Programa	**Proyectos Estratégicos en Cooperación**
Título del Proyecto	**iSARC-GENETICS** PLATAFORMA INTELIGENTE PARA LA DETECCIÓN TEMPRANA, AYUDA AL DIAGNÓSTICO Y SEGUIMIENTO DE LA SARCOPENIA
Entregable	E1.4 Resultados del estudio clínico de Sarcopenia EWGSOP + iSARC-GENETICS (M28)
Responsable	**Sabartech, Dawako**
Fecha	**Junio 2023**

1. INTRODUCCIÓN

T1.4: Protocolo de adquisición de datos clínicos para el diagnóstico de Sarcopenia según el Grupo de trabajo europeo sobre sarcopenia en personas mayores ("EWGSOP") + iSARC-GENETICS (Propietario: DAWAKO)

Para la validación de los biomarcadores de imagen y los biomarcadores genéticos en la memoria se incluyó el reclutamiento de 60 pacientes divididos en 3 grupos:

- ***20 pacientes de 30-60 años*** que acuden por patología banal que no afecta a miembros inferiores y sin limitaciones funcionales en su capacidad de marcha ni transferencias (epicondilitis, tendinitis de hombro, cervicalgia, etc.),
- ***20 pacientes mayores de 60 años*** que acuden al servicio de Rehabilitación por patología que no afecta a miembros inferiores y sin limitaciones funcionales en su capacidad de marcha ni transferencias (mismas patologías),
- ***20 pacientes mayores de 60 años*** que acuden a Rehabilitación a ser tratados específicamente por cuadros de fragilidad y pérdida de su funcionalidad (siguiendo criterios del Grupo de Trabajo europeo, pacientes que refieren caídas, debilidad, pérdida de velocidad en la marcha, dificultades para incorporarse de una silla, pérdida de peso o atrofia muscular).

A todos los pacientes reclutados se planteó realizar las siguientes determinaciones:

- Test funcionales: HANDGRIP, CHAIR STAND TEST, MARCHA, UP and GO
- SARC-F en los grupos mayores de 60 años
- UP and GO
- Protocolo ECO, como parámetros de calidad y cantidad muscular
- Análisis genético mediante tecnología de DNA Microarray.

Además se planteó recoger los siguientes parámetros:

- **Consentimiento informado**
- **Aprobación de comité de ética**
- **Variables:** edad, peso, altura, IMC, longitud de muslo, longitud de pierna, circometría muslo, circometría pierna.
- **Mediciones básicas:** grosor muscular, área de sección transversal, ángulo de penación, longitud fascicular, intensidad ECO

Para cumplir con los objetivos de la presente tarea se presentó el diseño del estudio al Comité de Ética de la Investigación de Pontevedra-Vigo-Ourense del SERVIZO GALEGO DE SAÚDE.

Una desviación en la presente tarea fue la no aprobación de los análisis genéticos por parte del Comité de Ética, el cual recogía el resto de mediciones asociadas a los biomarcadores de imagen. Para solventar la presente desviación SBT recogió muestras genéticas de un total de 34 voluntarios procedentes del estudio piloto en la Clínica Dawako.

Programa	**Proyectos Estratégicos en Cooperación**
Título del Proyecto	**iSARC-GENETICS** PLATAFORMA INTELIGENTE PARA LA DETECCIÓN TEMPRANA, AYUDA AL DIAGNÓSTICO Y SEGUIMIENTO DE LA SARCOPENIA
Entregable	E1.5 Generación del repositorio integrado biobanco de datos de Imágenes de Ultrasonidos y Genética (M28)
Responsable	**Dawako, Sabartech, CoMMLab**
Fecha	**Junio 2023**

1. INTRODUCCIÓN

T1.5: Generación del repositorio integrado biobanco de datos de Imágenes de Ultrasonidos y Genética (Propietarios: DAWAKO, SBT, COMMLAB)

Los datos de ultrasonidos se almacenan en un servidor para almacenamiento de imágen instalado en la nube en el entorno de iSARC-GENETICS y oculto tras el front-end para garantizar su seguridad. En este repositorio se almacenan todos los biomarcadores de imagen validados en una base de datos que contiene tanto las imágenes médicas como todos los análisis y biomarcadores asociados. Se ha creado una tabla de datos, llamada ***front_geneticstudy***, la cual contiene los resultados asociados al estudio genético.

Por su lado, los datos pertenecientes al análisis genético se almacenan en las bases de datos del sistema productivo de Sabartech. Ambas bases de datos, la de genética e imagen, se conectan vía API, y se intercambian la información a través de un archivo .json. La plataforma iSARC-GENETICS pide los siguientes datos al laboratorio que analiza el hisopo para obtener los datos genómicos.

Programa	**Proyectos Estratégicos en Cooperación**
Título del Proyecto	**iSARC-GENETICS** PLATAFORMA INTELIGENTE PARA LA DETECCIÓN TEMPRANA, AYUDA AL DIAGNÓSTICO Y SEGUIMIENTO DE LA SARCOPENIA
Entregable	E2.1. Resultados del Cribado Genético para identificar personas susceptibles de sufrir sarcopenia: Se mostrarán los resultados procedentes de los análisis de ADN extraído de muestras de saliva de los voluntarios participantes mediante la tecnología de DNA Microarray, procesados dentro de la PIPELINE genómica y mediante el Algoritmo genético (M31)
Responsable	**Sabartech**
Fecha	**Septiembre 2023**

1. INTRODUCCIÓN

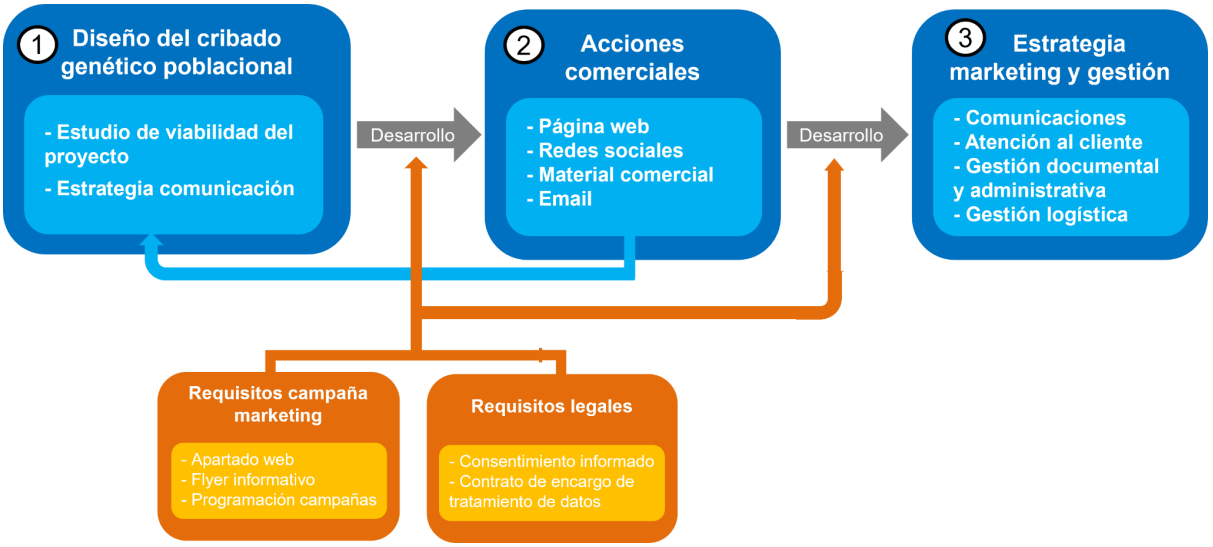
T2.1: Cribado genético en un entorno real de una muestra de población normal y heterogénea

