

Alumna: María Uriburu Gray
Mentor: Joan Gasull Jolis

Más que raras

Análisis de las Enfermedades Raras en Europa y su impacto regional en España



Agenda



¿Qué son las enfermedades raras?
Definiciones necesarias



Objetivos del estudio



Colección de datos



Tratamiento de datos



Resultados



Conclusiones



¿Qué son las enfermedades raras?

Enfermedad Rara

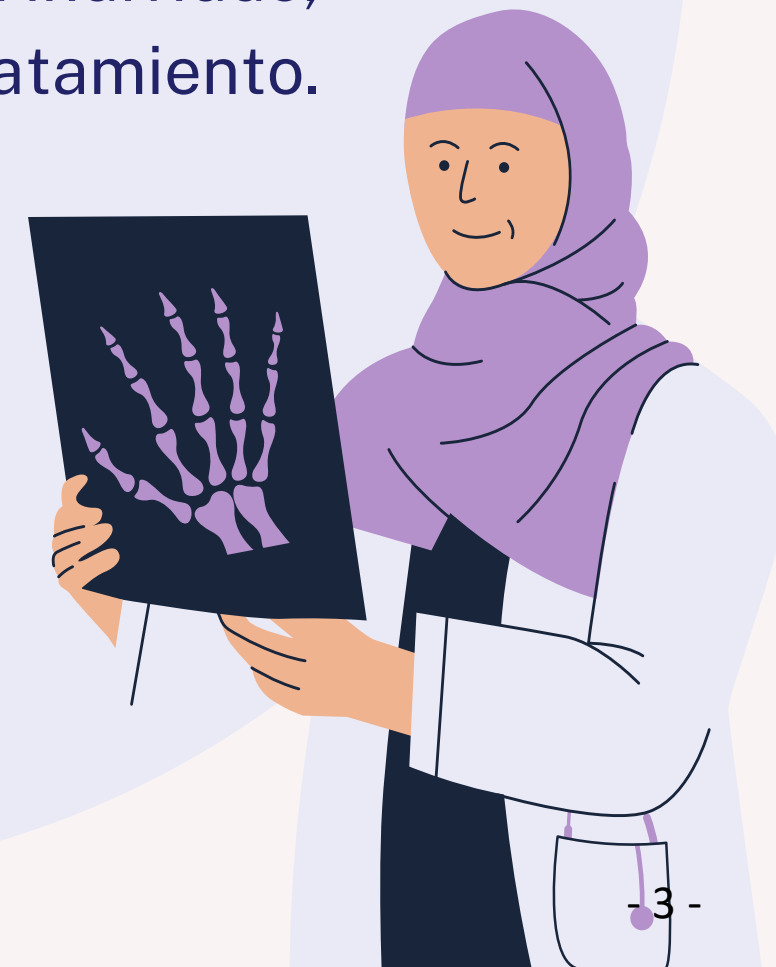
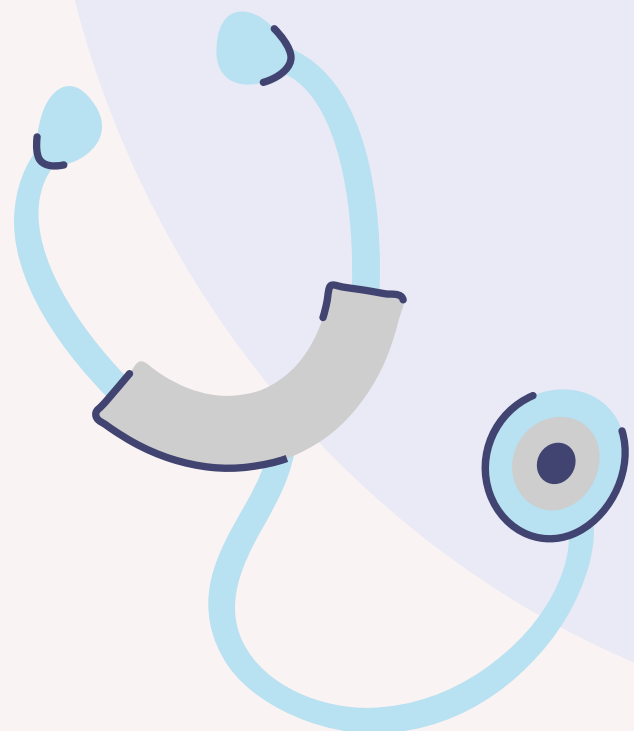
Aquella que, con peligro de muerte o invalidez crónica, tiene una prevalencia menor de 5 casos por cada 10.000 habitantes.

Fenotipos

Características observables de una enfermedad en un individuo, como signos, **síntomas** y respuestas al tratamiento.

Incidencia vs Prevalencia

La **incidencia** mide el número de **nuevos casos** de una enfermedad en una población y tiempo determinado, mientras que la **prevalencia** mide el total de personas que **tienen** una enfermedad.



Objetivos del estudio

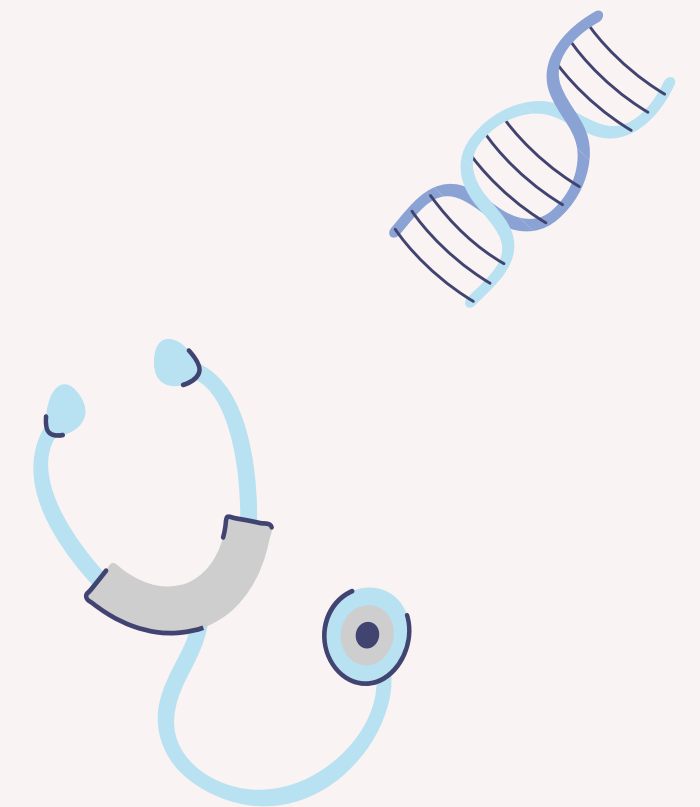
Identificar patrones que contribuyan a la investigación y gestión sanitaria de las Enfermedades Raras

1. Estudio de la visibilidad científica de las ER
¿Cuál es la situación de las enfermedades raras en la comunidad científica? ¿Qué tanto se habla de ellas en la literatura especializada?

2. Estudio de la complejidad de las ER
¿Existe relación entre el número de genes implicados y el número de fenotipos que presentan?

3. Estudio regional de la carga de ER en España
¿Existe alguna relación entre la concentración de diagnósticos de ER y la cantidad de población?

4. Estudio de la afectación por género.
¿Existe una mayor predisposición a padecer estas enfermedades según el género? En caso afirmativo, ¿cuál es más afectado: hombres o mujeres?



Colección de Datos

API PubMed

Acceso al registro de publicaciones científicas que contienen “rare disease” en el título y año de publicación.

API Orphadata

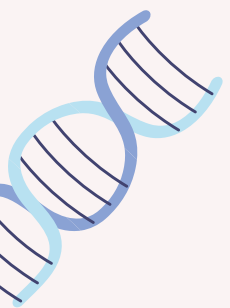
Acceso al registro completo de Enfermedades Raras. Contiene información de la epidemiología, historia clínica, prevalencia, genes y fenotipos.

Informes ReeR

Informes del Registro Estatal de Enfermedades Raras con información de los casos notificados por cada comunidad autónoma.

Datos INE

Población total por Comunidades Autónomas en 2018.

