

# Iniciación práctica al análisis de datos OMOP

## ¿Cómo empezar?



Anna Palomar

[apalomar@idiapjgol.org](mailto:apalomar@idiapjgol.org)

Marzo 2025

- ¿Por dónde empiezo?
  - OHDSI
  - OHDSI Spanish node
  - ¿Dónde encajo?
- Estudios en OMOP: pasos a seguir
  - Una sola base de datos
  - Estudios en red
- Retos y buenas prácticas

# ¿Cómo empezar? – OHDSI

- El libro de OHDSI - <http://book.ohdsi.org>
  - Se mantiene y **actualiza continuamente** por la comunidad
  - **Repositorio central** de conocimiento
  - Contenido:
    - Comunidad OHDSI
    - Explicación CDM y vocabularios estándar
    - Herramientas de código abierto
    - Calidad de datos
    - Estudios en red



**OHDSI**  
OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS

# ¿Cómo empezar? – OHDSI

Deberás registrarte para obtener una cuenta



## Join the OHDSI forum

OHDSI Tip: Follow topics to receive emails when new posts are added

## Introduce yourself!

Let the community know you're here by introducing yourself in the forum or at a community meeting

## Join an OHDSI meeting

Sit in on our weekly community meetings



## Join the OHDSI research network

By leading a study across the network

OR

By converting data to the OMOP Common Data Model

## Join an working group

Or start your own work group!



## Join the Journey

Help improve medical decision making today!

## Provide feedback

Identify and evaluate ways to use real-world evidence to inform decision making



OHDSI

Reuniones semanales  
Martes de 11 AM-12 PM (ET)

Múltiples working groups, veamos algunos...

# ¿Cómo empezar? – OHDSI grupos de trabajo

[Africa Chapter](#)

[APAC](#)

[ATLAS/WebAPI](#)

[Clinical Trials](#)

[Common Data Model](#)

[CDM Survey Subgroup](#)

[CDM Vocabulary Subgroup](#)

[Dentistry](#)

[Early-Stage Researchers](#)

[Electronic Animal Health Records](#)

[Eye Care & Vision Research](#)

[FHIR and OMOP](#)

[Generative AI & Analytics in Healthcare \(GAIA\)](#)

[GIS – Geographic Information System](#)

[HADES](#)

[Health Equity](#)

[Healthcare Systems](#)

[Industry](#)

[Latin America](#)

[Medical Devices](#)

[Medical Imaging](#)

[Methods Research](#)

[Natural Language Processing](#)

[Network Data Quality](#)

[Oncology](#)

[Open-Source Community](#)

[Patient-Level Prediction](#)

[Perinatal and Reproductive Health](#)

[Phenotype Development & Evaluation](#)

[Psychiatry](#)

[Rehabilitation](#)

[Registry](#)

[Steering Group](#)

[Surgery and Perioperative Medicine](#)

[Themis](#)

[Vaccine Vocabulary](#)



# ¿Cómo empezar? – Simposio europeo de OHDSI



 **OHDSI EUROPE'25 SYMPOSIUM**

Organized by 

 **5-7 July 2025**

 **Old Prison - Hasselt University, BELGIUM**

 **REGISTER NOW**

Envío abstracts – **31 Marzo 2025**

# ¿Cómo empezar? – OHDSI Spanish Node

- OHDSI Spanish Node

<https://www.ohdsi-europe.org/index.php/national-nodes/spain>



Dr. Talita Duarte Salles  
IDIAP Jordi Gol



Dr. Miguel Angel Mayer  
Consorti Parc de Salut Mar  
Barcelona (PSMAR)

## Objetivos

1. Adoptar y **promover la misión**, visión y valores de **OHDSI**.
2. **Facilitar la coordinación** entre los miembros de la red española y con otros nodos europeos e internacionales de OHDSI.
3. Contribuir a la coordinación y comunicación con las **instituciones políticas nacionales**.
4. Participar en la **comunidad de OHDSI**.

Formar parte



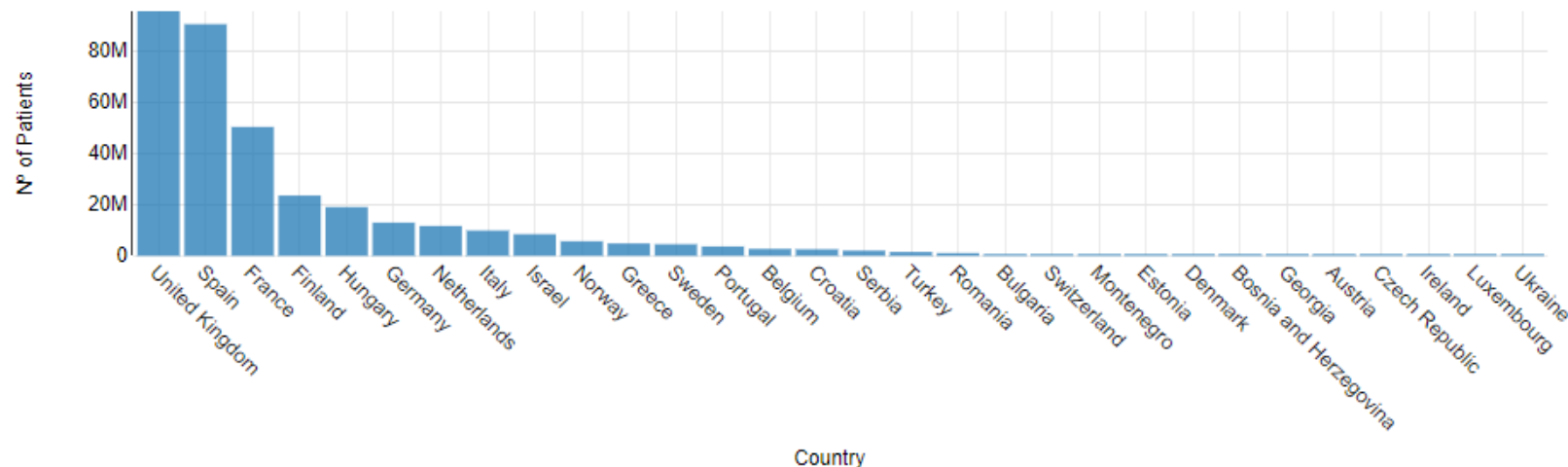
[https://docs.google.com/forms/d/e/1FAIpQLSdD3z5qMi9aqzO1ihDjKhaxcr-kXZKNn\\_HQBfhyAmiyRFTHzw/viewform](https://docs.google.com/forms/d/e/1FAIpQLSdD3z5qMi9aqzO1ihDjKhaxcr-kXZKNn_HQBfhyAmiyRFTHzw/viewform)