El objetivo que debe cumplirse en esta actividad es la obtención de información genética de un número representativo de individuos de una población lo más heterogénea posible, en la que representa la sociedad actual. Dichos datos son necesarios para la validación del algoritmo genético que permita identificar y estratificar a las personas susceptibles de sufrir sarcopenia de forma precoz en base a su genética y la gravedad de la enfermedad. Este paquete de trabajo se ha completado en su totalidad por lo que su objetivo se ha cumplido.

El tamaño muestral del cribado genético se ajustó a 940 voluntarios, que junto con los 34 pacientes de la Clínica Dawako empleados para el proceso de validación, el tamaño ascendió hasta un total de **974 individuos**. El tamaño muestral obtenido cumple con los objetivos establecidos en el proyecto que, calculados en base a la estimación del tamaño muestral considerando la prevalencia de sarcopenia (10%) y una precisión del 2%, se determinó en un total de 864 personas.

Para el diseño del estudio de cribado poblacional de la sarcopenia y el reclutamiento de voluntarios Sabartech contrató a la empresa Overgenes. Las acciones de Overgenes están englobadas en la **tarea 2.1 (Estudio de investigación poblacional y cribado genético de Sarcopenia)** y se dividen en 3 fases:

- **Fase 1**: diseño del cribado genético poblacional. Se realizó un estudio de la viabilidad del proyecto y se diseñó la estrategia de comunicación entre los clientes de Overgenes.
- **Fase 2**: acciones comerciales a través de la página web, redes sociales y por email.
- **Fase 3**: continuar con la estrategia de marketing y gestión de documentación administrativa para participar en el estudio.



Overgenes también se encargó de solucionar todas las consultas relacionadas con el proyecto y ofrecer las aclaraciones correspondientes tanto a voluntarios, clientes y distribuidores.

Se llevaron a cabo gestiones logísticas de envío y recepción de muestras. A lo largo del cribado surgieron distintas incidencias logísticas que fue necesario subsanar. Algunos de los análisis genéticos realizados no cumplieron con los estándares de calidad mínimos y se tuvieron que repetir con nuevas muestras para asegurar la robustez de los mismos.

2. CRIBADO GENÉTICO

Para la captación de voluntarios participantes en el estudio del cribado genético de personas con predisposición genética a desarrollar sarcopenia se llevaron a cabo las siguientes acciones:

- **Diseño del cribado genético población**. Se realizó un estudio de la viabilidad del proyecto y se diseñó la estrategia de comunicación entre los clientes de Overgenes.
- **Acciones comerciales** a través de la página web, redes sociales y por email.
- **Gestión de documentación administrativa** para participar en el estudio.

3. RESULTADOS DEL CRIBADO GENÉTICO

Para analizar el conjunto de 940 muestras incluídas en el cribado genético se ha empleado la implementación del algoritmo en los motores de producción de Sabartech, que ha permitido tanto su análisis como posterior estudio. Previamente el algoritmo genético ha sido validado basándonos en los resultados obtenidos con los pacientes de la Clínica Dawako.

El proceso de cribado genético ha sido fundamental para estudiar el comportamiento del algoritmo genético en un número representativo de la sociedad, seleccionando para ello una muestra heterogénea obtenida de manera aleatoria. De esta manera se permite validar el algoritmo genético y estudiar su capacidad para identificar y estratificar a las personas con susceptibilidad a sufrir Sarcopenia.

En la Tabla 1 se muestran los descriptivos de los resultados obtenidos a partir de las muestras del cribado genético, incluyendo la media del riesgo genético, su desviación típica en el conjunto de muestras y finalmente la clasificación de riesgo más representado. Como se puede apreciar en una muestra representativa y heterogénea de la población el riesgo más destacado es el bajo, lo cual es esperable dado que la prevalencia de la población es de aproximadamente un 10%. Concretamente, a partir de los datos del estudio se obtiene una prevalencia calculada del 24.47% (230 individuos de 940), considerando aquellos individuos con una predisposición genética de moderada a alta.

Por otra parte, tal y como se indica en la descripción, el estilo de vida influye en el desarrollo de la enfermedad y en su severidad. De acuerdo a esta premisa, se ha considerado el ajuste multiescala teniendo en cuenta el Índice de Masa Corporal (IMC), observando que al incluir esta variable el riesgo aumenta a moderado.

ESTADÍSTICO	VALOR	
	ÁMBITO GENÉTICO	ÁMBITO MULTIESCALA
MEDIA	0.360	0.534
DESVIACIÓN TÍPICA	0.078	0.229
MODA	RIESGO BAJO	RIESGO MODERADO

Tabla 1. Estadísticos obtenidos en el cribado masivo.

Seguidamente se muestran unos gráficos que resumen los resultados obtenidos. De esta manera en la Figura 1 se aprecia como el riesgo genético en el conjunto de muestras del cribado sigue una distribución normal con una ligera cola hacia la derecha, indicando que prácticamente ningún individuo del cribado presenta una predisposición genética inferior a 0.3, lo cual es esperable dada la prevalencia de la enfermedad. Esta distribución de riesgo se altera al incluir el parámetro multiescala (Figura 2), donde el riesgo tiende a centrarse en los rangos más elevados y bajos respectivamente, condicionando de gran manera la predisposición a presentar Sarcopenia. No obstante se puede apreciar un pico en el riesgo máximo, que se corresponde con los pacientes con obesidad. A la vista de los resultados se puede concluir que tanto el riesgo genético, como el riesgo multiescala presentan una variabilidad adecuada.

