

Instrucción 01/2025 de Técnica de normalización de ficheros del Registro Poblacional de Cáncer en Andalucía de la Dirección General de Salud Pública y Ordenación Farmacéutica.

1. Introducción

1.1. Objeto de la Instrucción

La presente Instrucción Técnica de normalización de ficheros del Registro Poblacional de Cáncer en Andalucía (RPCA) tiene por objeto establecer las condiciones para la aportación de datos de las diversas fuentes de información de dicho Registro. Dichas fuentes y los datos relacionados que deben proporcionar, se encuentran explicitados respectivamente en el Anexo I y II del Decreto 586/2019 de 29 de octubre.

El seguimiento de las pautas indicadas en este documento permitirá un mejor procesamiento, carga, análisis y visualización de los datos aportados, además de estandarizar y minimizar los tiempos en todas las fases del registro, incluyendo la detección y subsanación de anomalías.

1.2. Marco normativo

Mediante el Decreto 297/2007, de 18 de diciembre, se creó el RPCA, con la finalidad de conocer la magnitud de la enfermedad en su área geográfica, ayudar en la planificación y gestión sanitaria, identificar grupos de alto riesgo, así como realizar y colaborar en estudios de investigación sobre el cáncer, todo ello en aras de asegurar una mejor tutela de la salud de la población andaluza.

Son funciones del RPCA la obtención, procesamiento, análisis y evaluación de la información obtenida sobre los casos de cáncer, custodiar y controlar el acceso a los datos clínicos epidemiológicos y demográficos, conforme a la normativa vigente en materia de protección de datos de carácter personal. Por otra parte, también pretende regular y facilitar la información que resulte necesaria tanto para la planificación de los servicios sanitarios como para la evaluación de la atención sanitaria en el ámbito de las enfermedades tumorales.

En 2019, a la luz de la experiencia acumulada desde la puesta en marcha del RPCA, se entendió necesaria una nueva regulación del mismo para realizar una mejora en su organización y gobernanza, incluyendo entre sus finalidades la contribución a la eficacia del Sistema Sanitario Público de Andalucía y un aumento de la eficiencia en el tratamiento de la información. En este contexto, se publica el actual marco normativo Decreto 586/2019 de 29 de octubre.


1.3. Ámbito de aplicación

La presente Instrucción Técnica de Normalización de ficheros del RPCA está dirigida a las Delegaciones Territoriales que actualmente aportan información relacionada con la actividad de Anatomía Patológica (AP), tanto la procedente de Centros hospitalarios públicos como de los privados.

Esta Instrucción se complementa con las peticiones dirigidas a los Centros directivos responsables de otras fuentes de información que también nutren el RPCA, en relación a:

- Conjunto Mínimo Básico de Datos (CMBD)
- Base Poblacional de Salud (BPS)
- Cribados oncológicos
- Registro de Mortalidad



Puede verificar la integridad de este documento mediante la lectura del código QR adjunto o mediante el acceso a la dirección https://ws050.juntadeandalucia.es/verificarFirma indicando el código de VERIFICACIÓN			
FIRMADO POR	MANUEL FERNANDEZ ZURBARAN	07/08/2025	
VERIFICACIÓN	Pk2jm9QMTTMHPTA3X94SNUDV6C4KXM	PÁG. 1/8	



2. Requisitos generales de la información

2.1 Implementación del procedimiento de normalización de ficheros

Para una correcta ejecución del proceso ETL (*Extract, Transform, Load*), se precisa la incorporación de todos los registros de AP, tanto de Centros públicos como de los privados.

2.2 Tipos de datos requeridos

Los datos que deben proporcionarse se encuentran explicitados en el Anexo II del Decreto 586/2019 de 29 de octubre. No obstante, actualmente sólo se recogerán los datos recogidos en el apartado 3.2 de la presente Instrucción.

