



fundación  
MUTUALIDAD  
ABOGACÍA

escuela  
de pensamiento

# DESARROLLO DE UN ALGORITMO PREDICTIVO DE EDAD BIOLÓGICA BASADO EN HÁBITOS DE VIDA Y BIOMARCADORES GENÉTICOS Y NO GENÉTICOS

Coordinado por:  
Antonio López Farré y María Durbán

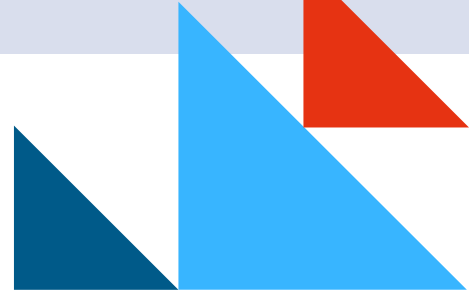
Con la colaboración de:

uc3m | Universidad  
Carlos III  
de Madrid



Instituto  
Gerontológico

<b>1. INTRODUCCIÓN</b>	<b>3</b>
<b>2. LA EDAD BIOLÓGICA COMO INDICADOR</b>	<b>4</b>
2.1 Aproximaciones al cálculo de la edad biológica	5
<b>3. POTENCIALES CONDICIONANTES DE LA EDAD BIOLÓGICA</b>	<b>6</b>
<b>4. DESCRIPCIÓN DE LA BASE DE DATOS</b>	<b>12</b>
<b>5. DESARROLLO DEL ALGORITMO PREDICTIVO</b>	<b>20</b>
5.1 Fase 1: Imputación de la edad biológica	20
5.2 Fase 2: Cálculo de scores	29
<b>6. CONCLUSIONES E INVESTIGACIÓN FUTURA</b>	<b>32</b>



## INTRODUCCIÓN

Este trabajo ha sido posible gracias al apoyo y financiación recibida de **Fundación Mutuality de la Abogacía**, a través de su Escuela de Pensamiento. Ha sido un proyecto colaborativo entre la Facultad de Medicina de la Universidad Complutense de Madrid (UCM) y el Departamento de Estadística de la Universidad Carlos III de Madrid (UC3M), coordinado por los catedráticos Antonio López Farré (UCM) y María Durbán (UC3M). El equipo de investigación ha estado formado por los siguientes investigadores:

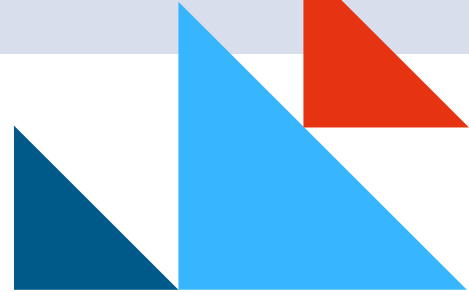
### Equipo UCM

Antonio López Farré  
Jose Javier Zamorano León  
Khaoula Zekri Nechar

### Equipo UC3M

Stefano Cabras  
Ignacio Cascos  
Bernardo D'Auria  
María Durbán  
Vanesa Guerrero

Este proyecto surge como una respuesta a la necesidad de encontrar nuevas medidas del envejecimiento que vayan más allá de la edad cronológica. Según el INE, el índice de envejecimiento de la población española, (relación entre población mayor de 65 años y menor de 15), ha pasado en 20 años del 0,908 en 1997 al 1.183 en 2017. Este envejecimiento será lógicamente más evidente en regiones especialmente afectadas por una menor tasa de natalidad. Sin embargo, la esperanza de vida al nacimiento se está incrementando por lo que la edad de fallecimiento se está retrasando, siendo en España la esperanza de vida en el momento actual de unos 83 años. La esperanza de vida viene creciendo de forma exponencial en los países desarrollados desde principios del siglo XIX, hasta el punto de que la edad cronológica no es necesariamente un buen indicador de nuestra edad real. De ahí surge la necesidad de buscar indicadores alternativos, por ejemplo, ya que sabemos que hay factores, como la nutrición, el estilo de vida y el historial médico, que influyen sobre la mortalidad y, por



tanto, en la esperanza de vida de un individuo. Es en este marco en el que se encuadra este proyecto cuyo objetivo fundamental es obtener una estimación fiable de la edad biológica de un individuo mediante un algoritmo estadístico que permita una implementación sencilla, y que además se construya a partir de una serie de variables que pueden recogerse sin excesivo coste. El algoritmo permitirá, a través de sus scores, determinar qué variables influyen y si lo hacen positiva o negativamente en la edad biológica.

## **2. LA EDAD BIOLÓGICA COMO INDICADOR REAL DEL ENVEJECIMIENTO**

El envejecimiento es un proceso continuo y complejo que se caracteriza por una disminución progresiva de las capacidades físicas, mentales y reproductivas, que conduce a la pérdida de función, a una mayor susceptibilidad a diferentes enfermedades y, en última instancia, el final de la vida.

Sin embargo, el proceso de envejecimiento puede diferir sustancialmente entre diferentes individuos, y podría ser modificado en función de la adherencia a hábitos saludables. En este sentido, algunas personas parecen frágiles y requieren asistencia médica a sus 70 años, mientras que otras permanecen independientes de la asistencia y parecen escapar de un deterioro fisiológico importante hasta edades incluso centenarias. Por eso, aunque la edad cronológica es un indicador comúnmente utilizado para el envejecimiento, sin embargo, la esperanza de vida muestra una variación considerable entre las personas con edades cronológicas iguales o similares. Esta variabilidad en la esperanza de vida se debe a factores genéticos, de hábitos de vida, factores socioeconómicos y culturales que modifican la edad biológica o real del organismo de cada individuo.

La edad biológica se refiere a la situación actual del sujeto en relación con su ciclo vital potencial. Podría definirse como el desgaste real de las energías producto del paso de los años. Tiene en cuenta los cambios físicos y biológicos que se van produciendo en las estructuras celulares, de tejidos, órganos y sistemas. En otras palabras, la edad biológica es la que se corresponde con



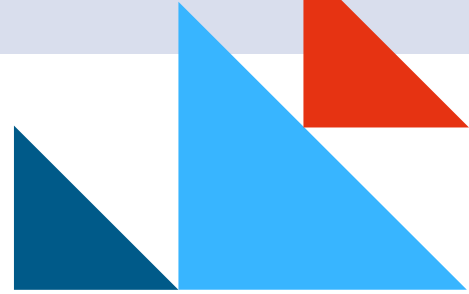
el estado funcional de nuestros órganos comparados con patrones estándar para una edad. Es un concepto fisiológico del estado del envejecimiento de nuestro organismo. La literatura científica nos dice que la diferencia observable entre la edad cronológica y la biológica puede variar en hasta más menos 12 años. Por ello, podríamos encontrarnos con una situación óptima en la que un individuo de 70 años cronológicos tenga una edad biológica de 58 años, o por el contrario, que ese mismo individuo de 70 años cronológicos tenga una edad biológica de 82 años.

## 2.1 Aproximaciones al cálculo de la edad biológica

Es solo cuestión de tiempo que, en medicina, la edad biológica (la edad real de nuestras células y órganos), sustituya a la edad cronológica y se vaya imponiendo como medidor de envejecimiento. En este sentido, se han intentado establecer algunas formas de determinar la edad biológica de un individuo. Por ejemplo, existen diversas calculadoras, muchas incluso accesibles a través de internet, basadas únicamente en cuestionarios de hábitos de vida. Sin embargo, este es un abordaje simplista que no considera el efecto del factor genético o la base fisiológica del individuo sobre la edad biológica. La edad biológica, una vez que se determine un algoritmo que de verdad se acerque a su cálculo real y sea de mayor exactitud, será sin duda el medidor del riesgo asociados a la naturaleza humana, como la salud, la dependencia y la predicción individualizada de la longevidad.

En el momento actual, no existe un indicador o combinación de estos, que pueda considerarse como “estándar oro” para la estimación del proceso de envejecimiento del organismo. No obstante, si se pueden seguir algunos criterios para considerar un biomarcador como válido para ser incluido en el cálculo de la edad biológica:

- **Que cambien con la edad;**
- **Que sean estables en su medida entre individuos;**
- **Que proporcione información sobre la condición funcional del organismo y de sus sistemas metabólicos y regulatorios.**



### 3. POTENCIALES CONDICIONANTES DE LA EDAD BIOLÓGICA

Por todo ello, la aproximación más adecuada para la estimación de la edad biológica es aquella en la que se valoren de forma conjunta las diferentes características sociodemográficas, genéticas, fisiológicas y hábitos de vida que pueden condicionar el estado de salud del organismo. Estos potenciales condicionantes de la edad biológica pueden clasificarse en factores modificables y no modificables.

