



Financiado por
la Unión Europea
NextGenerationEU

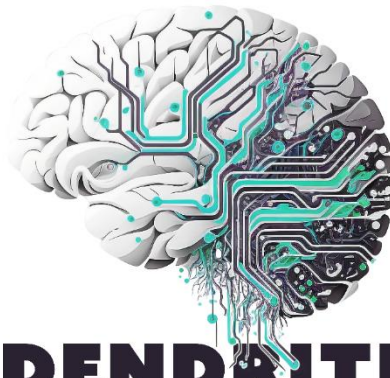


Plan de
Recuperación,
Transformación
y Resiliencia



Plan de Gestión de Datos (PGD)

Medicina Personalizada (MedPer) en la detección precoz del Deterioro Cognitivo (DC) preclínico. Desarrollo de un modelo predictivo de riesgo.



DENDRITE

meDicina pErsoNalizada en la DetECCIÓN
pRecoz del deterloro cogniTivo prEclínico

| PROYECTO | |
|----------------------|---|
| Número de referencia | [PMP22/00084] |
| Acrónimo | [DENDRITE] |
| Financiador | Instituto de Salud Carlos III |
| Convocatoria | Proyectos de Investigación de Medicina Personalizada (PMP). Convocatoria 2022 |
| Duración | 01/01/2023 – 31/12/2025 |

| PLAN DE GESTIÓN DE DATOS | |
|----------------------------|--------------|
| Fecha creación | [2024/05/08] |
| Versión actual | [PGD_v2] |
| Fecha última actualización | 2025/08/25 |

CONTROL DE CAMBIOS

| Revisión | Fecha | Descripción | Página |
|----------|------------|--|--------|
| 1.0 | 2024/05/08 | Publicación de la versión inicial del Plan de Gestión de Datos | Todas |
| 2.0 | 2025/08/25 | Adaptación al formato Horizonte Europa | Todas |

ACRÓNIMOS Y ABREVIACIONES

AP: Atención Primaria

CC: Coordinador clínico

CGD: Comité de Gestión de Datos

CIBER: Centro de Investigación biomédica en Red

CRD-e: Cuaderno de Recogida de Datos electrónico

CSV: Valores Separados por Comas (Comma-Separated Values)

DC: Deterioro Cognitivo

DCMI: Dublin Core Metadata Initiative

DOI: Identificador de Objeto Digital (Digital Object Identifier)

EGA: Archivo Europeo de Genomas (European Genome-phenome Archive)

EI: Equipo Investigador

GCA: Grupo Clínico Asociado

HCE: Historia Clínica Electrónica

IA: Inteligencia Artificial

IBSAL: Instituto de Investigación Biomédica de Salamanca

ISCI: Instituto de Salud Carlos III

IP: Investigador Principal

MMSE: Minimental State Examination

MOCA: Montreal Cognitive Assessment

OA: Acceso Abierto (Open Access)

