores de 60 años se distribuyen según las distribuciones por edad de inmigrantes y emigrantes de la población noruega en 1996 si bien la intensidad es la de entradas y salidas de España en cada año de 1900 a 1930. Para la estimación de los migrantes de las cohortes comprendidas entre los 40 y 59 años se dispone solamente de un único grupo de edad de 20 a 59 años. En este caso se ha tenido que recurrir a la simulación matemática de comportamientos migratorios según la edad. Para el establecimiento de los parámetros de la simulación se ha comprobado que mientras el saldo migratorio general y el del grupo de 20 a 59 son negativos en las tres décadas; el saldo para mayores de 59 años es positivo. Es decir, las entradas tienen una distribución por edades más equilibrada que las salidas, en donde la concentración de emigrantes es mayor en las edades más jóvenes. Esto implica que pese a que tanto en entradas como en salidas el mayor peso corresponde a las edades más jóvenes dentro del grupo 20-59, hay una mayor proporción de edades mayores del grupo en las entradas que en las salidas. Hemos de aclarar en este punto que esta tendencia es acorde al conocimiento de los grandes flujos de retorno de emigrantes de estas edades durante el periodo (Sánchez Alonso, 1990) y al saldo migratorio positivo resultante de nuestros datos a partir de 60 años. Una vez se dispone de las entradas y las salidas por grupo de edad y año, se calculan los saldos migratorios por décadas y cohortes que se emplean como migración neta en el método de cohortes intercensales.

Por último, en este primer enfoque hemos considerado independientemente cada período intercensal. Pese a que las posibles distorsiones son acumulables de una década a otra para el conjunto de los treinta años, dada la complejidad del fenómeno, en esta fase se tratan por separado por motivos de operatividad analítica.

En la tabla 1 se muestra la aplicación del método de cohortes intercensales. El ejemplo corresponde a la población total de España para el período intercensal 1910-1920, donde puede verse: las poblaciones observadas que definen el periodo intercensal ; las

TABLA 1
Periodo Intercensal 1910-1920. Población total de España

| Grupos de edad de las cohortes en 1910 | Población observada en 1910 | Defunciones por cohortes durante el periodo intercensal | Saldo migratorio en las cohortes durante el periodo intercensal | Grupos de edad de las cohortes en 1920 | Población estimada para 1920 | Población observada en 1920 | Razón de población observada a la población estimada |
|--|-----------------------------|---|---|--|------------------------------|-----------------------------|--|
| | (A) | (B) | (C) | | $[(A)-(B)+(C)]=D$ | (E) | $[(E)/(D)]=F$ |
| 40-44 | 1075365 | 139002 | 43837 | 50-54 | 980200 | 923755 | 0.9424 |
| 45-49 | 1107498 | 162306 | 42271 | 55-59 | 987464 | 917571 | 0.9292 |
| 50-54 | 893553 | 250065 | 36009 | 60-64 | 731996 | 651839 | 0.8904 |
| 55-59 | 833956 | 197566 | 34443 | 65-69 | 618334 | 582162 | 0.9415 |
| 60-64 | 600482 | 291377 | 5136 | 70-74 | 314240 | 299074 | 0.9517 |
| 65-69 | 542113 | 308551 | 4494 | 75-79 | 238056 | 228302 | 0.9590 |
| 70-74 | 263995 | 266229 | 1926 | 80-84 | -309 | 77454 | / |
| 75-79 | 194465 | 191488 | 642 | 85-89 | 3619 | 39742 | / |
| 80-84 | 69708 | 121200 | 385 | 90-94 | -51107 | 10672 | / |
| 85-89 | 38519 | 43903 | 193 | 95-99 | -5191 | 3404 | / |
| 90+ | 14193 | 3754 | 128 | 100+ | 10568 | 960 | / |

defunciones por cohortes entre los censos; los saldos migratorios de las cohortes durante el período; la población estimada para el segundo censo; y la razón de la población observada en el segundo censo a la población estimada para el segundo censo. Esta razón se considera un indicador de consistencia entre las fuentes de datos. Este procedimiento se ha aplicado a cada período a la población total. En la tabla 2 se resumen todos los indicadores de consistencia para cada la grupo de edad grupo, y período intercensal. Si bien se aprecia una mejora en la consistencia entre las fuentes con el paso del tiempo, la pauta interna a cada período, es muy similar en el conjunto de los 30 años, dentro de sus niveles relativos de mortalidad.

El indicador de consistencia es inferior a la unidad en los tres períodos en los seis primeros grupos de edad con errores inferiores al 10 por ciento. Por lo tanto en estos grupos de edad la población estimada es superior en tamaño a la observada en el segundo censo en menos de un 10 por ciento. En los dos primeros periodos, los grupos de 55 a 64 años son los menos consistentes, mientras que en el tercero son los grupos menos consistentes ascienden de 65 a 74. Como puede comprobarse mirando a los niveles máximos de error por periodo, estos seis primeros grupos tienen una evolución cronológica positiva en lo que a la diferencia entre los niveles mas altos y más bajos de consistencia entre los grupos de edad se refiere. En el decenio 1900-1910 esta diferencia alcanza niveles próximos al 8 por ciento, que adquiere niveles inferiores al 6 por ciento en 1910-1920, para situarse cercano al 3 por ciento en 1920-1930. En las tres etapas encontramos que a partir de los 80-84 años de edad en el segundo censo hay grupos de cohortes cuyas defunciones intercensales, o bien son más numerosas que los propios contingentes de población de los que han de ser sustraídas, o bien están muy próximas en tamaño a las mismas poblaciones de riesgo en el primer censo. Así pues, los contingentes de población estimada para los grupos a partir de los 80 años de edad en el segundo censo resultan negativos o tan reducidos que la razón de la población observada a la población estimada —empleada aquí como indicador de consistencia entre las fuentes—, pierde cualquier capacidad explicativa.