# ¿Cómo empezar? – OHDSI Spanish Node

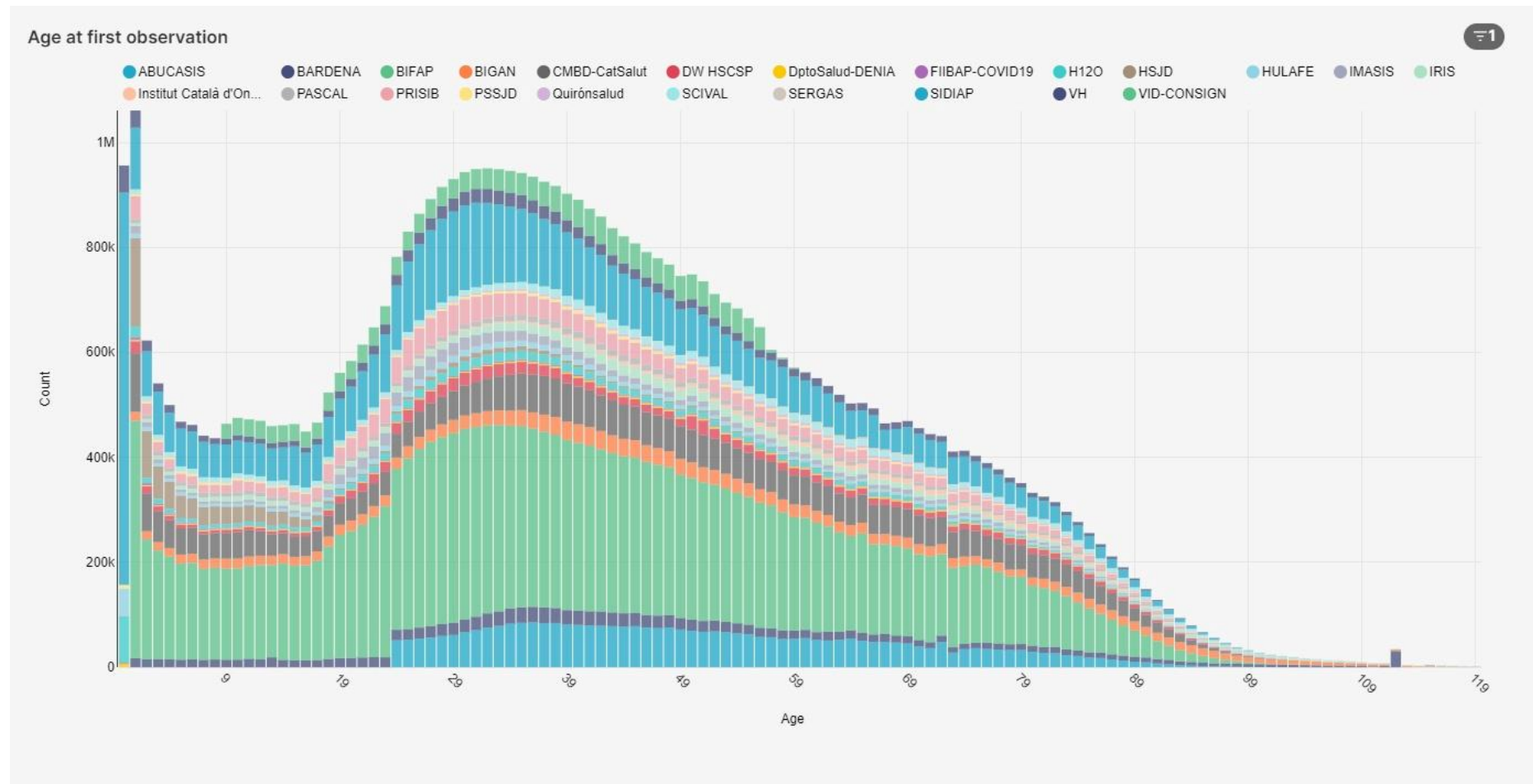
- **25 instituciones colaboradoras:** servicios de salud regionales, institutos de investigación, hospitales y pequeñas y medianas empresas certificadas por EHDEN.
- Este nodo abarca registros de salud de **más de 75 millones de pacientes** de **15 regiones de España**.

Patients by Country



# ¿Cómo empezar? – OHDSI Spanish Node

- **Distribución equitativa por sexo (52.7% mujeres) con una edad media de 42 años en la primera observación.**



Iniciación práctica al análisis de datos OMOP  
Grupo Real World Epidemiology  
10 - 13 Marzo 2025



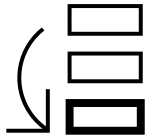
---

## Enlaces de interés

### Real World Data (RWD) and Common Data Models (CDM)

- Methodological Considerations when analyzing and interpreting Real-World data
  - [Link to Methodological Considerations](#)
- Common Problems, Common Data Model Solutions
  - [Link to Common Data Models Solutions](#)