#### FACTORES POTENCIALMENTE RELACIONADOS CON LA EDAD BIOLÓGICA, SUSCEPTIBLES DE SER INCLUIDOS EN EL MODELO DE CÁLCULO DE EDAD BIOLÓGICA

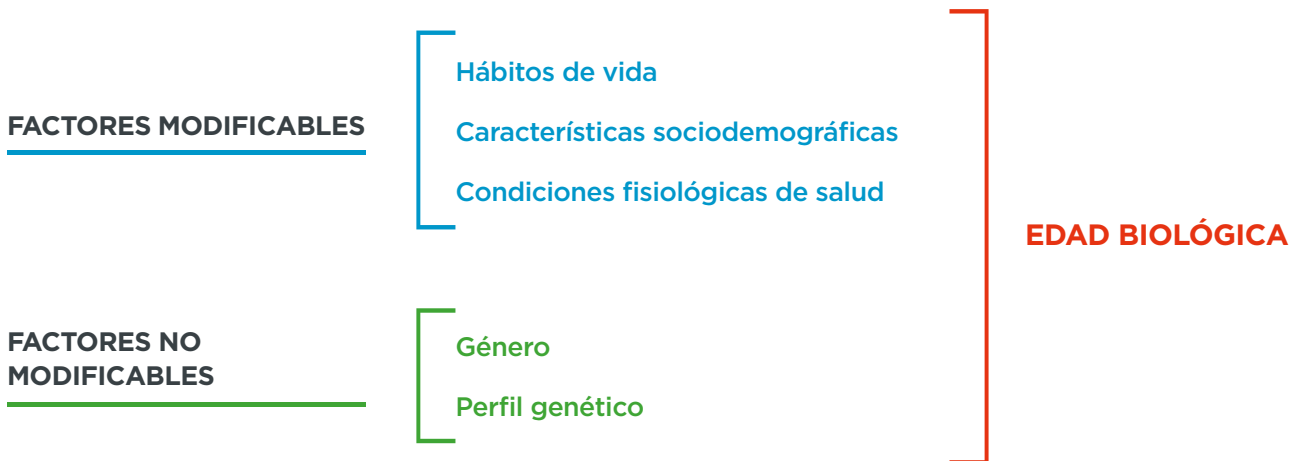
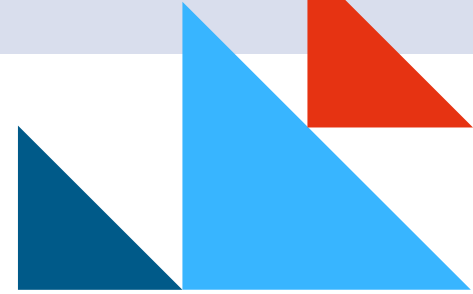


Figura 1: Factores que influyen en la edad biológica



## **a) Factores modificables**

Entre los factores conductuales que pueden ser modificados por el individuo destacan:

### ▶ **Marcadores de hábitos de vida**

Los beneficios sobre la salud de los hábitos de vida, fundamentalmente la alimentación, el ejercicio físico y el descanso, son bien conocidos en la sociedad moderna.

- La dieta Mediterránea en el caso de nuestro país es un conocido protector de la salud, fundamentalmente la cardiovascular, pero también de la conservación de la funcionalidad celular. Un trabajo reciente describió algunas adaptaciones importantes inducidas por la dieta mediterránea que podrían modificar la duración de la salud en las personas. Por ello, una alimentación saludable es un factor crucial para mantener una salud óptima y, por tanto, mantener una reducida edad biológica.
- La actividad física y el descanso adecuado parecen ser otros factores decisivos sobre el envejecimiento fisiológico. Sirva como ejemplo una encuesta que incluyó 92,000 personas en Inglaterra y que mostró que la realización de ejercicio físico es una actividad que generalmente se reduce con el paso de los años. Las personas sedentarias mayores de 50 años tienen dos veces un riesgo mayor de muerte comparado con aquellas personas que mantienen un nivel elevado de ejercicio físico. El mantenimiento de un estilo de vida físicamente activo hasta la edad media y avanzada se asocia con una mejor salud en la vejez y mayor longevidad. Por otro lado, un descanso adecuado se ha relacionado con beneficios metabólicos, disminuyendo los niveles de stress e incluso mejorando la regulación hormonal, las cifras de tensión arterial y la función cognitiva.

### ▶ **Marcadores sociodemográficos**

Hay otros factores que evidentemente van a influenciar sobre diferentes aspectos directamente relacionados con la salud del envejecimiento, en este sentido se han descrito que el nivel de ingresos y educativo, así como la clase social y el estado civil son predictores de riesgo de diferentes enfermedades crónicas (enfermedades Cardiovasculares u oncológicas) así como a adherencia de diferentes sistemas de prevención primaria (vacunación) y secundaria



(sistemas de cribados de enfermedades oncológicas). Entre estos factores tendremos que incluir factores socioeconómicos y sociodemográficos como son:

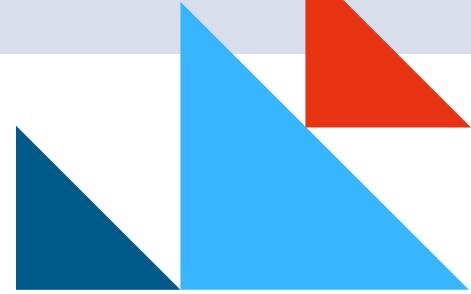
- Nivel de ingresos y de renta
- Nivel de educación
- Clase social
- Lugar de residencia
- Estado civil
- Religiosidad, entre otros

#### ► **Condiciones fisiológicas o de salud**

Aunque un factor importante en el cálculo de la edad biológica debe ser el estado de salud, la valoración del concepto salud en las personas que van teniendo más edad es muy compleja. Quizás la valoración del estado saludable de una persona, es uno de los fallos más recurrentes en las fórmulas y algoritmos establecidos para el cálculo de la edad biológica. En este sentido, la mayoría de los estudios que desarrollan fórmulas para el cálculo de la edad biológica, tienden a simplificar todos los factores incluidos en ellas, además de las enfermedades. En este sentido, muchas de las fórmulas desarrolladas utilizan para su diseño el concepto de sujetos saludable o casi sanos para excluir los impactos de las enfermedades.

Algunos investigadores incluso excluyeron a todas las personas no saludables y seleccionaron solamente a las sanas solamente. Por ejemplo, se excluyeron los voluntarios que habían sido diagnosticados con hipertensión, diabetes, trauma, insuficiencia renal, trastorno mental, apoplejía, cáncer, insuficiencia cardíaca, enfermedad coronaria, arteriopatía periférica, enfermedades pulmonares o antecedentes de enfermedad en los últimos 6 meses. Por eso, las fórmulas construidas sobre esta “base saludable” podría representar el proceso de envejecimiento normal o incluso funcionar como un modelo de predicción de enfermedad, pero en ningún caso puede ser considerada como fórmula de cálculo de la edad biológica en la población general de cualquier edad cronológica.





## **b) Factores modificables**

### ▶ **Género**

En la actualidad, es ampliamente aceptado que las mujeres son más longevas, viven más años que los hombres, argumentándose principalmente por la protección hormonal frente a diferentes enfermedades que presentan las mujeres durante su período fértil. Sin embargo, si se analizan los años de buena salud en función del género, parece que las diferencias no son tan grandes. En este sentido, recientes estudios parecen evidenciar que las diferencias de “esperanza de vida en buena salud” entre hombres y mujeres son pequeñas, mostrándose un comportamiento bimodal en el hombre: los hombres presentan probabilidad de envejecimiento saludable similares al género femenino, pero también mayor probabilidad de muerte debido al mantenimiento de hábitos de vida menos saludables mientras las mujeres presentan un estilo de vida menos arriesgado.

Estos resultados evidencian que las diferencias de género respecto al envejecimiento se deben explicar por la combinación de componentes fisiológicos, psicológicos y comportamentales o conductuales de manera conjunta.

### ▶ **Efectos del perfil genético sobre la edad biológica**

Desde que en 1953 los doctores Watson y Crick descubrieran la estructura y las características del ADN, se han generado una sucesión de hallazgos que han puesto de manifiesto la importancia de la genética en el campo de la salud. Actualmente, tenemos a nuestro alcance la información genética más relevante de cada uno de nosotros. Este valioso conocimiento no sólo se puede aplicar para el diagnóstico y/o tratamiento de enfermedades, sino también en población sana para modular nuestro estilo de vida con el objetivo de vivir más y mejor.

Actualmente, estamos inmersos en una nueva era, en la era de la “atención personalizada”, en el que conocer las características genéticas individuales juega un papel fundamental en la elección de hábitos saludables y promoción de la salud.

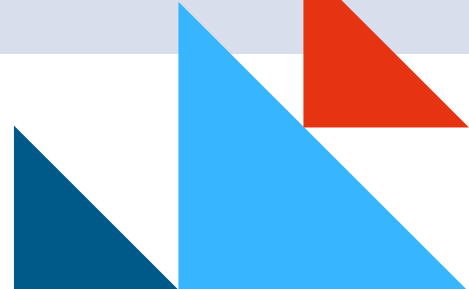


En los últimos años, el factor genético es uno de los que mayor interés ha despertado en la comunidad científica como factor determinante y condicionante del envejecimiento. En este sentido, es reconocido que los individuos más longevos suelen provenir de familias donde sus antecesores también han vivido mucho más que la media de la población correspondiente. Hoy ya conocemos algunos genes que están más activos en la edad juvenil pero que con la edad pierden actividad. Incluso genes que regulan hábitos de vida que se modifican con el envejecimiento, como aquellos genes implicados con los procesos de asimilación de alimentos, regulación del apetito, ciclo del sueño, etc. Tal es la importancia del aspecto genético en durante el envejecimiento que se han desarrollado conceptos como genética gerontológica o genética de la longevidad. Esta nueva disciplina médica analiza un conjunto de genes estrechamente relacionados con los principales procesos relacionados con el envejecimiento con el fin de predecir la capacidad funcional fisiológica basado en la información genética (perfil genético) de cada persona.