2.3 Periodicidad de envío

Los ficheros se proporcionarán con una periodicidad anual, de modo que las fuentes deberán aportar la información solicitada del año completo a registrar. Los requerimientos de estos ficheros se realizarán por parte de las Delegaciones Territoriales a los responsables de los Centros hospitalarios: documentalistas y/o servicios de AP.

2.4 Unidad de registro: Anatomía Patológica

En lo que respecta a los ficheros de AP, cada Centro deberá proporcionar la información referente a las muestras, atendiendo a los criterios del *Anexo I* de la presente Instrucción¹.

3. Estructura y formato de los ficheros de Anatomía Patológica

En este apartado se detallan la estructura y el formato de los ficheros para su incorporación en la base de datos del RPCA, siendo especialmente relevante para los encargados del proceso de exportación de los ficheros.

3.1 Denominación y formato de ficheros

El formato seleccionado para el intercambio de datos será CSV (*Comma-Separated Values*), codificado en UTF-8. Este formato consiste en un fichero de texto plano cuya primera fila contendrá los nombres de las columnas (indicadas en el apartado 3.2), separados por el carácter *pipe* (“|”). A continuación, se incluirán tantas filas como registros de datos se deban exportar, separando los distintos campos de cada registro, igualmente, mediante el carácter *pipe*.

Los saltos de línea incluidos dentro de campos de texto deberán codificarse utilizando el formato CRLF (*carriage return + line feed*, $\backslash r \backslash n$), conforme al estándar de codificación de fin de línea en sistemas Windows.

Asimismo, todos los campos incluidos en el fichero deberán estar entre comillas dobles (“”), según lo especificado en el estándar RFC 4180. Igualmente, según las especificaciones en este estándar, si un texto de un campo contiene a su vez comillas dobles, éstas deben duplicarse. Por ejemplo, si el texto es el siguiente: bloques “A5-A17” seriación de cuello, el valor del campo exportado al fichero CSV quedaría de la siguiente forma: “bloques ““A5-A17”” seriación de cuello”.

En cuanto al nombre de los ficheros generados de AP, mantendrán el formato “AP_AAAA_NOMBRE_HOSPITAL.csv”, siendo AAAA el año al que hacen referencia los datos y NOMBRE_HOSPITAL el Centro al que corresponde la información. En ningún caso debe contener espacios u otros caracteres, salvo guion bajo (“_”). Ej.: AP_2025_SAN_CECILIO.csv

¹ Aunque en los criterios de Caso Registrable se recoge que sólo se registra la localización del tumor primario y no el lugar de las metástasis, en el fichero de Anatomía Patológica deben enviarse, también, los tumores cuyo código de comportamiento sea /6 (metástasis).


Puede verificar la integridad de este documento mediante la lectura del código QR adjunto o mediante el acceso a la dirección https://ws050.juntadeandalucia.es/verificarFirma indicando el código de VERIFICACIÓN			
FIRMADO POR	MANUEL FERNANDEZ ZURBARAN	07/08/2025	
VERIFICACIÓN	Pk2jm9QMTTMHPTA3X94SNUDV6C4KXM	PÁG. 2/8	



3.2 Nombres y definiciones de las columnas


Se anexa a esta Instrucción un fichero CSV² (*Anexo II*) de ejemplo con la estructura requerida. La siguiente tabla contiene las especificaciones de datos y variables, tanto las obligatorias como las opcionales, siendo estas últimas: TNM_T, TNM_N, TNM_M, Estadiaje, CONCEPT_ID. El procesamiento de las variables opcionales se prevé que se lleve a cabo próximamente.

² Abrir con editor de texto plano para evitar alteración de formato.