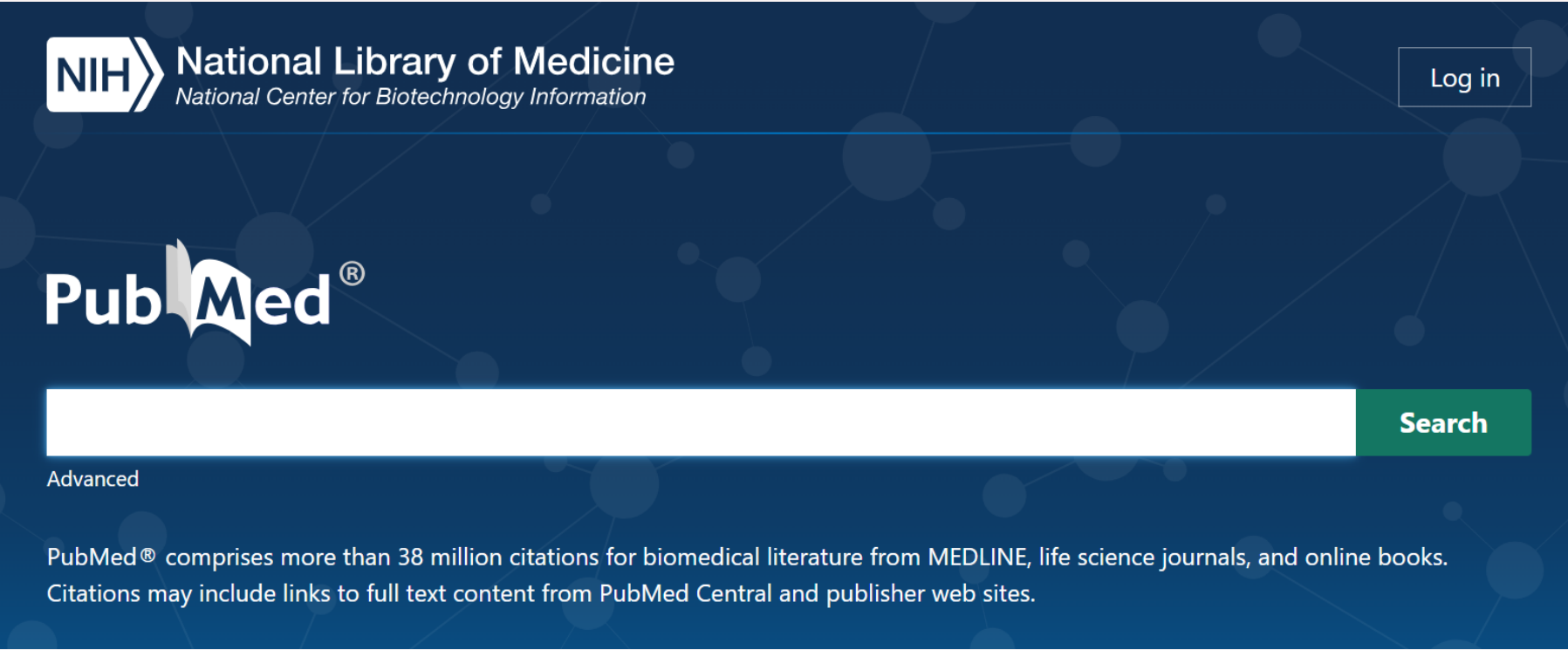




Tratamiento de Datos

API PubMed

Base de datos gratuita de la Biblioteca Nacional de Medicina de EE.UU.

- Librerías Requests y ElementTree para acceder y extraer datos de la API (formato XML).
- Dataset limpio: 2617 filas y 5 columnas.
- Interés en el N.º de publicaciones por año.

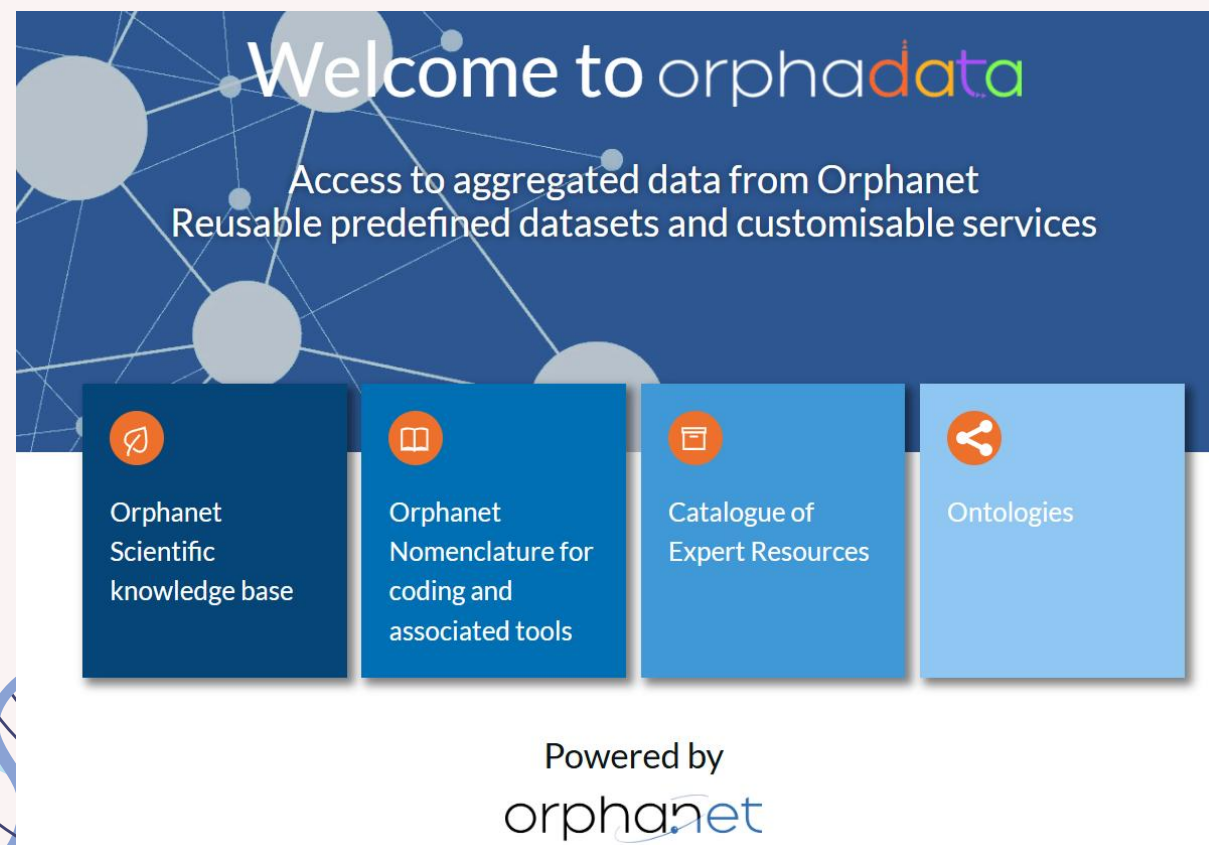


# pmid	# pub_year	# pubtype	# journal		
Missing: 0 (0%) Distinct: 2391 (91%)	Missing: 2 (<1%) Distinct: 2381 (91%)	Missing: 0 (0%) Distinct: 79 (3%)	Missing: 0 (0%) Distinct: 1184 (45%)		
 <div>2381 Distinct values</div>	 <div>Journal Article 28% Case Reports; Journal Article 25% Journal Article; Research Su... 8% Other 39%</div> <div>1184 Distinct values</div>				
0	19973918	A Case of a Rare Disease of the Long	1909	Journal Article	Proceedings of the Royal Society of I
1	19979883	A Rare Disease in Two Brothers.	1917	Journal Article	Proceedings of the Royal Society of I
2	30438855	A Rare Disease of the Conjunctiva, w	1933	Journal Article	Glasgow medical journal
3	20282579	Sandifort's Observaciones, chapter I,	1946	Biography; Journal Article	Bulletin of the history of medicine
4	20985704	Chronic erosive, granulomatous, atro	1946	Journal Article	Gastroenterology
5	15410547	Is early osteochondritis of the patella	1950	Journal Article	Acta chirurgiae orthopaedicae et tra
6	14822046	[Rare disease of the cornea caused b	1951	Journal Article	Ceskoslovenska oftalmologie
7	13026929	[An unusual case of a rare disease, g	1952	Journal Article	La Presse medicale
8	13012938	[A rare disease caused by accident: M	1952	Journal Article	Minerva medica

Tratamiento de Datos

API Orphadata

Orphanet es una base de datos online de acceso libre dedicada a las enfermedades raras y los medicamentos huérfanos. Orphadata es una plataforma que proporciona datos a la comunidad científica.