#### ▶ **¿Qué es un polimorfismo?, y ¿un SNP?**

De forma muy sencilla y abreviada, podríamos definir la información genética o genoma como una secuencia lineal y ordenada de la combinación de cuatro letras o nucleótidos (“A”, “T”, “G” y “C”). En este sentido, esta secuencia de letras es prácticamente idéntico entre todos los individuos, tan sólo menos del 1% del genoma varía en cada uno de nosotros. Estas variaciones en la secuencia lineal entre los individuos es lo que explica las diferencias entre todos nosotros, no sólo las diferencias físicas sino también las diferencias en el riesgo de sufrir enfermedades, las respuestas a los tratamientos farmacológicos, e incluso la predisposición a envejecer cronológicamente de forma más o menos saludable, es decir, envejecer cronológicamente con edades biológicas inferiores o superiores a la edad cronológica.

Los polimorfismos son alteraciones genéticas que se encuentran en una elevada proporción de la población. Los polimorfismos son muy frecuentes en la población y pueden tener diferentes efectos, algunos son inertes y no aportan beneficios ni perjuicios, otros inducen



beneficios a la salud mientras que otros polimorfismos podrían inducir efectos adversos a procesos relacionados con el envejecimiento y que, por tanto, predispondría a un individuo a presentar mayor edad biológica que la correspondiente a su edad cronológica.

El tipo de polimorfismo más frecuente en la población es el polimorfismo de un solo nucleótido (SNP, single nucleotide polymorphism) que consiste en la sustitución de un único nucleótido o letra por otra. Por ejemplo, se cambia una letra “G” por la letra “A”, cambiándose la secuencia en un único punto. Se ha calculado que los SNPs aparecen de media cada 100-300 nucleótidos a lo largo del genoma, por lo que un individuo presentará millones de SNPs. La presencia de diferentes polimorfismos o variantes genéticas en los distintos genes es lo que va a determinar las necesidades específicas de cada persona y, por tanto, que una persona engorde más o menos en función de la alimentación, que pueda vivir más, que un ejercicio físico le beneficie más que otro, etc.

Existen genes identificados y polimorfismos genéticos de un único nucleótido (en sus siglas en inglés SNP) que han comenzado a relacionarse con el envejecimiento y la forma de envejecer. En este sentido, entre otros polimorfismos, se han analizado diferentes polimorfismos relacionados con el efecto del ejercicio físico en la pérdida de grasa visceral, peso corporal y riesgo de enfermedades crónicas graves. También se analizaron polimorfismos en genes que condicionan la capacidad cognitiva y habilidades motoras, así como polimorfismos que condicionan la palatibilidad y, por tanto, el tipo de alimentos que prefiere consumir el portador del polimorfismo. Este abordaje genético es una de las principales fortalezas del abordaje conjunto que se realiza en el presente proyecto.



## 4. DESCRIPCIÓN DE LA BASE DE DATOS

La base de datos final constaba de 911 individuos de entre 16 y 87 años, el 60% eran mujeres y el 40% hombres, aproximadamente la mitad de ellos (45%) pertenecen a la Mutuality de la Abogacía y el resto se incluyeron en el estudio a través de consulta de medicina de familia. Todos los individuos contestaron cuestionarios de alimentación, hábitos de vida, características socio-económicas y culturales. Para las preguntas de factores ambientales y hábitos de vida se utilizarán los cuestionarios homologados de:

1. Adherencia a la Dieta Mediterránea (Modificado de: Trichopoulou A et al. N Engl J Med 2003; 348: 2.599-2.608)
2. Cuestionario Internacional de Actividad Física (IPAQ). Este cuestionario está propuesto por la OMS y aporta información sobre gasto energético estimado en 24 horas, en las distintas áreas de la vida diaria. El cuestionario IPAQ clasifica el grado de actividad física en bajo, moderado y vigoroso.
3. Cuestionario sobre factores de naturaleza socioeconómica y cultural en el que también se incluirán preguntas específicas sobre edad, sexo, capacidad de relación, características antropométricas (peso, talla) etc.

Las respuestas a los cuestionarios dieron lugar a más de 100 variables cualitativas y cuantitativas. Algunas de ellas fueron descartadas al tener un elevado número de datos faltantes, además de crear nuevas variables que permitían resumir la información obtenida de una manera más útil, por ejemplo, se creó la variable número de comorbilidades a partir de las respuestas de los participantes a 11 cuestiones en las que se preguntaban si padecían una determinada enfermedad. Además, se descartaron las variables lugar de residencia, al proceder el 95% de los particulares de la Comunidad de Madrid, y nivel de estudios, debido al posible sesgo que podríamos encontrar por el hecho de que todos los participantes que pertenecen a la Mutuality de la Abogacía tenían estudios universitarios.



Además, se tomó una muestra de la mucosa de la boca para la medida de los biomarcadores. Para ello, la toma se realizó frotando con un hisopo la cara interna de los carrillos derecho e izquierdo. El hisopo se rodó posteriormente sobre unas tarjetas comerciales preparadas para la recogida de este tipo de muestra. En esta muestra de mucosa bucal también se determinaron los biomarcadores de naturaleza genética y no genética. En particular, se genotiparon polimorfismos asociados con:

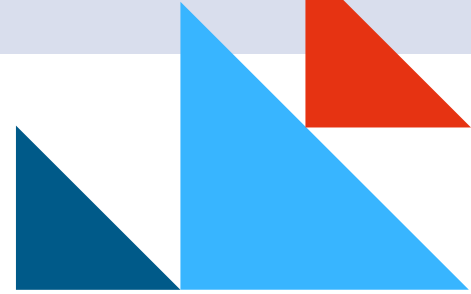
- **Biogenesis mitocondrial: PGC1alfa**
- **Metabolismo de hidratos de carbono: UCP2**
- **Capacidad de ejercicio aerobico o anaerobico: HIF1**
- **Funcionalidad cognitiva: Gen BDNF, codificante para el factor neurotrofico cerebral. Gen APOE, gen codificante para la apolipoproteina E.**

Además, se analizaron los siguientes biomarcadores no genéticos:

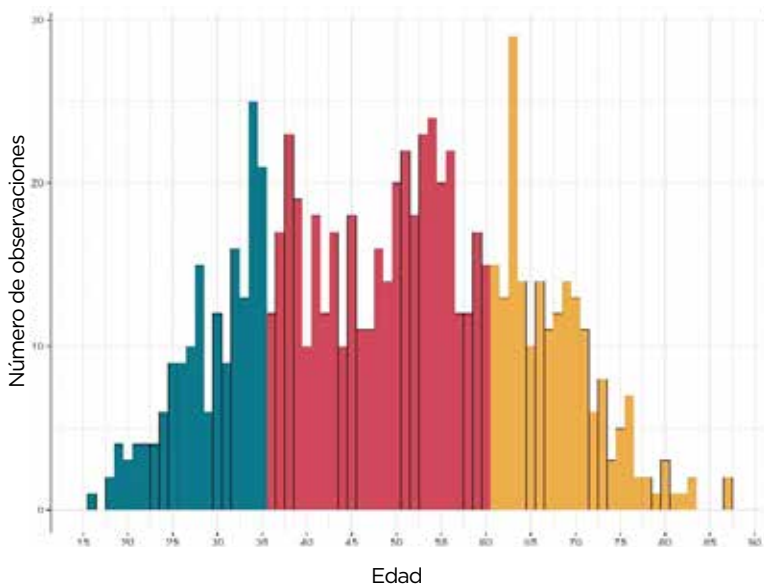
- **Biomarcadores relacionados con el numero y funcionalidad mitocondrial: Actividad citrato sintasa, actividad citocromo C oxidasa**
- **Biomarcador de estres: Cortisol**
- **Biomarcador de senescencia celular: Actividad beta-galactosidasa**

Los polimorfismos de los genes se determinaron mediante PCR cuantitativa utilizando sondas TaqMan y la medida enzimática se determinó mediante Kits comerciales. La determinación de expresión de proteína se se llevó a cabo mediante Western blot y Dot blot utilizando anticuerpos mononucleares o polinucleares específicos.

A continuación, pasamos a hacer un análisis descriptivo de algunas de las características de la muestra de individuos, en concreto se muestra una variable relacionada con el estado de



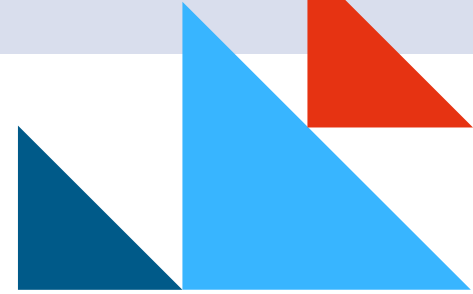
salud (número de comorbilidades), tres relacionadas con los hábitos de vida (IMC, horas de sueño y estrés) y una relacionada con la alimentación (ingesta de grasa animal semanal). Primero mostramos la distribución de las edades de los individuos, los cuales se han dividido en tres grupos de edad, menores de 36 años, entre 36 y 60 y mayores de 60 años. Se ha establecido esta estratificación de la atendiendo a los resultados presentados en la revista Nature Medicine (Nat Med 2019 Dec;25(12):1843-1850). En este estudio se muestra que el envejecimiento fisiológico no ocurre a un ritmo uniforme, sino que traza una trayectoria con tres puntos de inflexión en el ciclo de vida humano. El primero de estos puntos ocurre alrededor de los 34 años, el segundo tiene lugar alrededor de los 60, y el tercero, a los 78. En nuestro caso no hemos utilizado el último tramo de edad ya que el número de participantes en el estudio a partir de los 78 años es limitado.