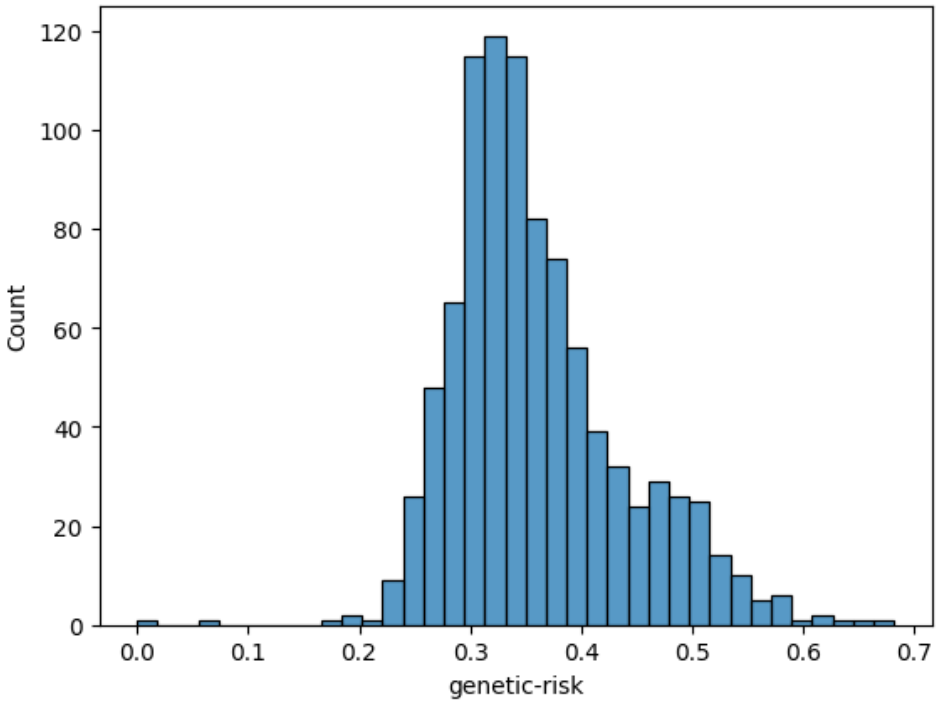


Figura 1. Histograma con la distribución del riesgo genético obtenido durante el cribado masivo.

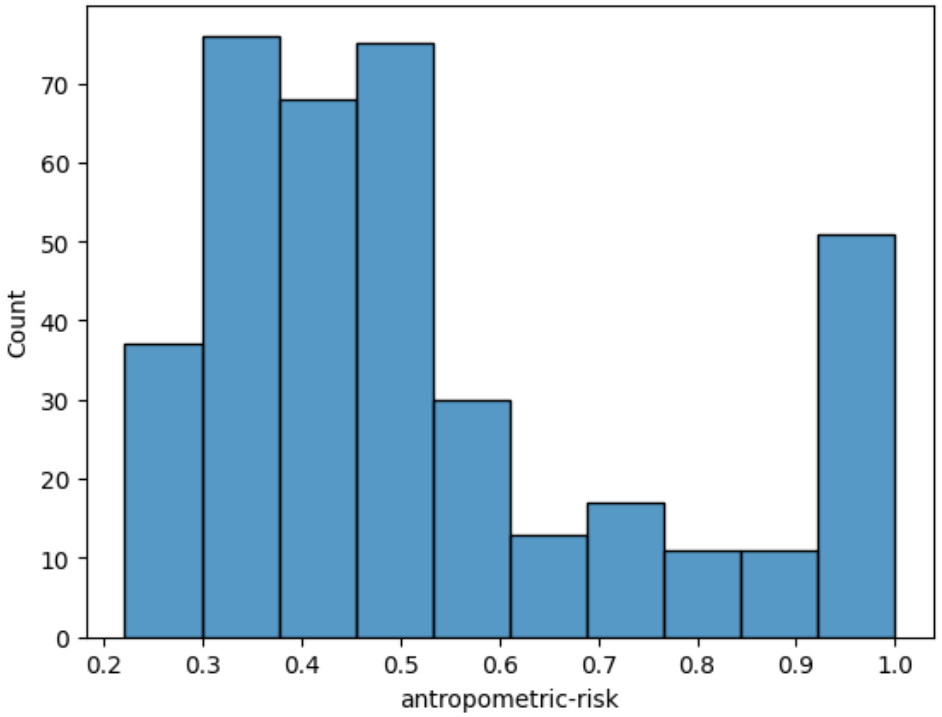


Figura 2. Histograma con la distribución del riesgo multiescala obtenido durante el cribado masivo.

El detalle de cómo se distribuye el riesgo genético estratificando por el IMC se aprecia en las Figuras 3 (por proporción de muestras, para homogeneizar la diferencia entre el número de individuos con normopeso, sobrepeso y obesidad) a 4 (por conteo de número de muestras). En estas se observa que los individuos con obesidad tienden a disponer de mayor riesgo genético. Remarcar que finalmente se eliminaron del modelo las variantes genéticas relacionadas con *traits* directamente relacionados con la gestión del peso corporal.

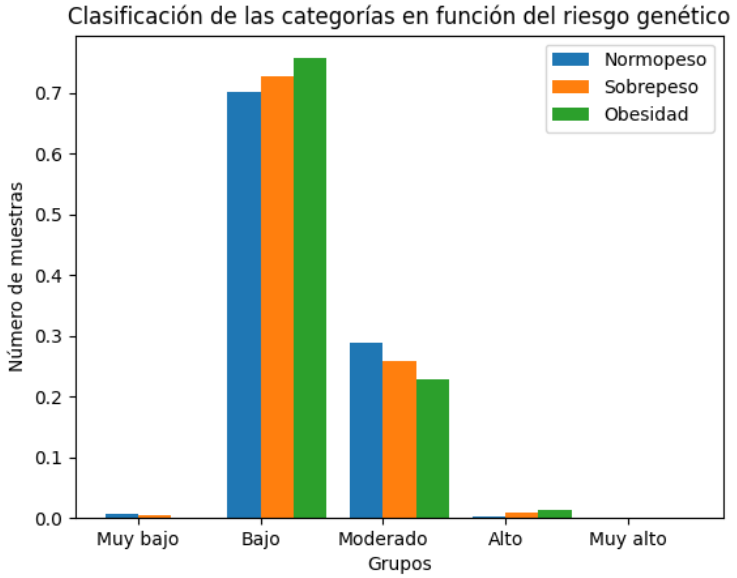


Figura 3. Gráfico de barras con la distribución de categorías de riesgo genético en función del índice de masa corporal.