TABLA 2
Indicadores de consistencia por edad y periodo

| <i>Población inicial de las cohortes de edades</i> | <i>Población final de las cohortes de edades</i> | <i>Indicador de consistencia 1900-10</i> | <i>Indicador de consistencia 1900-20</i> | <i>Indicador de consistencia 1900-30</i> |
|--|--|--|--|--|
| 40-44 | 50-54 | 0,9867 | 0,9424 | 0,9499 |
| 45-49 | 55-59 | 0,9084 | 0,9292 | 0,947 |
| 50-54 | 60-64 | 0,9074 | 0,8904 | 0,9386 |
| 55-59 | 65-69 | 0,9767 | 0,9415 | 0,9351 |
| 60-64 | 70-74 | 0,9792 | 0,9517 | 0,9377 |
| 65-69 | 75-79 | 0,9673 | 0,959 | 0,9652 |
| 70-74 | 80-84 | / | / | / |
| 75-79 | 85-89 | / | / | / |
| 80-84 | 90-94 | / | / | / |
| 85-89 | 95-99 | / | / | / |
| 90+ | 100+ | / | / | / |

Como se ha indicado anteriormente, las diferencias entre las poblaciones observadas y estimadas pueden basarse en varios motivos. Por una parte un plausible subregistro de defunciones, por otra parte el subrecuento de población, y por ultimo las pautas de declaración errónea de la edad. Sin embargo, dado el desequilibrio en los niveles de consistencia que encontramos entre los distintos grupos de edad, la declaración errónea de la edad sería el factor que estaría interfiriendo en nuestros resultados en primera instancia. Lo que por otra parte han mostrado ya otros estudios realizados en diversos países y épocas (Alter 1990), (Dechter, Preston 1991), (Mari Bhat 1995), o (Preston, Elo, Stewart 1997). Ciertamente las diferencias entre las poblaciones en los cinco últimos grupos de edad, están motivadas por la tendencia a declarar a la alta la edad de los individuos. Muchos de los habitantes y defunciones registrados en los grupos de cohortes comprendidas a partir de los 70 años, pertenecen en realidad a contingentes de personas y óbitos de grupos inferiores cuya edad es exagerada. Si bien a estas edades se pudiera estar exagerando la edad en las dos principales fuentes empleadas, población y defunciones, parece que la frecuencia de la exageración fuera mayor en las defunciones. Esto estaría en parte motivado por el hecho que en las defunciones la edad es declarada por alguien distinto al individuo afectado, puesto que un mayor subrecuento de población respecto al de

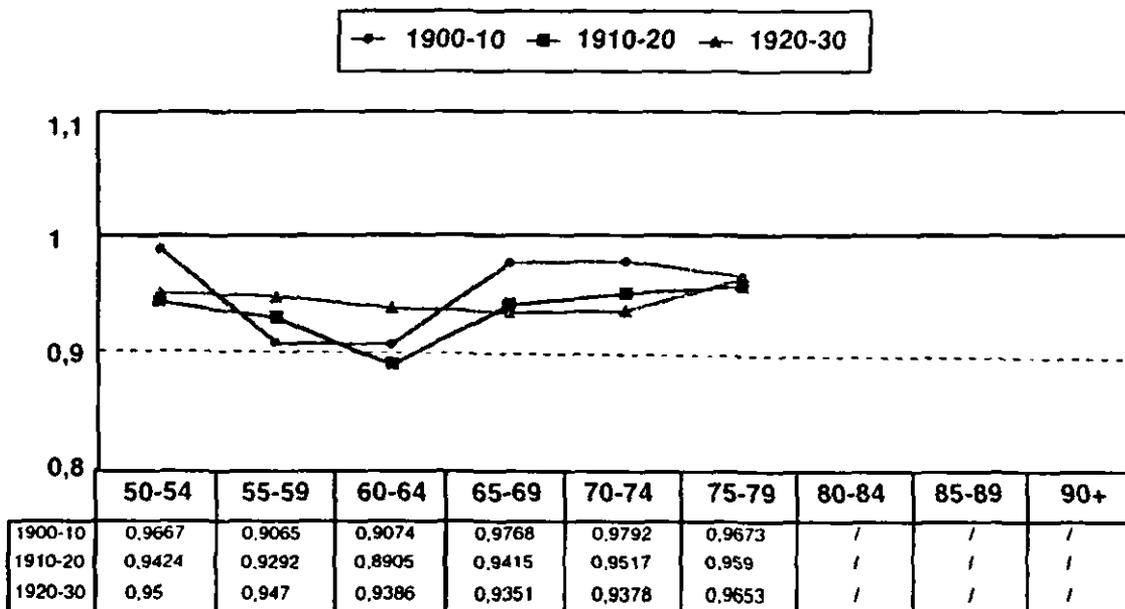
defunciones pese a que se encuentre interfiriendo en alguna medida, parece improbable que lo hiciera de forma tan abultada. El grupo 60-64 podría considerarse el punto a partir del cual la edad se comienza a declarar errónea y mayoritariamente a la alta.

A pesar de que la tendencia de declaración errónea de la edad es mayoritariamente hacia edades superiores, se han encontrado estadísticas demográficas en que la edad de individuos de las edades más avanzadas ha sido con frecuencia declarada en edades inferiores. En realidad la declaración errónea de la edad ocurre en ambas direcciones. Sin embargo, como consecuencia de que la distribución de las edades más avanzadas va disminuyendo de forma muy rápida, la base para transferencias hacia arriba dentro de un grupo de edad es mucho mayor que la base de transferencias hacia abajo fuera del mismo grupo. De esta forma, aunque la dirección neta de la declaración errónea de la edad fuera hacia abajo, podría haber más individuos que entran erróneamente en el grupo de edad desde el grupo inferior hacia arriba, que los que salen del mismo hacia abajo (Preston, Elo, Stewart 1997).