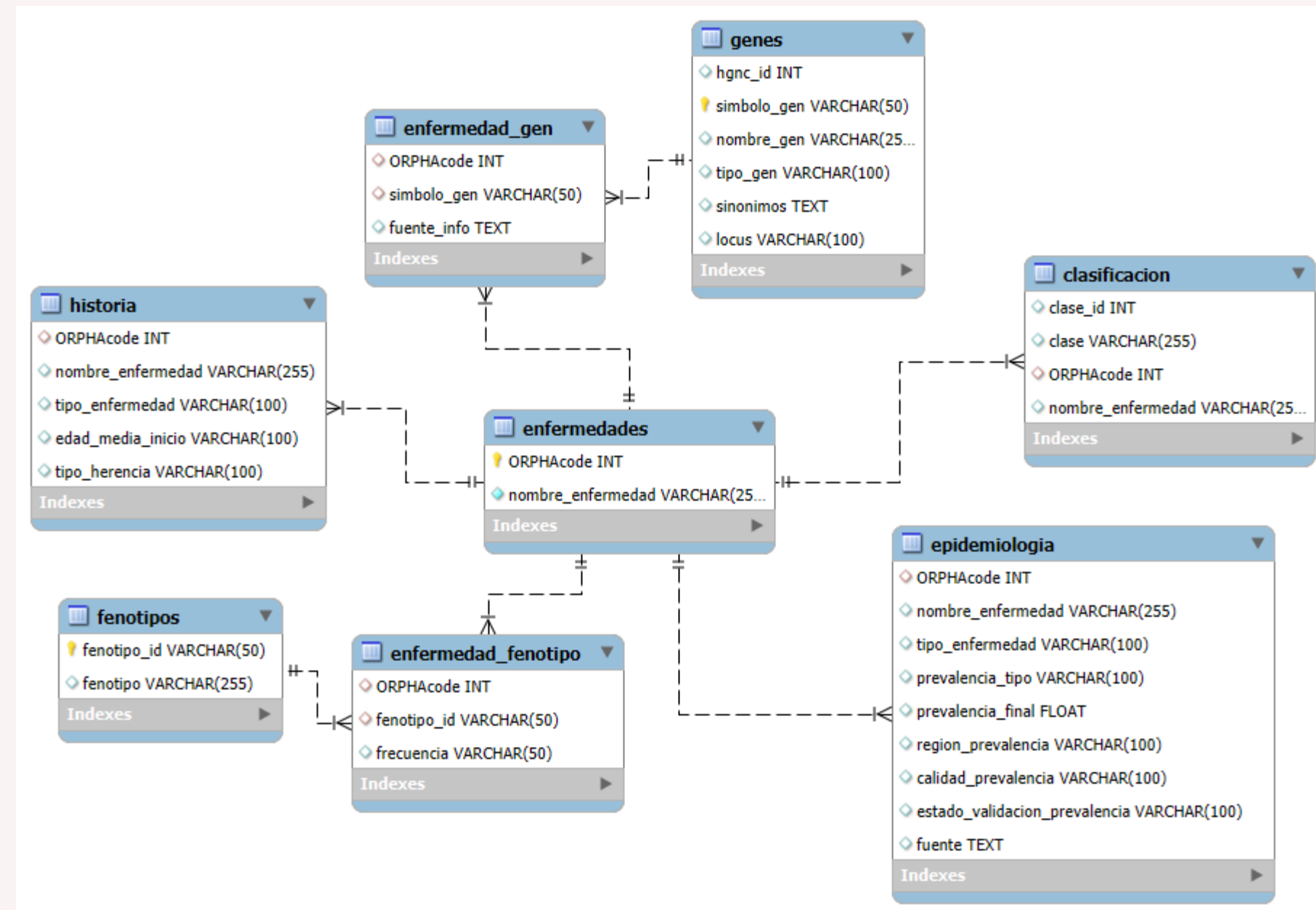


Welcome to orphadata

Access to aggregated data from Orphanet
Reusable predefined datasets and customisable services

- Orphanet Scientific knowledge base
- Orphanet Nomenclature for coding and associated tools
- Catalogue of Expert Resources
- Ontologies

Powered by orphanet

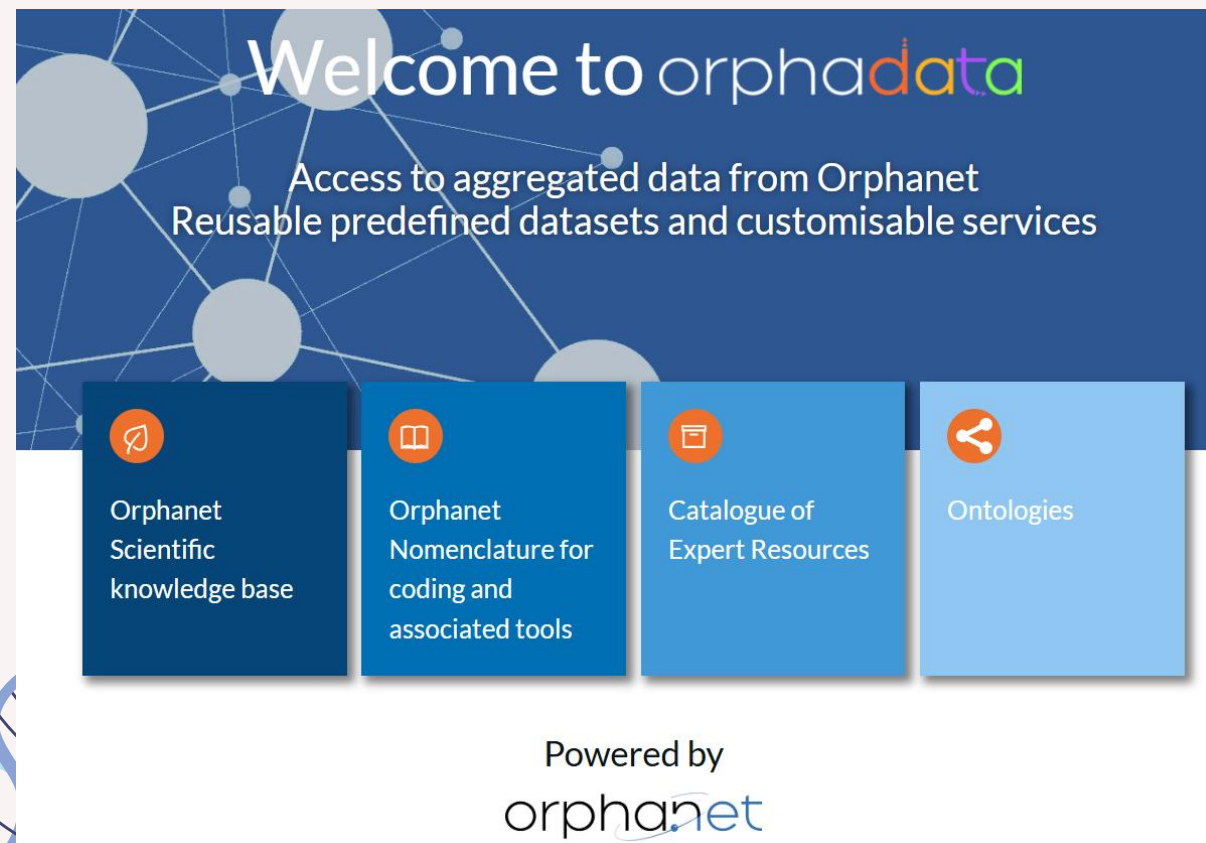


- Requests para acceder a la API (formato JSON).
- Reestructuración de tablas.
- Limpieza de datos.
- Tablas utilizadas: **ENFERMEDADES**, **GENES** y **FENOTIPOS**.

Tratamiento de Datos

API Orphadata

Orphanet es una base de datos online de acceso libre dedicada a las enfermedades raras y los medicamentos huérfanos. Orphadata es una plataforma que proporciona datos a la comunidad científica.