# ¿Dónde encajo?

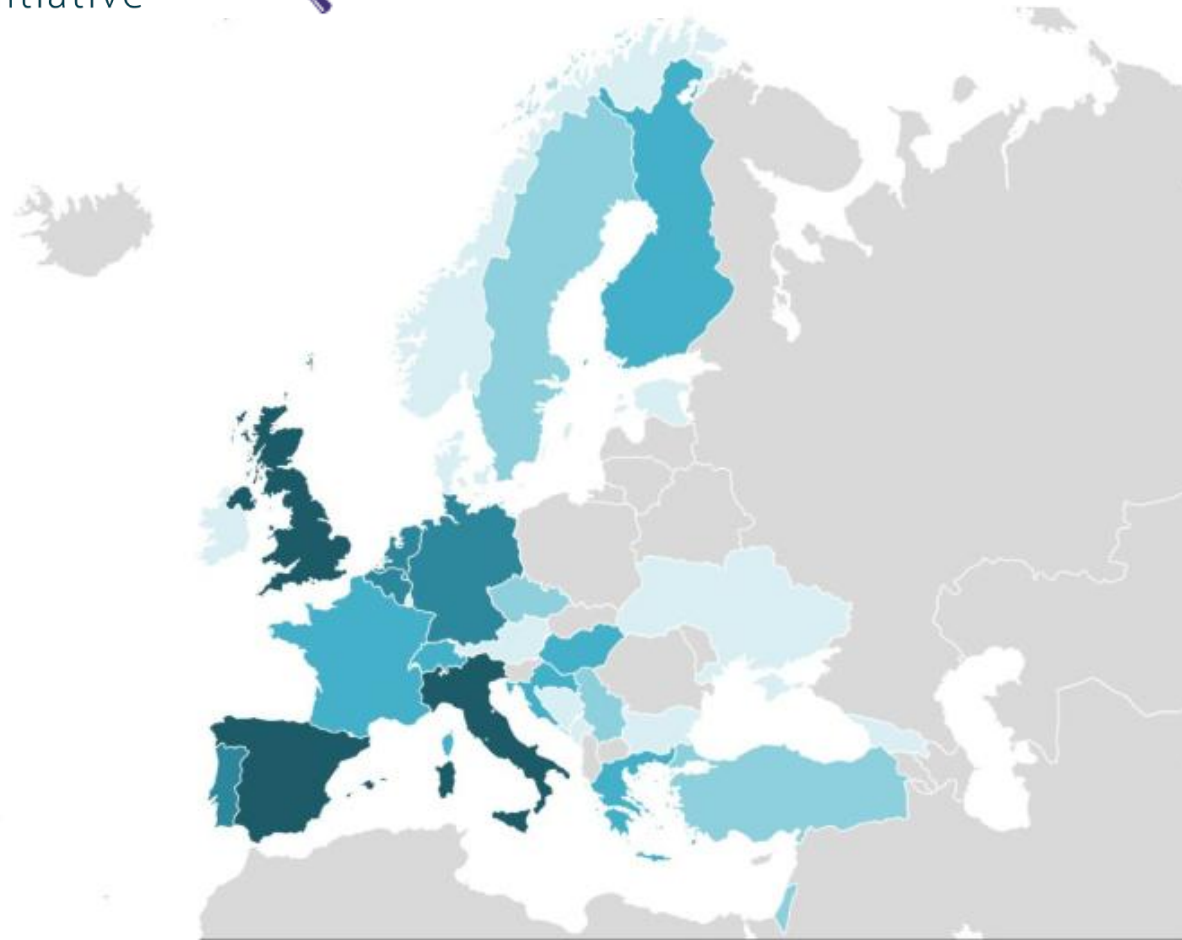


**Mapear** datos a OMOP



**Aprender** más sobre OHDSI y sus herramientas → [EHDEN academy](https://ehdenacademy.org/)

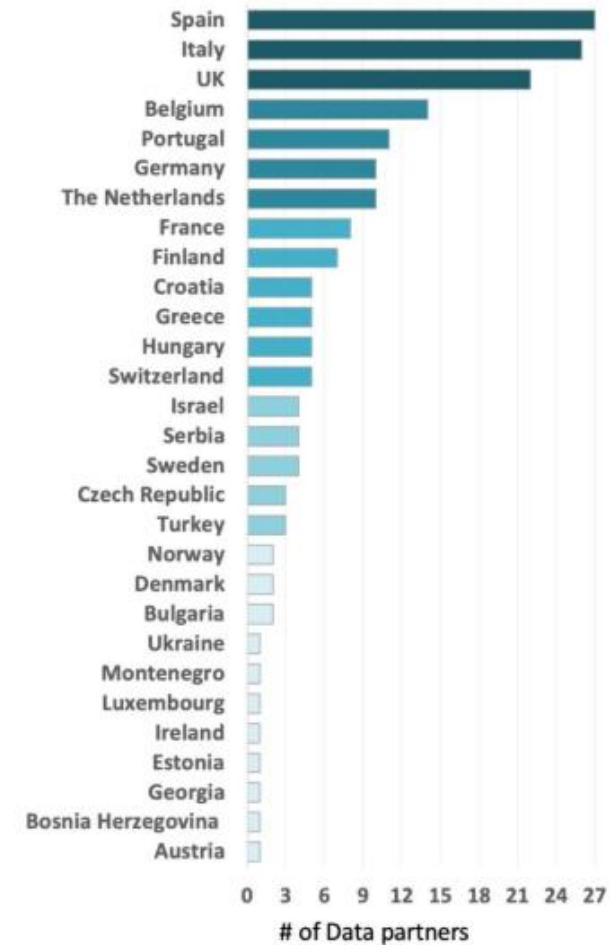
# ¿EHDEN?



Geographic spread of data partners. The shade of blue indicates the # of data partners in that country (darker = more)

Applications (n=563)

 Awarded applications (n=187)





Tool

OMOP CDM and  
Standardised  
Vocabularies

The structure of the common  
data model and its vocabularies.



Tool

ATLAS

Design and execute analyses on  
observational data.



Tool

Infrastructure

Install and configure the OHDSI  
infrastructure.



Tool

Extract, Transform and  
Load

Map raw observational data to  
the OMOP CDM.



Tool

Introduction to Usagi &  
Code Mappings for an  
ETL

Introduction to the Usagi tool,  
importing codes, review and  
output the code mapp...



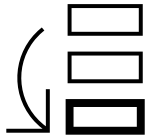
Tool

OHDSI-in-a-Box

Deploy a single instance  
implementation of OHDSI tools  
and sample data.



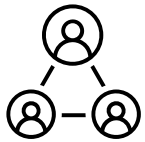
# ¿Dónde encajo?