GRÁFICO 1

Indicador de consistencia entre fuentes

TOTAL



En el gráfico 1 se observa la progresión del indicador de consistencia según avanza la edad en los tres períodos analizados. Como ya observamos en la tabla 2, los valores del indicador de consistencia entre fuentes son inferiores a la unidad, en este caso representado por la línea gruesa, y por tanto la población estimada para el segundo censo es superior a la observada en el mismo. Si bien en la primera y segunda décadas la pauta es similar, con menores oscilaciones en la segunda, la tercera década es la que mantiene un mayor grado de paralelismo con la línea que representa la consistencia perfecta. El mantenimiento de niveles fronterizos a 0.94 implica una mayor regularidad de las pautas de error en el tramo de edades de 50 a 79 años.

En edades superiores el indicador deja de ser significativo y por lo tanto no se refleja en el gráfico. A modo de sinopsis podemos manifestar que a partir de los 60-64 años se comienza a dar una declaración errónea neta de la edad hacia edades superiores. Esta exageración a la alta viene a compensar en alguna medida el mismo fenómeno en las edades inmediatamente sucesivas, y estas a su vez, hacen incoherente la interpretación en los últimos grupos empleados. No obstante hemos de señalar dos especificidades, en primer lugar, que la primera década muestra que esta exageración neta podría comenzar desde los 55-59 años, y por último, que en el período 1920-30 los niveles de consistencia se mantienen poco por debajo de la unidad en todos los grupos observables, conformando una línea prácticamente paralela a la que delimita la consistencia perfecta. Esto último no implica la ausencia de la exageración neta, sino un grado más moderado y regular de la misma, junto a una probable mejora en la recogida de la información y en los niveles de educación de la población afectada.

3.2. Localización de la edad de origen de la exageración

Con el propósito de detectar en que grupos de edad comienza a exagerarse la edad, hemos aplicado de nuevo el método de cohortes intercensales mediante la acumulación inversa de contingentes de población, cohortes de difuntos, y saldos migratorios por cohortes. De esta forma nuestras cohortes carecen ahora de límite de edad

superior y el límite inferior aumenta en cinco años de edad en cada tramo. Con esto cada cohorte de edades es sucesivamente decreciente en tamaño a intervalos de esa misma duración, y contiene los individuos, defunciones y migrantes netos de la cohorte inmediatamente superior más todos los que abarca el intervalo de edad que va desde su propio límite inferior al límite inferior de la siguiente cohorte de edades. Con esto consigue controlar cualquier transferencia de individuos, defunciones y migrantes netos existente entre los grupos de edad comprendidos en el acumulado. Sin embargo, al ir descartando sucesivamente cohortes quinquenales por la parte inferior, y si consideramos que la pauta predominante de transferencia de individuos por declaración errónea de la edad es hacia arriba según avanza la edad, los acumulados contendrán mayor tamaño del que les corresponde, dado que reciben casos correspondientes a edades más tempranas.

En la tabla 3 se puede observar un ejemplo de la aplicación del método de cohortes intercensales a las distribuciones de cohortes acumuladas decrecientes. El ejemplo corresponde a la población de España en el periodo 1910-1920. De forma muy similar a como ocurre en las cohortes quinquenales en los tres períodos considerados, en la tabla 4 puede observarse como las defunciones intercensales registradas a partir de los 80+ años (siempre en el segundo censo) son superiores en número a la población de riesgo en el primer censo, susceptible de fallecer durante el período. En los indicadores de consistencia, a diferencia que en las cohortes quinquenales y exceptuando los dos primeros acumulados, la población observada en el segundo censo en las cohortes acumuladas es de mayor tamaño que la estimada para los mismos. Razones superiores a la unidad se dan en los tres períodos estudiados y en todos los acumulados para los que los datos empíricos permiten el cómputo. La tendencia ascendente del indicador de consistencia por encima de la unidad a partir de los 60+ años, refleja que según avanza la edad mayor es el peso que representa el número de defunciones pertenecientes a edades anteriores al límite inferior del acumulado. Un factor a tener en cuenta a la hora de determinar la correlación positiva de la declaración errónea de la edad a la alta con el incremento

TABLA 3
Cohortes acumuladas decrecientes 1910-1920. Población total

| Grupos de edad de las cohortes en 1910 | Población observada en 1910 | Defunciones por cohortes durante el periodo intercensal | Saldo migratorio en las cohortes durante el periodo intercensal | Grupos de edad de las cohortes en 1920 | Población para estimada en 1920 | Población observada en 1920 | Razón de población observada a la población estimada |
|--|-----------------------------|---|---|--|---------------------------------|-----------------------------|--|
| | (A) | (B) | (C) | | $[(A)-(B)+(C)] = D$ | (E) | $[(E)/(D)] = F$ |
| 40+ | 5633847 | 1975441 | 169462 | 50+ | 3827868 | 3734933 | 0.9757 |
| 45+ | 4558481 | 1836439 | 125626 | 55+ | 2847668 | 2811178 | 0.9871 |
| 50+ | 3450983 | 1674133 | 83355 | 60+ | 1860205 | 1893608 | 1.0179 |
| 55+ | 2557430 | 1476567 | 47346 | 65+ | 1128209 | 1241768 | 1.0179 |
| 60+ | 1723474 | 1226502 | 12903 | 70+ | 509876 | 659606 | 1.1006 |
| 65+ | 1122993 | 935125 | 7768 | 75+ | 195636 | 360532 | 1.2936 |
| 70+ | 580880 | 626574 | 3274 | 80+ | -42420 | 132230 | / |
| 75+ | 316885 | 360344 | 1348 | 85+ | -42111 | 54777 | / |
| 80+ | 122420 | 168856 | 706 | 90+ | -45730 | 15035 | / |
| 85+ | 52712 | 47656 | 321 | 95+ | 5377 | 4364 | / |
| 90+ | 14193 | 3753 | 128 | 100+ | 10568 | 960 | / |

de la edad, es el rápido descenso del tamaño de los grupos de edad más avanzada. Estos grupos son en ocasiones tan pequeños que incluso un pequeño número de intrusos procedentes de edades inferiores representa una alta proporción del total del grupo, especialmente a partir de los 90 años (Alter 1990). Hemos de indicar que dada la reducida proporción que representa el contingente del saldo migratorio en cada acumulado, obviamos su influencia a la hora de explicar las pautas. No obstante, influyen en el mismo sentido que la población cuando el saldo es positivo, y en el mismo sentido que las defunciones cuando el saldo es negativo.