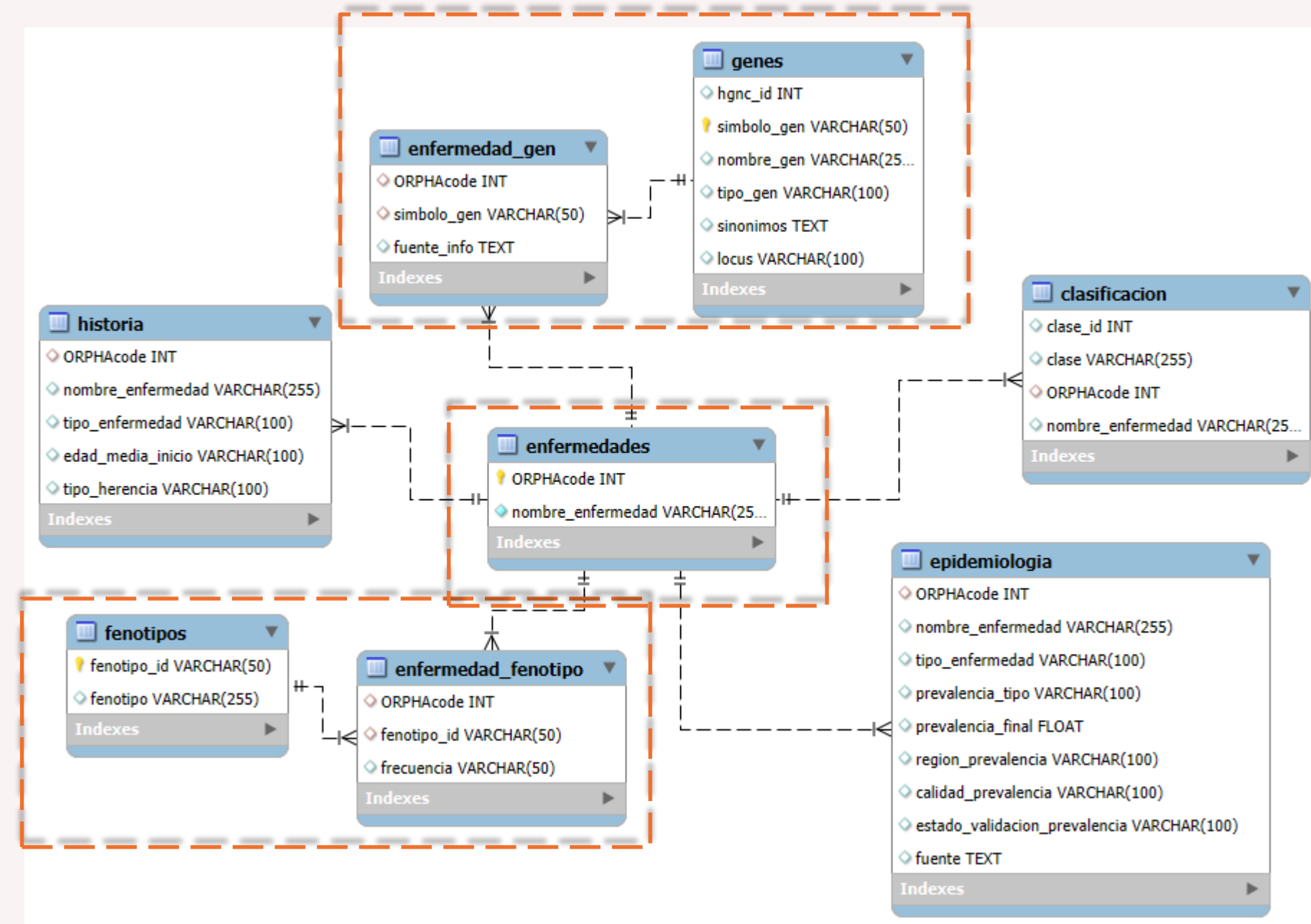


Welcome to orphadata

Access to aggregated data from Orphanet
Reusable predefined datasets and customisable services

- Orphanet Scientific knowledge base
- Orphanet Nomenclature for coding and associated tools
- Catalogue of Expert Resources
- Ontologies

Powered by orphanet



- Requests para acceder a la API (formato JSON).
- Reestructuración de tablas.
- Limpieza de datos.
- Tablas utilizadas: **ENFERMEDADES**, **GENES** y **FENOTIPOS**.

Tratamiento de Datos

ReeR

El Registro Estatal de Enfermedades Raras es un sistema de información del Ministerio de Sanidad, a través del Instituto de Salud Carlos III, **registra todos los casos de enfermedades raras en España.**

Red de registros autonómicos (RAER) transmiten los datos al ReeR central

INE

Instituto Nacional de Estadística, es el organismo encargado de producir las **estadísticas oficiales** de España.

- **PDFPlumber** es una librería de Python para extraer datos y tablas de archivos PDF.
- **Información limitada.** Faltan años, Comunidades Autónomas y faltan enfermedades.
- Tablas utilizadas:
 - **Casos notificados 2010 – 2018 por CC.AA.**
 - **Población total 2018 por CC.AA. (INE)**
 - **Casos según género 2020 – 2022.**
 - **Enfermedades según género 2022.**



MINISTERIO DE SANIDAD

Castellano

Buscar

Q

Ministerio

Áreas

Prensa y comunicación

Sanidad en datos

Servicios a la Ciudadanía

Participación Pública

Traducir

a

Está usted en:

Inicio

>

Áreas

>

Profesionales

>

Vigilancia en salud pública

Quiénes somos

Alertas de salud pública de actualidad

Vigilancia en salud pública

Sistema Nacional de Alerta Precoz y Respuesta Rápida (SIAPR)

Vigilancia en salud pública

¿Qué es la vigilancia en salud pública?

La vigilancia en salud pública es el conjunto de actividades destinadas a recoger, analizar, interpretar y difundir información relacionada con el estado de la salud de la población y los factores que la condicionan, con el objeto de fundamentar las actuaciones de salud pública (Ley 33/2011, de 4 de octubre de 2011, General de Salud Pública).

Análisis

Actuaciones

Registros

Situación de las enfermedades raras en España

Informe ReeR 2024: Situación de las Enfermedades Raras en España (2010-2021)

Informe ReeR 2023: Situación de las Enfermedades Raras en España (2010-2020)

Informe ReeR 2022: Situación de las Enfermedades Raras en España (2010-2019)

Informe ReeR 2021: Situación de las Enfermedades Raras en España (2010-2018)

Manual de procedimientos del Registro Estatal de Enfermedades Raras 2025

ReecR
Red Española de
Enfermedades Raras

3. RESULTADOS

3.1. Generales

En el ReecR, se han registrado 63.074 casos con una o varias de las 29 enfermedades raras incluidas en el informe, lo que corresponde a 47.364 personas vivas a 1 enero de 2022 (33.514 hombres, 23.849 mujeres y 1 indeterminado) (Tablas 1 y 3) y 15.611 personas fallecidas en el periodo entre el 1 de enero de 2010 y 31 de diciembre de 2021 (Tablas 2 y 4).

Tabla 1. Número de casos vivos a 1 de enero de 2022 según sexo y entidad clínica considerada.

Enfermedad Rara	VIVOS		Total
	Hombres	Mujeres	
Aniridiosis múltiple congénita	189	194	383
Aniridiosis múltiple congénita distal	48	24	72
Astasia de Friedreich	305	24	670
Azorra muscular espinal proximal	305	365	631
Ceroid bilia primario	337	294	3.952
Complegio escleroso tuberoso	376	3.576	2.277
Distrofia renal	1.091	1.196	2.259
Distrofia miotónica de Steinert	1.356	903	4.097
Enfermedad de Fabry	1.968	2.129	587
Enfermedad de Gaucher	281	306	293
Enfermedad de Huntington	151	141	1.836
Enfermedad de Niemann-Pick	151	1.029	84
Enfermedad de Rendu-Osier	807	29	1.711
Enfermedad de Wilson	55	1.007	1.062
Eclimosis lateral amiotrófica	704	477	2.503
Fenilketonuria	585	1.052	1.588
Fibrosis quística	1.451	853	3.568
Hemofilia A	735	1.792	677
Hipodistrofismo congénito	1.776	437	3.662*
Osteogénesis imperfecta	3.244	1.313	2.276
Retinitis pigmentaria y Retinosis pigmentaria sintomáticas	963	782	1.867
Síndrome de Angelman	675	2.758	5.082
Síndrome de Beckwith-Wiedemann	2.324	195	389
Síndrome de Bard-Bell-Weidmann	194	166	842
Síndrome de Beal	176	166	842

*Informa ReecR 2022

ReeR 2021

ción de las

ades Raras en

10-2018

Primer informe

3. RESULTADOS

3.1 Generales

Se obtuvieron un total de corresponden a 12 CC.AA. del año de captación.

