

I XORNADAS DE USUARIOS DE EN GALICIA

10 de Outubro

Centro de Novas Tecnoloxías de Galicia, Santiago

Investigación con R nun ámbito universitario e interdisciplinar. A visión dende o GRID[ECMB].

*Laura Calaza Díaz e Roberto Domínguez Gómez
Personal de xestión e organización*

GRID[ECMB][#]

*Grupo interdisciplinar de Estatística,
Computación, Medicina e Bioloxía*

Investigadora principal:

Carmen Cadarso-Suárez

Afiliación:

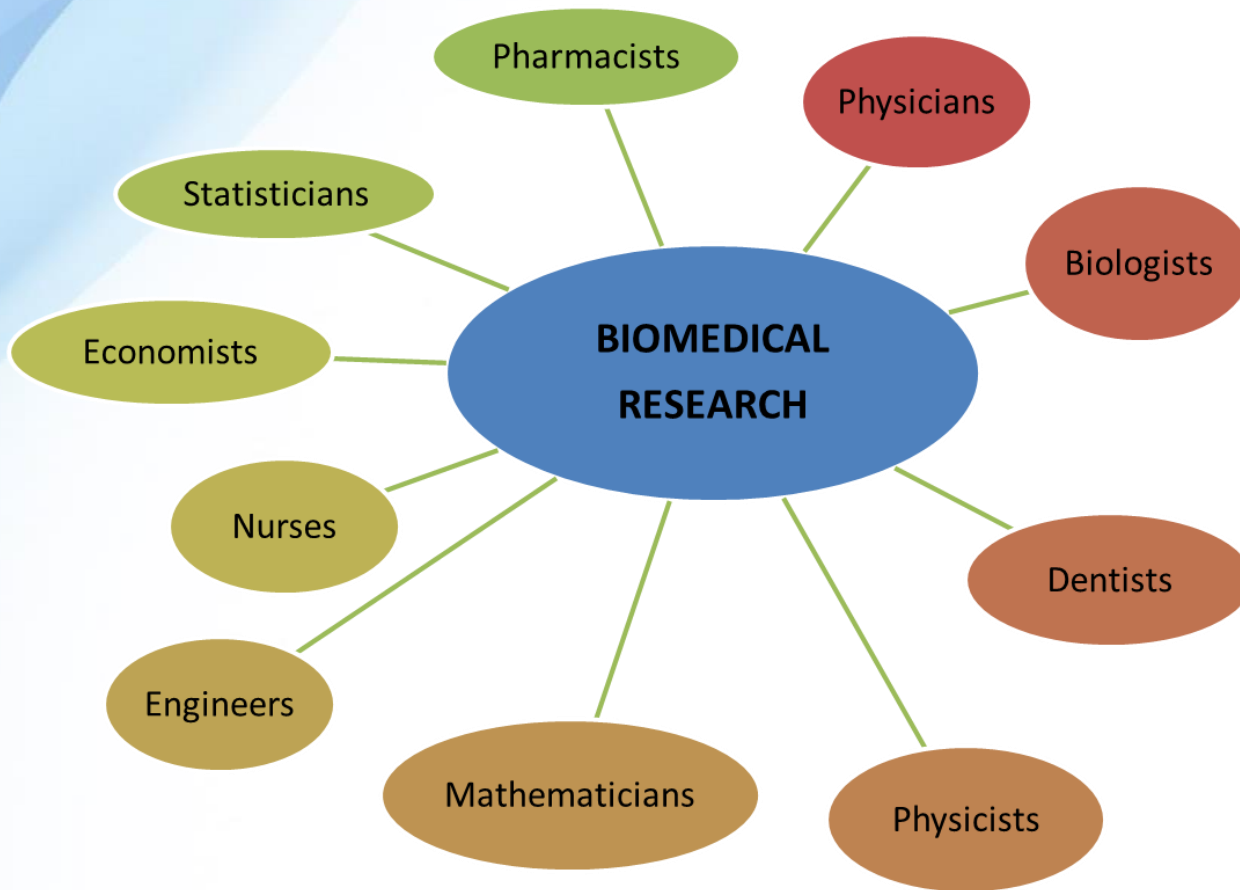
Centro de Investigación en Medicina Molecular e Enfermedades
Crónicas (CIMUS)