TABLA 4

Indicadores de consistencia de los acumulados decrecientes por edad y periodo

| <i>Población inicial de las cohortes de edades</i> | <i>Población final de las cohortes de edades</i> | <i>Indicador de consistencia 1900-10</i> | <i>Indicador de consistencia 1900-20</i> | <i>Indicador de consistencia 1900-30</i> |
|--|--|--|--|--|
| 40+ | 50+ | 0,9986 | 0,9757 | 0,9869 |
| 45+ | 55+ | 1,0029 | 0,9871 | 0,9994 |
| 50+ | 60+ | 1,0560 | 1,0179 | 1,0250 |
| 55+ | 65+ | 1,1574 | 1,1006 | 1,0739 |
| 60+ | 70+ | 1,3988 | 1,2936 | 1,2216 |
| 65+ | 75+ | / | / | / |
| 70+ | 80+ | / | / | / |
| 75+ | 85+ | / | / | / |
| 80+ | 90+ | / | / | / |
| 85+ | 95+ | / | / | / |
| 90+ | 100+ | / | / | / |

Podemos confirmar la existencia de un factor perturbador principal a partir de los 60 años, combinación a su vez de varios factores: la menor declaración errónea de edades a la alta en la población respecto a la declaración errónea de edades a la alta en las defunciones. Sin embargo, los acumulados anteriores a los 60+ años acusan un déficit de defunciones lo que implicaría una o ambas de las siguientes pautas: una mayor declaración errónea de edades de individuos a la baja proveniente de edades superiores a los 60 años respecto a la declaración errónea de edades de defunciones en el

mismo sentido, o una menor declaración errónea de edades a la alta en la población proveniente de edades inferiores a los 50 años, respecto a la misma declaración errónea de edades en las defunciones. La experiencia de otros trabajos centrados en otras regiones y periodos considera poco probable esta última pauta mientras confirma la existencia de la primera. Más adelante daremos cuenta de la existencia de la misma a principios de siglo en España. Si bien, podemos considerar los 60+ años como el punto de inflexión donde se pasa de un predominio de declaración errónea de edades a la baja y mayor en los datos de población, a un predominio de declaración errónea de edades a la alta y mayor en los datos de defunciones; el porcentaje de desviación de la unidad como consistencia perfecta es menor a partir de los 50+ años que a partir de los 60+ años en los periodos 1900-10 y 1920-30, y que a partir de los 65+ años en 1910-20.

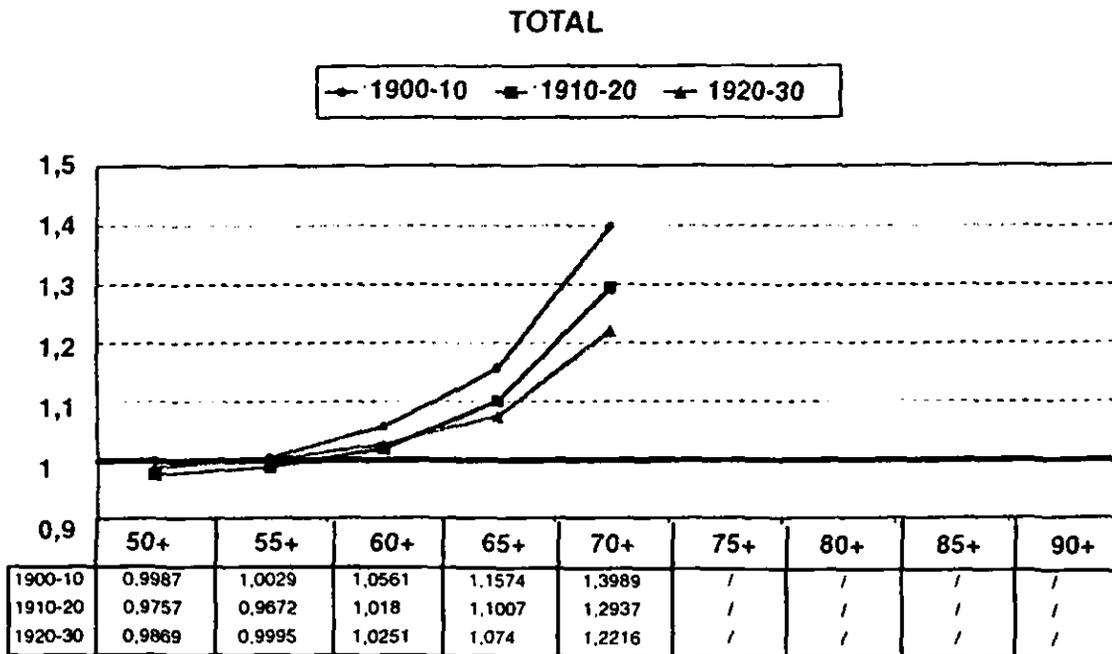
Las pautas representadas en el gráfico 2 son claras a la hora de confirmar lo que ya apreciábamos en la tabla 4. En los tres periodos se da una relación positiva entre el exceso de defunciones comprendidas en el acumulado y la edad. Es decir, cuanto más elevado es el límite inferior del acumulado, mayor es el número de defunciones pertenecientes a edades inferiores con relación al tamaño de la población inicial del período intercensal, especialmente a partir de los 60+ años. Asimismo, podemos corroborar como a partir de los 50+ años, con errores inferiores al 2.5 por ciento, encontramos que la practica totalidad de las defunciones correspondientes a las cohortes que tenían 40+ años en el primer censo, se encuentran registradas, y por tanto las inconsistencias encontradas para las cohortes superiores a 60 años son principalmente causadas por pautas de declaración errónea de la edad entre las cohortes comprendidas sobre ese límite y no tanto por un posible subregistro de población o defunciones.