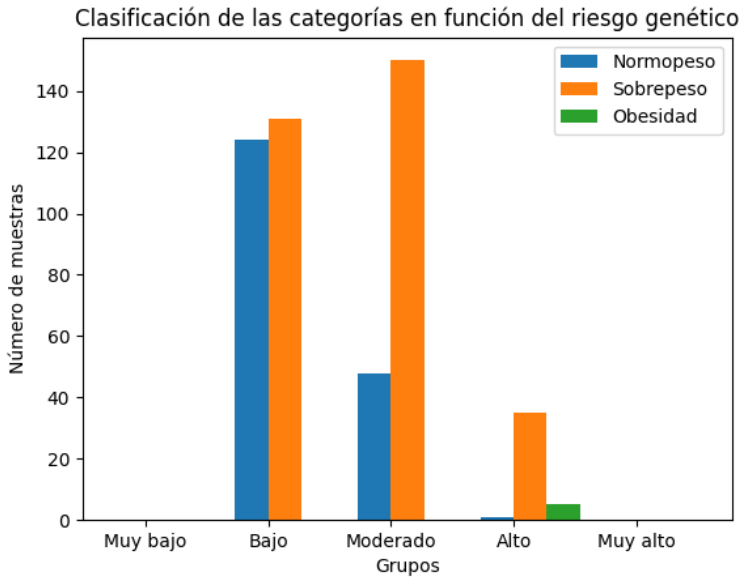


Figura 4. Gráfico de barras con la distribución de categorías de riesgo antropométrico en función del índice de masa corporal.

De acuerdo a Yuan & Larsson (2023), otro parámetro antropométrico influyente en la predisposición de riesgo a desarrollar Sarcopenia es el género del individuo. El detalle de cómo se distribuye el riesgo genético estratificando por el género se aprecia en las Figuras 5 (por conteo de número de

muestras) a 6 (por proporción de muestras, para homogeneizar la diferencia entre el número de individuos hombres y mujeres). De estas imágenes se puede extrapolar que las mujeres presentan tienden a disponer de un riesgo bajo-moderado y los hombres acumulan la mayor parte de resultados con riesgo alto.

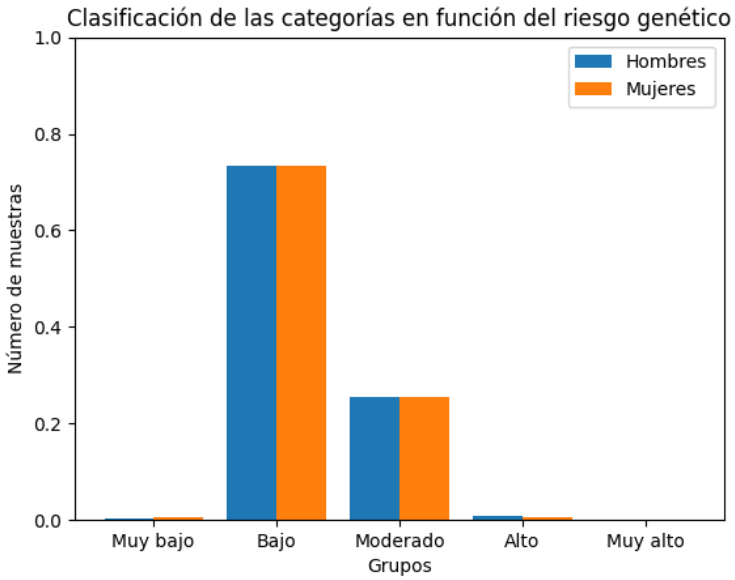


Figura 5. Gráfico de barras con la distribución de categorías de riesgo genético en función del sexo.

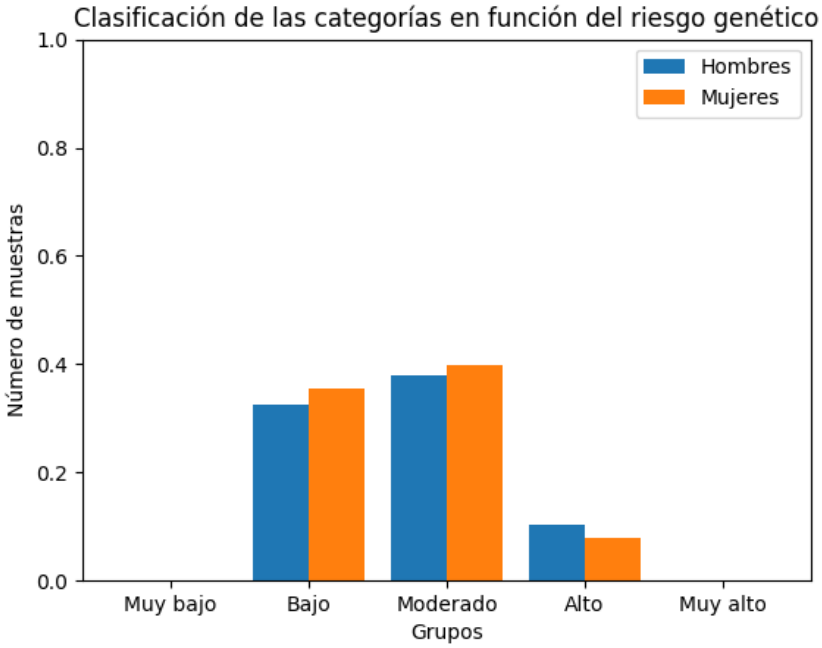


Figura 6. Gráfico de barras con la distribución de categorías de riesgo antropométrico en función del sexo.

Finalmente resaltar la no presencia de resultados genéticos con riesgo muy bajo o muy alto considerando únicamente el factor genético.

4. BIBLIOGRAFÍA

- Yuan, S., & Larsson, S. C. (2023). Epidemiology of sarcopenia: Prevalence, risk factors, and consequences. *Metabolism*, *144*, 155533. doi:10.1016/j.metabol.2023.155533

Actuació cofinançada per la Unió Europea a través del Programa Operatiu del Fons Europeu de Desenvolupament Regional (FEDER) de la Comunitat Valenciana 2014-2020

Programa	**Proyectos Estratégicos en Cooperación**
Título del Proyecto	**iSARC-GENETICS** PLATAFORMA INTELIGENTE PARA LA DETECCIÓN TEMPRANA, AYUDA AL DIAGNÓSTICO Y SEGUIMIENTO DE LA SARCOPENIA
Entregable	E3.6. Herramienta predictiva de análisis de sarcopenia basada en modelos de inteligencia artificial y su validación clínica (M21)
Responsable	**Commlab, Sabartech, Dawako**
Fecha	**Noviembre 2022**

1. INTRODUCCIÓN

T3.6: Herramienta predictiva de análisis de sarcopenia basada en modelos de inteligencia artificial y su validación clínica

El objetivo fundamental de este entregable, es presentar el modelo que se ha diseñado e implementado para poder predecir la existencia de sarcopenia en pacientes de forma precoz, utilizando marcadores basados en imagen médica y en genética.

Se ha diseñado un framework basado en Machine Learning, que está diseñado para aprender la relación entre los biomarcadores obtenidos de las secuencias de imagen de ultrasonidos, y los biomarcadores genéticos. De este modo, inferimos los umbrales para cada uno de los biomarcadores asociadas a sarcopenia.

