

# HULA R! Potencialidades de R na investigación sanitaria

Adrián Casanova, Andrés Blanco, Inés Sambade, Rocío Valín e Pilar Rodríguez-Ledo

Unidade de Investigación Clínica de Lugo (UNICL), Hospital Universitario Lucus Augusti (HULA), Fundación Pública Galega Instituto de Investigación Sanitaria de Santiago de Compostela (FIDIS), Lugo, España

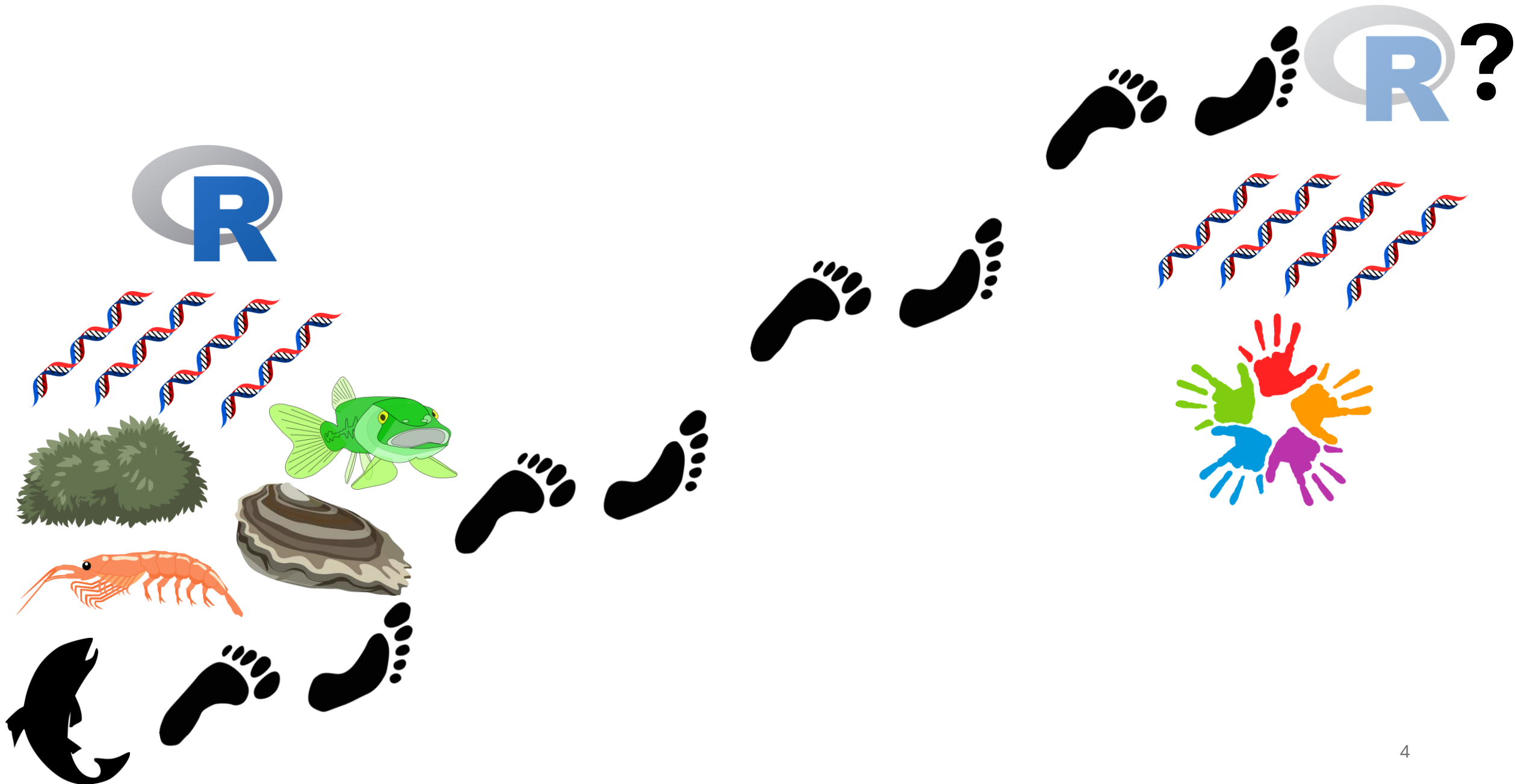
# Índice

**Introducción**

**Utilidades de R**

**Conclusiones**

# INTRODUCCIÓN



## El Hula abre su unidad de investigación y recibe 1,3 millones para desarrollarla

El conselleiro de Sanidade inauguró las instalaciones, que sumarán seis trabajadores gracias a una subvención del Instituto Carlos III. Contará con una unidad de telemedicina y con un secuenciador de nueva generación



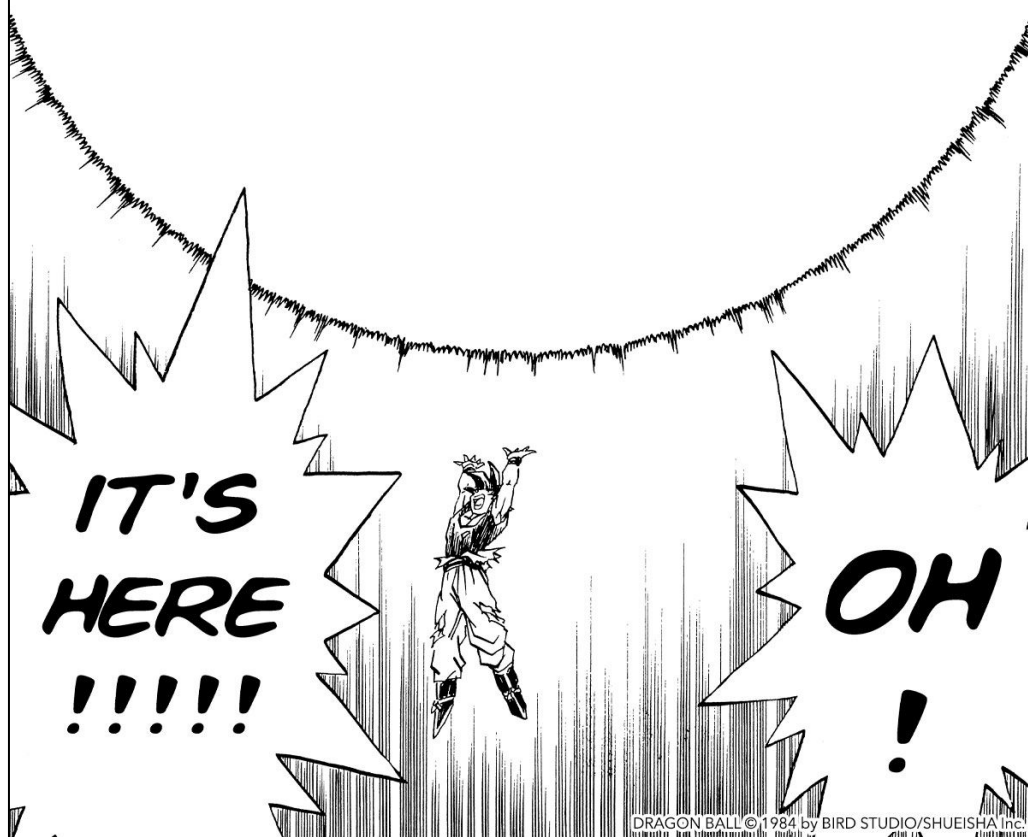
Instante en que el conselleiro de Sanidade corta la cinta. SEBAS SENANDE



MARÍA PIÑEIRO X

20/12/24 | 9:00 Actualizado: 20/12/24 | 9:18 | TIEMPO DE LECTURA: 5 MIN.