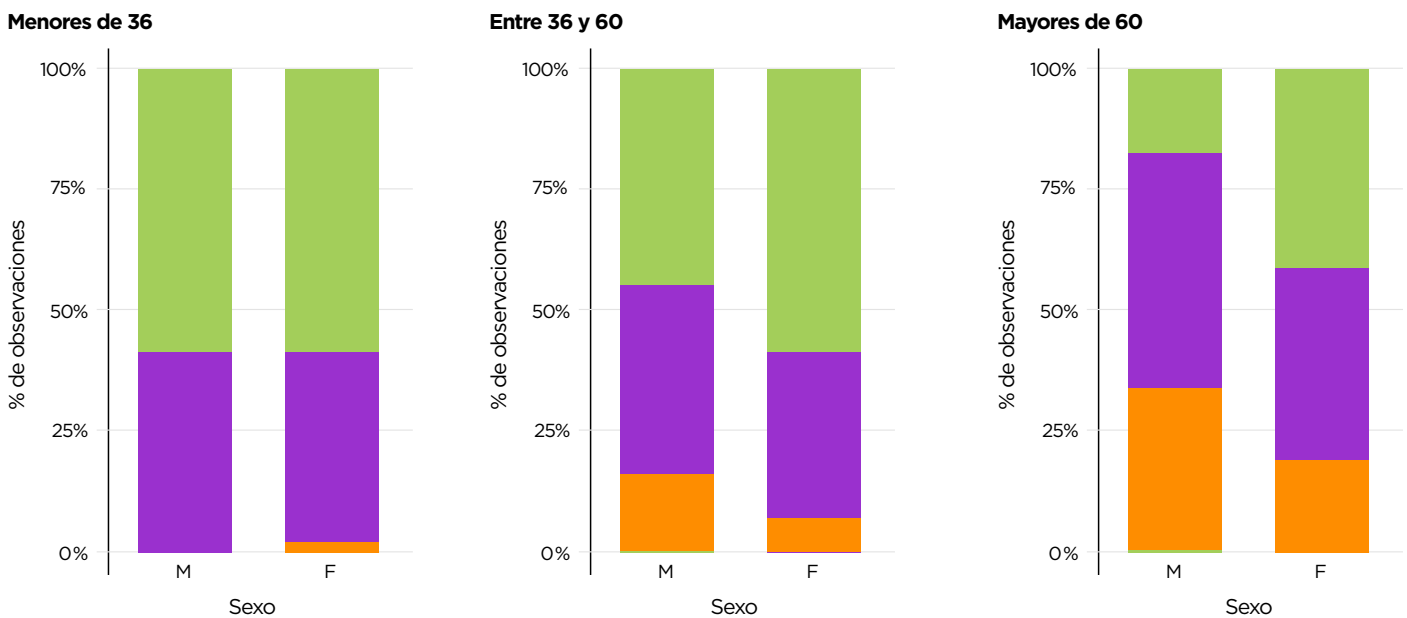
#### Franja de edad

- menor de 36 años
- entre 36 y 60 años
- mayor de 60 años

Figura 2: Distribución de participantes por franja de edad



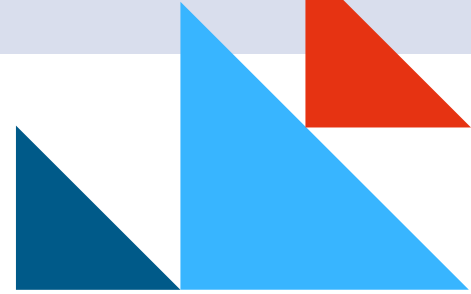
Podemos observar que el grupo mayoritario corresponde a la franja intermedia, y el que tienen menos participantes es el de menores de 16 años. La siguiente figura muestra la distribución del número de comorbilidades (0, 1, 2 o más) en cada franja de edad y según el sexo. Se observa cómo al aumentar la edad hay más personas con al menos una comorbilidad, además, en todas las franjas de edad los hombres tienen, en proporción, más comorbilidades que las mujeres.



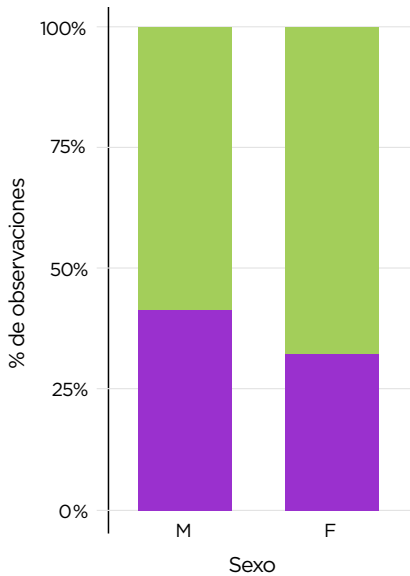
Número de comorbilidades: ■ 0 ■ 1 ■ 2 o más

Figura 3: Porcentaje de personas con 0, 1 y 2 o más comorbilidades por franja de edad y sexo

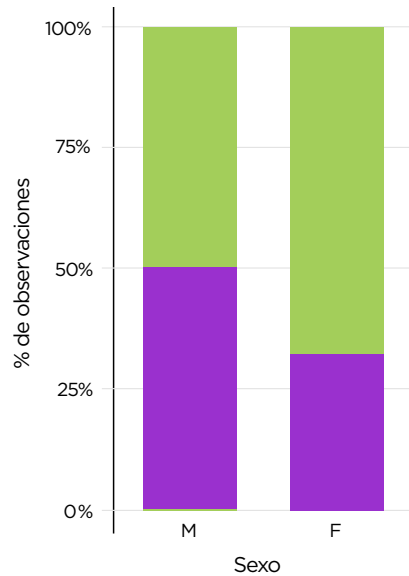
La Figura 4 muestra el porcentaje de individuos que tienen o no el Índice de Masa Corporal (IMC) normal (entre 18.5 y 14.9) o fuera de lo normal, en cada franja de edad y según el sexo. Se observa cómo a medida que aumenta la edad, tener un IMC fuera de lo normal es más frecuente y se sigue el patrón de las comorbilidades: es más frecuente que los hombres tengan un IMC fuera de lo normal.



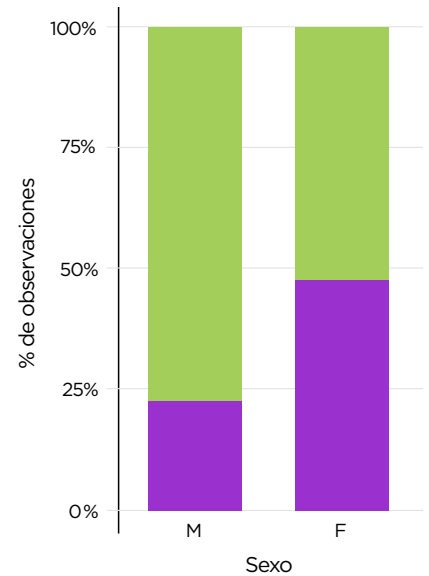
#### Menores de 36



#### Entre 36 y 60



#### Mayores de 60



IMC fuera de lo normal: ■ No ■ Si

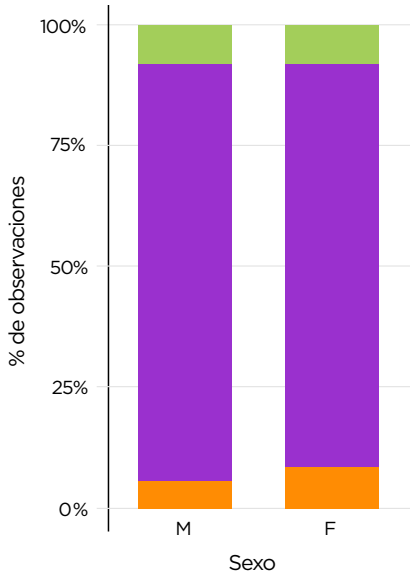
Figura 4: Porcentaje de individuos con IMC dentro o fuera de lo normal, por franja de edad y sexo

En cuanto al número de horas de sueño, podemos observar que la distribución es bastante homogénea entre sexos para todas las franjas de edad, siendo la franja de sueño entre 6 y 8 horas la más frecuente en todos los rangos de edad. Se observa un aumento, esperable, del porcentaje de personas que duermen más de 8 horas en los mayores de 60 años.