PI: Propiedad Intelectual

PGD: Plan de Gestión de Datos

T@M: Test de Alteración de Memoria

UC3M: Universidad Carlos III de Madrid

UCM: Universidad Complutense de Madrid

UVIGO: Universidad de Vigo

Índice de contenido

| | | |
|-------|--|----|
| 1. | Resumen de los datos | 5 |
| 1.1 | Propósito de la generación de datos y su relación con los objetivos del proyecto | 5 |
| 1.2 | Tipo, formato y origen de los datos generados y recopilados | 5 |
| 1.3 | Reutilización de los datos..... | 7 |
| 1.4 | Tamaño esperado de los datos | 7 |
| 1.5 | Utilidad de los datos..... | 8 |
| 2. | Fair data | 8 |
| 2.1 | Localización de los datos, incluyendo los metadatos (Findable Data)..... | 8 |
| 2.1.1 | Identificadores persistentes y únicos..... | 8 |
| 2.1.2 | Suministro de metadatos | 8 |
| 2.1.3 | Estándares para la creación de metadatos | 8 |
| 2.1.4 | Palabras clave | 8 |
| 2.1.5 | Recolección..... | 8 |
| 2.2 | Accesibilidad de los datos (Accesible Data) | 9 |
| 2.2.1 | Repositorios..... | 9 |
| 2.2.2 | Datos..... | 9 |
| 2.2.3 | Metadatos | 11 |
| 2.3 | Datos interoperables (Interoperable Data)..... | 11 |
| 2.3.1 | Vocabularios, estándares o metodologías para facilitar la interoperabilidad | 11 |
| 2.4 | Reutilización de los datos (Reusable Data) | 11 |
| 2.4.1 | Garantizar la reutilización..... | 11 |
| 2.4.2 | Tipo de licencia..... | 11 |
| 2.4.3 | Uso por terceras personas de los datos tras su finalización | 11 |
| 2.4.4 | Descripción de los procesos que aseguren la calidad de los datos | 12 |
| 2.4.5 | Período de tiempo durante el cual los datos permanecerán reutilizables | 12 |
| 3. | Asignación de recursos..... | 12 |
| 3.1 | Coste destinado a adaptar los datos a los principios FAIR y a la preservación a largo plazo del dataset..... | 12 |
| 3.2 | Responsabilidades de la gestión de datos | 12 |
| 4. | Seguridad | 13 |
| 5. | Aspectos éticos..... | 13 |
| 6. | Otros aspectos | 14 |
| 6.1 | Referencias..... | 14 |

1. Resumen de los datos

1.1 Propósito de la generación de datos y su relación con los objetivos del proyecto

El objetivo principal del proyecto MedPer_DC es obtener un modelo predictivo para la detección preclínica del deterioro cognitivo en población de 55-70 años de edad, permitiendo así la implantación efectiva de la Medicina Personalizada de Precisión en Atención Primaria. El estudio involucra la participación de 6 Comunidades Autónomas y 7 centros de Atención Primaria en diferentes ubicaciones geográficas. Es una cohorte observacional, multicéntrica y no intervencional con métodos mixtos convergentes. Se desarrollará combinando una fase cuantitativa observacional y prospectiva con una fase cualitativa que involucrará grupos de discusión.

Durante el estudio se recogerán datos multi-ómicos (moleculares, proteómicos, genómicos), asistenciales, sociales, ambientales, de voz y de conducta (provenientes de monitorización activa y pasiva) junto con información clínica complementaria. Todos estos datos se analizarán mediante diversas técnicas de análisis basadas en Inteligencia Artificial y en modelos estadísticos para sintetizar un perfil clínico, genético y molecular, de comportamiento y voz, y determinar su potencial para converger en un modelo predictivo global de DC.

1.2 Tipo, formato y origen de los datos generados y recopilados

El proyecto generará y recopilará distintos tipos de conjuntos de datos, organizados en las siguientes categorías:

1.2.1 Datos clínicos procedentes de la Historia Clínica Electrónica (HCE)

| | |
|---------------------------|---|
| Título del Dataset | Datos clínicos procedentes de la Historia Clínica Electrónica |
| Acrónimo | HCEDTS |
| Descripción | Diagnósticos activos y prescripción farmacológica activa. |
| Tipo de datos | Datos observacionales secundarios |
| Formato | CSV |
| Origen | HCE de los participantes en el estudio. |

1.2.2 Datos clínicos y sociodemográficos recogidos en el punto de asistencia

| | |
|---------------------------|--|
| Título del Dataset | Datos clínicos provenientes de monitorización en el punto de asistencia |
| Acrónimo | CLINDTS |
| Descripción | Cuestionarios sobre factores sociales y hábitos de vida, pruebas cognitivas (MMSE, MoCA, T@M, MiniCog, FotoTest) y pruebas funcionales (velocidad de la marcha). |
| Tipo de datos | Datos primarios observacionales |
| Formato | CSV |
| Origen | Registro directo de parámetros clínicos en el punto de asistencia a los participantes del estudio. |