## *Big data* biolóxico



Fonte: Toriyama, Akira. *Dragon Ball*, capítulo 515. Shueisha (1995).

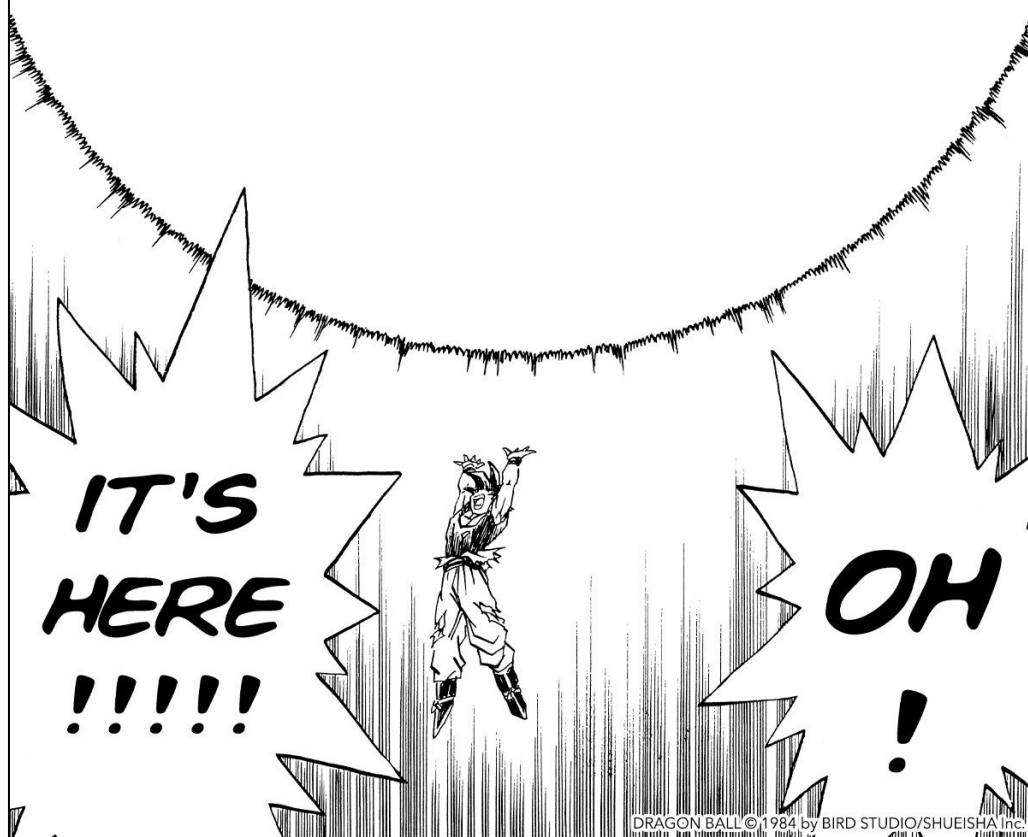
Medicina



traslacional



## *Big data* biolóxico

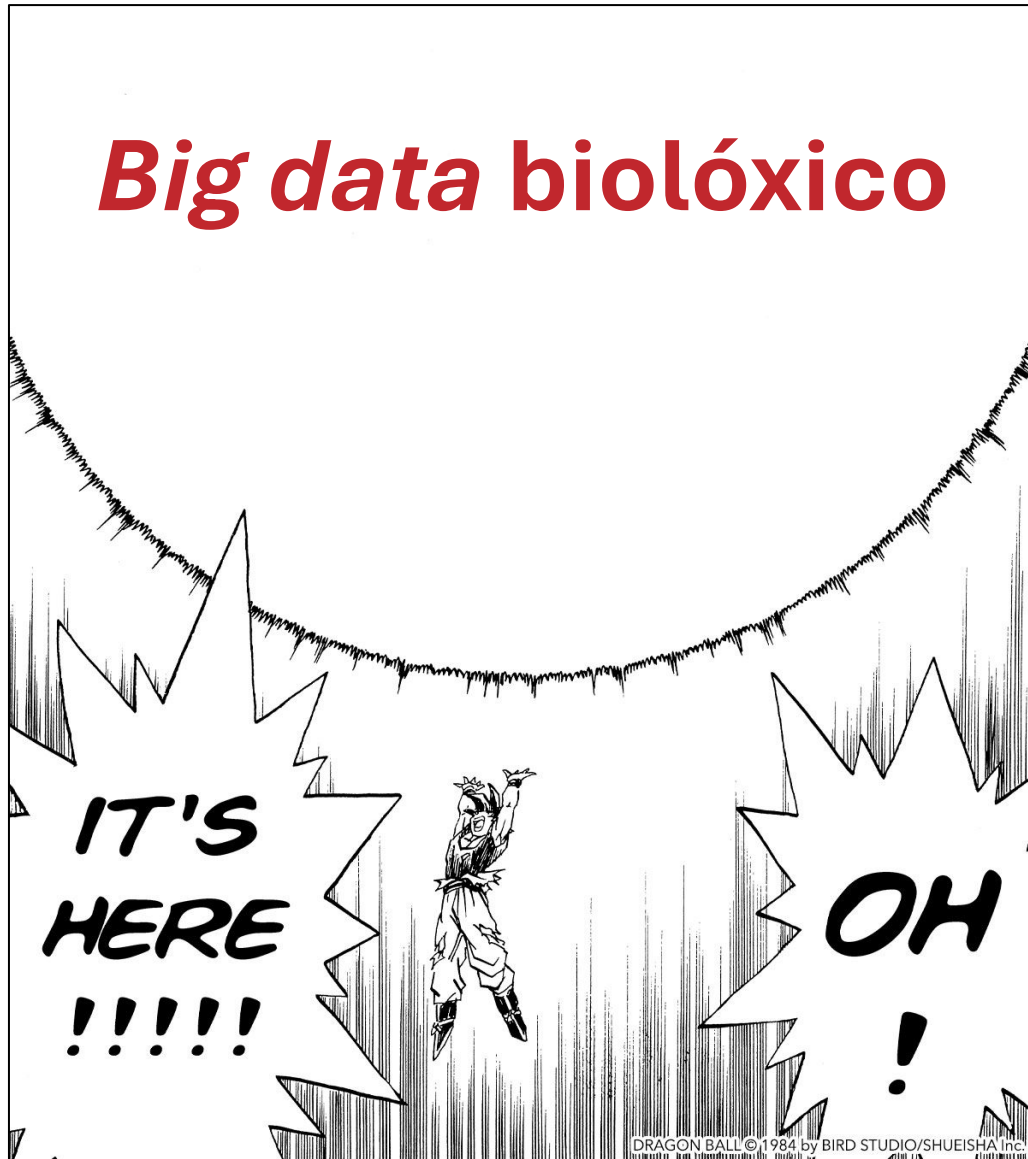


Fonte: Toriyama, Akira. *Dragon Ball*, capítulo 515. Shueisha (1995).

Como?



## ***Big data biolóxico***



Fonte: Toriyama, Akira. *Dragon Ball*, capítulo 515. Shueisha (1995).