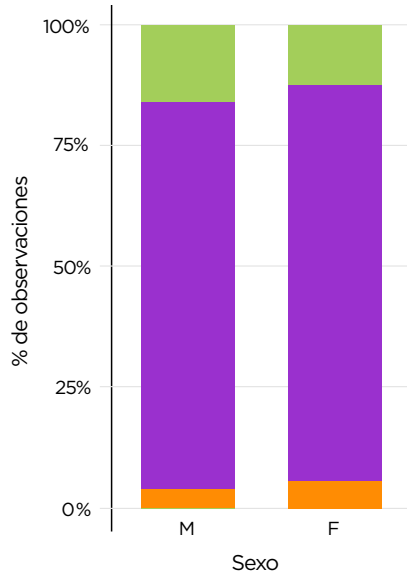




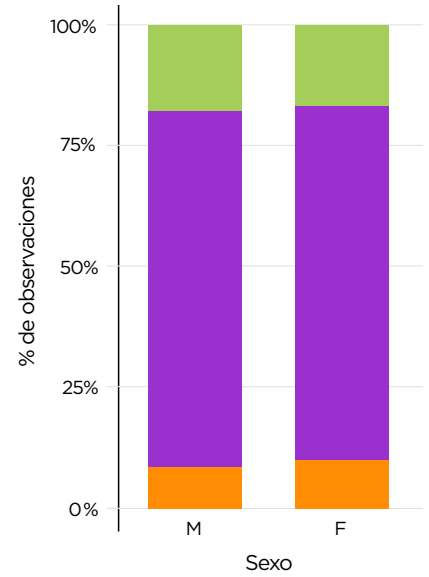
Menores de 36



Entre 36 y 60



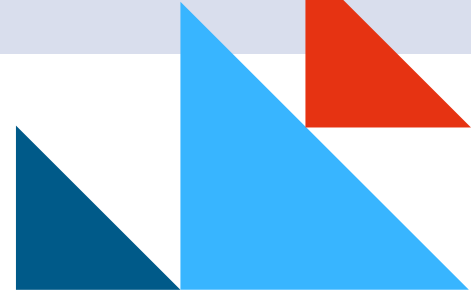
Mayores de 60



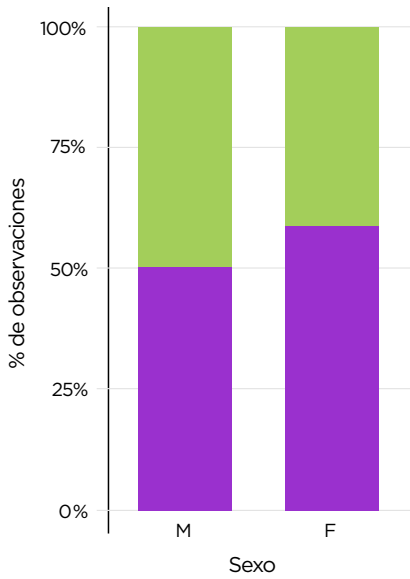
Dormir h. días: ■ Menos de 6 h. ■ Entre 6 y 8 h. ■ Mas de 8 h.

Figura 5: Distribución de participantes atendiendo al número de horas de sueño, franja de edad y sexo

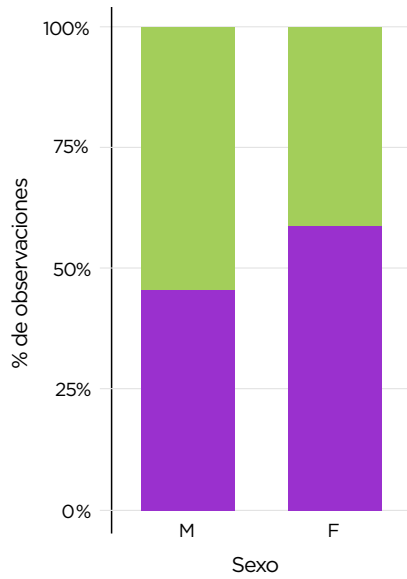
En lo relativo al estrés, existen marcadas diferencias entre sexos en todas las franjas de edad, siendo las mujeres las que padecen mayores niveles de estrés. Es especialmente significativa la diferencia en la franja de los mayores de 60 años, si bien, en esta franja de edad los niveles de estrés son inferiores a los de etapas anteriores de la vida.



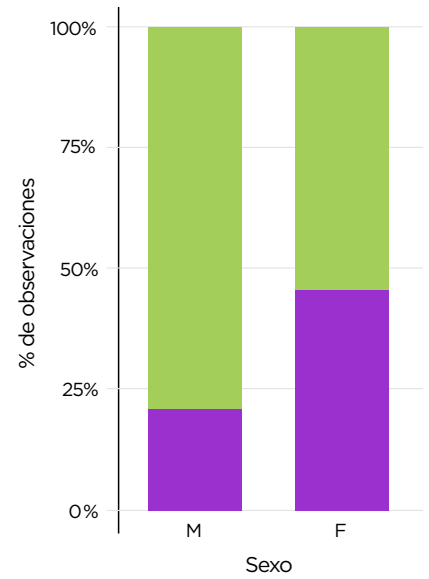
Menores de 36



Entre 36 y 60



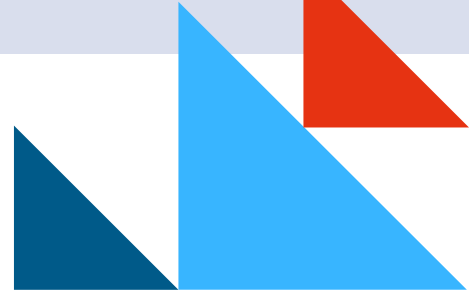
Mayores de 60



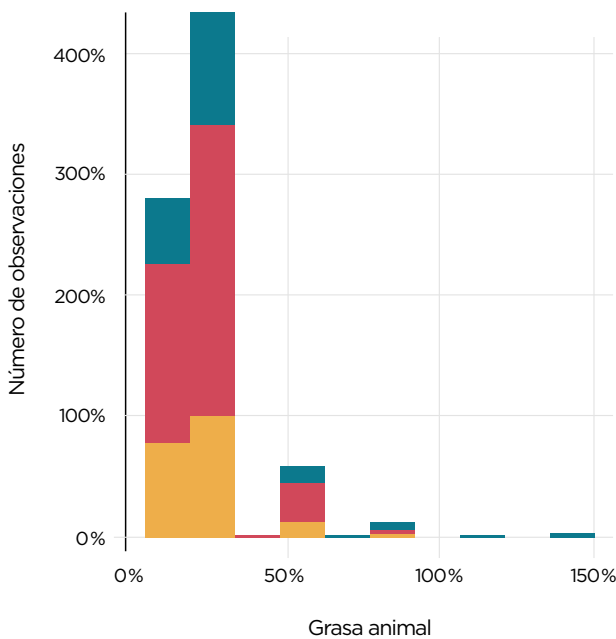
Tiene stress: ■ No ■ SI

Figura 6: Porcentaje de individuos que padecen o no estrés por sexo y franja de edad

Finalmente, se incluye el gráfico que muestra el consumo semanal de grasa por franja de edad. Se observa que en general el consumo es bajo o moderado en todas las franjas, aunque algunos individuos presentaron un consumo muy elevado.

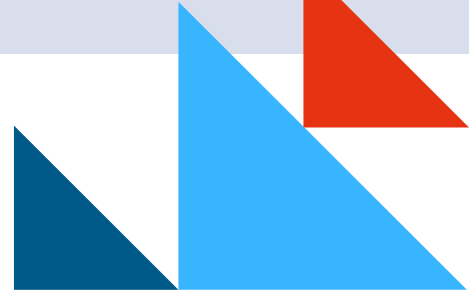


Finalmente, se incluye el gráfico que muestra el consumo semanal de grasa por franja de edad. Se observa que en general el consumo es bajo o moderado en todas las franjas, aunque algunos individuos presentaron un consumo muy elevado.



**Franja de edad:** ■ Menor de 36 años ■ Entre 36 y 60 años ■ Mayor de 60 años

Figura 7: Consumo semanal de grasa animal por grupo de edad



## 5. DESARROLLO DEL ALGORITMO PREDICTIVO

La construcción del algoritmo se ha desarrollado en dos fases, en una primera fase se ha imputado a cada individuo de la muestra su edad biológica y en una segunda fase se ha desarrollado un modelo de regresión multivariante que nos permite obtener los scores que predicen la edad biológica. A continuación, describimos en detalle cada una de estas fases.

### 5.1 Fase 1: Imputación de la edad biológica

Dado que la edad biológica es una variable no observable es necesario establecer un criterio que permita asignar a cada persona su edad biológica. En este caso se ha asumido que la edad biológica es la que se corresponde con el estado funcional de nuestros órganos comparados con patrones estándar para una edad cronológica determinada, por lo tanto, la premisa en la que nos hemos basado para esta fase es la siguiente: la edad biológica de un individuo será la edad cronológica de la clase de individuos de misma edad a la que más se parece, según las variables que representan el estado funcional.

La aplicación de esta definición requiere, en primer lugar, el cálculo lo que llamaremos el individuo medio, es decir, aquel individuo, que, para cada edad, representa a la media de los individuos con la misma edad cronológica. En segundo lugar, una vez definido los individuos medios, para cada individuo se calcula la edad biológica como la edad cronológica del individuo medio que más se le parece. El proceso de imputación de las edades biológicas se puede resumir como en la Figura 8.