Facultade de Medicina, Universidade de Santiago de Compostela



Investigación biomédica

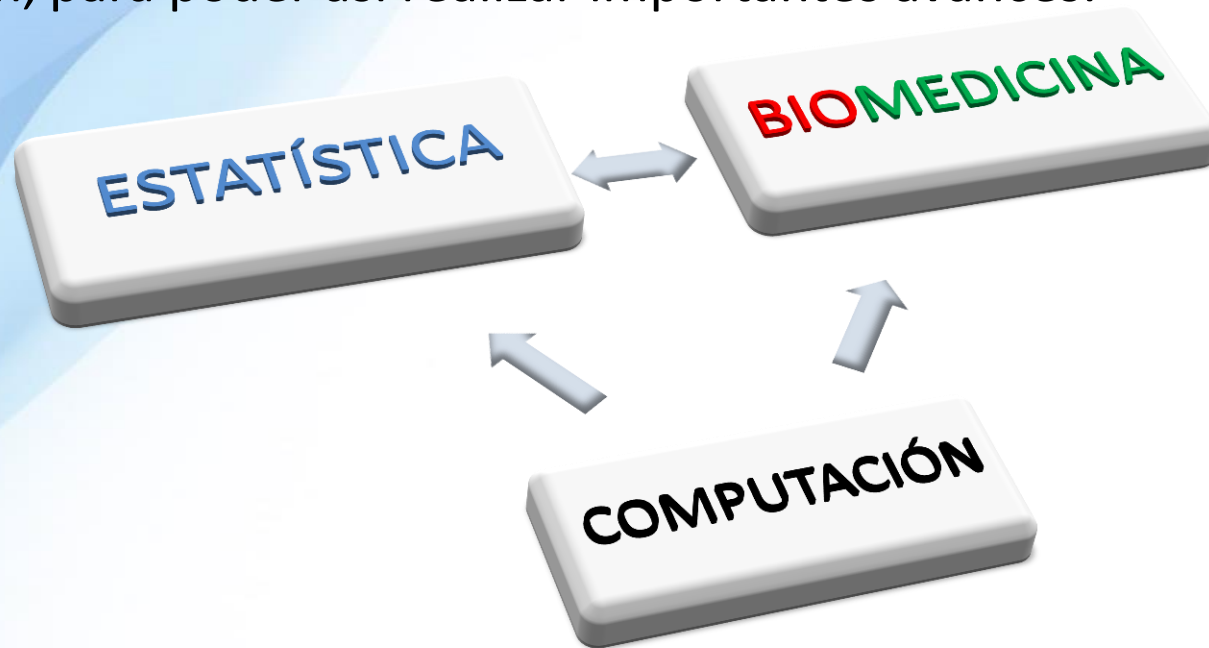
- A actual investigación biomédica require da colaboración con equipos multidisciplinares que abarquen especialistas en diferentes campos.



Os bioestadísticos xogan un papel importante...

A bioestatística actual

- O grupo trata de compaxinar as áreas de Estatística, Biomedicina e Computación, para poder así realizar importantes avances.



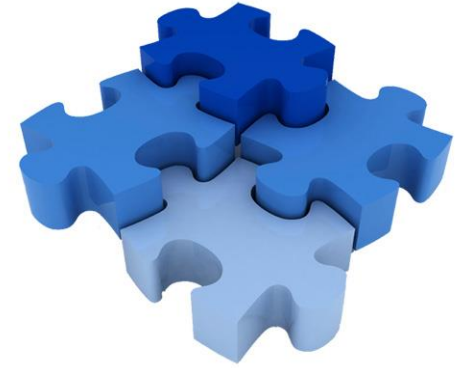
- Unha **boa aplicación da Bioestatística** require profesionais con:
 - Unha sólida formación en Matemáticas
 - Capacidade de desenvolver novos métodos estatísticos
 - Coñecementos básicos de epidemioloxía, ensaios clínicos, ...
 - **Competencia na creación de software (ex. co software R)**

Principais actividades

Máis de 20 anos de experiencia na Bioestatística en:



Objetivos



- **Investigación metodolóxica** en Estatística con aplicacións en diferentes campos biomédicos
- Desenvolvemento de **software** libre, eficiente e de fácil uso para investigadores noutras áreas.
- **Transferencia** de coñecemento.