**Mapear** datos a OMOP → [EHDEN academy](#)



**Aprender** más sobre OHDSI y sus herramientas → [EHDEN academy](#)



Buscar **colaboradores** → [OHDSI Researchers Topic](#) Forum OHDSI

# OHDSI Researchers Topic Forum OHDSI



Sign Up

Log In



topics of interest to the Research Network.

Launching the Generative AI and Analytics in Healthcare  
Workgroup

gaia



19

940

1d

Network Study: Seeking Data Partners in Rheumatology



10

440

12d

Network study: Concept Prevalence



31

2.8k

27d

Covariate Matching Based on Record Count

cdm, atlas, estimationmethods



0

14

Jan 24

Standardizing procedures (call for non-US based  
collaborators)



18

4.5k

Jan 13

Topics

More

Categories

General

Implementers

Developers

Researchers

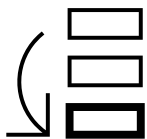
CDM Builders

All categories

Tags

cdm

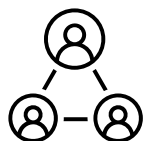
# ¿Dónde encajo?



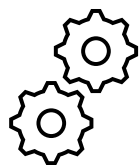
**Mapear** datos a OMOP → [EHDEN academy](#)



**Aprender** más sobre OHDSI y sus herramientas → [EHDEN academy](#)



Buscar **colaboradores** → [OHDSI Researchers Topic](#) Forum OHDSI



**Contribuir** al conjunto de **herramientas analíticas** de OHDSI → <https://github.com/OHDSI/>



Q Type / to search

Overview Repositories 321 Projects 14 Packages People 12



## Observational Health Data Sciences and Informatics

725 followers <http://ohdsi.org>

Follow

### Repositories

Q Find a repository...

Type

Language

Sort

152 results for all repositories written in R sorted by last updated

Clear filter

#### PhenotypeR

Public

R 2 Apache-2.0 0 27 1 Updated 39 minutes ago



#### CohortConstructor

Public

R 2 Apache-2.0 0 47 4 Updated 1 hour ago



#### OmopSketch

Public

R 1 Apache-2.0 0 7 0 Updated 2 hours ago



#### ClinicalCharacteristics

Public

[under development] table shell approach to OMOP characterization

R 1 Apache-2.0 1 11 1 Updated 4 hours ago



### People



### Top languages

R JavaScript Java Dockerfile  
HTML

### Most used topics

hades strategus-module  
patient-level-prediction skeleton

Iniciación práctica al análisis de datos OMOP

# Estudios en OMOP pasos a seguir



# Estudios observacionales en OMOP

1. Una sola base de datos
2. Múltiples bases de datos → Estudios en red



## 1. Definir la **pregunta de investigación**

- **Caracterización:** ¿Cuántos pacientes...? ¿Con qué frecuencia ...?
- **Estimación a nivel poblacional:** ¿Qué tratamiento funciona mejor? ¿Cuál es el riesgo de X en Y?
- **Predicción a nivel de paciente:** ¿Cuál es la probabilidad de que este paciente...? ¿Quiénes son candidatos para...?

## 2. Protocolo

## 3. Revisar la **disponibilidad** y la **calidad** de los datos

## 4. Definir la **población** de estudio

## 5. Factibilidad y **diagnóstico**

## 6. Ejecutar estudio

## 7. Interpretar resultados y diseminar



# Flujo de trabajo

Caracterizar la población con COVID-19 y estimar la incidencia anual y supervivencia



FENOTIPADO

COHORT o DRUG  
EXPOSURE  
DIAGNOSTICS

CARACTERIZACIÓN  
INCIDENCIA  
SUPERVIVENCIA

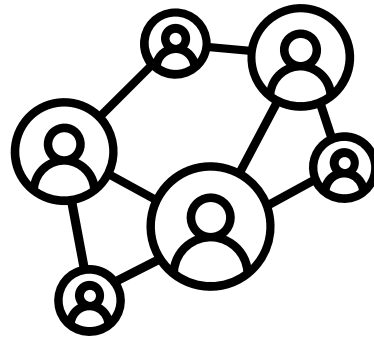
EVALUAR  
RESULTADOS

## ¿Por qué usarlos?

- Un estudio en una **sola base de datos** puede **no ser suficiente**, ya que los resultados pueden depender de **factores específicos del grupo de pacientes** (adherencia, genética, ambiente, estado de salud, etc.).
- Aumenta la **diversidad de fuentes y poblaciones**, evaluando si los resultados se **replican** en diferentes entornos.
- Identificar cómo los **factores "del mundo real"** afectan los hallazgos del estudio.

*Un estudio se considera en red cuando el código analítico se ejecuta en distintas bases. Esto es posible gracias al mapeo a OMOP CDM.*

Los estudios en red de OHDSI representan la culminación de una forma **transparente, consistente y reproducible** de llevar a cabo investigaciones con **múltiples datos en todo el mundo**.



## Enfoque Tradicional

- Se diseña el estudio con una **base de datos específica en mente**.
- Se **seleccionan códigos** de conceptos basándose en lo conocido del propio conjunto de datos.
- **Riesgo de sesgo** en la definición de cohortes al limitarse a un entorno o región específica.

## Enfoque estudio en red

- El estudio se diseña para ejecutarse en múltiples sitios globales, sin acceso directo a los datos de otras instituciones.
- Solo se comparten archivos de resultados, no datos brutos.
- Se requiere una **definición de cohortes exhaustiva** para cubrir la diversidad de entornos y bases de datos (CDM).