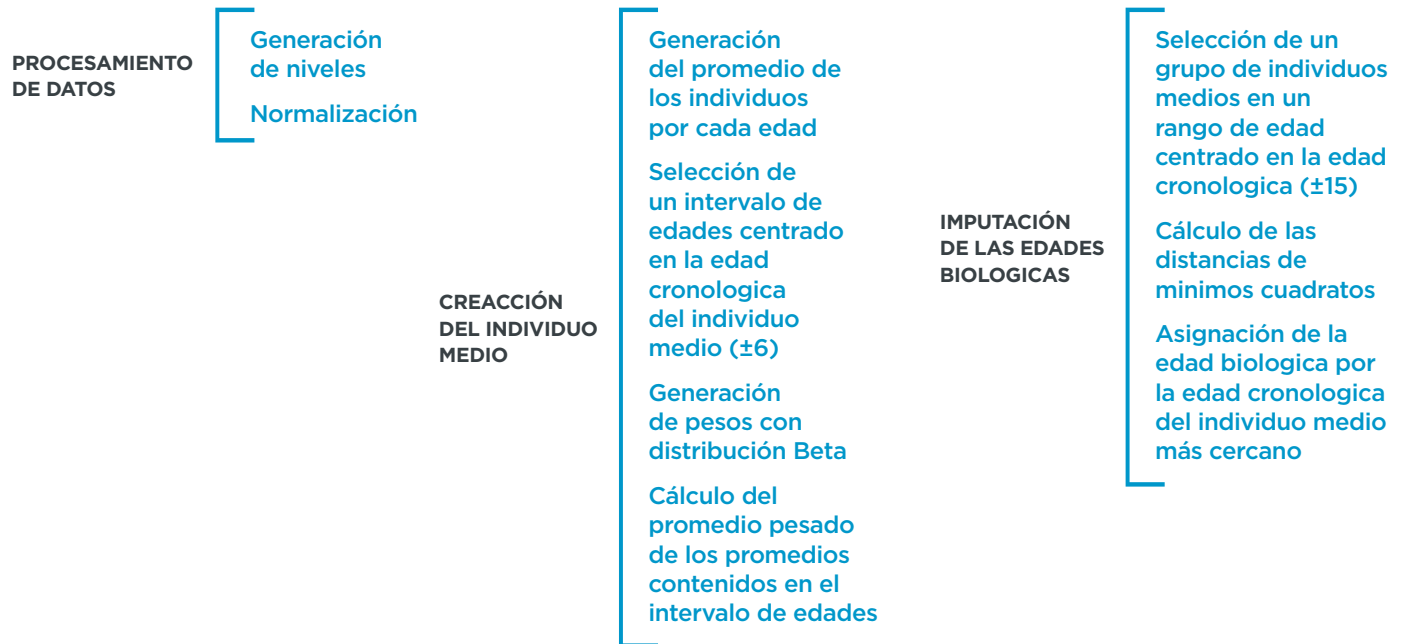


Figura 8: Etapas en la imputación de la edad biológica

A continuación, describimos en detalle las diferentes etapas de este proceso.

### a) Procesamiento de datos

La correcta imputación de las edades biológicas requiere en primer lugar la elección de las variables que caracterizan el estado funcional del individuo medio. La siguiente tabla muestra el conjunto de variables finalmente elegidas y que incluyen estilo de vida, hábitos alimenticios y biomarcadores de salud, se ha decidido no incluir las variables genéticas al no tener certeza de cuál es su impacto en el estado funcional de una persona.



## VARIABLES USADAS EN LA IMPUTACIÓN

Edad	Consumo de grasa animal semanal
Sexo	Consumo de grasa vegetal semanal
IMC	Consumo total de grasas
Consumo de alcohol	Cúantas horas al día pasó sentado
Consumo de frutas	Actividad de actividad física intensa a la semana
Consumo de verduras	Ejercicio físico moderado diario
Consumo de pescado	Tiempo dedicado a caminar semanalmente
Cúantas horas duerme al día	Total de actividad física semanal
Número de comorbilidades	

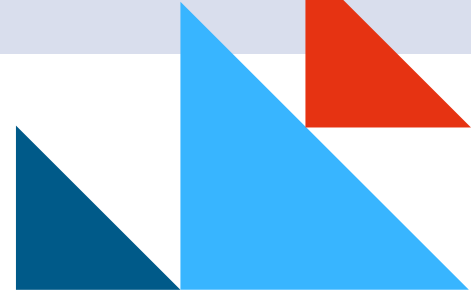
**Tabla 1:** Variables utilizadas en la imputación de la edad biológica

Dado que también es necesario calcular el valor medio de las variables categóricas, éstas se han redefinido como variables binarias. Además, todas las variables se han normalizado en el intervalo [0,1] según la siguiente fórmula:

$$x_n \leftarrow \frac{x_n - \min x_n}{\max x_n - \min x_n}$$

### b) Cálculo de los individuos medios

El individuo medio se calcula para cada edad cronológica en el intervalo de edades contenido en las bases de datos. Para ejemplificar la descripción de su cálculo, asumimos, por ejemplo, que queremos calcular el individuo medio correspondiente a una edad cronológica de 50 años.



A partir de esta edad se elige un intervalo de edades, centrado en esta, y de semi-longitud de  $\pm 6$  años, en nuestro ejemplo el intervalo va a ser [44,56]. Para cada edad en este intervalo, se genera una observación promedia que representa la clase de individuos con aquella edad cronológica.

Finalmente se calcula el individuo medio, promediando estos últimos promedios usando unos pesos (cuya suma es 1). Los pesos sirven para evitar que las clases de edades sin individuos no afecten el resultado, y para que la clase centrada, en nuestro ejemplo la de 50 años, tenga más peso que las clases perimetrales. La Tabla 2 muestra los pesos aplicados a las clases promedios del ejemplo, mientras la Figura 9 muestra los mismos pesos de forma gráfica.

Edad	Peso	Edad	Peso
44	0,01684115	51	0,11242604
45	0,04688211	52	0,10423305
46	0,07146108	53	0,09057806
47	0,09057806	54	0,07146108
48	0,10423305	55	0,04688211
49	0,11242604	56	0,01684115
50	0,11515702		

Tabla 2: Pesos asignados para el cálculo del individuo medio de edad 50 años

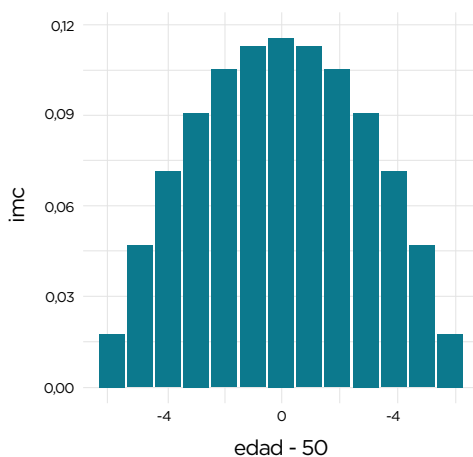
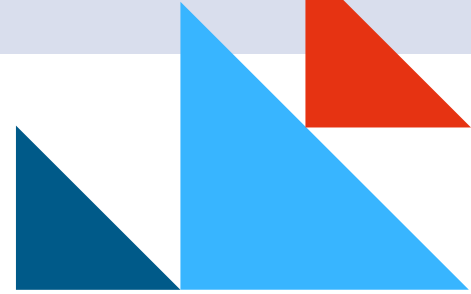


Figura 9: Representación gráfica de los pesos utilizados para el cálculo del individuo medio de edad 50 años



La Figura 10 resume, de forma iconográfica, el proceso de cálculo del individuo medio de edad 50 años.

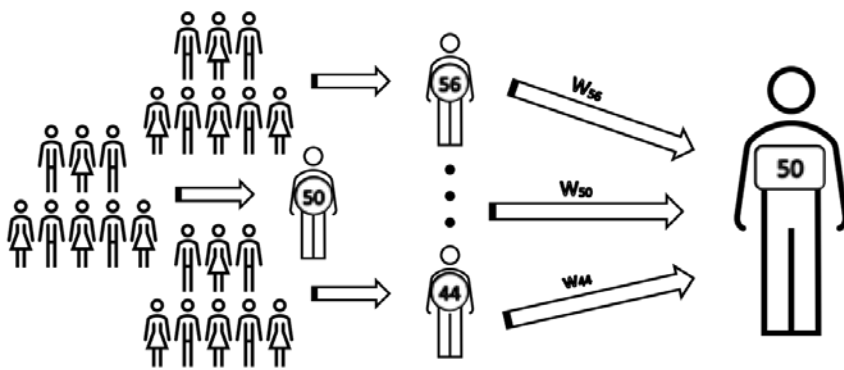
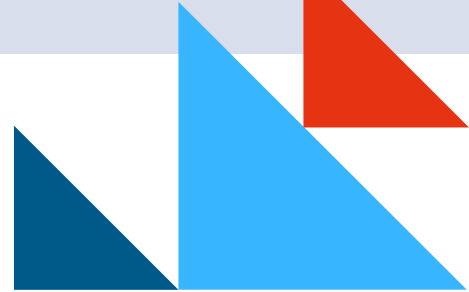


Figura 10: Proceso de cálculo del individuo medio de edad 50 años

### c) Cálculo de la distancia entre un individuo y un individuo medio

La imputación de la edad biológica para un individuo se hace asignándole la edad cronológica del individuo medio más cercano a él. La Figura 11 muestra este procedimiento. Dado un individuo a quien hay que asignar la edad cronológica, en primer lugar, se calcula un intervalo de edades de semi-longitud  $\pm 15$  años centrado en su edad. Por ejemplo, asumiendo un individuo de 50 años, el intervalo sería  $[35,65]$ . A continuación, se seleccionan los 31 individuos medios (ya que el tamaño del intervalo es 31), calculados anteriormente, cuyas edades se encuentran en este intervalo. Para cada individuo medio seleccionado, se calcula una distancia según la siguiente definición:





$$\text{Distancia}(x, y(n)) = \sum_{i \in \text{entorno de ind } n\text{-esimo}} (x_i - y_i(n))^2,$$

donde, x representa el individuo bajo análisis, mientras y(n) representa el n esimo individuo medio. El índice del sumatorio incluiría todas las variables descritas en la Tabla 1, con excepción de la edad. Una vez identificado el individuo medio con menor distancia, se asigna su edad cronológica, como edad biológica, al individuo x.