1.2.3 Datos de voz y contenido del habla

| | |
|---------------------------|--|
| Título del Dataset | Datos de voz y contenido del habla |
| Acrónimo | VOZDTS |
| Descripción | Grabaciones de audio en pruebas guiadas y variables acústicas/lingüísticas derivadas mediante algoritmos de IA. |
| Tipo de datos | Datos primarios observacionales |
| Formato | WAV (audio bruto) y CSV (datos derivados del análisis y distintos datos a nivel de palabra). |
| Origen | Datos recogidos a través de la aplicación EB2 (https://eb2.tech/), mediante grabaciones de voz realizadas por los participantes. |

1.2.4 Datos de comportamiento e interacción social

| | |
|---------------------------|---|
| Título del Dataset | Datos de comportamiento e interacción social |
| Acrónimo | BEHAVEDTS |
| Descripción | Cuestionarios sobre factores sociales (red de apoyo social, EQ-5D-5L), habilidades de interacción social evaluadas mediante escenarios teóricos de juegos, monitorización pasiva de patrones de comportamiento mediante aplicación móvil (pasos, uso de apps, sueño). |
| Tipo de datos | Datos primarios observacionales |
| Formato | CSV |
| Origen | Los datos serán registrados por los propios participantes del estudio en la aplicación EB2 (https://eb2.tech/). |

1.2.5 Datos de percepción social de riesgo de DC y determinantes sociales de la salud

| | |
|---------------------------|---|
| Título del Dataset | Datos de percepción social de riesgo de DC y determinantes sociales de la salud |
| Acrónimo | CUALIDTS |
| Descripción | Investigación cualitativa basada en grupos de discusión estratificados social y territorialmente, con el fin de explorar percepciones colectivas sobre el deterioro cognitivo y los determinantes sociales de la salud. |
| Tipo de datos | Datos primarios observacionales cualitativos |
| Formato | Los discursos se recogerán en audio (WAV), se transcribirán y se almacenarán en formato TXT/DOCX. |
| Origen | Grupos de discusión estratificados social y territorialmente. |

1.2.6 Datos multi-ómicos

| | |
|---------------------------|--|
| Título del Dataset | Datos multi-ómicos (moleculares, proteómicos, genómicos) |
| Acrónimo | OMICDTS |
| Descripción | Conjunto de datos experimentales obtenidos a partir de análisis genómicos (GWAS, SNPs, puntuaciones de riesgo poligénico), proteómicos (plasma, PBMCs, organoides cerebrales) y biosensores (plataformas electroquímicas, SIMOA). |
| Tipo de datos | Datos primarios experimentales |
| Formato | <p>Genómica: Archivos raw en formato FASTQ/BAM. Archivos procesados en VCF (variantes genéticas/SNPs asociados a DC).</p> <p>Proteómica: Archivos raw en formato RAW propietario de los espectrómetros de masas. Los datos procesados de proteómica cuantitativa y datos de proteómica dirigida se preservarán en CSV a nivel de datos cuantitativos y de concentración.</p> <p>Biosensores (SIMOA y electroquímicos): datos derivados en formato CSV (valores cuantitativos de concentración de marcadores proteómicos y genómicos).</p> |
| Origen | Análisis realizados en muestras biológicas de los participantes (plasma, PBMCs), mediante técnicas de genómica, proteómica y plataformas de biosensado. |

1.3 Reutilización de los datos

En este proyecto no se reutilizarán conjuntos de datos públicos o previamente generados por terceros; únicamente se utilizarán datos clínicos extraídos de la HCE de los participantes. La responsabilidad de gestionar y custodiar estos datos recae en las instituciones sanitarias donde se originan. La cesión de estos datos se llevará a cabo mediante procedimientos legalmente establecidos, en el marco del consentimiento informado, el cual especifica de forma clara los objetivos para los cuales se solicita el uso de los datos.

Los datos primarios generados durante la ejecución del proyecto serán procesados, estructurados y anonimizados para crear nuevos conjuntos de datos reutilizables, que se pondrán a disposición de la comunidad científica mediante su depósito en repositorios de acceso abierto.