Puede verificar la integridad de este documento mediante la lectura del código QR adjunto o mediante el acceso a la dirección https://ws050.juntadeandalucia.es/verificarFirma indicando el código de VERIFICACIÓN			
FIRMADO POR	MANUEL FERNANDEZ ZURBARAN	07/08/2025	
VERIFICACIÓN	Pk2jm9QMTTMHPTA3X94SNUDV6C4KXM	PÁG. 3/8	




Nombre columna	Descripción	Tipo de datos	Formato / Lista valores *
C_CENTRO	Código del hospital donde se toma la muestra.	Texto	000000000 (9 caracteres) Código de Centro Normalizado REGCESS, CCN (Anexo III).
NOMBRE	Nombre del paciente.	Texto	
APELLIDO1	Primer apellido del paciente.	Texto	
APELLIDO2	Segundo apellido del paciente.	Texto	
SEXO	Sexo del paciente.	Texto	[Desconocido Hombre Mujer]
FECHA_NAC **	Fecha de nacimiento del paciente.	Fecha	DD/MM/AAAA
NUHSA **	Número único de Historia de Salud de Andalucía.	Texto	[AN]0000000000 *** (12 caracteres)
NSS **	Número de la Seguridad Social.	Texto	000000000000 (12 dígitos)
DNI **	Número del Documento Nacional de Identidad o NIE.	Texto	A0000000L (9 caracteres)
DOMICILIO	Dirección del paciente.	Texto	
C_POBLACION	Código de la población o el municipio.	Texto	000009 (5 dígitos + 1 dígito de control) Codificación INE (Instituto Nacional de Estadística)
C_PROVINCIA	Código de la provincia.	Texto	00 (2 dígitos) Codificación INE
C_POSTAL **	Código postal.	Texto	00000 (5 dígitos)
DATOS_CLINICOS	Datos clínicos de la neoplasia.	Texto	
EDAD **	Edad en el momento de la toma de la muestra.	Número	Entero positivo

Puede verificar la integridad de este documento mediante la lectura del código QR adjunto o mediante el acceso a la dirección https://ws050.juntadeandalucia.es/verificarFirma indicando el código de VERIFICACIÓN			
FIRMADO POR	MANUEL FERNANDEZ ZURBARAN	07/08/2025	
VERIFICACIÓN	Pk2jm9QMTTMHPTA3X94SNUDV6C4KXM	PÁG. 4/8	



VALIDACION	Fecha de validación de la muestra.	Fecha	DD/MM/AAAA
TEXTO_COMPLEMENTARIO	Texto complementario.	Texto	
FECHA_ESTUDIO_COMPLEMENTARIO	Fecha del estudio complementario.	Fecha	DD/MM/AAAA
ESTUDIO	Identificador del estudio.	Texto	
MUESTRA	Letra asociada a cada una de las muestras de un estudio en orden alfabético.	Texto	[A-Z]
ORGANO	Órgano de donde ha sido extraída la muestra.	Texto	
TIPO_MUESTRA	Tipo de muestra.	Texto	
FECHA_TOMA_MUESTRA	Fecha de la toma de la muestra.	Fecha	DD/MM/AAAA
MACRO	Descripción macroscópica.	Texto	
MICRO	Descripción microscópica.	Texto	
DIAGNOSTICO	Diagnóstico del anatomopatólogo.	Texto	
NOTAS	Comentarios adicionales.	Texto	
TOPOGRAFICO	Literal del código topográfico asociado a la neoplasia.	Texto	
C_TOPOGRAFICO	Código topográfico asociado a la neoplasia.	Texto	
MORFOLOGICO	Literal del código morfológico asociado a la neoplasia.	Texto	
C_MORFOLOGICO	Código morfológico asociado a la neoplasia.	Texto	Codificación CIE-O según versión vigente del año.
VALORACIONES	Apreciaciones adicionales.	Texto	
TNM_T	Recoge la T del TNM		Codificación TNM según versión vigente del año.
TNM_N	Recoge la N del TNM		Codificación TNM según versión vigente del año.
TNM_M	Recoge la M del TNM		Codificación TNM según versión vigente del año.
C_ESTADIAJE	Estadio del tumor.		0. Tis

Puede verificar la integridad de este documento mediante la lectura del código QR adjunto o mediante el acceso a la dirección https://ws050.juntadeandalucia.es/verificarFirma indicando el código de VERIFICACIÓN			
FIRMADO POR	MANUEL FERNANDEZ ZURBARAN	07/08/2025	
VERIFICACIÓN	Pk2jm9QMTTMHPTA3X94SNUDV6C4KXM	PÁG. 5/8	



			I - II. Localizado III. Regional IV. Metastásico
CONCEPT_ID	Código del SNOMED CT		

Tabla 1: Campos en ficheros de Anatomía Patológica.