# Estudios en red: Ejemplos

## Long-Term Impact of the COVID-19 Pandemic on the Incidence of Depression Diagnoses and Antidepressant Prescribing: An International **Network** Cohort Study from Three European Countries

25 Pages • Posted: 6 Dec 2023

[Berta Raventós](#)

Fundació Institut Universitari per a la recerca a l'Atenció Primària de Salut Jordi Gol i Gurina (IDIAP)Gol)

[Carlen Reyes](#)

Fundació Institut Universitari per a la recerca a l'Atenció Primària de Salut Jordi Gol i Gurina (IDIAP)Gol)

[More](#)

## Characterising the background incidence rates of adverse events of special interest for covid-19 vaccines in eight countries: multinational **network** cohort study

*BMJ* 2021 ; 373 doi: <https://doi.org/10.1136/bmj.n1435> (Published 14 June 2021)

Cite this as: *BMJ* 2021;373:n1435

## The burden of post-acute COVID-19 symptoms in a multinational **network** cohort analysis

[Kristin Kostka](#), [Elena Roel](#), [Nhung T. H. Trinh](#), [Núria Mercadé-Besora](#), [Antonella Delmestri](#), [Lourdes Mateu](#),

[Roger Paredes](#), [Talita Duarte-Salles](#), [Daniel Prieto-Alhambra](#) , [Martí Català](#) & [Annika M. Jödicke](#)

## Comparative risk of thrombosis with thrombocytopenia syndrome or thromboembolic events associated with different covid-19 vaccines: international **network** cohort study from five European countries and the US

*BMJ* 2022 ; 379 doi: <https://doi.org/10.1136/bmj-2022-071594> (Published 26 October 2022)

Cite this as: *BMJ* 2022;379:e071594

Article | [Open access](#) | Published: 06 October 2020

## Deep phenotyping of 34,128 adult patients hospitalised with COVID-19 in an international **network** study

[Edward Burn](#), [Seng Chan You](#), [Anthony G. Sena](#), [Kristin Kostka](#), [Hamed Abedtash](#), [Maria Tereza F. Abrahão](#), [Amanda Alberga](#), [Heba Alghoul](#), [Osaid Alser](#), [Thamir M. Alshammari](#), [Maria Aragon](#), [Carlos Areia](#), [Juan M. Banda](#), [Jaehyeong Cho](#), [Aedin C. Culhane](#), [Alexander Davydov](#), [Frank J. DeFalco](#), [Talita Duarte-Salles](#), [Scott DuVall](#), [Thomas Falconer](#), [Sergio Fernandez-Bertolin](#), [Weihua Gao](#), [Asieh Golozar](#), [Jill Hardin](#), ... [Patrick](#)

[Ryan](#) [+ Show authors](#)

## Venous or arterial thrombosis and deaths among COVID-19 cases: a European **network** cohort study

[Edward Burn](#), DPhil • [Talita Duarte-Salles](#), PhD • [Sergio Fernandez-Bertolin](#), MSc • [Carlen Reyes](#), PhD •

[Kristin Kostka](#), MPH • [Antonella Delmestri](#), PhD • et al. [Show all authors](#)



# Estudios en red: ¿Cómo funciona?

**Colaboración entre investigadores** para responder a una pregunta de investigación.

Los *data partners* colaboran con datos y conocimiento en el estudio.



# Estudios en red: Pasos a seguir

1. Definir la pregunta de investigación
2. Protocolo
3. Revisar la disponibilidad y la calidad de los datos
4. Definir la población de estudio
5. Factibilidad y Diagnóstico
6. Finalización protocolo
7. Ejecutar estudio
8. Interpretar resultados y diseminar

# Estudios en red: Pasos a seguir

1. Definir la pregunta de investigación
- 2. Protocolo** ----->
3. Revisar la disponibilidad y la calidad de los datos
4. Definir la población de estudio
5. Factibilidad y Diagnóstico
6. Finalización protocolo
7. Ejecutar estudio
8. Interpretar resultados y diseminar

Cada *data partner* debe pasar las regulaciones específicas (comités científicos, éticos...)


# Estudios en red: Pasos a seguir

1. Definir la pregunta de investigación
2. Protocolo
- 3. Revisar la disponibilidad y la calidad de los datos**
4. Definir la población de estudio
5. Factibilidad y Diagnóstico
6. Finalización protocolo
7. Ejecutar estudio
8. Interpretar resultados y diseminar



- EHDEN – portal: <https://www.ehden.eu/>





Ana Palomar


- HOME
- CATALOGUE**
- DASHBOARD
- ACADEMY
- EHDEN
- PUBLICATIONS
- STATUS
- PORTAL
- ABOUT
- GET STARTED
- PROFILE
- SIGN OUT

## Databases / SIDIAP

- Fingerprint
- Database Dashboard**


### Database-Level Dashboard

- Demographics**
- Data Domains
- Data Provenance
- Observation Period
- Visit
- Concept Browser**
- Meta Data

Number of Patients	Gender	Gender						
7.89M	<div>gender</div> <table> <tr> <th></th> <th>number of records</th> </tr> <tr> <td>FEMALE</td> <td>3.96M</td> </tr> <tr> <td>MALE</td> <td>3.93M</td> </tr> </table>		number of records	FEMALE	3.96M	MALE	3.93M	<div>MALE FEMALE</div> 
	number of records							
FEMALE	3.96M							
MALE	3.93M							

- EHDEN – portal: <https://www.ehden.eu/>





Ana Palomar

HOME

CATALOGUE >

DASHBOARD

ACADEMY

EHDEN

PUBLICATIONS

STATUS

PORTAL

ABOUT

GET STARTED

PROFILE

SIGN OUT

## Database-Level Dashboard

Demographics Data Domains Data Provenance Observation Period Visit **Concept Browser** Meta Data

The concepts browser allows you to search for concepts by name or concept\_id in all the data sources you select. No exact number of patients or occurrences are provided but they are rounded up to the nearest multiple of 100. Note that the number of patients and occurrences are for that concept only and does not include the children of that code.

### Domain Filter

ENTITY TYPE

Type or Select [Entity Type]

### Concept Browser

Show 10 entries

Search 13256 records...

concept_id	concept_name	domain_id	drc	rc
3025315	Body weight	Measurement	52167200	52167200
3038553	Body mass index (BMI) [Ratio]	Measurement	36585900	36585900
3036277	Body height	Measurement	27084100	27084100



# Estudios en red: Pasos a seguir

1. Definir la pregunta de investigación
2. Protocolo
3. Revisar la disponibilidad y la calidad de los datos
- 4. Definir la población de estudio**
- 5. Factibilidad y Diagnóstico**
6. Finalización protocolo
7. Ejecutar estudio
8. Interpretar resultados y diseminar

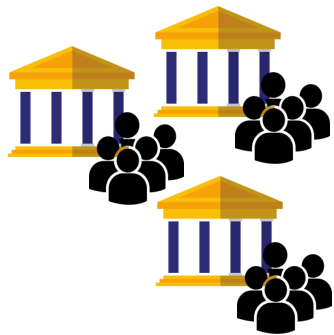
# Estudios en red: *Phenotyping*



**Definir *phenotypes***: Los investigadores principales hacen una **primera versión** que compartirán con los *data partners* para validarla en cada base de datos.