1.4 Tamaño esperado de los datos

El tamaño total de los datos generados dependerá del volumen final de participantes y de su posterior procesamiento. Se estima que los datos más voluminosos corresponderán a las grabaciones de voz y a los datos multi-ómicos. Se prevé que el tamaño total del conjunto de datos no exceda los 2 TB, con archivos individuales que no superaren los 1 GB.

1.5 Utilidad de los datos

Los datos se utilizarán para generar modelos predictivos capaces de detectar en fase preclínica el deterioro cognitivo e implantar de forma efectiva la medicina personalizada en Atención Primaria. Por otro lado, los conjuntos de datos se compartirán dentro del consorcio como base de trabajo para elaborar publicaciones científicas asociadas a la investigación. Más allá del consorcio, los conjuntos de datos podrán ser utilizados por investigadores para la reproducibilidad de los resultados o como base para nuevas investigaciones en el ámbito del diagnóstico preclínico del DC.

Estos datos contribuirán al Espacio Europeo de Datos de Salud, a través de la plataforma IMPaCT, fomentando su reutilización.

2. Fair data

2.1 Localización de los datos, incluyendo los metadatos (Findable Data)

2.1.1 Identificadores persistentes y únicos

Se asignarán identificadores persistentes DOI o Handle (según el repositorio empleado) a todos los conjuntos de datos generados o compartidos en el marco del proyecto.

2.1.2 Suministro de metadatos

Todos los conjuntos de datos irán acompañados de metadatos generados en formato estandarizado y legible por máquina como Dublin Core, en función de los requisitos del repositorio. Adicionalmente, se incluirá un archivo README en formato de texto libre (.txt) con metadatos en lenguaje natural que complemente los metadatos estructurados. Se incluirán, como mínimo:

- **Metadatos administrativos:** título, IPs, instituciones participantes, acrónimo, código de referencia, entidad financiadora y periodo del proyecto.
- **Metadatos descriptivos:** objetivos, metodología, tamaño muestral, estructura de los datos y descripción de los campos.
- **Metadatos de citación:** título, autores, año, resumen, palabras clave e identificador persistente.

2.1.3 Estándares para la creación de metadatos

Los metadatos se crearán siguiendo el esquema Dublin Core, reconocido internacionalmente, lo que garantiza su interoperabilidad y cumplimiento con los principios FAIR.

2.1.4 Palabras clave

Los metadatos incluirán palabras clave normalizadas extraídas de vocabularios controlados (MeSH/DeCS). Se deberán mencionar, como mínimo, las siguientes palabras clave: deterioro cognitivo, diagnóstico precoz y medicina de precisión (*cognitive Dysfunction, early diagnosis, precision medicine*).

2.1.5 Recolección

Los metadatos serán recolectables mediante el protocolo OAI-PMH, de forma que puedan ser indexados en plataformas federadas como OpenAIRE y RECOLECTA.

2.2 Accesibilidad de los datos (Accessible Data)

2.2.1 Repositorios

Dónde se depositarán los datos

Los conjuntos de datos generados durante el proyecto se depositarán en repositorios confiables y de acceso abierto, garantizando su conservación y preservación a largo plazo. Se priorizará el uso de:

- **Repositorios temáticos europeos**, en coordinación con IMPaCT Data e IMPaCT Genómica:
 - Los datos genómicos se depositarán en el [Archivo Europeo de Genomas \(EGA\)](#), acompañados de toda la información necesaria para su interpretación, incluyendo datos clínicos, protocolos de laboratorio y herramientas de análisis.
 - Los datos de proteómica se depositarán en [ProteomeXchange](#), una red internacional que garantiza el acceso y reutilización de datos de proteómica en formatos estandarizados.
- **Repositorios institucionales federados con RECOLECTA**: cada entidad participante será responsable de autoarchivar la documentación derivada del proyecto (informes, entregables, artículos, comunicaciones científicas, etc.) en su repositorio institucional, asegurando su indexación en la plataforma [RECOLECTA](#), lo que facilita su descubrimiento a nivel nacional y europeo.
- **Repositorios generales** como Zenodo o plataformas europeas alineadas con OpenAIRE para asegurar una mayor visibilidad y acceso.