Tabla 1. Número de casos notificados

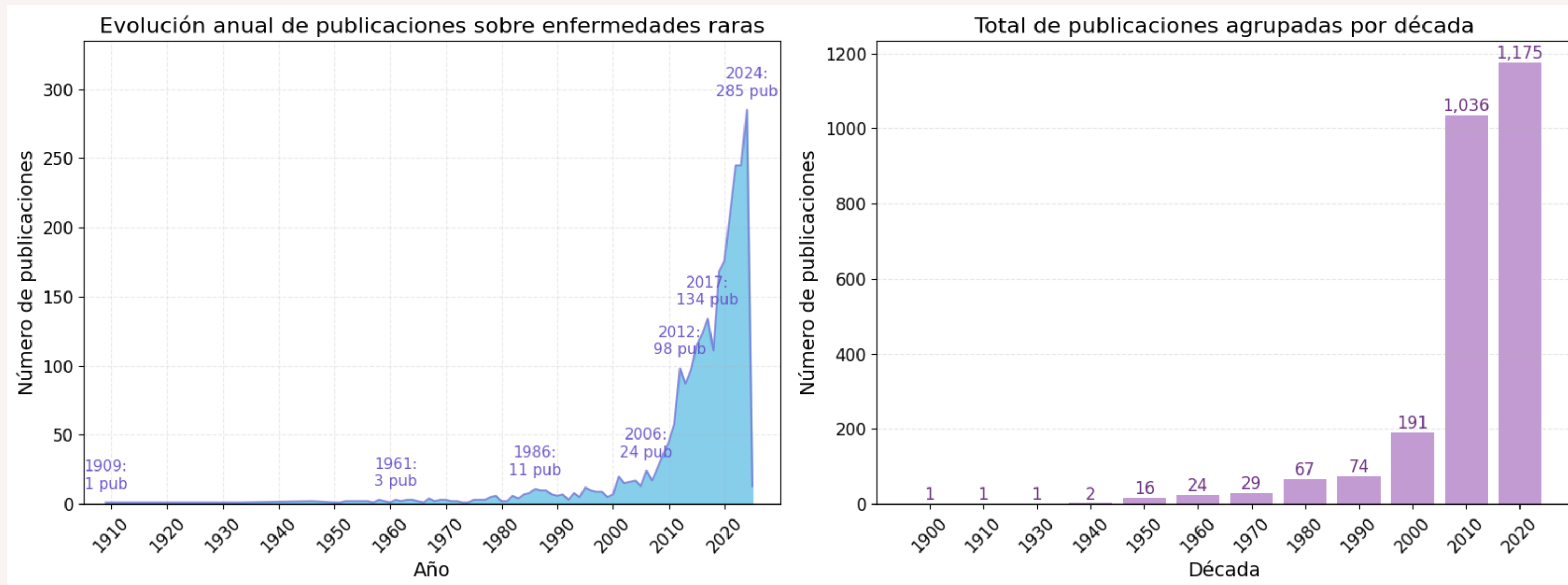
Región	Años de 2016	2019
Andalucía		
Aragón	3.262	469
Cantabria		
Castilla y León	7	130
Cataluña	470	333
Comunidad Valenciana	1.075	391
Galecia	720	311
Madrid, Comunidad de	222	45
Murcia, Región de	3.090	503
N Navarra, Comunidad Foral de	631	108
Pais Vasco	531	34
Rioja, La		
Total	10.065	2.117

Este informe epidemiológico hace referencia excluyendo aquellos que requieren revisión adicional debido a que se corresponden con personas afectadas de diciembre del 2018 (Tabla 2) y 6.004 han fallecido (Tabla 3).

9

Resultados – PubMed –

Extracción del número de publicaciones científicas con “*rare disease*” en el título



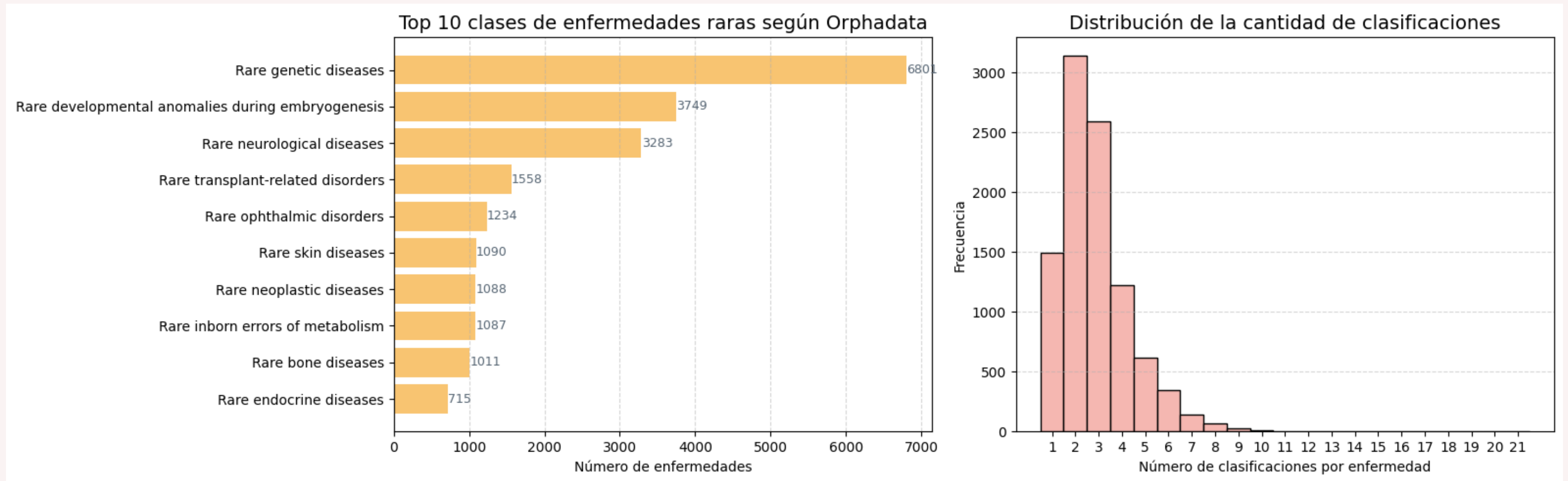
❑ La primera publicación de enfermedades raras fue en el año 1906.

❑ Aumento significativo del número de investigaciones publicadas en la década del 2010.

Resultados

- Orphadata -

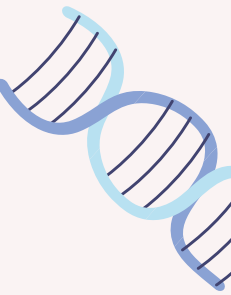
9673 registros de Enfermedades Raras distribuidas en 35 clases según Orphadata



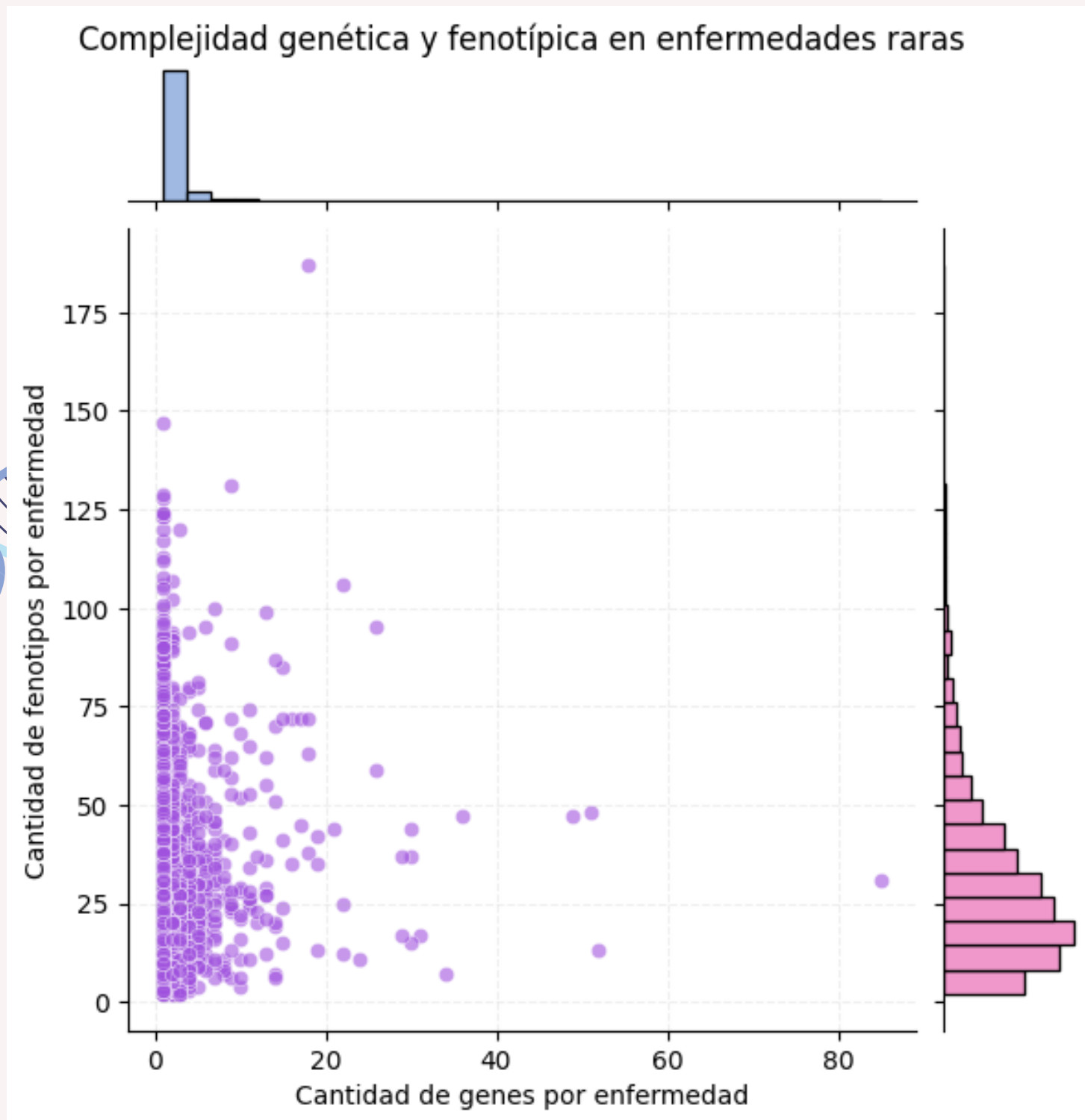
- ❑ La mayoría de las enfermedades tienen un origen genético.
- ❑ Las enfermedades pueden catalogarse en más de una clase, la mayoría pertenecen a 2 o 3 clases.