☐ Capacidade de computación



☐ Ferramentas bioinformáticas

## **Cales?**

# Por que traballar con R na investigación sanitaria?

- ❑ A súa **polivalencia bioinformática** grazas ao amplo abano de paquetes aplicables en diferentes campos **{gam}** como os creados, alomenos inicialmente, para campos específicos **{OptimalCutpoints}**
- ❑ Presenta unha masa crítica que garante a **viabilidade** deste proxecto a **longo prazo** e facilita a existencia de **multitude de recursos** (e.g., titoriais, libros)
- ❑ A dispoñibilidade de R en **todos os sistemas operativos** e no Centro de Supercomputación de Galicia (**CESGA**)
- ❑ Aliñación cos principios FAIR para o Software de Investigación (FAIR4RS)

scientific **data**OPEN  
ARTICLE

## Introducing the FAIR Principles for research software

Michelle Barker<sup>1</sup>✉, Neil P. Chue Hong<sup>2</sup>, Daniel S. Katz<sup>3</sup>, Anna-Lena Lamprecht<sup>4</sup>, Carlos Martínez-Ortiz<sup>5</sup>, Fotis Psomopoulos<sup>6</sup>, Jennifer Harrow<sup>7</sup>, Leyla Jael Castro<sup>8</sup>, Morane Gruenpeter<sup>9</sup>, Paula Andrea Martínez<sup>10</sup> & Tom Honeyman<sup>11</sup>

Research software is a fundamental and vital part of research, yet significant challenges to discoverability, productivity, quality, reproducibility, and sustainability exist. Improving the practice of scholarship is a common goal of the open science, open source, and FAIR (Findable, Accessible, Interoperable and Reusable) communities and research software is now being understood as a type of digital object to which FAIR should be applied. This emergence reflects a maturation of the research community to better understand the crucial role of FAIR research software in maximising research value. The FAIR for Research Software (FAIR4RS) Working Group has adapted the FAIR Guiding Principles to create the FAIR Principles for Research Software (FAIR4RS Principles). The contents and context of the FAIR4RS Principles are summarised here to provide the basis for discussion of their adoption. Examples of implementation by organisations are provided to share information on how to maximise the value of research outputs, and to encourage others to amplify the importance and impact of this work.

Barker M., Chue Hong N.P., Katz D.S., Lamprecht A-L., Martínez-Ortiz C., Psomopoulos F., Harrow J., Castro L.J., Gruenpeter M., Martínez P.A., Honeyman T. (2022). Introducing the FAIR Principles for research software. *Scientific Data* 9, 622.

<https://doi.org/10.1038/s41597-022-01710-x>

**FAIR4RS:**

Findable

Accessible

Interoperable

Reusable

for

Research

Software

1. **Atopables:** grazas a repositorios públicos como CRAN (*Comprehensive R Archive Network*)
2. **Accesibles:** xa que a súa natureza de código aberto e a súa gratuidade elimina barreiras de licenzas privativas ou económicas
3. **Interoperables:** pola súa capacidade para comunicarse con outras linguaxes coma Python (e.g., paquete de R **{reticulate}**), manexando ducias de formatos de datos de maneira máis estandarizada
4. **Reutilizables:** porque a filosofía de R promove que una investigación poida ser o punto de partida para outra

# UTILIDADES DE R

- ❑ A versión de **R** empregada foi a **4.5.1 “Great Square Root”** traballando nun Entorno de Desenvolvemento Integrado (EDI) proporcionado pola versión de **RStudio 2025.09.0+387**
- ❑ Os **datos** mostrados foron xerados **aleatoriamente** con Intelixencia Artificial (*Google Gemini*), se ben baseados en situacións realistas

```
datos_pacientes<-read.csv("datos_pacientes.csv",header = TRUE,sep = ",",dec = ".")
```

## FORMACIÓN

*Comandos, sen medo*

|    | datos_pacientes.csv                                                                                                          |
|----|------------------------------------------------------------------------------------------------------------------------------|
| 1  | "ID_Paciente","Idade","Sexo","Fumador","Actividade_Fisica","Colesterol_Total_mg_dl","Tension_Arterial_Sistolica","Patoloxia" |
| 2  | "PAC_1001",78,"Home","No","Moderada",183,155,"Diabetes"                                                                      |
| 3  | "PAC_1002",66,"Home","Si","Alta",181,136,"Hipertensión"                                                                      |
| 4  | "PAC_1003",30,"Home","No","Moderada",234,129,"Infarto miocardio"                                                             |
| 5  | "PAC_1004",54,"Home","No","Baixa",164,129,"Sen patoloxía aparente"                                                           |
| 6  | "PAC_1005",39,"Muller","No","Alta",261,119,"Sen patoloxía aparente"                                                          |
| 7  | "PAC_1006",65,"Muller","Exfumador","Alta",203,139,"Hipertensión"                                                             |
| 8  | "PAC_1007",47,"Home","No","Baixa",197,121,"Sen patoloxía aparente"                                                           |
| 9  | "PAC_1008",78,"Home","No","Moderada",215,164,"Hipertensión"                                                                  |
| 10 | "PAC_1009",76,"Muller","No","Moderada",201,124,"Sen patoloxía aparente"                                                      |
| 11 | "PAC_1010",53,"Muller","Exfumador","Moderada",196,118,"Sen patoloxía aparente"                                               |
| 12 | "PAC_1011",36,"Muller","No","Baixa",244,132,"Sen patoloxía aparente"                                                         |
| 13 | "PAC_1012",65,"Home","Si","Alta",233,111,"Infarto miocardio"                                                                 |
| 14 | "PAC_1013",54,"Muller","No","Moderada",161,146,"Sen patoloxía aparente"                                                      |
| 15 | "PAC_1014",66,"Muller","No","Alta",212,123,"Sen patoloxía aparente"                                                          |
| 16 | "PAC_1015",75,"Home","No","Baixa",189,133,"Sen patoloxía aparente"                                                           |
| 17 | "PAC_1016",49,"Muller","No","Baixa",184,131,"Sen patoloxía aparente"                                                         |
| 18 | "PAC_1017",55,"Muller","No","Alta",168,130,"Sen patoloxía aparente"                                                          |
| 19 | "PAC_1018",79,"Muller","No","Moderada",213,182,"Hipertensión"                                                                |
| 20 | "PAC_1019",76,"Muller","No","Moderada",195,118,"Sen patoloxía aparente"                                                      |
| 21 | "PAC_1020",32,"Muller","No","Moderada",215,147,"Sen patoloxía aparente"                                                      |
| 22 | "PAC_1021",70,"Muller","Exfumador","Baixa",193,155,"Hipertensión"                                                            |
| 23 | "PAC_1022",54,"Muller","Si","Baixa",180,142,"Hipertensión"                                                                   |
| 24 | "PAC_1023",56,"Muller","No","Baixa",238,129,"Sen patoloxía aparente"                                                         |
| 25 | "PAC_1024",65,"Home","Exfumador","Moderada",207,151,"Diabetes"                                                               |
| 26 | "PAC_1025",66,"Muller","No","Alta",268,124,"Infarto miocardio"                                                               |
| 27 | "PAC_1026",60,"Muller","No","Moderada",164,118,"Sen patoloxía aparente"                                                      |
| 28 | "PAC_1027",74,"Home","No","Baixa",201,142,"Diabetes"                                                                         |