Liñas actuais de investigación

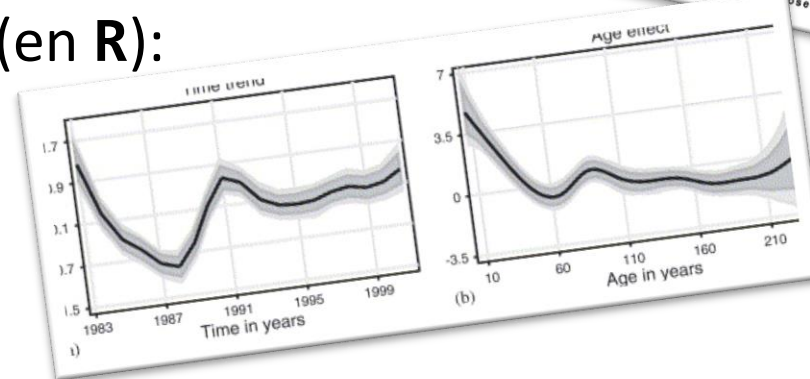
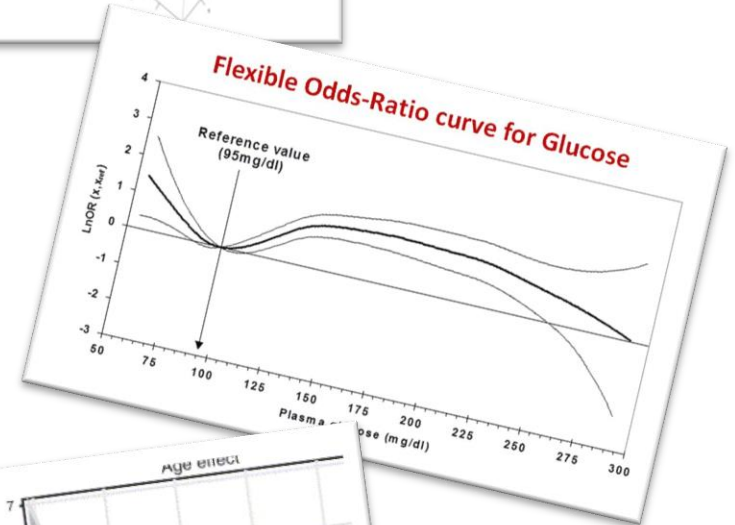
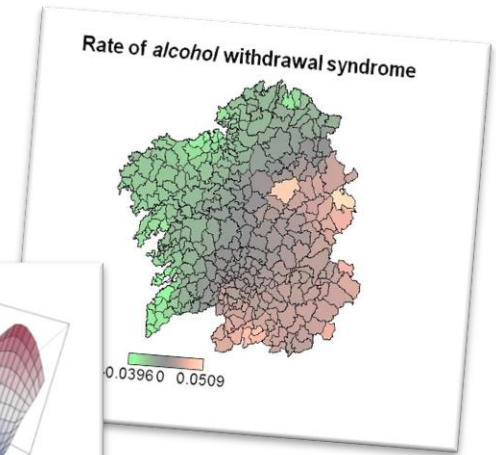
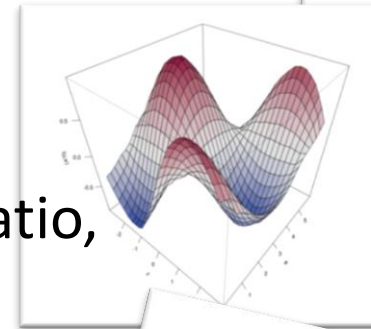
Generalized Additive Models (GAMs)