2. DISEÑO DE MODELO DE IA

El sistema que se ha implementado está basado en árboles de decisión. Puesto que inicialmente no sabemos *a priori* que biomarcadores son más predictivos para identificar la Sarcopenia, se ha llevado a cabo una selección de características.

Actuació cofinançada per la Unió Europea a través del Programa Operatiu del Fons Europeu de Desenvolupament Regional (FEDER) de la Comunitat Valenciana 2014-2020

SELECCIÓN DE CARACTERÍSTICAS

Dado que contamos con un conjunto inicial de 40 características, es conveniente realizar una selección de las más relevantes antes de utilizar el árbol de decisión. Para este propósito, hemos utilizado el método SelectKBest de la biblioteca scikit-learn.

El método SelectKBest es una técnica de selección de características que se utiliza para identificar las K características más relevantes dentro de un conjunto de datos. Este método evalúa cada característica individualmente y asigna una puntuación a cada una de ellas. Cuanto mayor sea la puntuación, más relevante se considera la característica para el objetivo del análisis.

Tras el proceso de filtrado, se determinó que las características más influyentes en la predicción de sarcopenia son:

1. isarcIndex
2. BMI
3. Thigh Thickness
4. Sex
5. no_risk_homocygotes
6. risk_heterocygotes
7. risk_homocygotes
8. risk_sarcopenia
9. sarcopenia_clasification

ÁRBOLES DE DECISIÓN

El árbol de decisión es un modelo de predicción utilizado en el ámbito del aprendizaje automático que se utiliza para resolver problemas de clasificación o regresión. En nuestro caso, hemos utilizado un árbol de decisión para predecir la probabilidad de sarcopenia en pacientes.

Un árbol de decisión está compuesto por nodos y ramas. Cada nodo representa una característica o atributo del conjunto de datos, y cada rama representa una posible salida o resultado basado en esa característica. Los nodos se dividen en nodos de decisión y nodos hoja.

Los nodos de decisión toman decisiones basadas en una regla o condición específica. Por ejemplo, si la característica "isarcIndex" es mayor que un umbral determinado, el árbol seguirá por una rama determinada. Si es menor o igual al umbral, seguirá por otra rama. De esta manera, el árbol de decisión va dividiendo el conjunto de datos en función de las características más relevantes y sus umbrales.

Los nodos hoja representan las clasificaciones o resultados finales. En nuestro caso, un nodo hoja puede indicar si un paciente tiene alta probabilidad de sarcopenia o baja probabilidad de sarcopenia, por ejemplo. Estos nodos hoja son el resultado de las decisiones tomadas a lo largo del árbol.

En nuestro enfoque para predecir la probabilidad de sarcopenia, hemos utilizado la implementación del árbol de decisión provista por la biblioteca scikit-learn de Python, específicamente el método DecisionTreeClassifier.

Actuació cofinançada per la Unió Europea a través del Programa Operatiu del Fons Europeu de Desenvolupament Regional (FEDER) de la Comunitat Valenciana 2014-2020

Además de la construcción y predicción del árbol de decisión, scikit-learn también proporciona funciones para evaluar el rendimiento del modelo, como la precisión, el puntaje F1 y la matriz de confusión.

3. RESULTADOS

PROCESO DE ENTRENAMIENTO

El proceso de entrenamiento y prueba se lleva a cabo de la siguiente manera:

1. División de los datos: Inicialmente, separamos nuestro conjunto de datos en conjunto de entrenamiento y conjunto de prueba. Una proporción común es utilizar el 80% de los datos para entrenamiento y el 20% restante para prueba, aunque esto puede variar dependiendo del tamaño y la naturaleza de los datos.

2. Entrenamiento del árbol de decisión: Utilizamos el conjunto de entrenamiento para entrenar el árbol de decisión. Esto implica proporcionar al modelo las características de entrada y las clasificaciones correspondientes. El árbol de decisión ajusta sus parámetros internos para encontrar las reglas y umbrales óptimos que permitan clasificar los datos de manera efectiva.

3. Evaluación del rendimiento: Una vez entrenado el árbol de decisión, utilizamos el conjunto de prueba para realizar predicciones y evaluar su rendimiento. Comparamos las predicciones del modelo con las clasificaciones reales conocidas en el conjunto de prueba. Se calculan métricas correspondientes para medir el rendimiento y la capacidad de generalización del modelo.

VISUALIZACIÓN DE RESULTADOS

Para comprender mejor las decisiones tomadas por el árbol de decisión, hemos utilizado la herramienta Graphviz. Esta herramienta nos permite la visualización del árbol de decisión, resultando muy útil para comprender la estructura y las reglas de clasificación generadas por el modelo. Graphviz es una biblioteca de código abierto que permite crear diagramas gráficos de estructuras de datos, como árboles de decisión en nuestro caso.

En el siguiente ejemplo se ha entrenado un árbol de decisión seleccionando las características ‘isarcIndex’, ‘BMI’, ‘Tigh Thickness' y ‘Sex’, con el objetivo de predecir si un paciente padece o no sarcopenia. Se ha conseguido clasificar con éxito el 80% de los casos de prueba.

El resultado es un gráfico que muestra la estructura del árbol de decisión, incluyendo los nodos de decisión, los umbrales y las ramas que conducen a los nodos hoja. La visualización del árbol permite comprender cómo se toman las decisiones basadas en las características seleccionadas, y cómo se clasifican los datos en las diferentes ramas del árbol.

Actuació cofinançada per la Unió Europea a través del Programa Operatiu del Fons Europeu de Desenvolupament Regional (FEDER) de la Comunitat Valenciana 2014-2020