Cuando proceda, los datos generados serán integrados en la infraestructura de la Cohorte IMPaCT, lo que garantizará su descubrimiento y reutilización. Asimismo, las muestras biológicas sobrantes (plasma y PBMC) se conservarán en el Biobanco del Instituto de Salud Carlos III para su posible reutilización científica futura bajo condiciones controladas. Por último, se fomentará que las publicaciones científicas derivadas del proyecto se depositen, siempre que sea posible, en repositorios de acceso abierto como Open Research Europe.

Características de los repositorios seleccionados

Los repositorios seleccionados deberán cumplir con ciertos criterios esenciales para garantizar la gestión responsable y sostenible de la información:

- Ser fiable y sostenible a largo plazo.
- Contar con una política de acceso abierto.
- Soportar protocolo de recolección OAI-PMH para la interoperabilidad con recolectores como OpenAIRE y RECOLECTA.
- Permitir la asignación de identificadores persistentes.
- Incluir mecanismos de control de versiones.
- Aplicar políticas de preservación a largo plazo y formatos abiertos.

2.2.2 Datos

Condiciones de acceso

Los datos anonimizados y validados estarán disponibles en los repositorios especificados, en acceso abierto, siempre que la naturaleza de los datos lo permita. En

caso de que existan restricciones legales, éticas o de privacidad, los datos se compartirán bajo acceso controlado, previa solicitud motivada y evaluación por el Comité de Gestión de Datos (CGD).

Los datos en formato audio (.wav) no se compartirán por motivos de privacidad y seguridad. La anonimización no se considera una alternativa, ya que el tamaño de la muestra, especialmente en el caso de los grupos de discusión, permite extraer conclusiones sobre la identificación de los participantes. Se almacenarán en los servidores del ISCIII y serán destruidos en un plazo máximo de un año tras la finalización del proyecto.

Periodo de embargo a los datos

En cumplimiento con la normativa actual relativa a la actividad investigadora financiada mayoritariamente con fondos públicos, se requiere que tanto las publicaciones científicas como los conjuntos de datos asociados al proyecto sean depositados en repositorios de acceso abierto. Este depósito debe incluir la versión final aceptada de la publicación, junto con los datos asociados, y debe realizarse preferentemente al mismo tiempo que la publicación. En cualquier caso, el depósito no debe demorarse más de 12 meses después de la fecha de publicación para las áreas de ciencias y de 6 meses para las áreas de ciencias sociales y humanidades.

Acceso durante y tras la finalización del proyecto

Durante la ejecución del proyecto, los datos se almacenarán de forma confidencial y pseudoanonimizada en la plataforma REDCap. El acceso al sistema por parte del equipo investigador estará limitado por localización geográfica participante, y se realizará mediante identificación personal con claves intransferibles.

Una vez depurados, normalizados y anonimizados, los conjuntos de datos estarán disponibles en repositorios de acceso abierto o controlado. Paralelamente, los conjuntos de datos completos se conservarán en los servidores del Instituto de Salud Carlos III, como soporte a las instituciones involucradas, cumpliendo con sus políticas internas de seguridad y confidencialidad.

Restricciones de acceso a los datos

Durante la ejecución del proyecto, el equipo investigador solo podrá solicitar acceso a los datos en caso de encontrarse en situación de alta como investigador. Tras el alta de un nuevo investigador en el equipo investigador, se considera que sus derechos de autoría se inician pasados seis meses de su fecha de alta. Tras la baja de un investigador, se considerará que sus derechos de autoría prevalecen durante los 6 meses posteriores a su fecha de baja, perdiéndose a partir de ese periodo. Todos los investigadores con acceso a los datos deberán firmar el *Acuerdo sobre la confidencialidad y la propiedad intelectual* para que esto sea efectivo. Se restringe el acceso a los datos a las personas pertenecientes al Grupo Clínico Asociado (GCA).