Curvas ROC para Diagnose e Prognose

Análisis de Supervivencia

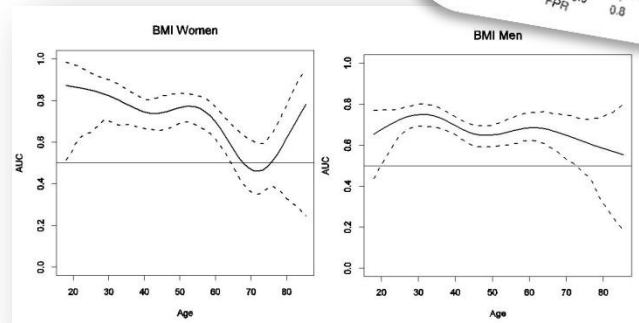
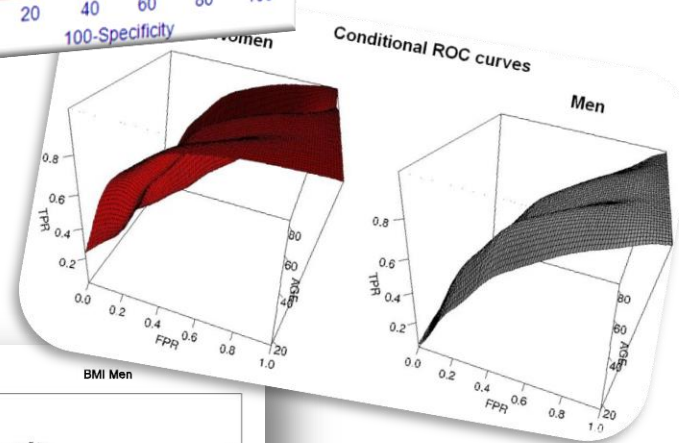
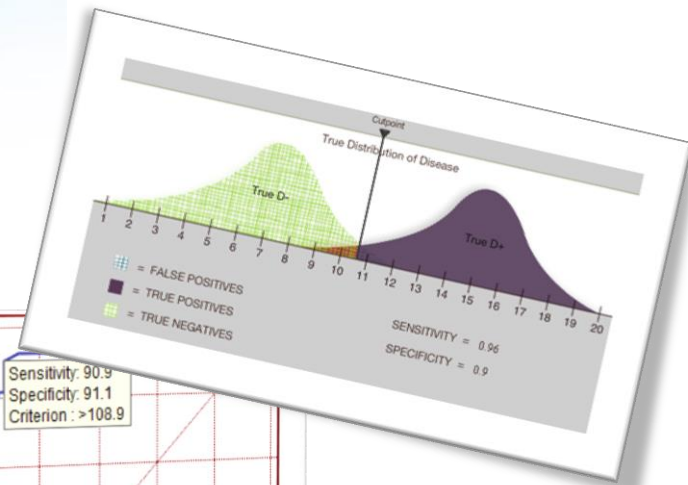
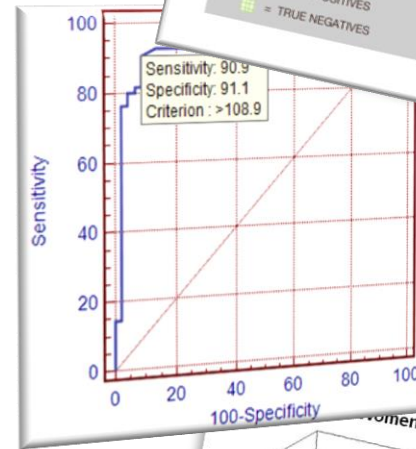
Generalized Additive Models (GAMs)

- GAM incluíndo interaccións
- Inferencia con bootstrap
- Medidas flexibles de efectos (Odds-Ratio, Risco Relativo) para marcadores contínuos
- Structured Additive Models (STARs) incluíndo efectos espaciais
- Paquetes de Software (en R):
 - `np.OR.mgcv`
 - `boosting`