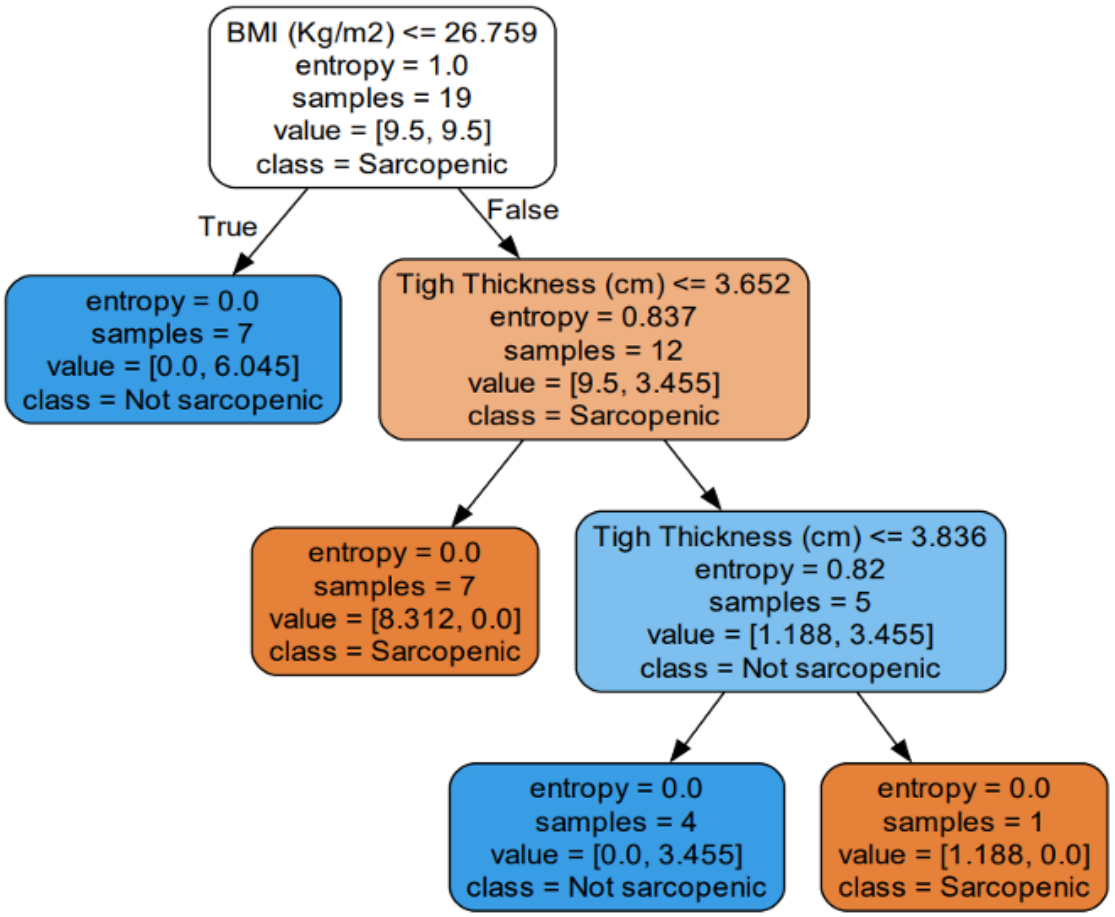


Figura 1. Resultados del árbol de decisión con umbrales incluidos. Los valores anaranjados reflejan características de sarcopénicas para cada biomarcador utilizado en el ejemplo.

Programa	**Proyectos Estratégicos en Cooperación**
Título del Proyecto	**iSARC-GENETICS** PLATAFORMA INTELIGENTE PARA LA DETECCIÓN TEMPRANA, AYUDA AL DIAGNÓSTICO Y SEGUIMIENTO DE LA SARCOPENIA
Entregable	**E.4.1** Informe sobre el diseño del biobanco de datos clínicos y biomarcadores
Responsable	**CoMMLab, Dawako, Sabartech**
Fecha	**Mes 15: junio 2021**

1. INTRODUCCIÓN

El objetivo fundamental de las herramientas desarrolladas en el proyecto iSarc-Genetics es permitir la explotación de algoritmos avanzados de cálculo, con el fin de extraer biomarcadores de pacientes sarcopénicos a diferentes escalas.

Para ello se requiere de una base de datos de pacientes sobre la que se pueda desarrollar y testar todos los algoritmos implementados, tanto a escala subcelular (genómica), como a escala de tejido (imágenes de ultrasonidos). Por lo tanto, uno de los objetivos del proyecto es el diseño de una base de datos que permita almacenar, datos de imágenes de pacientes, junto a todos sus descriptores y biomarcadores asociados, y una base de datos genómica, que albergue los marcadores genéticos, genes y otros datos de relevancia. Ambas bases de datos deben disponer de una API, que permita la extracción de información para un determinado paciente, que además estará relacionada mediante un código único.

2. BASE DE DATOS Y API PARA IMAGEN MÉDICA

Para el almacenamiento de datos clínicos y sus biomarcadores, se ha implantado una base de datos relacional debido al uso del front-end orthanc para el almacenamiento y acceso a la base de datos médicos de ultrasonidos, que se almacena en formato estándar DICOM.

En esta base de datos se han definido los campos necesarios para estructurar la información por usuario (médico o clínica). Para cada usuario, se dispone de una estructura basada en pacientes y estudios, que a su vez puede estar formados por instancias. Cada instancia, representa una adquisición realizada con el ecógrafo y por tanto dispone de todos los datos asociados a la adquisición.

Desde el front-end utilizado como interfaz, la API se comunica con todos los interfaces web para permitir la subida y almacenamiento de datos, así como todas las funciones de anonimización, y procesado de la información para la obtención de los biomarcadores clínicos.

La base de datos se ha ampliado con tablas adicionales específicas para el almacenamiento de los biomarcadores clínicos. Cada entrada está asociada a una imagen captada con el ecógrafo, donde se pueden almacenar biomarcadores MAP (Muscle Architecture Parameters) de textura, e información intermedia de procesado para acelerar los cálculos o poder reproducirlos de nuevo de una forma fácil.

El almacenamiento de los datos tiene en cuenta el formato utilizado para la comunicación con el front-end, que en este caso es JSON. De este modo es fácil actualizar y recuperar datos que pueden ser usados para tareas de visualización y de validación.

3. BASE DE DATOS Y API PARA DATOS GENÓMICOS

Sabartech ha desarrollado un sistema de bases de datos relacionales, una de ellas para almacenar los datos que nutrirán al algoritmo genético y contiene los campos descritos en el documento E.1.2, y otra con la GENE-SCORE final y la significación clínica del mismo para cada paciente. Así mismo cada resultado irá codificado y anonimizado de acuerdo con la normativa de Protección de Datos.

Ambas bases de datos estarán alojadas en los servidores de producción y preproducción de Sabartech, donde se aloja a su vez la Pipeline Genómica y los motores responsables de calcular la predisposición de riesgo genético asociado a sarcopenia para cada individuo (documento E.3.5). Finalmente, mediante un sistema de comunicación Sabartech devolverá el resultado solicitado en un fichero en formato JSON, junto con otra información de interés (descrito con mayor detalle en el E.4.2) para combinar la parte de imagen y la genética.

Actuació cofinançada per la Unió Europea a través del Programa Operatiu del Fons Europeu de Desenvolupament Regional (FEDER) de la Comunitat Valenciana 2014-2020

Programa	**Proyectos Estratégicos en Cooperación**
Título del Proyecto	**iSARC-GENETICS** PLATAFORMA INTELIGENTE PARA LA DETECCIÓN TEMPRANA, AYUDA AL DIAGNÓSTICO Y SEGUIMIENTO DE LA SARCOPENIA
Entregable	E4.3. Desarrollo software de la plataforma basada en algoritmos de IA en la nube. (M26)
Responsable	**CoMMlab, Sabartech**
Fecha	**Abril 2023**

1. INTRODUCCIÓN

T4.3: Desarrollo de plataforma basada en algoritmos de IA en la nube

INTRODUCCIÓN AL SISTEMA PROPUESTO

El objetivo de este paquete de trabajo es el desarrollo e integración de la Plataforma de análisis y cuantificación de sarcopenia en la nube.

