



Horizon Europe

Data Management Plan Template

Version 1.0
20 Junio 2026

HISTORY OF CHANGES		
Version	Publication date	Changes
1.0	20.06.2026	▪ Initial version

Action Number: PI24CIII-00047

Action Acronym: [DETEMOCAES]

Action title: [Distribución Espacio-Temporal de la Mortalidad por Cáncer en España por Edad y Sexo (DETEMOCAES)]

Date: [20/06/2026]

DMP version: [1.0]

1. Resumen de datos

El proyecto DETEMOCAES tiene como objetivo analizar la distribución espacio-temporal de la mortalidad por cáncer y otras causas de muerte en España, así como evaluar la influencia de factores socioeconómicos, demográficos, ambientales y territoriales sobre dicha distribución.

Para ello, el proyecto utiliza principalmente registros individuales de mortalidad facilitados por el Instituto Nacional de Estadística (INE), incluyendo información sobre causas de defunción, características demográficas y variables geográficas autorizadas para su uso en investigación científica. Asimismo, se emplean datos poblacionales, indicadores socioeconómicos y territoriales procedentes de organismos oficiales, información ambiental, cartografía digital y otras fuentes de información complementarias necesarias para los análisis epidemiológicos y espaciales previstos.

Los datos se utilizan para desarrollar análisis descriptivos, estudios espacio-temporales, modelización estadística avanzada, análisis bayesianos, estudios de desigualdades geográficas, elaboración de atlas de mortalidad y desarrollo de herramientas de visualización y difusión de resultados.

Los formatos manejados incluyen TXT, CSV, PX, RDS, XLSX, SHP, GeoPackage, GeoJSON, PDF y scripts desarrollados principalmente en lenguaje R. Como resultado del proyecto también se generan bases de datos derivadas, indicadores agregados, mapas, tablas de resultados, informes científicos, manuscritos y aplicaciones interactivas para la visualización de resultados.

Los datos individuales de mortalidad son suministrados por el Instituto Nacional de Estadística bajo condiciones específicas de acceso y uso para investigación científica, estando sujetos a la normativa de secreto estadístico y a las condiciones de seguridad establecidas por dicho organismo.

El volumen esperado de información se sitúa en el rango de decenas a cientos de gigabytes, incluyendo tanto los datos originales como los productos derivados generados durante el proyecto.

Los resultados obtenidos serán de utilidad para investigadores, administraciones públicas, profesionales de salud pública y responsables de planificación sanitaria interesados en el estudio de la mortalidad, las desigualdades territoriales y la epidemiología espacial del cáncer.

2. Datos FAIR

2.1. Hacer que los datos sean localizables, incluyendo los metadatos

Los datos se organizan mediante estructuras jerárquicas de directorios y sistemas de nomenclatura estandarizados que permiten identificar el origen de la información, el periodo temporal, el ámbito geográfico y la fase analítica correspondiente.

Los metadatos asociados incluyen información sobre la fuente de los datos, ámbito temporal, nivel geográfico, variables incluidas, procedimientos de transformación aplicados, versiones de los scripts utilizados y fecha de generación de los distintos productos derivados.

Los conjuntos de datos derivados generados por el proyecto estarán documentados mediante diccionarios de variables, documentación metodológica y descripciones técnicas que faciliten su interpretación y reutilización.

Cuando proceda, los resultados y productos científicos derivados del proyecto podrán depositarse en repositorios institucionales o científicos que proporcionen identificadores persistentes y mecanismos de localización estables.

2.2. Hacer que los datos sean accesibles

Los datos individuales de mortalidad facilitados por el Instituto Nacional de Estadística permanecen almacenados en infraestructuras informáticas seguras del Instituto de Salud Carlos III y únicamente son accesibles para el personal autorizado expresamente incluido en las solicitudes aprobadas por el INE.

El acceso a dichos datos se realiza exclusivamente con fines científicos dentro del proyecto autorizado y de acuerdo con las condiciones establecidas por el Instituto Nacional de Estadística y la normativa vigente sobre secreto estadístico.

Los resultados generados por el proyecto se difundirán exclusivamente de forma agregada y anonimizada, garantizando en todo momento la imposibilidad de identificar directa o indirectamente a personas individuales.

Los metadatos descriptivos, documentación metodológica, publicaciones científicas y productos derivados podrán hacerse accesibles a la comunidad científica de acuerdo con las políticas institucionales y las condiciones de uso de los datos originales.

2.3. Hacer que los datos sean interoperables

Los datos se gestionan utilizando formatos ampliamente aceptados por la comunidad científica y los sistemas de información geográfica, incluyendo TXT, CSV, XLSX, PX, RDS, SHP, GeoPackage y GeoJSON.

Las variables geográficas se referencian utilizando las divisiones territoriales oficiales vigentes en cada momento y las codificaciones estandarizadas utilizadas por los organismos oficiales productores de información.

Los análisis se desarrollan principalmente utilizando lenguaje R y software estadístico ampliamente utilizado en epidemiología espacial, análisis geográfico y modelización estadística avanzada.