**Ejecutar validación**: Se comparte un repositorio con las definiciones de las cohortes y el código necesario (e.g. CohortDiagnostics) para **estudiar y validar cada una de las definiciones**. Cada *data partner* ejecuta el código y **revisa internamente** los resultados.



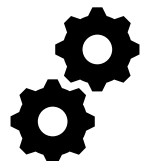
**Validación *phenotypes***: Los *data partners* comparten los resultados con los investigadores principales. ***Data partners* e IP trabajan juntos para mejorar las definiciones.**

Proceso iterativo (2-3 reruns)

# Estudios en red: Pasos a seguir

1. Definir la pregunta de investigación
2. Protocolo
3. Revisar la disponibilidad y la calidad de los datos
4. Definir la población de estudio
5. Factibilidad y Diagnóstico
6. Finalización protocolo
- 7. Ejecutar estudio**
- 8. Interpretar resultados y diseminar**

# Estudios en red: Ejecución estudio



**Crear código analítico:** Los investigadores principales desarrollan el código que va a ser compartido con los *data scientist* de los centros colaboradores.



**Lanzar código:** Cada *data partner* ejecuta el código de estudio de forma local en su base de datos, y **revisa** los resultados de forma interna.



**Revisión resultados:** Los *data partners* comparten los **resultados agregados** (NO información a nivel individual) con los investigadores principales.

Pueden **surgir errores en el código** (e.g. distintas distribuciones de SQL), se pueden obtener resultados inesperados (e.g. diferencias en el mapeo).

**Proceso iterativo (2-3 reruns)**

- Shiny app

For each cohort in each database comparing vaccinated with unvaccinated people, covariate balance was calculated before and after propensity score weighting, with adequate balance achieved for all covariates after weighting, except for GP practice in CPRD AURUM and CPRD GOLD (for cohorts two and three only; results are accessible in the [interactive web application](#)). Propensity scores and overlap weight distributions for the primary outcome are included in the [appendix \(pp 132–133\)](#). Baseline characteristics for cohorts one to four are shown in [table 1](#) for CPRD AURUM, and in the appendix for CPRD GOLD, SIDIAP, and CORIVA (pp 16–17, 26–29, 38–41, and 48–51), with all tables showing comparability of study participants when vaccinated and unvaccinated groups were compared after weighting. Negative control outcome analyses suggested some residual confounding after weighting; therefore, in the subsequent sections we report empirically calibrated estimates and uncalibrated estimates were reported only in the web application.

Menu

Databases

Cohort details

Propensity scores

Negative control outcomes

Estimates

Kaplan-Meier plots

Censor data

Details of the databases that participated in the study

Identifier 'cdm\_name' is the one used in the multiple selection panels of the shiny

	cdm_name	cdm_source_name	cdm_version	cdm_holder	cdm_release_date	vocabulary_version	person_cnt	observation_period_cnt
1	CPRD AURUM	cdm_aurum_p21_000557	5.3	OMOP	2018-06-14	v5.0 29-AUG-22	10,592,564	10,592,564
2	CORIVA	CORIVA_covid_data_from_Feb_2023_NEW	5.3.1	UT		v5.0 19-NOV-21	438,433	438,433
3	CPRD GOLD	cdm_gold_202201	5.3	OMOP		v5.0 09-APR-22	16,958,587	16,958,587
4	SIDIAP	SIDIAP	5.3.1	SIDIAP	2022-11-14	v5.0 22-JUN-22	8,265,343	8,265,343

Showing 1 to 4 of 4 entries

Iniciación práctica al análisis de datos OMOP

# Retos y buenas prácticas



# Retos y buenas prácticas

**R:** ¿Cómo **compartir** el código analítico, jsons de cohortes, shiny app, etc. ?

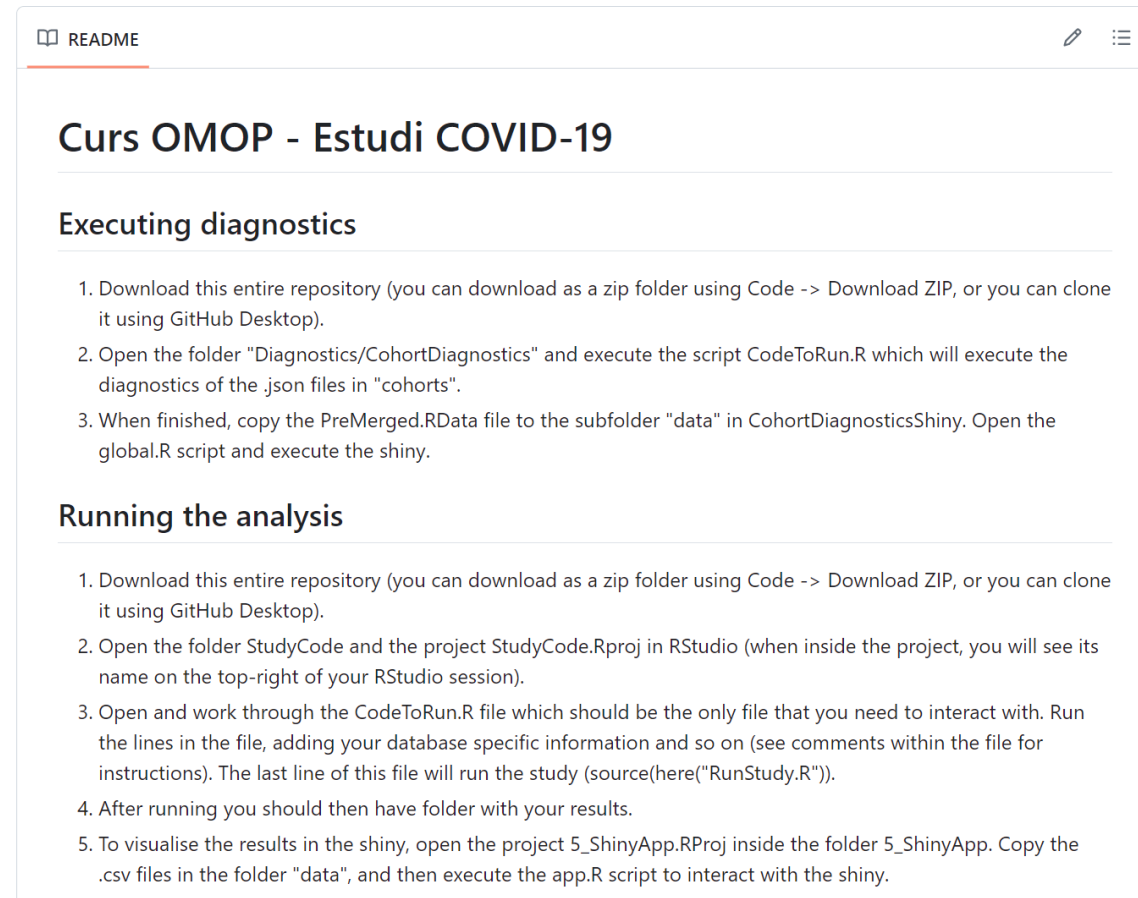
**BP:** Utilizar un repositorio de **GitHub** (como el del curso)