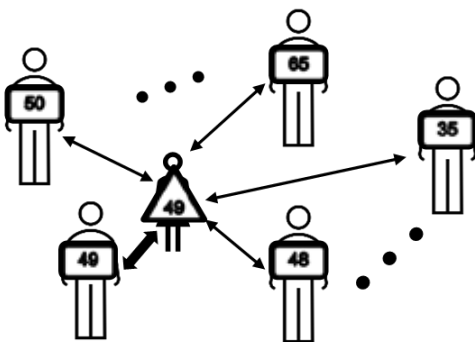
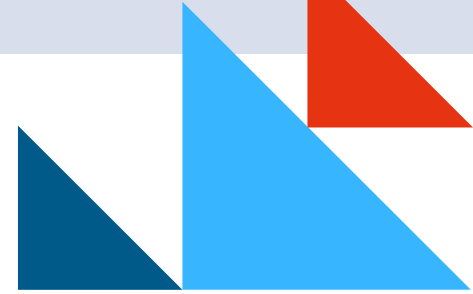


Figura 11: Asignación de la edad biológica mediante el criterio de distancia mínima

Como alternativa a la utilización del individuo medio para imputar la edad biológica, también puede considerarse la imputación según profundidad. En Estadística Multivariante, el término profundidad hace referencia al grado de centralidad de una observación respecto de una nube de puntos. Así, cuanto más central sea la observación, mayor será su profundidad. Si partimos de una única característica cuantitativa, por ejemplo, la estatura y consideramos la población de varones españoles adultos, el valor 176 cm (estatura media) será muy profundo, mientras que la profundidad de otros valores como 170 cm ó 180 cm no será tan grande y la de 160 cm ó 190 cm será aún menor. La situación es más interesante si consideramos más de una característica, por ejemplo, la estatura y el peso.



Dentro de los varones españoles adultos, la observación asociada a un individuo de 176 cm y 76 kg será muy profunda, mientras que la de otro de 190 cm y 90 kg será bastante menos profunda y la de uno de 190 cm y 60 kg será muy poco profunda, dado que resulta muy anómalo que un individuo tan alto tenga tan poco peso. En conclusión, la profundidad refleja lo representativas que son las características asociadas a un individuo dentro del grupo al que pertenece. Un valor elevado para la profundidad debe interpretarse como que el individuo es muy representativo, es decir, recoge fielmente las características globales del grupo al que pertenece.

Para imputar la edad biológica de un individuo según profundidad, seleccionamos una submuestra con todos los individuos cuya edad biológica difiera de la del individuo de interés a lo sumo en una cantidad fija de años. A continuación, balanceamos esta submuestra para que todas las edades biológicas presentes en ella tengan la misma representación en cuanto a número de individuos. Finalmente, el valor imputado de la edad biológica para el individuo en consideración es aquella edad (biológica) que le haría más profundo (más representativo) respecto al resto de individuos si la consideramos en conjunto con el resto de sus características.

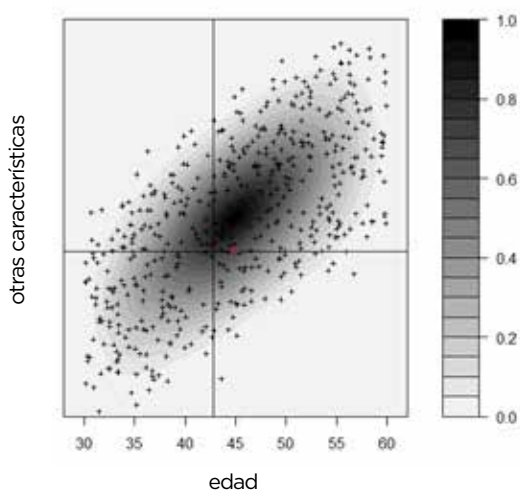
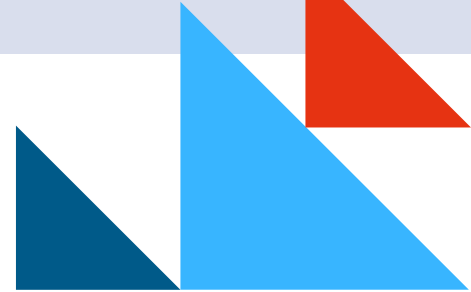


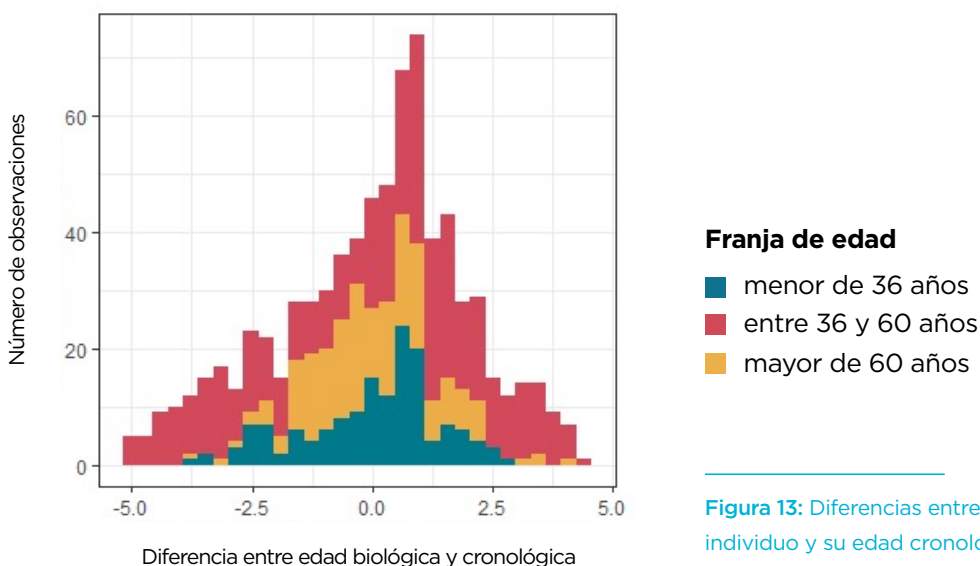
Figura 12: Asignación de la edad biológica según el criterio de profundidad



Para explicar de un modo gráfico la imputación según profundidad, en la Figura 12 hemos representado una muestra de 500 individuos. El eje X se corresponde con la edad biológica, mientras que el eje Y se corresponde con “otras características”. La parte más oscura del gráfico refleja los puntos más profundos. Si deseamos asignar la edad biológica de un individuo de 45 años (en rojo), tomamos su valor de la variable “otras características” y consideramos todas las posibles edades (manteniendo fijo el valor de “otras características”), por lo que nos movemos a lo largo de la línea horizontal de la gráfica. Al individuo que hemos considerado, le imputamos la edad en la que se alcanza la máxima profundidad (línea vertical). En nuestro ejemplo, hemos imputado una edad cronológica de prácticamente 43 años a un individuo cuya edad biológica eran 45. Naturalmente, cuantas más variables consideremos asociadas a cada individuo, más interesante será la imputación.

Después de comparar ambos enfoques para la imputación de la edad biológica nos hemos decantado por el primer enfoque, el cálculo del individuo promedio, al ser metodológicamente más sencillo.

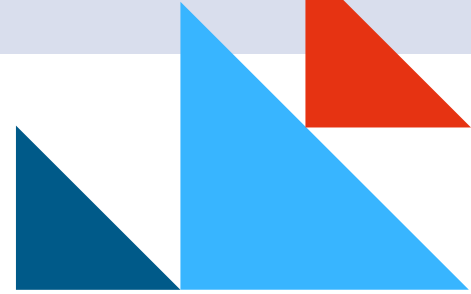
La Figura 13 muestra la diferencia entre la edad biológica imputada a cada individuo y su edad cronológica, donde el color indica las diferentes franjas de edad. A valores menores de esta diferencia corresponde una edad biológica menor de la edad cronológica del individuo.



**Franja de edad**

- menor de 36 años
- entre 36 y 60 años
- mayor de 60 años

**Figura 13:** Diferencias entre la edad biológica imputada a cada individuo y su edad cronológica por franja de edad



Vemos que en todos los rangos de edad la diferencia entre la edad biológica y la cronológica se encuentra entre -5 y 5. La literatura científica indica que estas diferencias pueden llegar a ser de hasta 15 o 20 años, en nuestro caso, dado que el estado de salud de ninguno de los participantes estaba extremadamente deteriorado no se han encontrado estos casos extremos. La Figura 14 muestra, mediante un diagrama de dispersión, cada individuo en función de la pareja de parámetros (edad cronológica, edad biológica), junto con la línea que marca la igualdad de ambas edades. También en este caso se han mantenido los colores para delimitar las tres franjas de edad consideradas en el análisis. Vemos como para todas las franjas de edad hay individuos con edad biológica superior e inferior a la cronológica, apreciándose una mayor dispersión en la franja de edad intermedia.