|    | ID_Paciente | Idade | Sexo   | Fumador   | Actividade_Fisica | Colesterol_Total_mg_dl | Tension_Arterial_Sistolica | Patoloxia              |
|----|-------------|-------|--------|-----------|-------------------|------------------------|----------------------------|------------------------|
| 1  | PAC_1001    | 78    | Home   | No        | Moderada          | 183                    | 155                        | Diabetes               |
| 2  | PAC_1002    | 66    | Home   | Si        | Alta              | 181                    | 136                        | Hipertensión           |
| 3  | PAC_1003    | 30    | Home   | No        | Moderada          | 234                    | 129                        | Infarto miocardio      |
| 4  | PAC_1004    | 54    | Home   | No        | Baixa             | 164                    | 129                        | Sen patoloxía aparente |
| 5  | PAC_1005    | 39    | Muller | No        | Alta              | 261                    | 119                        | Sen patoloxía aparente |
| 6  | PAC_1006    | 65    | Muller | Exfumador | Alta              | 203                    | 139                        | Hipertensión           |
| 7  | PAC_1007    | 47    | Home   | No        | Baixa             | 197                    | 121                        | Sen patoloxía aparente |
| 8  | PAC_1008    | 78    | Home   | No        | Moderada          | 215                    | 164                        | Hipertensión           |
| 9  | PAC_1009    | 76    | Muller | No        | Moderada          | 201                    | 124                        | Sen patoloxía aparente |
| 10 | PAC_1010    | 53    | Muller | Exfumador | Moderada          | 196                    | 118                        | Sen patoloxía aparente |
| 11 | PAC_1011    | 36    | Muller | No        | Baixa             | 244                    | 132                        | Sen patoloxía aparente |
| 12 | PAC_1012    | 65    | Home   | Si        | Alta              | 233                    | 111                        | Infarto miocardio      |
| 13 | PAC_1013    | 54    | Muller | No        | Moderada          | 161                    | 146                        | Sen patoloxía aparente |
| 14 | PAC_1014    | 66    | Muller | No        | Alta              | 212                    | 123                        | Sen patoloxía aparente |
| 15 | PAC_1015    | 75    | Home   | No        | Baixa             | 189                    | 133                        | Sen patoloxía aparente |
| 16 | PAC_1016    | 49    | Muller | No        | Baixa             | 184                    | 131                        | Sen patoloxía aparente |
| 17 | PAC_1017    | 55    | Muller | No        | Alta              | 168                    | 130                        | Sen patoloxía aparente |
| 18 | PAC_1018    | 79    | Muller | No        | Moderada          | 213                    | 182                        | Hipertensión           |
| 19 | PAC_1019    | 76    | Muller | No        | Moderada          | 195                    | 118                        | Sen patoloxía aparente |
| 20 | PAC_1020    | 32    | Muller | No        | Moderada          | 215                    | 147                        | Sen patoloxía aparente |
| 21 | PAC_1021    | 70    | Muller | Exfumador | Baixa             | 193                    | 155                        | Hipertensión           |
| 22 | PAC_1022    | 54    | Muller | Si        | Baixa             | 180                    | 142                        | Hipertensión           |
| 23 | PAC_1023    | 56    | Muller | No        | Baixa             | 238                    | 129                        | Sen patoloxía aparente |
| 24 | PAC_1024    | 65    | Home   | Exfumador | Moderada          | 207                    | 151                        | Diabetes               |
| 25 | PAC_1025    | 66    | Muller | No        | Alta              | 268                    | 124                        | Infarto miocardio      |
| 26 | PAC_1026    | 60    | Muller | No        | Moderada          | 164                    | 118                        | Sen patoloxía aparente |
| 27 | PAC_1027    | 74    | Home   | No        | Baixa             | 201                    | 142                        | Diabetes               |

{base}

summary(datos\_pacientes)

{skimr}

skim(datos\_pacientes)

{dyplr} {gtsummary}

datos\_pacientes %>% select(Idade, Sexo, Fumador, Colesterol\_Total\_mg\_dl, Patoloxia) %>% tbl\_summary(by = Patoloxia)

```
## ID_Paciente      Idade      Sexo      Fumador
## Length:100      Min.    :30.00  Length:100  Length:100
## Class :character 1st Qu.:46.50  Class :character Class :character
## Mode :character  Median :57.00  Mode :character Mode :character
##                Mean   :56.32
##                3rd Qu.:68.25
##                Max.   :79.00
## Actividade_Fisica  Colesterol_Total_mg_dl  Tension_Arterial_Sistolica
## Length:100        Min.    :134.0      Min.    :105.0
## Class :character  1st Qu.:190.5      1st Qu.:126.0
## Mode :character  Median :207.0      Median :134.5
##                Mean   :209.2      Mean   :137.2
##                3rd Qu.:232.2      3rd Qu.:147.0
##                Max.   :274.0      Max.   :182.0
## Patoloxia
## Length:100
## Class :character
## Mode :character
##
##
```

|                            |           |                 |        |       |       |          |            |        |                                                                                           |
|----------------------------|-----------|-----------------|--------|-------|-------|----------|------------|--------|-------------------------------------------------------------------------------------------|
| Data summary               |           |                 |        |       |       |          |            |        |                                                                                           |
| Name                       |           | datos_pacientes |        |       |       |          |            |        |                                                                                           |
| Number of rows             |           | 100             |        |       |       |          |            |        |                                                                                           |
| Number of columns          |           | 8               |        |       |       |          |            |        |                                                                                           |
|                            |           |                 |        |       |       |          |            |        |                                                                                           |
| Column type frequency:     |           |                 |        |       |       |          |            |        |                                                                                           |
| character                  |           | 5               |        |       |       |          |            |        |                                                                                           |
| numeric                    |           | 3               |        |       |       |          |            |        |                                                                                           |
|                            |           |                 |        |       |       |          |            |        |                                                                                           |
| Group variables            |           | None            |        |       |       |          |            |        |                                                                                           |
| Variable type: character   |           |                 |        |       |       |          |            |        |                                                                                           |
| skim_variable              | n_missing | complete_rate   | min    | max   | empty | n_unique | whitespace |        |                                                                                           |
| ID_Paciente                | 0         | 1               | 8      | 8     | 0     | 100      | 0          |        |                                                                                           |
| Sexo                       | 0         | 1               | 4      | 6     | 0     | 2        | 0          |        |                                                                                           |
| Fumador                    | 0         | 1               | 2      | 9     | 0     | 3        | 0          |        |                                                                                           |
| Actividade_Fisica          | 0         | 1               | 4      | 8     | 0     | 3        | 0          |        |                                                                                           |
| Patoloxia                  | 0         | 1               | 8      | 22    | 0     | 4        | 0          |        |                                                                                           |
| Variable type: numeric     |           |                 |        |       |       |          |            |        |                                                                                           |
| skim_variable              | n_missing | complete_rate   | mean   | sd    | p0    | p25      | p50        | p75    | p100 hist                                                                                 |
| Idade                      | 0         | 1               | 56.32  | 14.97 | 30    | 46.5     | 57.0       | 68.25  | 79     |
| Colesterol_Total_mg_dl     | 0         | 1               | 209.20 | 28.72 | 134   | 190.5    | 207.0      | 232.25 | 274    |
| Tension_Arterial_Sistolica | 0         | 1               | 137.23 | 16.42 | 105   | 126.0    | 134.5      | 147.00 | 182  |

| Characteristic         | Diabetes<br>N = 17 <sup>1</sup> | Hipertensión<br>N = 22 <sup>1</sup> | Infarto miocardio<br>N = 15 <sup>1</sup> | Sen patoloxía aparente<br>N = 46 <sup>1</sup> |
|------------------------|---------------------------------|-------------------------------------|------------------------------------------|-----------------------------------------------|
| Idade                  | 62 (37, 72)                     | 66 (51, 72)                         | 57 (39, 65)                              | 55 (47, 65)                                   |
| Sexo                   |                                 |                                     |                                          |                                               |
| Home                   | 11 (65%)                        | 11 (50%)                            | 9 (60%)                                  | 22 (48%)                                      |
| Muller                 | 6 (35%)                         | 11 (50%)                            | 6 (40%)                                  | 24 (52%)                                      |
| Fumador                |                                 |                                     |                                          |                                               |
| Exfumador              | 2 (12%)                         | 5 (23%)                             | 1 (6.7%)                                 | 7 (15%)                                       |
| No                     | 15 (88%)                        | 14 (64%)                            | 10 (67%)                                 | 36 (78%)                                      |
| Sí                     | 0 (0%)                          | 3 (14%)                             | 4 (27%)                                  | 3 (6.5%)                                      |
| Colesterol_Total_mg_dl | 203 (183, 228)                  | 208 (194, 226)                      | 238 (207, 246)                           | 201 (184, 225)                                |