Resultados

- Orphadata -



Análisis de la dependencia del número de genes y fenotipos asociados a una ER



❑ Correlación de Spearman = 0.0392 → No se observa.

❑ Genes por enfermedad:

Discapacidad intelectual no sindrómica (ORPHAcode: 528084) está asociada a **108 genes** distintos.

❑ Enfermedades por gen:

Genes **LMNA** y **TP53** implicados en **21 enfermedades** cada uno.

Resultados

- Ornhadata -

¿Sabías que...?

¡Es posible consultar información genética directamente desde Python!

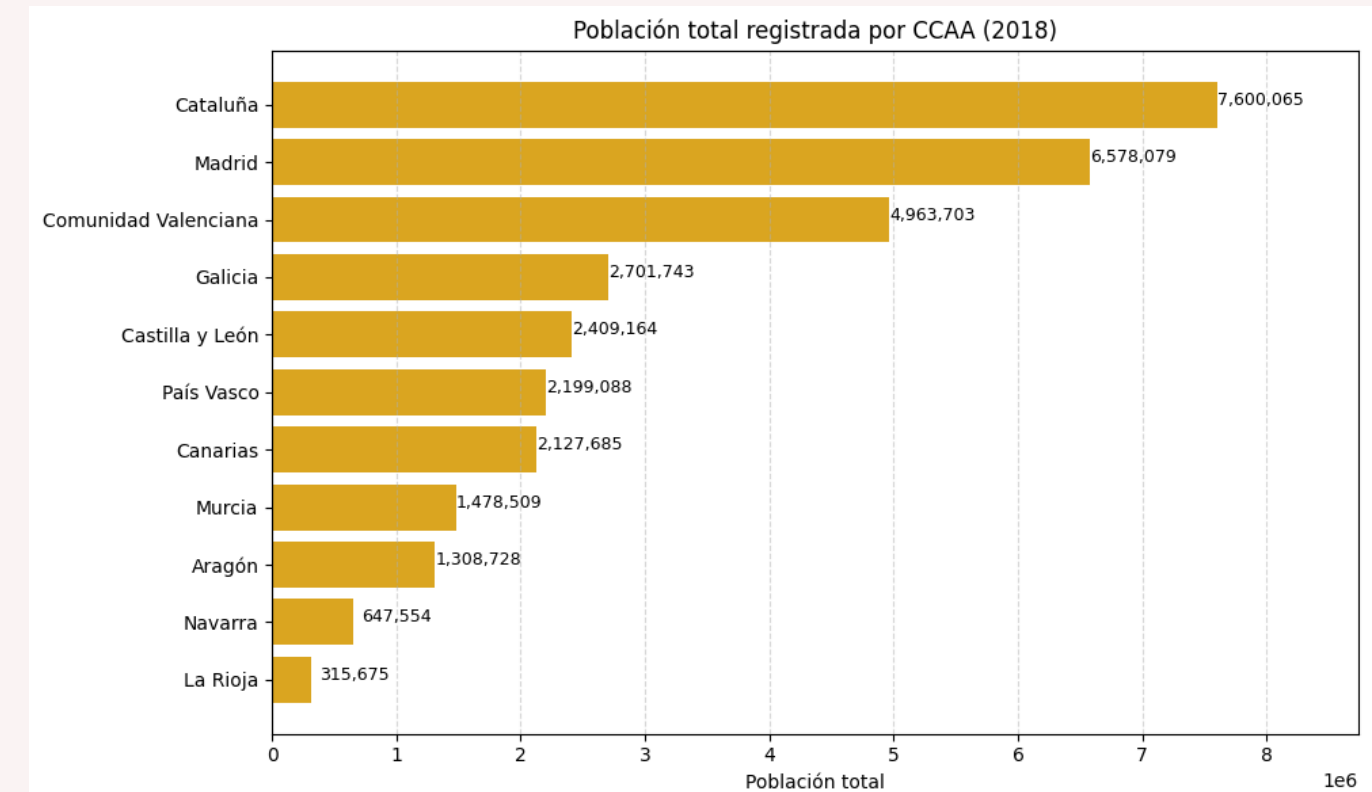
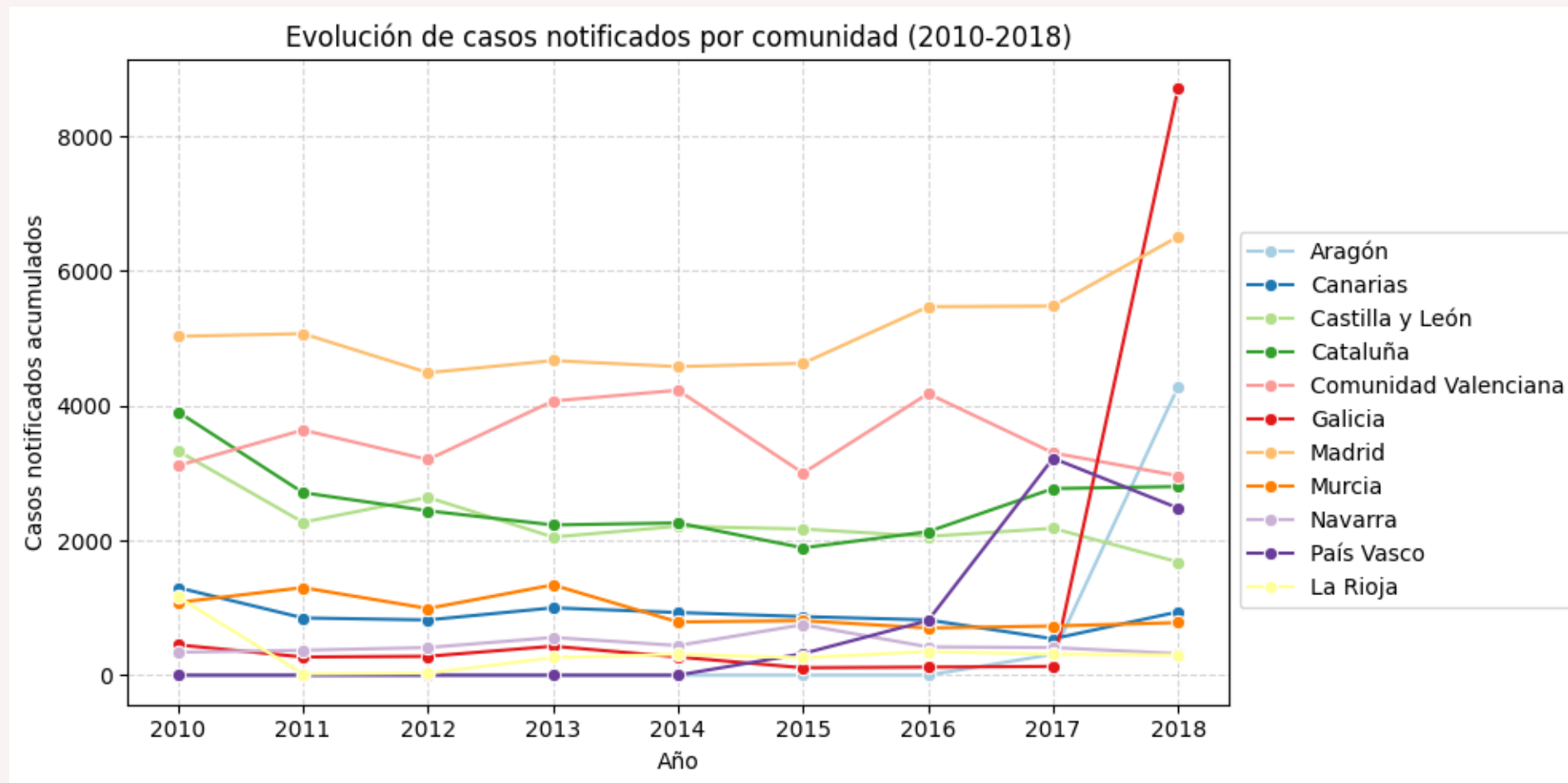
Existe una librería llamada BioPython que permite acceder a bases de datos biomédicas como el NCBI (National Center for Biotechnology Information) y obtener información de genes, proteínas y secuencias genómicas directamente desde tu código.

```
1 from Bio import Entrez
2 Entrez.email = "A.N.Other@example.com"
```