* En la columna 'Formato/Lista Valores' sólo se especifica la máscara de entrada en campos que exigen estandarización.

** Dada la ausencia de valores para los campos NUHSA y NSS en registros de AP de Centros Privados, se requiere la cumplimentación obligatoria de DNI, FECHA_NAC o EDAD, C_POSTAL y C_PROVINCIA en estos ficheros.

*** Caracteres que definen máscaras de entrada:

- 0 Obligatorio (dígito).
- 9 Opcional (dígito).
- L Obligatorio (letra).
- A Obligatorio (dígito o letra).

[] Cadenas de texto fijas (Ej.: en el campo NUHSA, los valores comenzarán sin excepción por la cadena "AN" seguida de 10 dígitos (Ej.: AN0123456789)).

3.3 Registro de casos

Una vez generadas las columnas con los nombres indicados en el apartado anterior, se procederá a añadir los registros. Cada fila se corresponde con una muestra.

3.4 Codificación estandarizada de la morfología

El código morfológico (C_MORFOLOGICO) debe codificarse en CIE-O (hasta el 2020 CIE-O 3.1, a partir del 2021 CIE-O 3.2) y concordar con el texto aportado en MORFOLOGICO. A medida que se implementen nuevas versiones se trasladará el año de aplicación correspondiente.

3.5 Consideraciones acerca de los ficheros de Anatomía Patológica

A continuación, se recogen algunas consideraciones a tener en cuenta respecto al formato, estructura y contenido de estos ficheros:

- Las cabeceras se denominarán tal como se indica en la *Tabla 1*.
- No se dará validez a ficheros en los que convivan valores de campos diferentes en una misma columna. Ej.: APELLIDO1 + APELLIDO2 en una misma columna.
- Sólo se admitirán ficheros que contengan un único valor en los campos C_MORFOLOGICO y C_TOPOGRAFICO.
- Se evitarán textos introductorios redundantes, especialmente en: DATOS_CLINICOS, DIAGNOSTICO, MACRO, MICRO, NOTAS y VALORACIONES. Ej.: "Descripción macroscópica/microscópica:", "Diagnóstico:", "Nota:", "Comentario:".

Puede verificar la integridad de este documento mediante la lectura del código QR adjunto o mediante el acceso a la dirección https://ws050.juntadeandalucia.es/verificarFirma indicando el código de VERIFICACIÓN			
FIRMADO POR	MANUEL FERNANDEZ ZURBARAN	07/08/2025	
VERIFICACIÓN	Pk2jm9QMTTMHPTA3X94SNUDV6C4KXM	PÁG. 6/8	



- Los valores en campos de tipo 'Fecha' (FECHA_NAC, VALIDACION, VALIDACION_COMPLEMENTARIO, FECHA_TOMA_MUESTRA) se ajustarán en todo caso al formato: DD/MM/AAAA. Por ejemplo: "01/01/2020".
- Se verificará la concordancia entre el volumen de muestras de AP recogidas en los ficheros y la carga hospitalaria de los servicios prestados en los Centros de su provincia.
- Se revisará la coherencia de datos tanto a nivel de Cartera de servicios del Centro hospitalario como a nivel clínico. Por ejemplo: No pueden aparecer intervenciones que no se ofertan en la Cartera de servicios del Centro.

4. Procedimiento de envío de ficheros

En este apartado se detallan los requisitos que las Delegaciones Territoriales deben cumplir para el envío de los ficheros. En caso de incumplimiento, no será posible el tratamiento de los datos ni su registro en el RPCA, por lo que se requerirá su reenvío especificando las modificaciones necesarias.