Siguiendo las directrices definidas en el proyecto se ha desplegado una aplicación basada en el sistema cloud de Microsoft Azure y siguiendo una arquitectura de microservicios empleando el servicio de Kubernetes.Según se especificó en el Entregable 4.2, la arquitectura basada en microservicios ha permitido dividir los diferentes módulos de software, tanto del front-end como del back-end, en contenedores independientes en muchos casos.

Por consiguiente, conociendo los requisitos de la aplicación y considerando las necesidades del equipo de desarrollo para esta etapa del proyecto se diseña una arquitectura de microservicios, que implementa la tecnología de los contendores para el desarrollo y el despliegue de la solución en un entorno de nube privada o pública.

VALIDACIÓN DEL SISTEMA

En resumen, las pruebas de rendimiento realizadas demostraron que el servicio es capaz de soportar la creación de pacientes por parte de múltiples usuarios concurrentes.

Resumen de las pruebas de rendimiento:

En la prueba 1, el servicio fue capaz de validar satisfactoriamente las imágenes de los 10 usuarios concurrentes sin errores significativos. Los tiempos promedio de login, validación de pacientes y ejecución total de pruebas fueron aceptables y el uso de recursos se mantuvo en niveles razonables.

En la prueba 2, el servicio fue capaz de validar la mayoría de los estudios de los 20 usuarios concurrentes, aunque dos usuarios presentaron errores durante la validación. Los tiempos promedio de login, validación de pacientes y ejecución total de pruebas fueron aceptables y el uso de recursos se mantuvo en niveles razonables.

En la prueba 3, el servicio presentó limitaciones significativas en términos de carga de trabajo. A pesar de que se lograron validar los estudios de 16 usuarios concurrentes, cerca de la mitad de los usuarios presentaron errores durante la validación y tres usuarios no pudieron realizar el login. El tiempo promedio de validación de pacientes y ejecución total de pruebas fueron aceptables, pero el tiempo promedio de login fue ligeramente mayor que en las pruebas anteriores. El uso de recursos alcanzó niveles preocupantes, con un pico máximo de uso de CPU del 90.5% y el pico más bajo de memoria RAM disponible de sólo 169 MB.

En general, el servicio demostró tener limitaciones en términos de carga de trabajo en algunos escenarios, y se requiere una mejora en la capacidad de manejar errores y una mejor gestión de los recursos para garantizar una experiencia satisfactoria para todos los usuarios.

Consideraciones para la prueba:
Para el desarrollo de esta prueba se consideraron los siguiente imagines médicas:

- *Right Transversal Biceps Brachii*
- *Left Transversal Biceps Brachii*
- *Right Transversal Biceps Femoris*
- *Left Transversal Biceps Femoris*
- *Right Transversal Biceps Gastrocnemius*
- *Left Transversal Biceps Gastrocnemius*

Se han seleccionado únicamente imágenes de tipo ”Transversal” debido a que, en el caso de haber más de una imagen transversal o longitudinal, la plataforma ubica aleatoriamente la posición en el área de validación. Al no poder distinguir entre ambos tipos de estudios por medio del Bot, se decidió utilizar solamente imágenes transversales. Además, el número de imágenes seleccionadas ha sido suficiente para detectar errores en la plataforma.

La prueba se configuró con una carga de trabajo de 20 unidades de carga (u/c). Cada usuario se ejecutó durante un tiempo aleatorio entre 1 y 120 segundos para evitar sobrecargas.

Conclusiones: Los resultados indican que el servicio es capaz de validar los estudios de 18 u/c de 20.
En cuanto a los tiempos empleados por los usuarios que completaron satisfactoriamente la prueba podemos ver que el tiempo medio para hacer login es de 0.085 minutos.

Actuació cofinançada per la Unió Europea a través del Programa Operatiu del Fons Europeu de Desenvolupament Regional (FEDER) de la Comunitat Valenciana 2014-2020

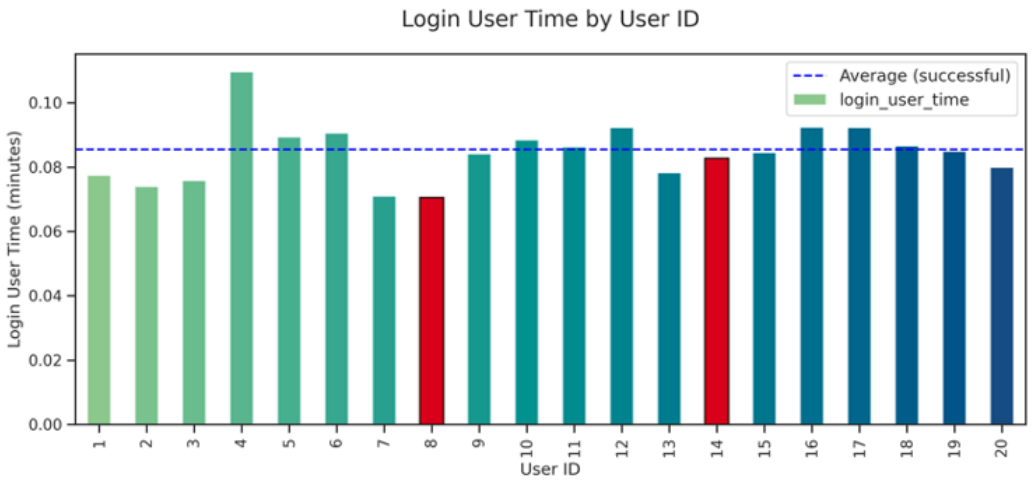


***Figura 1.** Tiempos para hacer login con 20 u/c.*

Programa	**Proyectos Estratégicos en Cooperación**
Título del Proyecto	**iSARC-GENETICS** PLATAFORMA INTELIGENTE PARA LA DETECCIÓN TEMPRANA, AYUDA AL DIAGNÓSTICO Y SEGUIMIENTO DE LA SARCOPENIA
Entregable	**E.4.4** Informe sobre la evaluación de aseguramiento de la calidad y asuntos regulatorios de la plataforma
Responsable	**Dawako, Sabartech**
Fecha	**Mes 28: julio 2023**

1. INTRODUCCIÓN

DAWAKO y SABARTECH deciden comenzar a comercializar la plataforma iSARC-GENETICS en el mercado europeo. Para ello hay que cumplir con el Reglamento de productos sanitarios (EU) 2017/745.