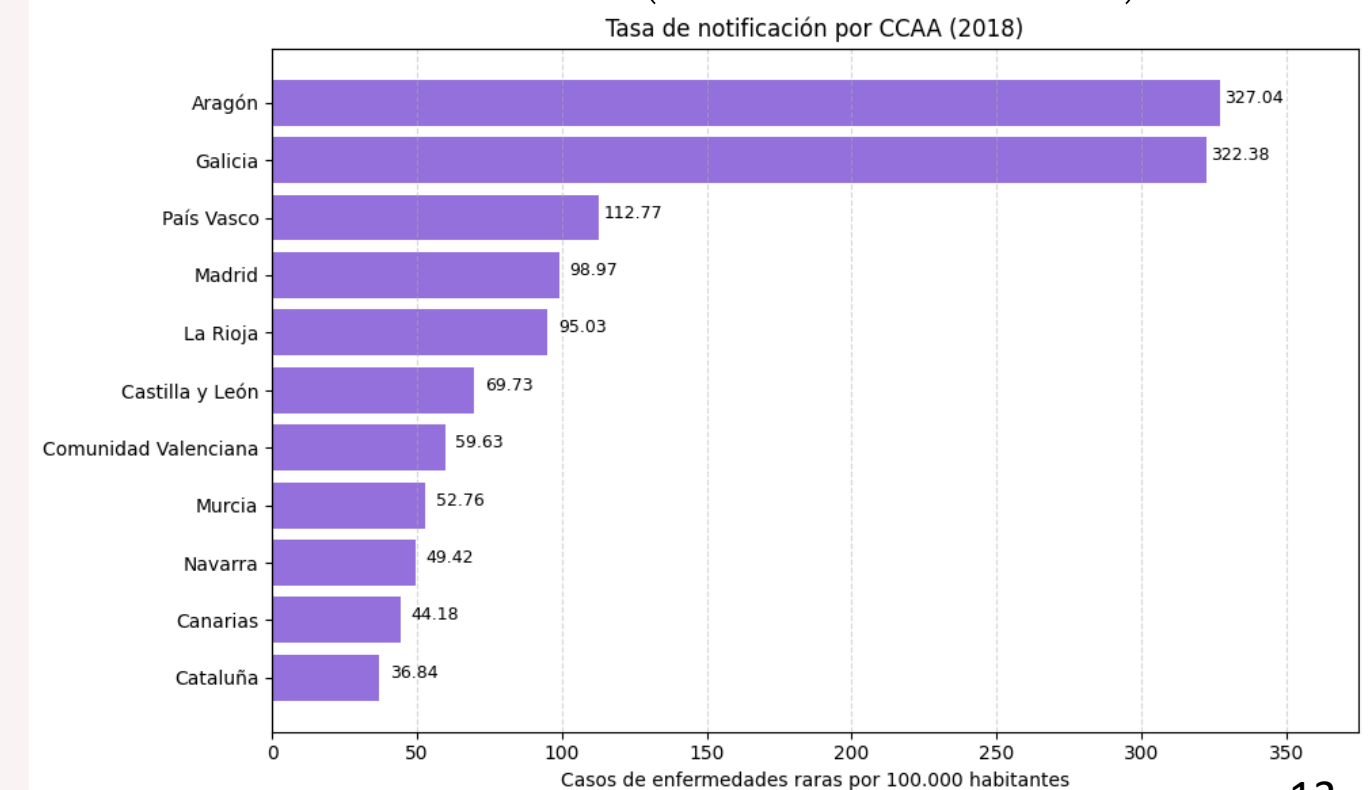
Gene Symbol	Full Name	Chromosome	Summary	NCBI Link
LMNA	lamin A/C	1	The protein encoded by this gene is part of the nuclear lam	https://www.ncbi.nlm.nih.gov/gene/
TP53	tumor protein p53	17	This gene encodes a tumor suppressor protein containing 1	https://www.ncbi.nlm.nih.gov/gene/

Resultados - Reek -

Incidencia de las ER en España (2010 - 2018)



$$\text{Tasa casos ER por 100.000 hab.} = \left(\frac{\text{Casos ER notificados en 2018}}{\text{Población total}} \right) \times 100.000$$

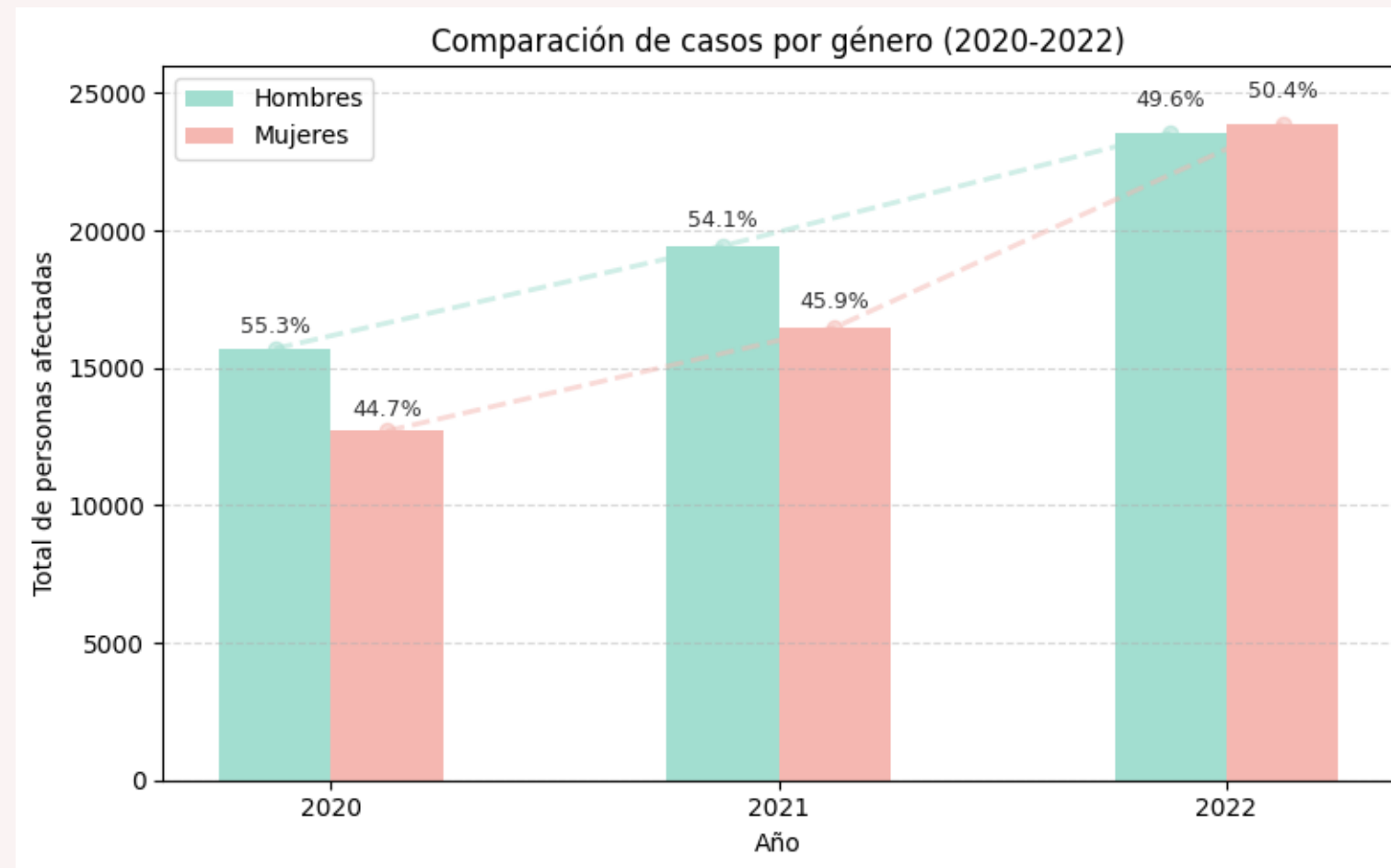


- ❑ Aumento de casos notificados en Galicia, Aragón y Madrid.
- ❑ Las comunidades con mayor tasa de notificación de ER por habitante son Aragón y Galicia, pero las más pobladas son Cataluña y Madrid.
- ❑ No hay relación directa entre cantidad de diagnósticos y población.