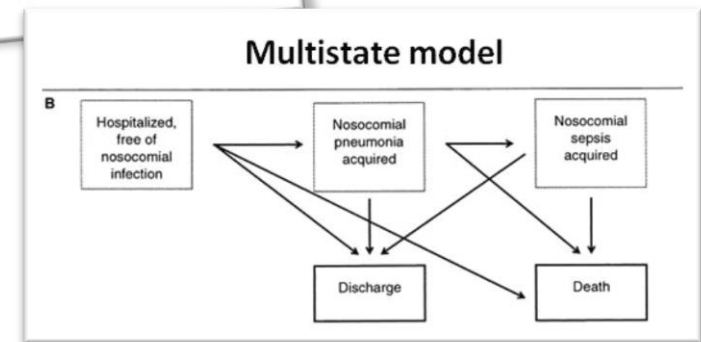
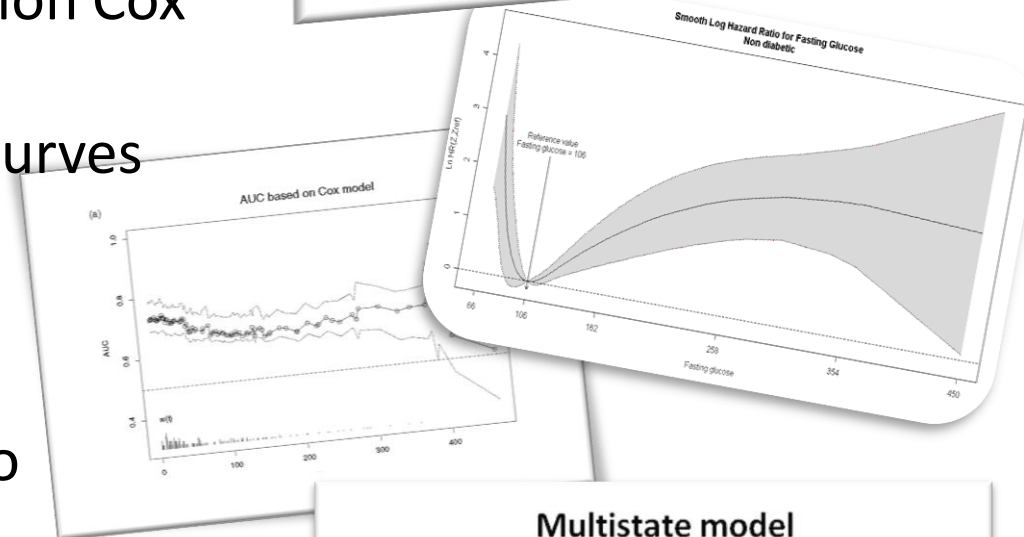
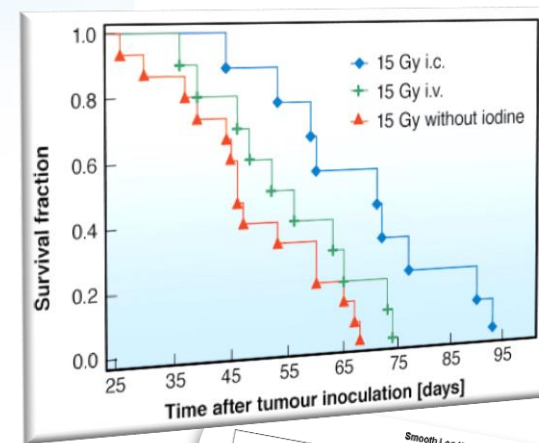
Análise de curvas ROC

- Curvas ROC. Area Under the Curve (AUC)
- Análise de regresión ROC
- Selección óptima de puntos de corte para biomarcadores continuos
- Paquetes de software (en R):
 - ROCregression
 - Optimalcutpoints



Análise de Supervivencia

- Análise de Supervivencia incluindo efectos variables no tempo
- Modelos flexibles de regresión Cox
- Smooth Hazard Ratio (HR) curves
- Modelos multiestado
- Curvas ROC dependentes do tempo
- Paquetes de software (en **R**):
 - SmoothHR
 - tdc.msm
 - PEMselect



Porque é importante o desenvolvemento de software libre?

SITUACIÓN 1



Quero aplicar GAMs con
robustez, tes datos que
conteñan atípicos?

Teño que mirar
...
SI!!!!

Cantos datos!!!
Necesitamos
software!!!