Es también evidente en el gráfico 2 como a lo largo del periodo se reducen los errores en la declaración incorrecta de la edad acercándose a la línea que representa la consistencia perfecta, donde el decenio 1920-30 tiene las desviaciones mínimas y es la más cercana a la unidad. La mejora en los procedimientos de recogida de información por los organismos competentes a lo largo de los 30 pri-

meros años del siglo XX, junto a una mejora en los niveles de alfabetización de los individuos de estas edades justificarían esta progresión hacia una mejor consistencia entre las fuentes derivada de la reducción de la cantidad de edades mal declaradas.

GRÁFICO 2

Indicador de consistencia entre fuentes por acumulados



3.3. Segundo Acercamiento: Corrección de las distribuciones por edad

Una vez observados los resultados poco satisfactorios de los datos brutos desagregados linealmente, pasamos a una segunda fase con el objeto de aproximar las estructuras por edad de las fuentes originales a las estructuras por edad reales, en ausencia de edades mal declaradas. Para ello se ha desestimado: el uso de matrices correctoras de las pautas basadas en otras regiones o épocas, ya que no resultan corregir los datos españoles durante este periodo; así como simulaciones matemáticas de pautas de declaración errónea de la edad, puesto que no disponemos de pruebas suficientes para reconocer regularidades en dichas pautas y generar los pará-

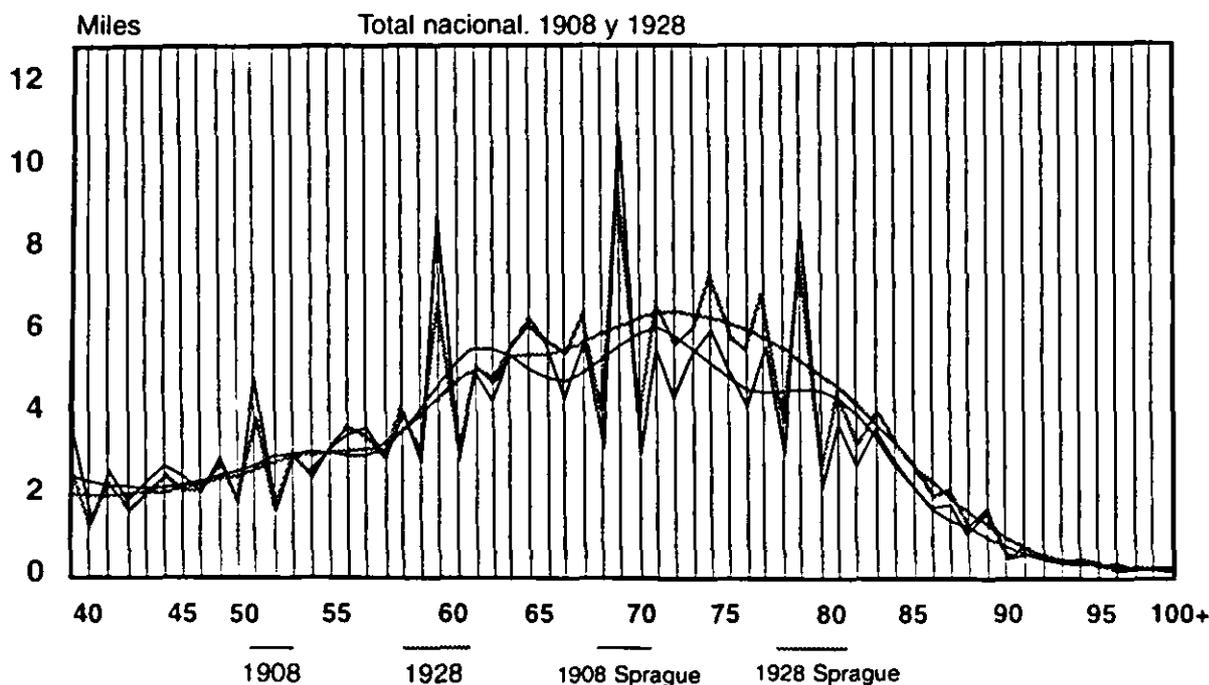
metros de la simulación. Sin embargo, en esta ocasión se han desagregado los datos empleando los coeficientes de Sprague y se ha desarrollado una extensión al método de Cohortes Intercensales para corregir los problemas derivados de las pautas de declaración errónea de la edad en las estructuras por edad. Posteriormente se han calculado factores de corrección por edad de las estructuras de las fuentes originales para permitir el cálculo de indicadores de mortalidad más fiables.

3.4. Efecto de los coeficientes de Sprague en las estructuras por edad

En el primer acercamiento a los datos se emplearon métodos de desagregación lineal de los grupos de edad de las fuentes originales, no obstante, en este segundo acercamiento se ha considerado más apropiada la desagregación de los mismos mediante el empleo de los coeficientes de Sprague. La desagregación lineal, basada en un solo polinomio, distribuye de forma equitativa el número total de acontecimientos del grupo de edad entre número de edades simples que forman el grupo. Sin embargo, los coeficientes de Sprague (1881), como método de interpolación osculatoria, están basados en una fórmula que superpone dos polinomios en una sola ecuación. De esta forma, se evitan rupturas repentinas entre los grupos de edad, suavizando la unión entre los grupos. Este último método es apropiado puesto que como puede observarse en el gráfico 3, la principal tendencia de declaración errónea de edades es hacia las terminadas en 0 y en 5, precisamente las edades que separan unos grupos de edad de otros. En nuestro caso los coeficientes de Sprague son de gran utilidad, puesto que corrigen en gran medida esa pauta de error. Así, en vez de distribuirse linealmente entre todas las edades que forman parte de su mismo grupo de edad de forma unidireccional y desproporcionada, parte de los contingentes declarados en edades terminadas en 0 y en 5 se distribuyen entre las edades colindantes del grupo de edades al que pertenecen y del grupo de edades con el que hacen contacto, a donde muchos realmente pertenecen. El gráfico 3 muestra las estructuras por edades simples, antes y después de aplicar los coeficientes de Sprague. Aunque tan solo muestra dos años modales, se han aplicado a todas las distri-

buciones disponibles (excepto migraciones) en todos los años del periodo. Tanto es así, que incluso las distribuciones de las que disponíamos en edades simples en la fuente original han sido agregadas y desagregadas posteriormente con los coeficientes de Sprague para suavizar la acumulación en los años terminados en 0 y 5.