Comité de acceso a datos

El acceso a los conjuntos de datos por parte de las entidades participantes será regulado por el Comité de Gestión de Datos, que será responsable de evaluar las propuestas de explotación científica de los datos, incluidas publicaciones, comunicaciones, trabajos

académicos o estudios secundarios. Las solicitudes deberán enviarse por correo electrónico a PerMedCD@isciii.es, adjuntando el formulario del *Acuerdo de confidencialidad y propiedad intelectual* (Anexo 4), debidamente cumplimentado.

2.2.3 Metadatos

Accesibilidad a los metadatos

Los metadatos estarán disponibles incluso si el conjunto de datos no puede ser compartido abiertamente. Se publicarán bajo licencia Creative Commons Zero (CC0) e incluirán información que permitirá a los usuarios identificar y localizar el conjunto de datos, como el DOI correspondiente, a través del cual se podrá acceder al conjunto de datos.

Disponibilidad de los metadatos

Los metadatos cumplirán con los principios FAIR y seguirán esquemas estándar reconocidos como Dublin Core. Se garantizará su disponibilidad a largo plazo, incluso si el dataset asociado deja de estar accesible.

2.3 Datos interoperables (Interoperable Data)

2.3.1 Vocabularios, estándares o metodologías para facilitar la interoperabilidad

Todos los conjuntos de datos generados seguirán estándares de datos y metadatos reconocidos:

- Terminologías normalizadas como SNOMED CT y LOINC.
- Palabras clave con vocabularios controlados MeSH/DeCS.
- Esquemas de metadatos Dublin Core (DCMI).
- Representación de fechas ISO 8601.
- Formatos abiertos como CSV, TXT, JSON, XML.

Cuando proceda, se aplicarán los estándares definidos por la infraestructura IMPaCT, facilitando así la integración con iniciativas nacionales y su futura conexión con el Espacio Europeo de Datos de Salud (EHDS).

2.4 Reutilización de los datos (Reusable Data)

2.4.1 Garantizar la reutilización

Cada conjunto de datos publicado incluirá un fichero README, metadatos estructurados según Dublin Core, referencias a publicaciones y, cuando sea posible, acceso a scripts de análisis para facilitar su reutilización.

2.4.2 Tipo de licencia

Los datos se publicarán bajo licencia CC-BY-NC-SA 4.0, que permite distribuir, remezclar, adaptar y crear a partir del material en cualquier medio o formato, únicamente con fines no comerciales, siempre que se dé atribución al creador y que las adaptaciones se compartan bajo los mismos términos.

2.4.3 Uso por terceras personas de los datos tras su finalización

Los datos estarán disponibles para su reutilización en acceso abierto siempre que la naturaleza de los datos lo permita.

2.4.4 Descripción de los procesos que aseguren la calidad de los datos

El CGD garantizará la evaluación de la calidad de los datos. Los conjuntos de datos no serán compartidos hasta superar este proceso de control de calidad.

2.4.5 Período de tiempo durante el cual los datos permanecerán reutilizables

Los datos publicados seguirán siendo accesibles y reutilizables hasta que los repositorios en los que hayan decidido depositarse dejen de dar soporte de calidad (en el caso de Zenodo, los conjuntos de datos pueden reutilizarse con una garantía mínima de 20 años).

3. Asignación de recursos

3.1 Coste destinado a adaptar los datos a los principios FAIR y a la preservación a largo plazo del dataset

No hay costes asociados a los mecanismos descritos para cumplir con los principios FAIR y preservar la base de datos a largo plazo.

3.2 Responsabilidades de la gestión de datos

La gestión de datos del proyecto se organiza a través de distintos niveles de responsabilidad y equipos definidos en cada centro participante.