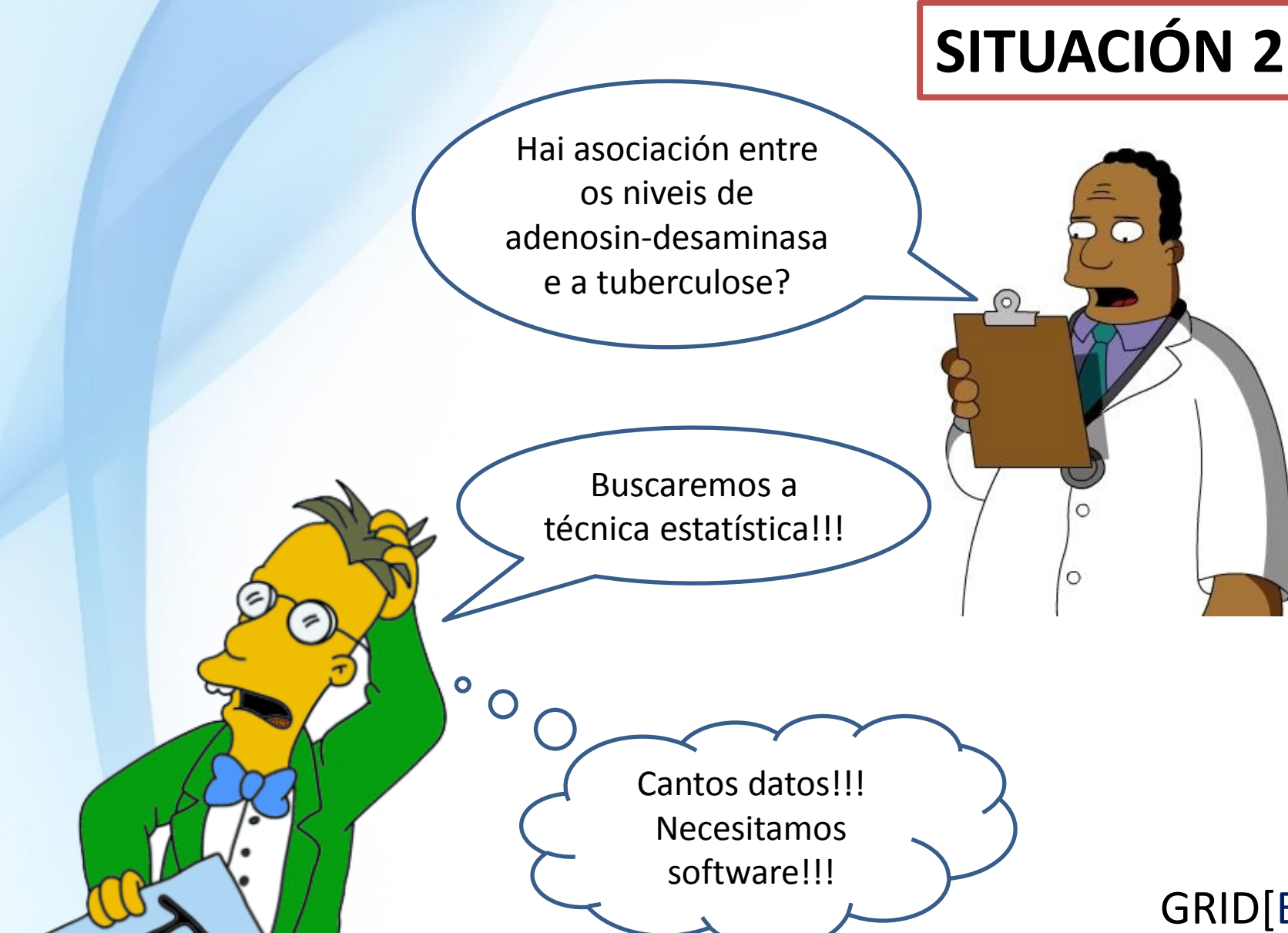


Porque é importante o desenvolvimento de software livre?



Porque é importante o desenvolvemento de software libre?

SITUACIÓN 2



Hai asociación entre os niveis de adenosin-desaminasa e a tuberculose?

Buscaremos a técnica estatística!!!

Cantos datos!!!
Necesitamos software!!!

Porque é importante o desenvolvimento de software livre?



Desenvolvemento de software libre

Epidat

Programa dirixido a epidemiólogos e outros profesionais da saúde para o manexo de datos.

EpiLinux

Sistema operativo, orientado á utilización de ferramentas de análise epidemiolóxico e bioestatístico.

Distribución completa de Linux que inclúe:

- Ferramentas ofimáticas
- Software de Internet
- Software para a realización de estudos de bioestatística, análise de datos e epidemioloxía, entre eles, **R!**

Desenvolvemento de paquetes en R

PMICALC

José I. Muñoz-Barús, María Sol Rodríguez-Calvo, José M. Suárez-Peñaranda, Duarte N. Vieira, Carmen Cadarso-Suárez, Manuel Febrero-Bande

Código R-based freeware, un paquete que estima o intervalo post-mortem (PMI) en cadáveres de morte recente, usando dous diferentes modelos de regresión: Modelos Aditivos (AM) e Support Vector Machine (SVM).