GRÁFICO 3
Distribución de defunciones por edades simples



3.5. Desarrollo del método de corrección por equilibrio entre eventos

Una vez observada la ausencia de progresiones regulares en las pautas de declaración errónea de la edad comunes a los tres periodos según avanza la edad, y tras la desagregación de los datos mediante los coeficientes de Sprague, se ha recurrido al desarrollo de un método de corrección que tenga en cuenta la particularidad de estas pautas en cada periodo, en cada grupo de edad, y en cada tipo de eventos: población, cohortes de defunciones y migración neta de las cohortes. Si bien se ha denominado método de corrección ha de interpretarse que el método es más una aproximación a las estructuras reales, controlando las alteraciones en los contin-

gentes de eventos con las distribuciones originales. Con esta intención, se han considerado como eventos a todas las unidades que componen los contingentes de población, las cohortes de defunciones y los saldos migratorios de las cohortes, que forman parte de un sistema cerrado basado en dos supuestos razonables. En primer lugar se supone la distribución de los eventos de población de 1930 como correcta. Y en segundo lugar se considera que el conjunto de la cohorte de 40+ años en el primer censo tiene registradas la práctica totalidad de las defunciones y la migración neta de la cohorte durante el periodo intercensal, si bien no todas figuran en el grupo de edad correspondiente. El primero de los supuestos está basado en el resultado claramente más favorable que mantenía la consistencia entre las fuentes en 1920-30, donde tanto los datos de varones y mujeres como los totales por grupos y por acumulados, mostraron la menor irregularidad y la mayor cercanía a la consistencia perfecta. El segundo de los supuestos tiene su fundamento en el análisis por Cohortes Intercensales de los acumulados decrecientes, en que la cohorte de 40+ años era la que menor diferencia tenía respecto de la unidad que representa la consistencia perfecta. Con estos dos supuestos acotamos las dos líneas temporales del sistema que establece el método de Cohortes Intercensales. Por una parte el tiempo/duración en 1930, y por otra parte el tiempo/edad en los 40 años.

Pasamos ahora a hacer una concisa descripción de los pasos seguidos en el método de equilibrio entre eventos tomando como ejemplo su aplicación sobre la población de mujeres en el periodo intercensal 1920-30. Después de desagregar los datos de población y defunciones mediante los coeficientes de Sprague, aplicamos a los mismos el método de cohortes intercensales descrito anteriormente. Suponiendo que la distribución de la población en 1930 es correcta, consideramos como eventos de igual categoría a todos los contingentes de las distribuciones de población en 1920, de las cohortes de defunciones, y de los saldos migratorios de las cohortes.

Una vez se conoce el número total de eventos en el sistema, es importante saber el exceso o la falta de eventos en cada grupo de

edad para poder saber si para el total del sistema de 40+ años existe una desviación de falta o exceso de eventos. (Estas desviaciones totales de eventos no han excedido en ningún caso una proporción del 0.02 del número total de eventos). Dadas unas desviaciones de eventos tan reducidas, se ha procedido a incluirlas o excluirlas del sistema, según fuera pertinente, considerando que el riesgo de contribuir a la desviación total es proporcional al número de eventos de los contingentes por edad y tipo de suceso en las distribuciones iniciales. La inclusión o exclusión de eventos se realiza teniendo en cuenta tres consideraciones: si hay una falta o exceso de eventos en el total del sistema, si hay falta o exceso de eventos en cada grupo de edad, y dependiendo del tipo de suceso. Una última consideración hace referencia a que los eventos de población siguen la dirección opuesta a los eventos de defunciones, y que los eventos de los saldos migratorios siguen la misma dirección que la población cuando son positivos y la misma dirección que las defunciones cuando son negativos.

Tras realizar estos procedimientos, se alcanza un equilibrio entre los excesos y las faltas de eventos por grupos de edad en el sistema. Por lo tanto, si añadimos eventos a los grupos de edad donde faltan y descontamos eventos de los grupos de edad donde sobran, justo en la que medida en que es necesario como para que en cada grupo de edad el método de cohortes intercensales dé como resultado poblaciones estimadas idénticas a las observadas para el segundo censo, lo que estamos haciendo de forma subyacente es transferir eventos erróneamente ubicados en unos grupos de edad a los grupos de edad a los que realmente pertenecen. La cuestión en este punto es como distribuir la desviación por grupo de edad entre los tres de tipos de eventos que intervienen en cada cohorte de edad. Hemos considerado que una distribución proporcional al peso de cada tipo de evento en cada grupo de edad, es el menos arbitrario de los. Así, el riesgo de declarar erróneamente la edad en cada tipo de evento es proporcional al tamaño de su contingente de eventos dentro de cada grupo de edad.

Finalmente, si aplicamos de nuevo el método de cohortes intercensales a las distribuciones resultantes de población inicial, defun-

ciones intercensales y saldos migratorios, observamos que cada grupo de edad tiene un indicador de consistencia igual a 1 que representa la consistencia perfecta. Este procedimiento se ha aplicado para el total de los 30 años, por décadas mediante la proyección retrospectiva del método de Cohortes Intercensales. Es decir, una vez se corrigen las distribuciones de población de 1920, a partir de las de 1930, se corrigen las de 1910 con las de 1920 y las de 1900 con las de 1910.

3.6. *Coefficientes de corrección de distribuciones de población y defunciones por edad durante el periodo*

Una vez corregidas las distribuciones por edad de los censos de población y las defunciones podemos generar unos coeficientes correctores de las fuentes originales por edad y periodo intercensal. Con este fin, hemos calculado el cambio porcentual del tamaño del contingente de cada grupo de edad de las distribuciones originales. La tabla 5 muestra los porcentajes de cambio en cada grupo y fecha censal de las distribuciones de población.