<sup>1</sup> Median (Q1, Q3); n (%)

*Comandos, sen medo:  
as análises sinxelas fanse sinxelas*

{psych}

## Alfa de Cronbach (0-1)

Avaliar a fiabilidade dun test a través da súa consistencia interna

| Código pregunta | Durante as dúas últimas semanas, con que frecuencia sentiu molestias por... |
|-----------------|-----------------------------------------------------------------------------|
| PA_1            | sentirse nervioso/a ou intranquilo/a?                                       |
| PA_2            | non poder controlar a súa preocupación?                                     |
| PA_3            | preocuparse demasiado por diferentes cousas?                                |
| PA_4            | ter dificultades para relaxarse?                                            |
| PA_5            | sentir medo coma se algo terrible puidera pasar?                            |

Táboa 1: Exemplo dalgunhas preguntas para testar a ansiedade. Baseado no test de *Mental Health America* (<https://screening.mhanational.org/screening-tools/test-de-ansiedad/>).

{psych}

```
fiabilidade<-alpha(datos_ansiedade)
print(fiabilidade)
```

R Console

data.frame 1 x 9

data.frame 2 x 3

data.frame 5 x 8

data.frame 5 x 7

Description: df [1 x 9]

|       | raw_alpha<br><dbl> | std.alpha<br><dbl> | G6(smc)<br><dbl> | average_r<br><dbl> | S/N<br><dbl> | ase<br><dbl> | mean<br><dbl> | sd<br><dbl> | median_r<br><dbl> |
|-------|--------------------|--------------------|------------------|--------------------|--------------|--------------|---------------|-------------|-------------------|
| 1 row | 0.9721812          | 0.9721892          | 0.9730396        | 0.8748662          | 34.95724     | 0.006351227  | 2.724         | 1.467409    | 0.8746135         |