4.1 Método de envío de datos

Se llevará a cabo mediante ficheros en formato CSV, a través de un canal que garantice la protección de los datos, dada su elevada sensibilidad, conforme a la normativa vigente en esta materia. Para ello, se ha habilitado un canal seguro con cifrado SSL, basado en un servidor SFTP (SFTPGO) que permite acceso restringido por Centro, con una estructura de carpetas jerarquizada.

Desde la Consejería de Salud y Consumo se proporcionarán credenciales a las diferentes Delegaciones Territoriales y Centros hospitalarios. Cada Delegación Territorial dispondrá de acceso a la totalidad de las carpetas correspondientes a los Centros de su provincia, mientras que cada Centro tendrá acceso únicamente a su propia carpeta.

Este sistema reemplaza el envío por correo electrónico, minimizando riesgos asociados al incumplimiento normativo.

Los usuarios accederán a la plataforma para cargar y descargar archivos, con permisos configurados en función de su ámbito competencial. Además, se ha incorporado un mecanismo de doble factor de autenticación para reforzar la seguridad del protocolo SSL.

Los detalles sobre el acceso, configuración y operativa del servidor están recogidos en *Anexo IV*.

La plataforma está disponible en la siguiente URL: <https://www.juntadeandalucia.es/salud/sftpgo-rpca/web/client>

4.2 Confirmación de recepción

Se enviará confirmación de recepción de los ficheros a la Delegación Territorial correspondiente, sin que esto implique la validación de su formato y estructura.

5. Validación de formato y estructura

Cada fichero remitido se someterá a un proceso de validación de formato y estructura para el posterior control de calidad de su contenido. Este procedimiento se realizará de modo centralizado en la Consejería de Salud y Consumo.

5.1. Comprobación de estructura y formato

Verificación de los requerimientos explicitados en los apartados anteriores sobre la estructura y el formato de los datos enviados. Entre ellos:

- Formato de archivo y sus campos.
- Codificación del texto.
- Unificación en una sola línea de toda la información referente a una misma muestra.
- Codificación de datos geográficos y Centros hospitalarios.

Puede verificar la integridad de este documento mediante la lectura del código QR adjunto o mediante el acceso a la dirección https://ws050.juntadeandalucia.es/verificarFirma indicando el código de VERIFICACIÓN			
FIRMADO POR	MANUEL FERNANDEZ ZURBARAN	07/08/2025	
VERIFICACIÓN	Pk2jm9QMTTMHPTA3X94SNUDV6C4KXM	PÁG. 7/8	



5.2 Resultados de las comprobaciones e informe

Una vez finalizados los procesos de validación, se emitirá a las Delegaciones Territoriales un informe con las apreciaciones. En función de los resultados, el contenido del fichero se incluirá en el RPCA o bien se requerirá un nuevo envío por parte de la Delegación Territorial correspondiente, tras subsanar los errores especificados. En todo caso, desde la Dirección General de Salud Pública y Ordenación Farmacéutica se emitirá confirmación de recepción e informe de validación para cada fichero, tanto en la primera versión como en sucesivas subsanaciones si fueran necesarias.

6. Anexos

Anexo I: Criterios de caso registrable en los ficheros de AP


Anexo II: Modelo de fichero con estructura plana (CSV)

Anexo III: Listado de centros con código normalizado REGCESS

Anexo IV: Manual de envío de ficheros a través de SFTPGO

DIRECTOR GENERAL DE SALUD PÚBLICA Y ORDENACIÓN FARMACÉUTICA

Manuel Fernández Zurbarán

Puede verificar la integridad de este documento mediante la lectura del código QR adjunto o mediante el acceso a la dirección https://ws050.juntadeandalucia.es/verificarFirma indicando el código de VERIFICACIÓN			
FIRMADO POR	MANUEL FERNANDEZ ZURBARAN	07/08/2025	
VERIFICACIÓN	Pk2jm9QMTTMHPTA3X94SNUDV6C4KXM	PÁG. 8/8	