A la hora de interpretar estos porcentajes hay que tener en cuenta que son resultados netos. Es decir, donde son negativos

TABLA 5
Cambio porcentual de las distribuciones de población por edad y periodo

| | 1900 | 1910 | 1920 | 1930 |
|-------|--------|--------|--------|------|
| 40-44 | 1,27 | 2,18 | -2,84 | 0 |
| 45-49 | -14,4 | -12,21 | -4,65 | 0 |
| 50-54 | -9,59 | -11,27 | 4,59 | 0 |
| 55-59 | 2,57 | -3,09 | -9,89 | 0 |
| 60-64 | 4,7 | 0,57 | -8,88 | 0 |
| 65-69 | 2,91 | 1,87 | 1,97 | 0 |
| 70-74 | 15,31 | 14,31 | 12,73 | 0 |
| 75-79 | 20,82 | 19,02 | 11,91 | 0 |
| 80-84 | 30,78 | 34,3 | 28,56 | 0 |
| 85-89 | 23,85 | 21,2 | 31,46 | 0 |
| 90-94 | -34,78 | -44,14 | -12,97 | 0 |
| 95-99 | -36,63 | -3,73 | -58,85 | 0 |
| 100+ | -46,11 | -32,07 | -62,82 | 0 |

implica que la salida de eventos declarados en esa edad ha sido mayor que las entradas. De igual forma, los porcentajes positivos implican que el balance entre entradas y salidas es favorable a las entradas. Por esto, a partir de cada grupo de edad, no es posible determinar cual es el volumen de eventos que ha sido transferido a cada uno de los otros grupos sino el balance entre el total de salidas a otros grupos y el total de entradas de otros grupos. Los cambios en la población 1930 aparecen como nulos puesto que ésta se supuso como correcta y se tomo como punto de partida para la corrección. En el resto de años censales puede observarse una pauta común en la estructura de los cambios netos entre grupos de edad, si bien no hay una regularidad evidente en la intensidad con que se dan los mismos. Esto último es más evidente a partir de los 90 años, donde los tamaños muestrales originales son tan reducidos que cualquier cambio en los mismos representa una alta proporción del total de eventos. En términos generales, la pauta a la que hacíamos referencia consiste en un intervalo de edades comprendido entre los 60 y 89 años, donde entran más individuos de los que salen, y dos tramos de edad, de 40 a 59 años y de 90 años en adelante, donde el balance de transferencias es favorable a las salidas de individuos. Esto implicaría, que en los datos originales de 40 a 59 años, habría una proporción de individuos que declaran ser mas jóvenes de lo que realmente son. Y que en el tramo de 90 años en adelante, habría una proporción de individuos que declaran tener más edad de la que realmente tienen.

Por lo tanto habría al menos dos pautas de declaración errónea de la edad en los censos. Una tendencia a quitarse años una vez superados los 60, y una tendencia a añadirselos antes de llegar a los 90. Estas pautas, además de poder estar causadas por la falta de cultura numérica de los individuos, y por una probable deficiencia de calidad en la recogida de la información, pueden tener origen en la valoración positiva que se tuviera de ser más joven y de ser longevo. Valoraciones cuyo punto de inflexión se encontraría entre los 60 y 89 años.

En la tabla 6 podemos observar los mismos porcentajes de cambio. Esta vez para las defunciones. En éstas, aunque también sin una regularidad en la intensidad de las pautas, predominan dos pautas generales de declaración errónea de la edad. Por una parte habría una proporción de difuntos de 90 y más años cuya edad declarada al morir estaría por debajo de los mismos, y por otra, habría una proporción de individuos por debajo de los 60 años cuya edad al morir es declarada por encima. En los grupos comprendidos entre 60 y 89 años la tendencia es favorable a las salidas de eventos aunque es muy probable que muchas de las transferencias se realicen entre esos mismos grupos.

TABLA 6

Cambio porcentual de las distribuciones de defunciones por edad y periodo

| | 1900-10 | 1910-20 | 1920-30 |
|-------|---------|---------|---------|
| 40-44 | -1,07 | -1,69 | 2,62 |
| 45-49 | 14,19 | 12,32 | 4,87 |
| 50-54 | 9,36 | 11,26 | -4,2 |
| 55-59 | -2,53 | 3,12 | 9,76 |
| 60-64 | -4,65 | -0,61 | 8,77 |
| 65-69 | -2,86 | -1,86 | -1,99 |
| 70-74 | -15,21 | -14,49 | -12,82 |
| 75-79 | -20,86 | -19,44 | -12,08 |
| 80-84 | -31,01 | -35,12 | -29,08 |
| 85-89 | -23,93 | -21,74 | -31,68 |
| 90 | 35,34 | 39,13 | 31,44 |

3.7. Tasas específicas de mortalidad

En la tabla 7 pueden observarse las tasas específicas de mortalidad corregidas con los coeficientes de corrección de distribuciones, y las mismas tasas sobre los datos originales de población y defunciones. Si comparamos las tasas corregidas respecto a las originales, en términos generales son algo más elevadas de 40 a 54 años, mantienen niveles semejantes de 55 a 69, y bastante más bajas de 70 a 80+ años. No obstante, las tasas corregidas muestran

un descenso progresivo en los niveles generales de mortalidad a lo largo del periodo, especialmente entre 1920 y 1930. Es también en este periodo cuando la mortalidad de 80+ años desciende a niveles inferiores al 200 por mil. En cuanto a los niveles de mortalidad por sexo, los varones mantienen en todas las fechas y edades niveles superiores a las mujeres. Asimismo, el descenso de las mujeres es más pronunciado que el de los varones a lo largo del periodo. Por todo, podemos concluir que en las primeras décadas del siglo XX el receso de las enfermedades infecciosas también produjo un sustancial descenso de la mortalidad de adultos y ancianos, especialmente de mayores de 70 años.