The screenshot shows a GitHub repository interface for 'Curso\_Introduccion\_OMOP\_2025', which is marked as 'Private'. The repository is owned by 'annapalomarc' and contains a folder named 'PatientProfiles'. The interface includes a 'Watch' button, a 'main' branch selector, and a search bar. Below the repository header, there is a table of recent commits.

Commit Message	Author	Time
Update README.md	annapalomarc	4 days ago
PatientProfiles	annapalomarc	2 hours ago
Update README.md	annapalomarc	yesterday
update	annapalomarc	yesterday
main	annapalomarc	2 weeks ago

**R:** ¿Cómo **guiarse** dentro de un repo que te han compartido?

**BP:** **README** en el GitHub con las **instrucciones** para ejecutar el código.



The screenshot shows a GitHub README file for a repository titled "Curs OMOP - Estudi COVID-19". The file is titled "README" in the top left corner. The content is organized into two main sections: "Executing diagnostics" and "Running the analysis", each with a list of numbered steps.

## Curs OMOP - Estudi COVID-19

### Executing diagnostics

1. Download this entire repository (you can download as a zip folder using Code -> Download ZIP, or you can clone it using GitHub Desktop).
2. Open the folder "Diagnostics/CohortDiagnostics" and execute the script CodeToRun.R which will execute the diagnostics of the .json files in "cohorts".
3. When finished, copy the PreMerged.RData file to the subfolder "data" in CohortDiagnosticsShiny. Open the global.R script and execute the shiny.

### Running the analysis








1. Download this entire repository (you can download as a zip folder using Code -> Download ZIP, or you can clone it using GitHub Desktop).
2. Open the folder StudyCode and the project StudyCode.Rproj in RStudio (when inside the project, you will see its name on the top-right of your RStudio session).
3. Open and work through the CodeToRun.R file which should be the only file that you need to interact with. Run the lines in the file, adding your database specific information and so on (see comments within the file for instructions). The last line of this file will run the study (source(here("RunStudy.R"))).
4. After running you should then have folder with your results.
5. To visualise the results in the shiny, open the project 5\_ShinyApp.Rproj inside the folder 5\_ShinyApp. Copy the .csv files in the folder "data", and then execute the app.R script to interact with the shiny.



# Retos y buenas prácticas

**R:** ¿Cómo **trabajar con un código** que no has escrito tu?

**BP:** Tener un **proyecto de R** y un único script “**CodeToRun**” o similar que sea el único con el que se tenga que interactuar para ejecutar el código.

	cohorts	11/12/2024 11:34	Carpeta de archivos	
	concept_sets	11/12/2024 11:34	Carpeta de archivos	
----->	 2_study.Rproj	04/12/2024 11:25	Archivo RPROJ	1 KB
----->	 codeToRun.R	04/12/2024 11:25	Archivo R	2 KB
	 README.md	04/12/2024 11:25	Archivo MD	7 KB
	 renv.lock	04/12/2024 11:25	Archivo LOCK	19 KB
	 RunStudy.R	04/12/2024 11:25	Archivo R	5 KB

# Retos y buenas prácticas

**R:** ¿Cómo trabajar con un código que no has escrito tu?

**BP: CodeToRun** definir las variables de R específicas de cada *data partner*

```
# The name/ acronym for the database
dbName <- "..."
```

```
# Database connection details
db <- dbConnect(
  RPostgres::Postgres(),
  dbname = server_dbi,
  port = port,
  host = host,
  user = user,
  password = password
)
```

Detalles de conexión a la base de datos

```
# The name of the schema that contains the OMOP CDM with patient-level data
cdmSchema <- "..."
```

```
# A prefix for all permanent tables in the database
writePrefix <- "..."
```

Prefijo de conexión.

```
# The name of the schema where results tables will be created
writeSchema <- "..."
```