R Console

data.frame 1 x 9

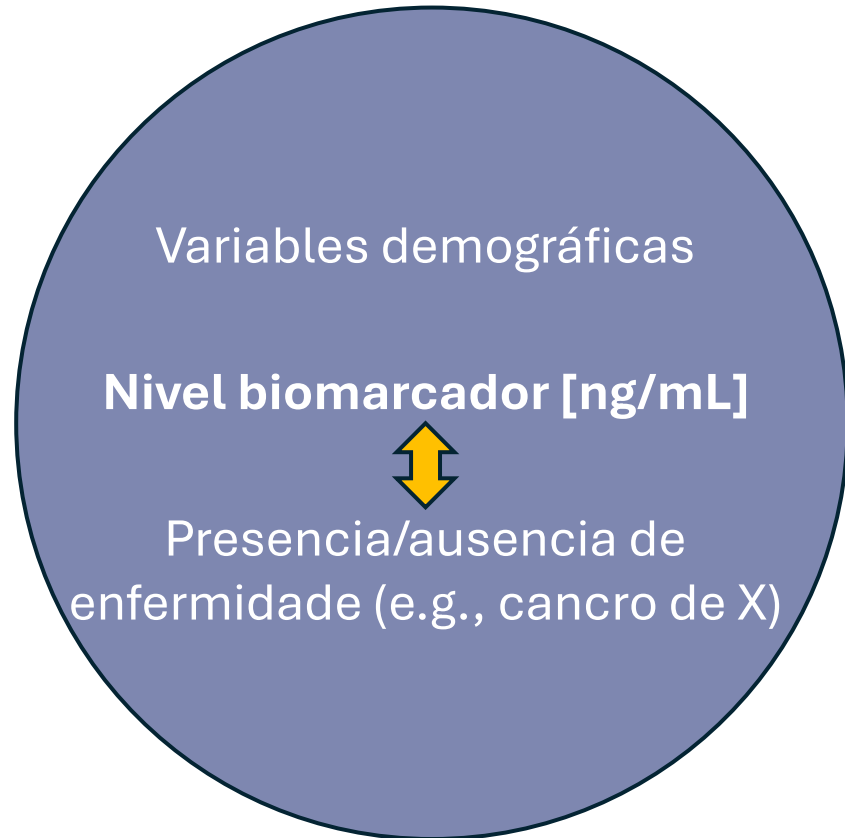
data.frame 2 x 3

data.frame 5 x 8

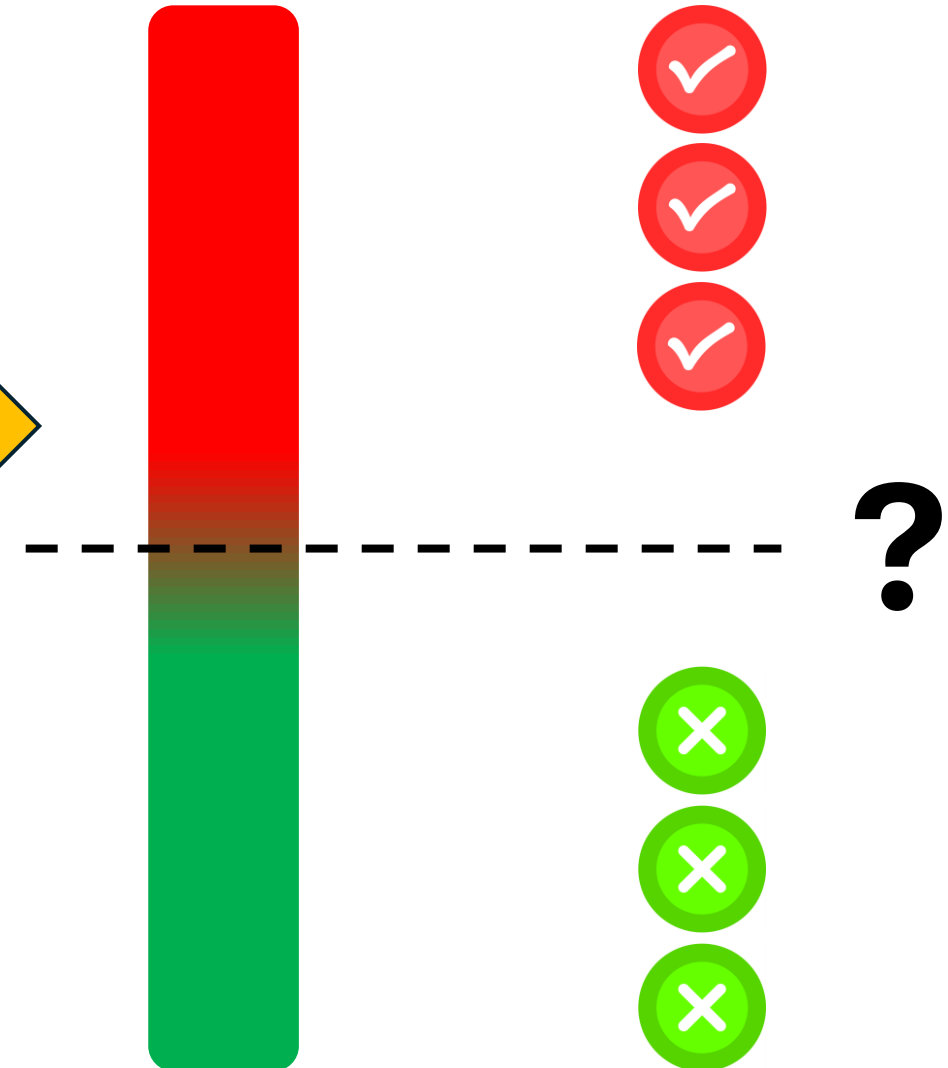
data.frame 5 x 7

Description: df [5 x 8]

|        | raw_alpha<br><dbl> | std.alpha<br><dbl> | G6(smc)<br><dbl> | average_r<br><dbl> | S/N<br><dbl> | alpha se<br><dbl> | var.r<br><dbl> | med.r<br><dbl> |
|--------|--------------------|--------------------|------------------|--------------------|--------------|-------------------|----------------|----------------|
| CA_1   | 0.9592230          | 0.9592385          | 0.9500387        | 0.8547197          | 23.53298     | 0.009534612       | 0.0014034156   | 0.8493454      |
| CA_2   | 0.9644917          | 0.9645244          | 0.9602317        | 0.8717472          | 27.18840     | 0.008375238       | 0.0019312109   | 0.8796160      |
| CA_3   | 0.9706431          | 0.9706746          | 0.9669283        | 0.8921838          | 33.10019     | 0.006931891       | 0.0008798797   | 0.9001322      |
| CA_4   | 0.9680468          | 0.9680556          | 0.9661143        | 0.8833967          | 30.30436     | 0.007601334       | 0.0015485295   | 0.8972928      |
| CA_5   | 0.9647071          | 0.9646885          | 0.9591441        | 0.8722837          | 27.31943     | 0.008290843       | 0.0013733221   | 0.8746135      |
| 5 rows |                    |                    |                  |                    |              |                   |                |                |

**N = 15 000 pacientes****REALIZAR BIOPSIA?****Novos pacientes (cancro?)**

Nivel biomarcador [ng/mL]



## {OptimalCutpoints}



### *Journal of Statistical Software*

October 2014, Volume 61, Issue 8.

<http://www.jstatsoft.org/>

## **OptimalCutpoints: An R Package for Selecting Optimal Cutpoints in Diagnostic Tests**

**Mónica López-Ratón**

Universidad de Santiago de Compostela

**María Xosé Rodríguez-Álvarez**

Universidade de Vigo

**Carmen Cadarso-Suárez**

Universidad de Santiago de Compostela

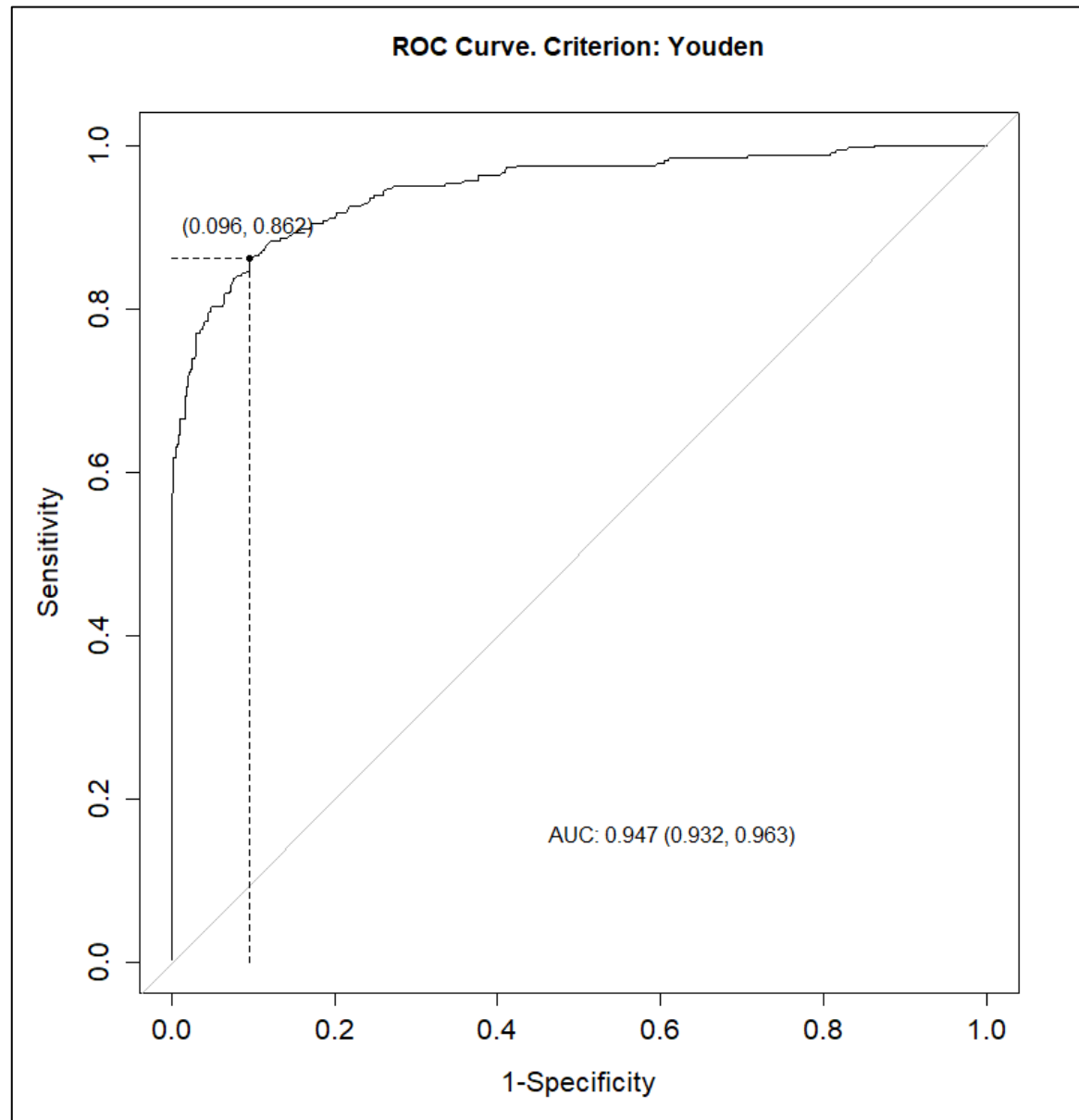
**Francisco Gude-Sampedro**

Complejo Hospitalario Universitario  
de Santiago de Compostela (CHUS)

## {OptimalCutpoints}

```
research_cutpoint<- optimal.cutpoints(X = "biomark_nivel", status = "presencia_cancro", data = datos_N15000,
tag.healthy = 0, methods = c("Youden", "ValueSe"), ci.fit = TRUE, conf.level = 0.95, control =
control.cutpoints(valueSe = 0.95))
```

|                                           | Estimación del 95% | Límite inferior IC 95% | Límite superior IC 95% |
|-------------------------------------------|--------------------|------------------------|------------------------|
| Umbral (ng/mL)   “valueSe = <b>0.95</b> ” | <b>3.39</b>        | 2.87                   | 3.65                   |
| Umbral (ng/mL)   “Youden”                 | <b>4.29</b>        | 4.14                   | 5.05                   |
| Sensibilidade                             | <b>0.86</b>        | 0.82                   | 0.90                   |
| Especificidade                            | 0.90               | 0.87                   | 0.92                   |
| Valor Predictivo Positivo (PPV)           | 0.81               | 0.77                   | 0.86                   |
| Valor Predictivo Negativo (NPV)           | 0.93               | 0.91                   | 0.95                   |



AUC (*A*rea *U*nder the *C*urve)

**94.7%** de que un individuo seleccionado ao azar do grupo con cancro teña unha concentración do biomarcador seleccionado máis elevada que un individuo seleccionado ao azar do grupo san.