TABLA 7

Tasas específicas de mortalidad. Corregidas y originales (por mil)

| <i>Corregido</i> | | | | |
|------------------|--------|--------|--------|--------|
| | 1900 | 1910 | 1920 | 1930 |
| 40-44 | 13,41 | 11,3 | 12,05 | 9,06 |
| 45-49 | 17,52 | 14,16 | 12,27 | 9,72 |
| 50-54 | 21,76 | 19,93 | 16,09 | 13,9 |
| 55-59 | 26,43 | 22,94 | 23,93 | 19,16 |
| 60-64 | 42,34 | 41,89 | 42,23 | 32,74 |
| 65-69 | 57,82 | 48,83 | 48,82 | 44,43 |
| 70-74 | 86,85 | 80,08 | 82,06 | 73,32 |
| 75-79 | 117,57 | 103,38 | 108,34 | 103,02 |
| 80+ | 218,93 | 217,52 | 207,59 | 194,87 |

| <i>Original</i> | | | | |
|-----------------|--------|--------|--------|--------|
| | 1900 | 1910 | 1920 | 1930 |
| 40-44 | 12,8 | 11,01 | 11,06 | 8,47 |
| 45-49 | 13,3 | 10,81 | 11,35 | 9,4 |
| 50-54 | 19,61 | 16,63 | 17,19 | 14,3 |
| 55-59 | 25,92 | 19,71 | 19,9 | 17,59 |
| 60-64 | 48,48 | 45,26 | 39,83 | 33,35 |
| 65-69 | 57,46 | 49,77 | 45,73 | 41,59 |
| 70-74 | 127,78 | 118,77 | 110,25 | 88,7 |
| 75-79 | 151,02 | 124,56 | 120,61 | 102,93 |
| 80+ | 290,2 | 303,09 | 296,87 | 227,02 |

Bibliografía

- ALTER, GEORGE (1990). «Old age mortality and age misreporting in the United States, 1900-1940». *Working Paper*, No. 24.. *Indiana University, Population Institute for Research and Training* : Bloomington, Indiana.
- 1995 «Trends in United States old age mortality, 1900-1935: evidence from railroad pensions.» pp. 304 -328, en KERTZER, DAVID I.; LASLETT, PETER (1995). «*Aging in the Past : Demography, Society, and Old Age.*» *Studies in Demography*, 7. Univ. California Press.
- BENNETT, N.G. y HORIUCHI (1984). «Mortality estimation from registered deaths in less developed countries», *Demography* 21, pp. 217-34.
- CACHINERO SÁNCHEZ, BENITO (1985). «Estimating levels of adult mortality in eighteenth-century Spain», *Historical Methods*, Vol. 18, 2, pp. 63-70.
- CASELLI, GRAZIELA (1991). «Long-term trends in European Mortality». *Studies on Medical and Population Subjects*, 56. London.
- COALE, ANSLEY J. y CASELLI, GRAZIELA (1990). «Estimation of the number of persons at advanced ages from the number of deaths at each age in the given year and adjacent years», *Genus*, Vol. 66, N° 1/2 pp. 1-23.
- CONDRAN, GRETCHEN A.; HIMES, CHRISTINE L.; PRESTON, SAMUEL H. (1991). «Old-age mortality patterns in low-mortality countries: an evaluation of population and death data at advanced ages, 1950 to the present.» *Population Bulletin of the United Nations*, No. 30, pp. 23-60. New York.
- DECHTER, A. y PRESTON, S. (1991). «Age misreporting and its effects on adult mortality in Latin America». *Population Bulletin of the United Nations*, No. 31/52, pp.1-17. New York.
- DOPICO, FAUSTO, y REHER, DAVID (1999). *El declive de la mortalidad en España, 1860-1930*, Monografía nº 1 de la Asociación de Demografía Histórica, Zaragoza.
- EWBANK, DOUGLAS C. (1981). «Age misreporting and age-selective underenumeration: Sources, patterns, and consequences for demographic analysis». *Committee on population and Demography*, Report No. 4. National Academy Press. Washington D.C.
- GÓMEZ REDONDO, ROSA (1992). *La mortalidad infantil española en el siglo XX*. Centro de Investigaciones Sociológicas. Monografía nº 123. Siglo XXI. Madrid.
- KANNISTO, VÄINÖ. (1990). *Mortality of the elderly in late 19th and early 20th century Finland*, Central Statistical Office of Finland. Helsinki.
- MARI BHAT, P. N. (1995). «Age misreporting and its impact on adult mortality estimates in South Asia» , *Demography India*, Vol. 24, 1, pp. 59-80.
- KERTZER, DAVID I.; LASLETT, PETER. (1995). «Aging in the Past: Demography, Society, and Old Age». *Studies in Demography*, 7. Univ. California Press.

- PRESTON, SAMUEL; COALE, ANSLEY (1982). «Age structure, growth, attrition and accesion: a new sinthesys», *Population Index* 48, pp. 217-59.
- PRESTON, S., ELO, I., ROSENWAIKE, I. y HILL, M. (1996). «African-American mortality at older ages: Results of a matching study», *Demography*, Vol. 33, N° 2, pp. 193-209.
- PRESTON, S., ELO, I. Y STEWART, Q. (1997). «Effects of age misreporting on mortality estimates at older ages», *Population Aging Research Center, University of Pennsylvania Working Paper Series* 98-01.
- REHER, D.; VALERO LOBO, A. (1995). *Fuentes de información demográfica en España*. Cuadernos Metodológicos, 13. CIS, Madrid.
- SÁNCHEZ ALONSO, BLANCA (1990). «Una nueva serie anual de la emigración española: 1882-1930». *Revista de Historia Económica*, 8, 1, pp. 133-170.
- SPRAGUE, T.B. (1881). «The explanation of a new formula for interpolation». *Journal of the Institute of Actuaries*, 22.
- VINCENT, PAUL (1951). «La mortalité des vieilliards», *Population*, Vol. 6, 2. pp. 181-204. París.