- **Investigadores Principales:** grupo liderado por la Responsable Científica del proyecto. Asumen la responsabilidad de:
 - Asegurar la calidad de los datos a nivel global y local.
 - Coordinar el análisis de la base de datos global.
 - Discutir resultados y elaborar informes finales.
 - Coordinar la difusión científica de los datos del proyecto.
- **Grupo Coordinador:** colabora estrechamente con los IPs, aportando experiencia especializada en distintas áreas clave (biomarcadores, análisis de datos, etc.).
- **Coordinación Clínica:** responsable de la gestión de datos en cada centro. Supervisa los circuitos de recogida de datos, controla la correcta recogida y custodia del consentimiento informado y asume la custodia del listado de códigos de disociación que permite la reidentificación de los participantes, en caso de que sea necesario por motivos de seguridad y salud. El listado será destruido una vez finalice el proyecto.
- **Comité de Gestión de Datos:** es responsable de coordinar y supervisar la implantación, revisión y actualización del PGD. Controlan y aprueban las solicitudes de acceso y reutilización de datos.
- **Equipo Investigador:** investigadores dados de alta en el proyecto y que figuran como tal a efectos de la agencia financiadora. Son responsables de la recogida de datos y de la información y firma del consentimiento informado.
- **Personal contratado:** sus responsabilidades serán similares a la del equipo investigador.

- Grupo Clínico Asociado: personal clínico que actúa como apoyo al equipo investigador en tareas relacionadas con la recogida de datos.

4. Seguridad

En relación con el tratamiento, comunicación y cesión de datos personales, se cumplirá con lo dispuesto en la Ley Orgánica 3/2018 de Protección de Datos Personales y garantía de los derechos digitales, así como en el Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo (RGPD). Los participantes podrán ejercer sus derechos de acceso, rectificación, supresión, limitación del tratamiento, portabilidad y oposición, según lo previsto en dichas normativas.

Durante la ejecución del proyecto, los datos se almacenarán en la plataforma REDCap, alojada en un servidor seguro del Instituto de Salud Carlos III. Los participantes serán identificados mediante un código pseudonimizado, garantizando la privacidad de sus datos personales. Solo el equipo investigador local tendrá acceso a la clave que permite la reidentificación, en caso necesario, por motivos de salud o seguridad.

El acceso a los datos estará restringido por centro participante y se realizará con identificación personal mediante claves intransferibles. Todo el personal autorizado para el tratamiento de datos deberá firmar un "Compromiso de confidencialidad y aceptación de las normas de seguridad".

El equipo investigador tratará los datos personales exclusivamente con fines científicos y de interés público. Si los resultados del estudio fueran susceptibles de ser publicados en revistas científicas, en ningún caso se incluirán datos personales de los participantes.

Para garantizar la seguridad de los datos en los servidores del Instituto de Salud Carlos III, se seguirán las siguientes directrices:

- Copias de seguridad periódicas para prevenir la pérdida de datos.
- Cifrado de información sensible para asegurar su confidencialidad.
- Etiquetado estandarizado de archivos y control de versiones que aseguren la trazabilidad de las modificaciones realizadas.

5. Aspectos éticos

El proyecto aplicará una política ética rigurosa, en cumplimiento con los principios establecidos por la Declaración de Helsinki, el Código de Núremberg, las Guías de Buena Práctica Clínica y la normativa legal vigente, incluyendo la Ley 14/2007 de Investigación Biomédica y el Convenio Europeo sobre Derechos Humanos y Biomedicina.

Antes del inicio de las actividades, todos los procedimientos serán evaluados y autorizados por los comités de ética de la investigación correspondientes a nivel nacional y local.

Se garantizará la voluntariedad de los participantes mediante la obtención de su consentimiento informado, que incluirá información relativa a los objetivos del estudio, el tratamiento de datos personales, su posible reutilización y preservación a largo plazo,

así como el derecho a abandonar el estudio en cualquier momento. Dicho consentimiento será custodiado en formato físico bajo llave en su centro de Atención Primaria de referencia y en formato digital en los servidores del Instituto de Salud Carlos III.

6. Otros aspectos

El presente Plan de Gestión de Datos se depositará en abierto en el repositorio institucional pertinente.

6.1 Referencias

Wilkinson MD, Dumontier M, Aalbersberg IJJ, Appleton G, Axton M, Baak A, et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data*. 2016;3:160018. doi:10.1038/sdata.2016.18.