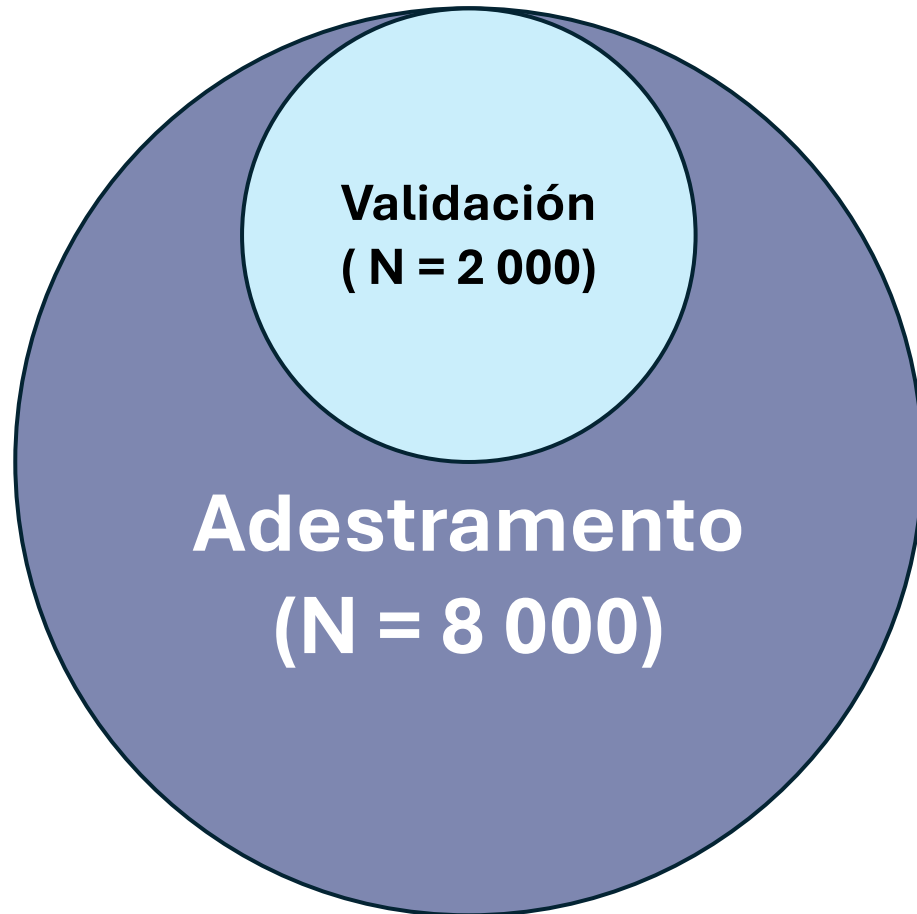
Aprendizaxe automática  
(*machine learning*)

- ❑ **Supervisada:** con datos de adestramento (~80%) e validación (~20%) {**caret, randomForest, tidymodels**}
- ❑ **Non supervisada:** clusterizando as entradas dos datos en base a patróns compartidos sen hipóteses previas {**dbscan, fda.usc**}
- ❑ **Por reforzo:** recompensando as respostas correctas {**ReinforcementLearning**}

{caret}

## Probabilidade de reingreso do paciente aos 90 días

**N = 10 000 pacientes**

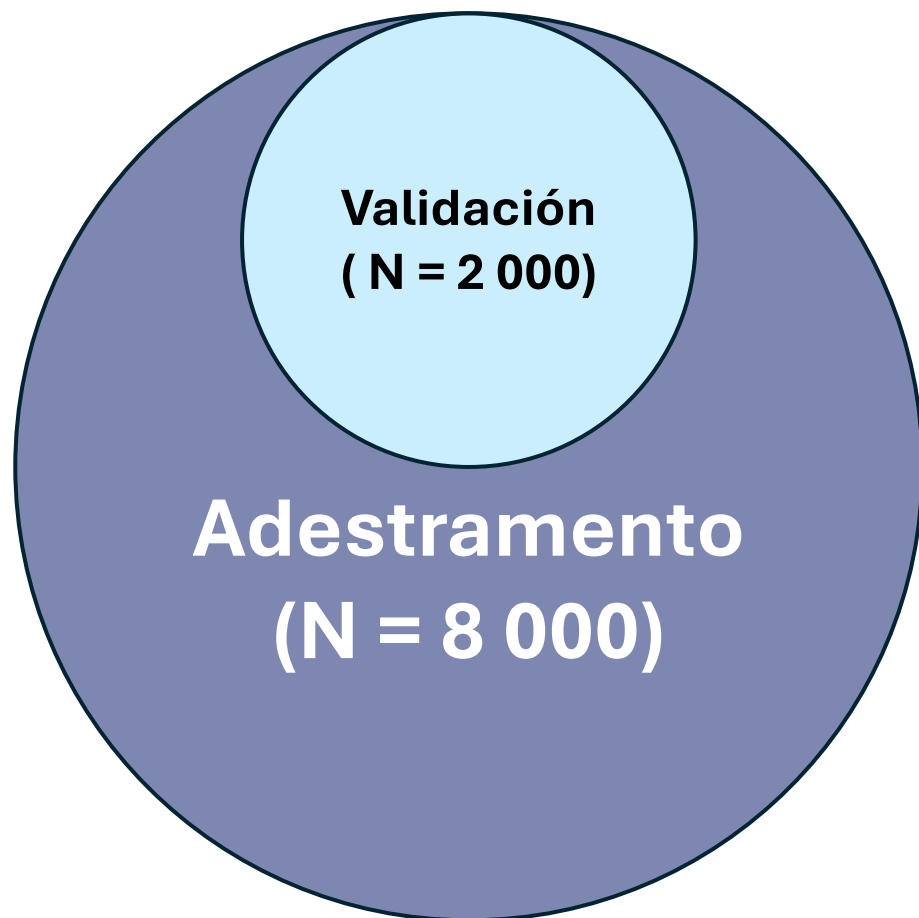


- ☐ Variables demográficas
- ☐ Patoloxía principal
- ☐ Comorbilidade
- ☐ Infección nosocomial
- ☐ Duración ingreso previo
- ☐ Días na UCI
- ☐ Reingreso aos 90 días (si/non)
- ☐ ...