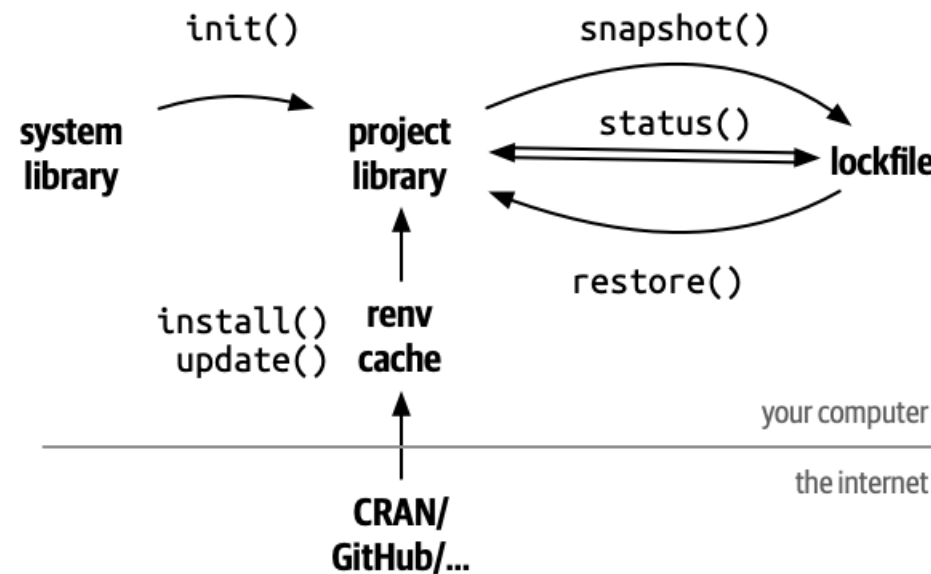
```
# minimum counts that can be displayed according to data governance
minCellCount <- 5
```

Numero de eventos mínimos para reportar resultados.

# Retos y buenas prácticas

**R:** ¿Qué pasa con las diferentes actualizaciones y **versiones** de los **paquetes**?

**BP:** El paquete de R “**renv**” permite **reproducir el entorno** R del proyecto en el que se ha desarrollado el código.



# Retos y buenas prácticas

**R:** ¿Cómo detectar el **error** en un **código** que no has escrito tu?

**BP:** Implementar herramientas para **seguir el progreso de la ejecución código** (e.g. paquete **log4r**).

```
1 INFO [2024-04-10 00:50:06] CREATE_LOGGER
2 INFO [2024-04-10 00:50:06] STEP 0 INITIAL SETTINGS ----
3 INFO [2024-04-10 00:50:07] STEP 1 INSTANTIATE COHORTS ----
4 INFO [2024-04-10 00:50:07]   - JSON cohorts
5 INFO [2024-04-10 00:50:26]   - Denominator cohort
6 INFO [2024-04-10 00:51:08] STEP 2 CALCULATE INCIDENCE ----
7 INFO [2024-04-10 00:52:47] STEP 3 CHARACTERISATION ----
8 INFO [2024-04-10 00:53:15] STEP 4 SURVIVAL ----
9 INFO [2024-04-10 00:53:41] STEP 5 ZIP RESULTS ----
10 INFO [2024-04-10 00:53:42] -- DONE! --
11
```

**R:** ¿Qué pasa si...? Se detecta un **error** al final del código, se hacen **cambios** en una parte del código, **se para la ejecución** por algún imprevisto (corte de luz)

**BP:** Separar en partes **independientes distintos procesos del análisis**.

```
79      # Jobs to Run
80      ejecutarInstanciarCohorts <- TRUE
81      ejecutarCaracteritzacio   <- TRUE
82      ejecutarIncidencia        <- TRUE
83      ejecutarSupervivencia     <- TRUE
```

**R:** ¿Puedo reaprovechar código analítico de estudios similares?

**BP:** Código abierto.

oxford-pharmacoepi / LongcovidVaccineEffectiveness

Issues Pull requests Actions Projects Wiki Security Insights

LongcovidVaccineEffectiveness Public

main 1 Branch 0 Tags

Go to file

Add file Code

nmercadeb update with revision codes a0b3835 · 3 months ago 8 Commits		
1_InstantiateCohorts	extract censoring data	9 months ago
3_PropensityScores	compute censor	9 months ago
4_OutcomeModel	revision	5 months ago
5_ShinyApp	update with revision codes	3 months ago
Data	Delete Data/dataSIDIAP directory	last year
.gitattributes	public code	last year
.gitignore	extract censoring data	9 months ago
CodeToRun.R	public code	last year
Longcovid_vaccination_effectiveness.Rproj	public code	last year
README.md	public code	last year
RunStudy.R	public code	last year
functions.R	compute censor	9 months ago

README

## The effectiveness of COVID-19 vaccines to prevent long COVID symptoms: staggered cohort study of data from the UK, Spain, and Estonia

Martí Català, Nùria Mercadé-Besora, Raivo Kolde, Nhung T H Trinh, Elena Roel, Edward Burn, Trishna Rathod-Mistry, Kristin Kostka, Wai Yi Man, Antonella Delmestri, Hedvig M E Nordeng, Anneli Uuskula, Talita Duarte-Salles, Daniel Prieto-Alhambra\*, Annika M Jödicke\*

### About

No description, website, or topics provided.

Readme  
Activity  
Custom properties  
0 stars  
3 watching  
0 forks  
Report repository

### Releases

No releases published  
[Create a new release](#)

### Packages

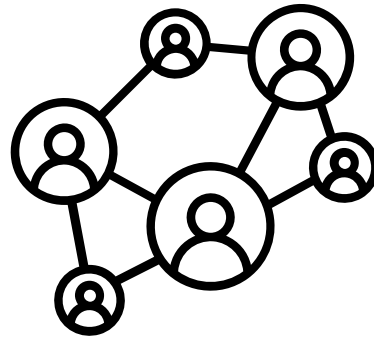
No packages published  
[Publish your first package](#)

### Contributors 2

catalamarti  
nmercadeb Nùria Mercadé-Besora

**R:** ¿Qué pasa si (cómo IP) vemos **resultados inesperados** en otra base de datos?

**BP:** Los *data partners* son quien mejor **conocen la base de datos** y sus **particularidades** (cómo se diagnostica cierta enfermedad en el país de origen, si el fármaco en estudio se prescribe o no, etc). Es clave el **trabajo en equipo** y la **comunicación**



# Retos y buenas prácticas

**R:** ¿Qué pasa si (cómo IP) vemos **resultados inesperados** en otra base de datos?

**BP:** Es muy útil hacer un ***study-a-thon***, un encuentro de 3-4 días dónde se discuten presencialmente los resultados de un estudio en red.





**¡Muchas gracias por vuestra atención!**