En primer lugar, procedemos a realizar la descripción técnica de la plataforma iSARC-GENETICS para poder clasificar el producto dentro del Reglamento. Así, DAWAKO, como compañía de dispositivos médicos, debe cumplir con las regulaciones de la UE (marcado CE) y en los EE. UU. (FDA), donde planea su futura comercialización. El objetivo es definir junto con una consultora especializada el plan detallado de certificación de la plataforma para la obtención del marcado CE (Organismo Notificado).

Por su lado, SABARTECH ya ha presentado la solicitud para obtener la licencia de funcionamiento de instalaciones de productos sanitarios que permita comercializar su tecnología, requisito indispensable para la obtención del marcado CE en territorio nacional. A su vez, está trabajando en la obtención del marcado CE a nivel europeo, cumpliendo así la normativa para la comercialización de productos sanitarios.

2. DESCRIPCIÓN TÉCNICA

La plataforma iSARC-GENETICS es una plataforma basada en la nube, en la inteligencia artificial (IA), de bioimágenes y genética, diseñada para trabajar con equipos de imágenes médicas (ultrasonido inalámbrico) y datos genéticos. La plataforma está diseñada para permitir el análisis cuantitativo de imágenes de ultrasonido del sistema musculoesquelético (MSK) y datos genómicos para brindar apoyo a la decisión clínica (CDSS) para diagnosticar la probabilidad de sarcopenia.

El software para la visualización de las imágenes médicas que cumple con el estándar DICOM y se ejecuta en cualquier ordenador, se basa en un procedimiento semiautomático de análisis cuantitativo de imágenes médicas remotas y un análisis genómico de los datos del sujeto a partir de una muestra tomada con un hisopo. La plataforma crea informes estructurados de forma automática para ser descargados directamente desde la aplicación.

Las características del software son las siguientes:

- Almacenamiento y visualización de imágenes DICOM
- Análisis de imágenes de ultrasonido (UI) y extracción de biomarcadores de imágenes de MSKU
- Análisis de datos genómicos y extracción de biomarcadores genéticos
- Herramientas de análisis y manipulación de datos
- Reporte estructurado de biomarcadores de imágenes MSKU y biomarcadores genómicos

3. REGLA CLASIFICACIÓN

Regla de clasificación aplicable según el Reglamento (UE) 2017/745 sobre Dispositivos Médicos:

Regla 11

El software destinado a proporcionar información que se utiliza para tomar decisiones con fines diagnósticos o terapéuticos se clasifica en la clase II a, excepto si dichas decisiones tienen un impacto que puede causar:

- La muerte o el deterioro irreversible del estado de salud de una persona, en cuyo caso se incluirá en la clase III; o
- Un deterioro grave del estado de salud de una persona o una intervención quirúrgica, en cuyo caso se clasifica en la clase II b.

El software destinado a monitorizar procesos fisiológicos se clasifica como clase II a, excepto si está destinado a monitorizar parámetros fisiológicos vitales, donde la naturaleza de las variaciones de esos parámetros es tal que podría resultar en un peligro inmediato para el paciente, en cuyo caso es clasificado como clase II b.

Todo el resto del software se clasifica como clase I.

Razón fundamental

La plataforma iSARC-GENETICS es un software destinado a proporcionar información que se utiliza para tomar decisiones con fines diagnósticos o terapéuticos y no puede causar:

- La muerte o un deterioro irreversible del estado de salud de una persona o,
- Un deterioro grave del estado de salud de una persona o una intervención quirúrgica.

La plataforma iSARC-GENETICS no está diseñada para monitorizar procesos fisiológicos.
Por lo tanto, la plataforma iSARC-GENETICS se clasifica como un producto sanitario **clase II a.**

Actuació cofinançada per la Unió Europea a través del Programa Operatiu del Fons Europeu de Desenvolupament Regional (FEDER) de la Comunitat Valenciana 2014-2020

Programa	**Proyectos Estratégicos en Cooperación**
Título del Proyecto	**iSARC-GENETICS** PLATAFORMA INTELIGENTE PARA LA DETECCIÓN TEMPRANA, AYUDA AL DIAGNÓSTICO Y SEGUIMIENTO DE LA SARCOPENIA
Entregable	E7.3. Validación del Algoritmo Genético (M30)
Responsable	**Sabartech**
Fecha	**Agosto 2023**

1. INTRODUCCIÓN

T7.5: Validación del diseño y sistema funcional final

Con los datos provenientes de los pacientes analizados en la Clínica Dawako se ha podido validar el algoritmo genético. Se ha validado con datos de 11 sarcopénicos, de los cuales 2 presentan sarcopenia severa, y 23 controles. Adicionalmente se han considerado los resultados del cribado genético de 940 pacientes para estudiar la distribución del riesgo genético.

Al evaluar cada uno de los bloques que conforman el riesgo genético: Alto Impacto, Bajo Impacto e Impacto científico, acorde a lo descrito en la fórmula definida en el Entregable E.3.5, se ha determinado que las variantes que predicen adecuadamente el riesgo a desarrollar Sarcopenia son las extraídas de los estudios GWAS (*Genome-wide association study)* procedentes del bloque de Bajo Impacto y las procedentes del bloque de Impacto Científico.

En el caso del bloque de Impacto Científico, gracias a la validación se ha podido determinar que este bloque ha sido capaz de predecir adecuadamente el riesgo después de haber sido realizada una limpieza de variantes en el mismo, habiéndo observado que las que mejor predicen el riesgo son las obtenidas mediante la herramienta de minería de texto.

Actuació cofinançada per la Unió Europea a través del Programa Operatiu del Fons Europeu de Desenvolupament Regional (FEDER) de la Comunitat Valenciana 2014-2020

Tanto en el bloque de Impacto Científico como en el de Bajo Impacto se ha evaluado el tipo de modelo de dominancia que se ha de suponer para predecir el riesgo con estas variantes, existiendo tres posibilidades:

- Un modelo dominante, donde es indiferente presentar el heterocigoto o el homocigoto de riesgo, ya que ambos incrementan el riesgo.
- Un modelo aditivo, en el cual el homocigoto de riesgo duplica el riesgo del heterocigoto.
- Un modelo recesivo, en el que únicamente incrementa el riesgo el homocigoto de riesgo.

Tras realizar diversas pruebas el modelo que mejor realiza la predicción es el modelo dominante.

Además de para realizar las correcciones realizadas anteriormente, la validación ha servido para determinar que el IMC es bastante importante en la predicción del riesgo de sarcopenia, lo que ha permitido evaluar y ajustar el componente multiescala del algoritmo.

No obstante, el riesgo genético de los pacientes de la Clínica Dawako es siempre moderado o alto. Dado que son muestras seleccionadas por presentar riesgo a sarcopenia u otras condiciones asociadas a la fragilidad. Es por ello que es necesario la inclusión de muestras heterogéneas, procedentes del cribado genético, donde se observa una mayor variabilidad genética entre pacientes. Para saber más detalles de los resultados, ver Entregable 2.1.