{caret}

## Probabilidade de reingreso do paciente aos 90 días

N = 10 000 pacientes



```
control <- trainControl(method = "cv", number = 10,  
summaryFunction = twoClassSummary, classProbs = TRUE,  
verboselter = TRUE, sampling = "up")
```

```
set.seed(123)
```

```
modelo_gbm <- train( reingreso_90dias ~ . - id_paciente, data =  
datos_entrenamiento, method = "gbm", metric = "ROC",  
trControl = control)
```

{caret}

## Probabilidade de reingresso do paciente aos 90 días

Valores predición

|              |                                                                                             |
|--------------|---------------------------------------------------------------------------------------------|
| <b>NO/NO</b> | <b>NO/SI</b>                                                                                |
| <b>1 082</b> | <b>4</b>  |
| <b>SI/NO</b> | <b>SI/SI</b>                                                                                |
| <b>19</b>    | <b>895</b>                                                                                  |

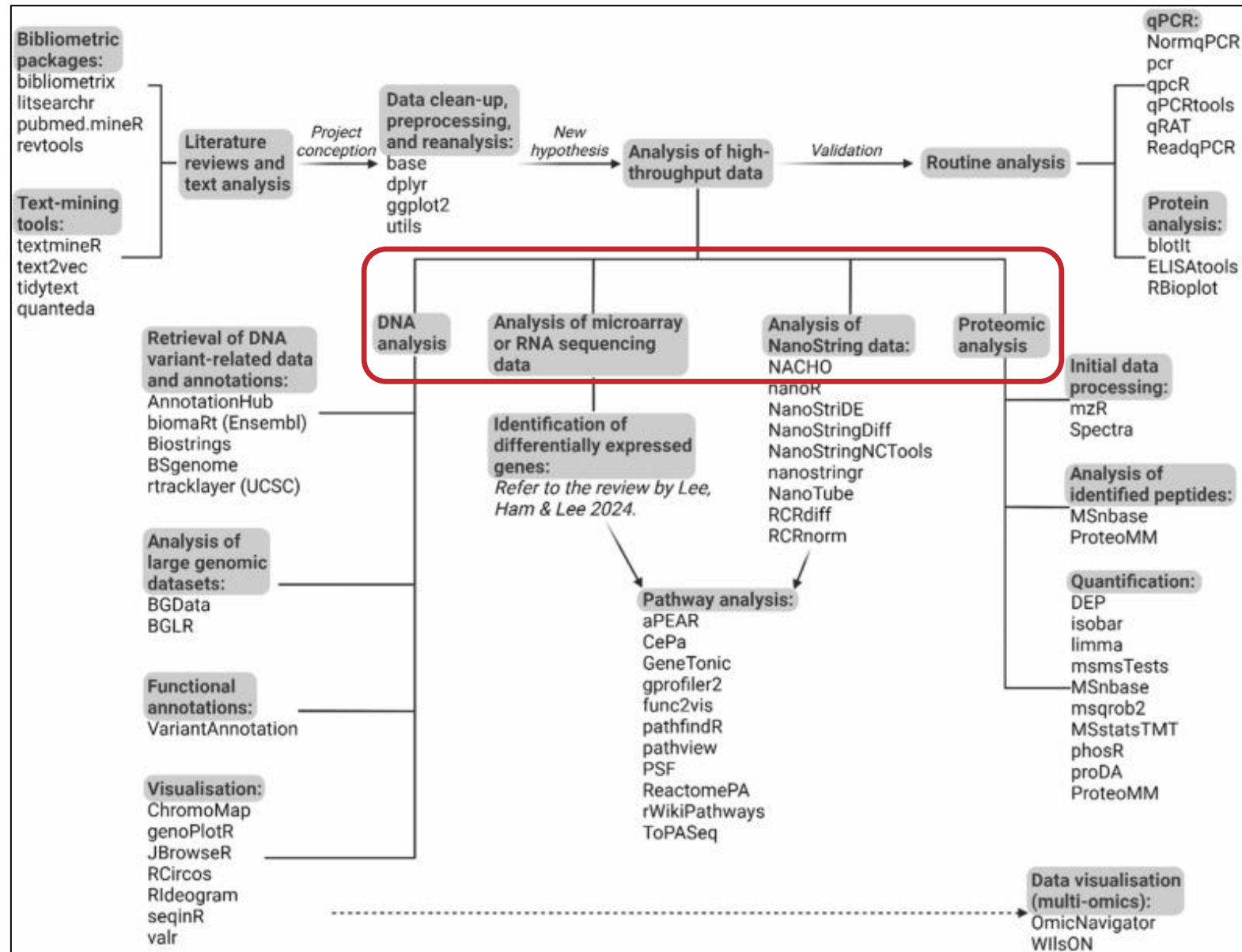
Valores reais

| Índices/coeficientes      | Valores |
|---------------------------|---------|
| Kappa                     | 0.9768  |
| Mcnemar's test $p$ -valor | 0.0035  |
| Sensibilidade             | 0.9956  |
| Especificidade            | 0.9827  |

{caret}

## Probabilidade de reingresso do paciente aos 90 días

```
> print(head(pacientes_alto_riesgo, 10))  
# A tibble: 10 × 4  
  id_paciente  edad patologia_principal prob_reingresso_Si  
    <int> <dbl> <fct>                <dbl>  
1    10099     64 Digestiva                1.000  
2    10058     73 Oncolóxica                1.000  
3    10017    100 Oncolóxica                1.000  
4    10029     70 Respiratoria                1.000  
5    10098     93 Cardiovascular                1.000  
6    10088     79 Outras                1.000  
7    10092     85 Neurolóxica                1.000  
8    10009     91 Neurolóxica                0.999  
9    10087     86 Neurolóxica                0.999  
10   10026     95 Oncolóxica                0.999
```



# CONCLUSIÓN

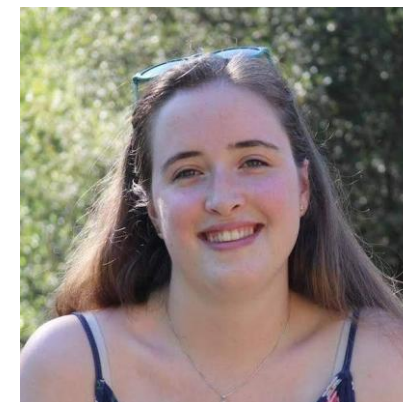
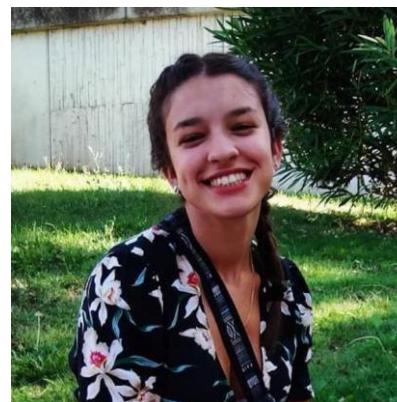
- ❑ Recomendamos o uso do entorno de R en **calquera campo de investigación**, incluído o sanitario
- ❑ Consideramos recomendable para aquelas persoas que **cambien de campo de traballo a exploración** previa das posibilidades de R para o deseño das súas **novas rutas bioinformáticas**

*Vaia onde vaia...*



*R estará conmigo*

# O melloR traballo, en equipo ( $p\text{-valor} < 0.0000000000000001$ )



O financiamento para o desenvolvemento da UNICL provén da convocatoria do Instituto de Salud Carlos III baixo o PERTE para la Salud de Vanguardia do Ministerio de Ciencia e Innovación no marco do Plan de Recuperación, Transformación y Resiliencia, regulada mediante la Orden CNU/709/2024, de 1 de julio (BOE nº 164, de 08/07/2024)