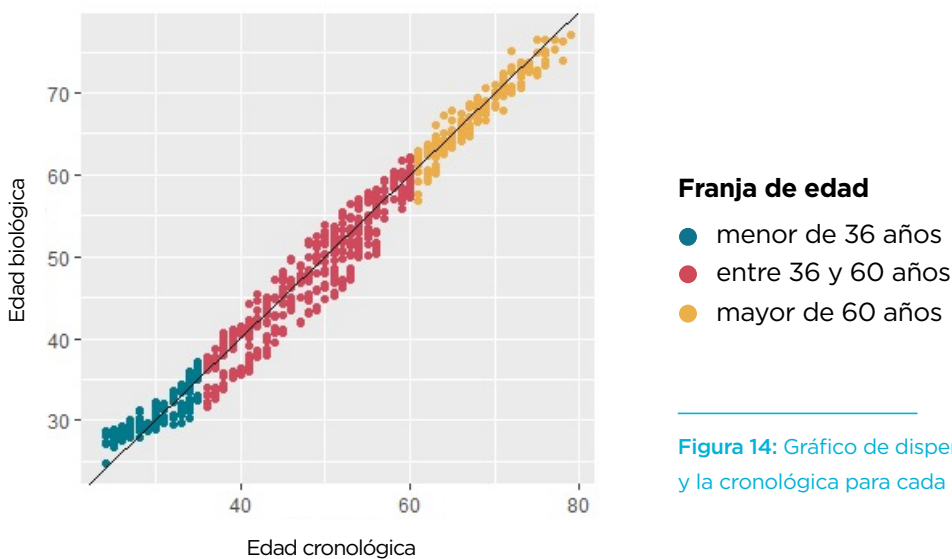
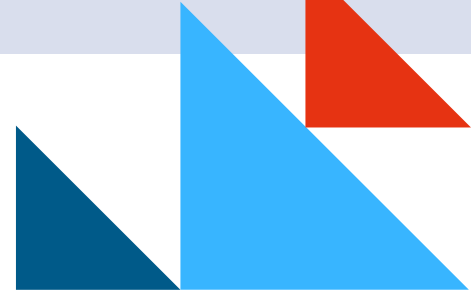


Figura 14: Gráfico de dispersión entre la edad biológica y la cronológica para cada franja de edad

Es necesario tener especial cuidado a la hora de imputar la edad biológica correspondiente a aquellos individuos que se encuentran en los extremos ya que solo es posible compararlos con individuos que tienen una edad superior (en el caso de los más jóvenes) o una edad inferior (en el caso de las personas de edad más avanzada).



## 5.2 Fase 2: Cálculo de scores

Si nuestro interés es simplemente estimar la edad biológica de una persona, el proceso descrito hasta ahora sería suficiente. Sin embargo, el objetivo principal de este proyecto es la obtención de scores que nos permitan:

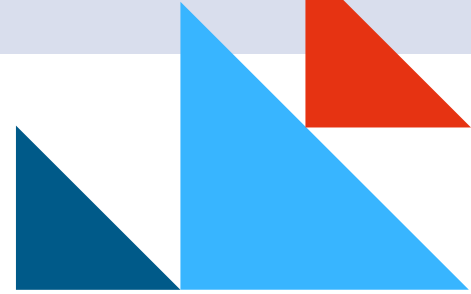
- a) determinar qué variables son las óptimas para estimar la edad biológica
- b) conocer el efecto favorable o adverso de cada una de ellas y en qué magnitud.

Para lograr estos objetivos se ha desarrollado un modelo de regresión multivariante utilizando como variable respuesta la diferencia entre la edad biológica y la cronológica y como predictores tanto las variables recogidas a través del cuestionario como los resultados del análisis genético. Se han estimado dos modelos: un modelo en el que no se incluyen las variables genéticas y un segundo modelo en el que se incluye la información genética a través del conteo de alelos de los genes estudiados.

Tras desarrollar los modelos de diferenciación entre la edad biológica y cronológica, se obtuvo que correlación entre la diferencia entre la edad biológica y cronológica observada y predicha fue de  $\approx 0.7$ . Se encontraron diferentes factores que condicionan significativamente la edad biológica de un individuo. Entre estos factores podemos distinguir:

**Género:** el género femenino reporta un beneficio respecto al masculino, ya que el hecho de ser mujer parece predisponer una disminución de la edad biológica en los modelos predictivos de edad biológica. Este resultado podría sustentarse en la protección hormonal que presentan las mujeres durante la edad fértil, que disminuye la incidencia de patologías cardiovasculares y retrasa su aparición respecto a los varones.

**Índice de Masa Corporal (IMC) y consumo de grasa animal:** según los modelos de edad biológica, los IMC con valores superiores a 24.9 y el consumo de grasas de origen animal predisponen a mayor edad biológica. En este sentido, las situaciones de sobrepeso y obesidad así como el consumo de grasa animal conllevan respuestas pro-inflamatorias y pro-



oxidativas que promueven alteraciones fisiológicas, favoreciendo la aparición de diferentes enfermedades crónicas asociadas al envejecimiento (enfermedades cardiovasculares, oncológicas y metabólicas).

**Presencia de comorbilidades:** la presencia de comorbilidades o patologías disminuye la funcionalidad celular, fisiológica y de reserva vital, por lo que padecer una o más enfermedades induce un mayor grado de envejecimiento biológico real. En concordancia, y ratificando la validez de estos modelos, el número de comorbilidades fue identificado como un indicador de mayor edad biológica.

**Descanso y stress:** un inadecuado descanso y estilo de vida asociado a situaciones de estrés conllevan alteraciones en la liberación hormonal (respuesta humoral), sistema nervioso, problemas digestivos, inmunológicos e incluso alteraciones en el sistema cardiovascular que disminuye su funcionalidad, aumentando la aparición de situaciones patológicas graves relacionadas con procesos de envejecimiento prematuro. En concordancia con los datos anteriores, los modelos de caracterización de edad biológica han identificado que no dormir entre 6-8 horas y presentar niveles altos de stress eran factores implicados en una creciente edad biológica.

**Consumo de alcohol:** sorprendentemente, el consumo de alcohol mostró un comportamiento bimodal en función del “tipo de alcohol” consumido. Los modelos de predicción de edad biológica revelaron que el consumo de alcohol destilado se asocia con mayor edad biológica, mientras que el consumo de alcohol fermentado parece ser un factor que podría disminuir la edad biológica. Este efecto diferencial parece sustentarse en los componentes beneficiosos de las bebidas fermentadas como los polifenoles con actividad antioxidante, antiinflamatoria y vasodilatadora.

**Factor genético:** finalmente, y como valor añadido y principal fortaleza del modelo predictivo de edad biológica, se ha incorporado el efecto del perfil genético en el cálculo de edad biológica, identificándose su papel predictor significativo sobre la edad biológica, aportando mayor precisión en la estimación de la edad biológica. En este sentido, se ha podido identificar la interacción del factor genético con el resto de factores condicionantes de edad

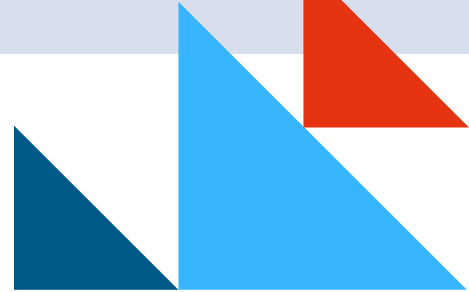


biológica identificado en el primer modelo. Cuando se profundiza en los aspectos genéticos individuales, se ha encontrado que la presencia de diferentes polimorfismos relacionados con la efectividad del ejercicio físico para la degradación de grasa visceral (rs4994), capacidad cognitiva (rs429358) y la palatibilidad (rs10246939), que regula la percepción de sabores dulces y amargos y, que por tanto, condiciona los gustos y la alimentación, son predictores que condicionan la edad biológica de cada individuo.

Como mencionamos anteriormente hemos optado por un modelo sencillo de implementar, en el que cada variable de las seleccionadas tenga asociado un score multiplicativo, por lo que solo es necesario conocer los valores que toman las variables para un individuo concreto.

Por ejemplo:

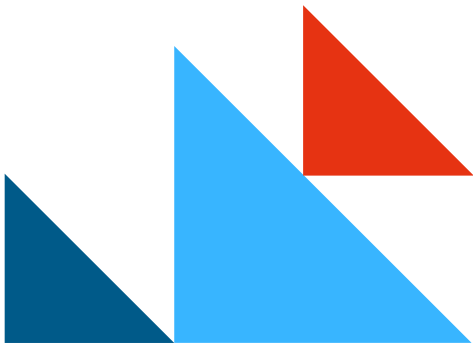
1. Una mujer de 52 años, que duerme más de 8 horas, no padece estrés, no tiene ninguna enfermedad, con un IMC dentro de lo normal, no consume ningún tipo de alcohol y tiene un consumo de grasa animal de 15gr al día tendría una edad biológica de 49 años.
2. Un hombre de 65 años que duerme menos de 6 horas, padece estrés, tiene 1 enfermedad, consume alcohol destilado, tiene un IMC superior a 25 y tiene un consumo diario de 120gr de grasa animal, tiene una edad biológica de 69 años.



## 6. CONCLUSIONES E INVESTIGACIÓN FUTURA

- ▶ La edad biológica es un indicador más preciso del estado real del envejecimiento del organismo que la edad cronológica, y viene definida por diferentes factores modificables (consumo de alcohol, consumo de grasas animales, obesidad, horas de descanso, nivel de estrés, presencia de sobrepeso/obesidad y número de enfermedades crónicas) y no modificables (género y perfil genético).
- ▶ El algoritmo predictivo de edad biológica desarrollado en el presente proyecto es el primer algoritmo que combina hábitos de vida y biomarcadores genéticos y no genéticos en un mismo modelo. Además, permite establecer con mayor precisión y exactitud la esperanza de vida para cada individuo de manera individualizada, pudiéndose diseñar programas de prevención basados en aspectos conductuales específicos con la finalidad de disminuir la edad biológica.
- ▶ El perfil genético parece ser un factor determinante y condicionante del envejecimiento, que regula el impacto de diferentes factores conductuales y modificables sobre la edad biológica.
- ▶ La creación de una plataforma Web donde se recojan datos de nuevos individuos permitirá mejorar el algoritmo de una forma dinámica y continua, pudiendo identificar nuevos factores que mejoren la predicción de la edad biológica y que puedan predisponer a un envejecimiento saludable o feliz (plenitud en la vejez).





*fundación*  
**MUTUALIDAD**  
**ABOGACÍA**

*escuela*  
*de pensamiento*

Con la colaboración de:

**uc3m** | Universidad  
**Carlos III**  
de Madrid



Instituto  
Gerontológico 