Cuando procede, se mantienen referencias explícitas a las fuentes originales de información, facilitando la integración de resultados con otros estudios y proyectos de investigación.

2.4. Incrementar la reutilización de los datos

Todos los análisis desarrollados en el proyecto se documentan mediante scripts reproducibles, documentación metodológica, informes técnicos y registros de procesamiento que permiten reconstruir los procedimientos realizados.

La documentación generada incluye información sobre los procesos de depuración, validación, agregación y análisis de los datos, así como las versiones de software utilizadas y los parámetros empleados en cada procedimiento.

La procedencia de los datos se documenta desde su recepción hasta la obtención de los resultados finales, manteniendo la trazabilidad de todas las transformaciones realizadas.

Los resultados agregados, indicadores derivados, mapas y productos científicos generados podrán ser reutilizados por otros investigadores respetando las condiciones de uso establecidas por las instituciones responsables de los datos originales y las políticas de acceso abierto aplicables.

3. Otros resultados de investigación

Además de los datos de investigación, el proyecto genera otros resultados científicos digitales que forman parte esencial de las actividades desarrolladas.

Entre ellos se incluyen:

- Scripts desarrollados en lenguaje R para análisis epidemiológicos, espaciales y espacio-temporales.
- Flujos de trabajo reproducibles para el tratamiento y análisis de datos de mortalidad.
- Procedimientos de modelización estadística y análisis bayesiano.
- Sistemas de integración de indicadores territoriales, ambientales y socioeconómicos.
- Cartografía temática y atlas de mortalidad.
- Informes técnicos y documentación metodológica.
- Manuscritos científicos, comunicaciones a congresos y material de difusión científica.
- Aplicaciones y servidores interactivos para la visualización de resultados.

Estos materiales se almacenan junto con la documentación asociada en las infraestructuras del Instituto de Salud Carlos III y constituyen elementos fundamentales para garantizar la reproducibilidad, transparencia y reutilización de los resultados obtenidos.

4. Asignación de recursos

Las actividades de gestión de datos se realizan utilizando las infraestructuras informáticas y los recursos tecnológicos disponibles en el Instituto de Salud Carlos III.

La obtención de determinados conjuntos de datos requiere la formalización de solicitudes específicas ante organismos productores de información, particularmente el Instituto Nacional de Estadística, de acuerdo con las condiciones establecidas

para el acceso a datos confidenciales con fines científicos.

La responsabilidad de la gestión de los datos recae en el equipo investigador del proyecto y en sus investigadores principales.

La preservación de la información se garantiza mediante los sistemas institucionales de almacenamiento, respaldo y recuperación disponibles en el Instituto de Salud Carlos III, así como mediante los procedimientos establecidos para la conservación y protección de información científica.

5. Seguridad de los datos

Los datos utilizados en el proyecto se almacenan y procesan exclusivamente en equipos y servidores institucionales gestionados por el Instituto de Salud Carlos III.

Los datos individuales facilitados por el Instituto Nacional de Estadística están sujetos a las obligaciones derivadas del secreto estadístico y únicamente pueden ser utilizados por el personal autorizado que haya suscrito los compromisos de confidencialidad correspondientes.

La transferencia, almacenamiento y tratamiento de los datos se realiza siguiendo las medidas de seguridad establecidas por el Instituto Nacional de Estadística y las políticas institucionales de seguridad de la información del Instituto de Salud Carlos III.

El acceso a la información está restringido a usuarios autorizados mediante mecanismos de autenticación y control de acceso. Asimismo, se aplican procedimientos de copia de seguridad, recuperación y protección frente a accesos no autorizados.

Una vez finalizado el periodo autorizado de utilización de los datos, los ficheros sujetos a las condiciones establecidas por el Instituto Nacional de Estadística serán destruidos o eliminados de acuerdo con los procedimientos establecidos por dicho organismo.

6. Aspectos éticos

El proyecto utiliza información procedente de registros oficiales de mortalidad y otras fuentes estadísticas y administrativas obtenidas conforme a los procedimientos legalmente establecidos.

Los datos individuales facilitados por el Instituto Nacional de Estadística están protegidos por la normativa sobre secreto estadístico y únicamente se utilizan para los fines científicos autorizados.

No se realizan intentos de identificación de personas ni cruces de información no autorizados. Los resultados generados se presentan exclusivamente de forma agregada y anonimizada.

Todas las actividades desarrolladas se realizan de acuerdo con la normativa vigente en materia de investigación científica, protección de datos, secreto estadístico y seguridad de la información.

7. Otros aspectos

El proyecto sigue los procedimientos institucionales establecidos por el Instituto de Salud Carlos III para la gestión, almacenamiento, transferencia y conservación de datos científicos.

Asimismo, se aplican las condiciones específicas de acceso y uso establecidas por el Instituto Nacional de Estadística para los datos confidenciales utilizados en el proyecto, incluyendo las obligaciones de confidencialidad, control de accesos, custodia de la información y destrucción de los ficheros una vez finalizado el periodo autorizado de utilización.

Actualmente no se prevén procedimientos adicionales de gestión de datos distintos de los establecidos por las instituciones responsables de la información utilizada y por el propio proyecto DETEMOCAES.