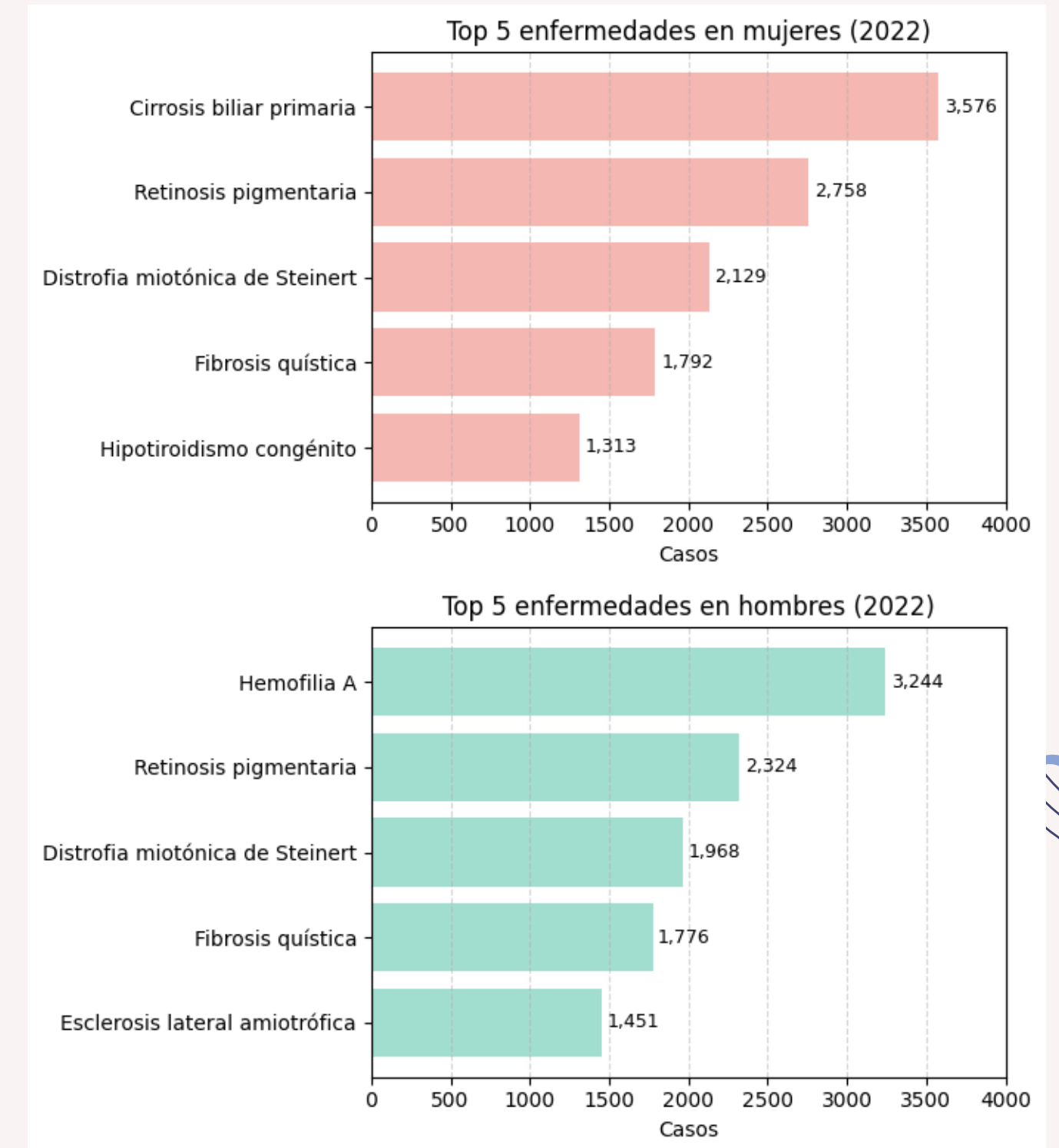
Resultados

- Reek -

Incidencia de las ER según género (2020 – 2022)

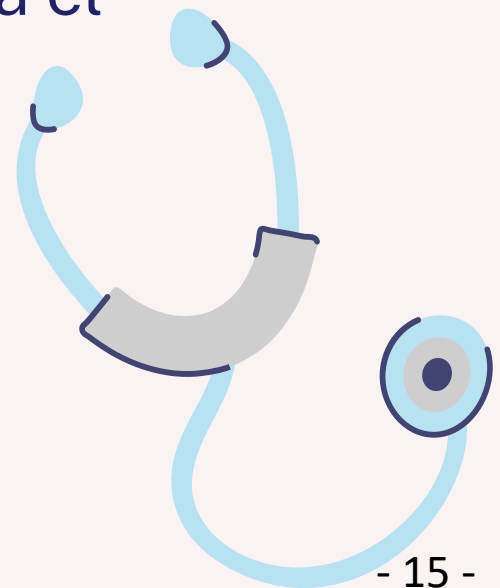


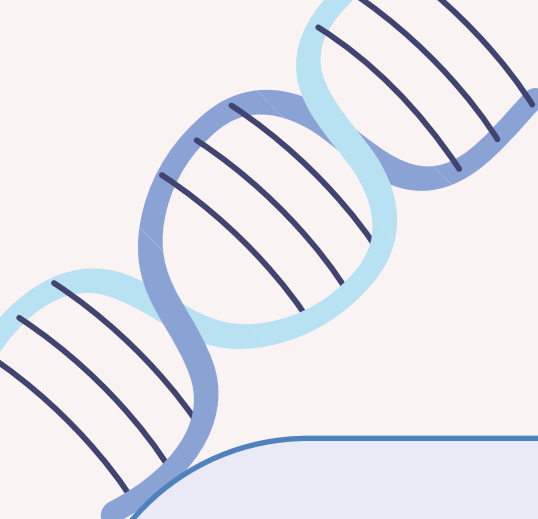
- ❑ Inicialmente los hombres presentaban una mayor incidencia (55%).
- ❑ En 2022 los diagnósticos en mujeres se incrementaron (50%).
- ❑ Tres de las cinco ER más notificadas presentan una incidencia similar.



Conclusiones


- ❑ Se confirma el **crecimiento** sostenido de la **investigación** en el ámbito de las **Enfermedades Raras**.
- ❑ **No existe** una **correlación** significativa entre el número de **genes** y **fenotipos** asociados a las Enfermedades Raras de origen genético.
- ❑ En España, las comunidades autónomas con **mayor población** no concentran necesariamente **más diagnósticos**.
 - ❑ Posibles **diferencias** en la **organización sanitaria** a nivel autonómico. Desigualdad estructural.
- ❑ El **análisis por género** reveló un **aumento** significativo de diagnósticos en **mujeres**, hasta el 50 % en 2022.
 - ❑ Posibles mejoras en el acceso al diagnóstico o una reducción de sesgos clínicos.





Trabajo futuro...

Limitaciones del estudio realizado

- Fuentes de datos incompletas:
 - API de Orphadata aún en fase de implementación.
 - Falta de homogeneidad en los criterios de notificación al Registro Estatal de enfermedades Raras entre las diferentes comunidades.
 - Acceso limitado a los datos.
- 

- Clasificación de enfermedades por sistema afectado (Aprendizaje Supervisado).

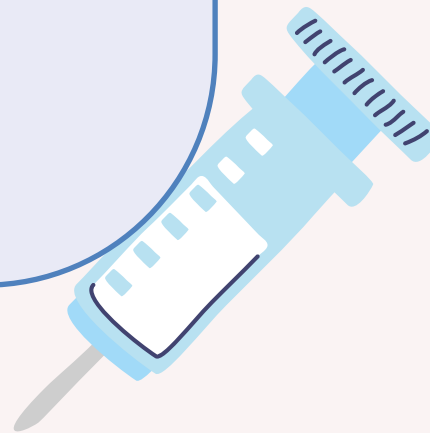
Predecir a qué sistema del cuerpo afecta una ER a partir del estudio de características (prevalencia, edad de inicio, fenotipos, etc.).

- Analizar las **interacciones** entre un fármaco huérfano y la proteína de un gen.

Por ejemplo: interacción CBD y proteínas de los genes causantes de las EED (SYNGAP1, SLC6A1, etc.).

- Estudiar rutas metabólicas e interferencias con proteínas similares.

- Estudiar fármacos químicamente similares.



¡Muchas gracias!