<http://eio.usc.es/pub/febrero/software/software.html>

p3state.msm

Luís Meira-Machado e Javier Roca-Pardiñas

Ofrece funcións para a estimación de modelos de regresión semi-paramétricos, entre outros.

<http://cran.r-project.org/web/packages/p3state.msm/p3state.msm.pdf>

TPmsm

Luís Meira-Machado

Para a obtención de estimacións paramétricas e semiparamétricas das probabilidades de transición en modelos multiestado con tres estados.

<http://CRAN.R-project.org/package=TPmsm>

Desenvolvemento de paquetes en R

genSurv

Luís Meira-Machado

Desenvolto para a simulación de datos de supervivencia para o modelo de mortalidade (de dous estados) e para o modelo illness-death progresivo.

<http://CRAN.R-project.org/package=genSurv>

smoothHR

Meira-Machado L., Araújo A., Cadarso-Suárez C., Gude F.

Ofrece unha estimación non paramétrica puntual das curvas hazard ratio (HR) para predictores continuos.

<http://cran.r-project.org/web/packages/smoothHR/smoothHR.pdf>

ROCRegression

Rodríguez-Álvarez, M.X., López de Ullibarri, I., Cadarso-Suárez, C.

Permite a avaliación do efecto de covariables na capacidade de discriminación, medida a través da curva ROC dun marcador contínuo.

Desenvolvemento de paquetes en R

npROCRegression

Rodríguez-Álvarez, M.X., Roca-Pardiñas, J., Cadarso-Suárez, C.

Incorporación de covariables dentro da análise ROC dende unha perspectiva non paramétrica.

OptimalCutpoints

Mónica López-Ratón e María Xosé Rodríguez-Álvarez

Selección de puntos de corte óptimos en test diagnósticos.

<http://CRAN.R-project.org/package=OptimalCutpoints>

catpredict

Barrio, I., Aróstegui, I., Rodríguez-Álvarez, M.X., Quintana J.M.

Categorización óptima de variables cuantitativas para a súa incorporación en modelos de predición.

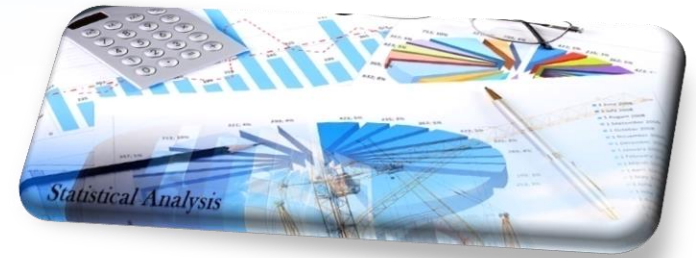
Transferencia de conocimiento con R

■ Formación a demanda:

- Fundamentos de Bioestadística
 - "Introducción a bioestadística para investigadores"
 - "Bioestadística Básica con R"
 - "Bioestadística con R"
- Cursos especializados en técnicas estadísticas.
 - "Modelos lineales generalizados (GLM) y modelos aditivos generalizados (GAM). Aplicaciones con R"
 - "Modelos de regresión múltiple en R. GLM y GAM"
 - "Curso de Modelos Mixtos utilizando R"
 - "Survival Analysis Using R"
 - "Modelos Mixtos utilizando R"
- Software estadístico.
 - "Curso de iniciación al software estadístico R"



Consultoría estatística



- **Soporte estatístico** a diferentes grupos universitarios.
- Realízanse **servizos de consultoría e asesoramento estatístico**:
 - Análise de datos
 - Estudos específicos
 - Redacción de artigospara os cales é fundamental un bo manexo do **software R**.

Estruturas a través das cales damos soporte




BIostatnet
Rede Interdisciplinar de Bioestatística

GRID[ECMB]#
Interdisciplinary Group of Statistics, Computation, Medicine and Biology

BIO
STA
TECH

SPIN-OFF académica

Interdisciplinary Center of Biostatistics of the USC (ICBUSC)



Estruturas a través das cales damos soporte



BIostatnet
Rede Interdisciplinar de Bioestatística

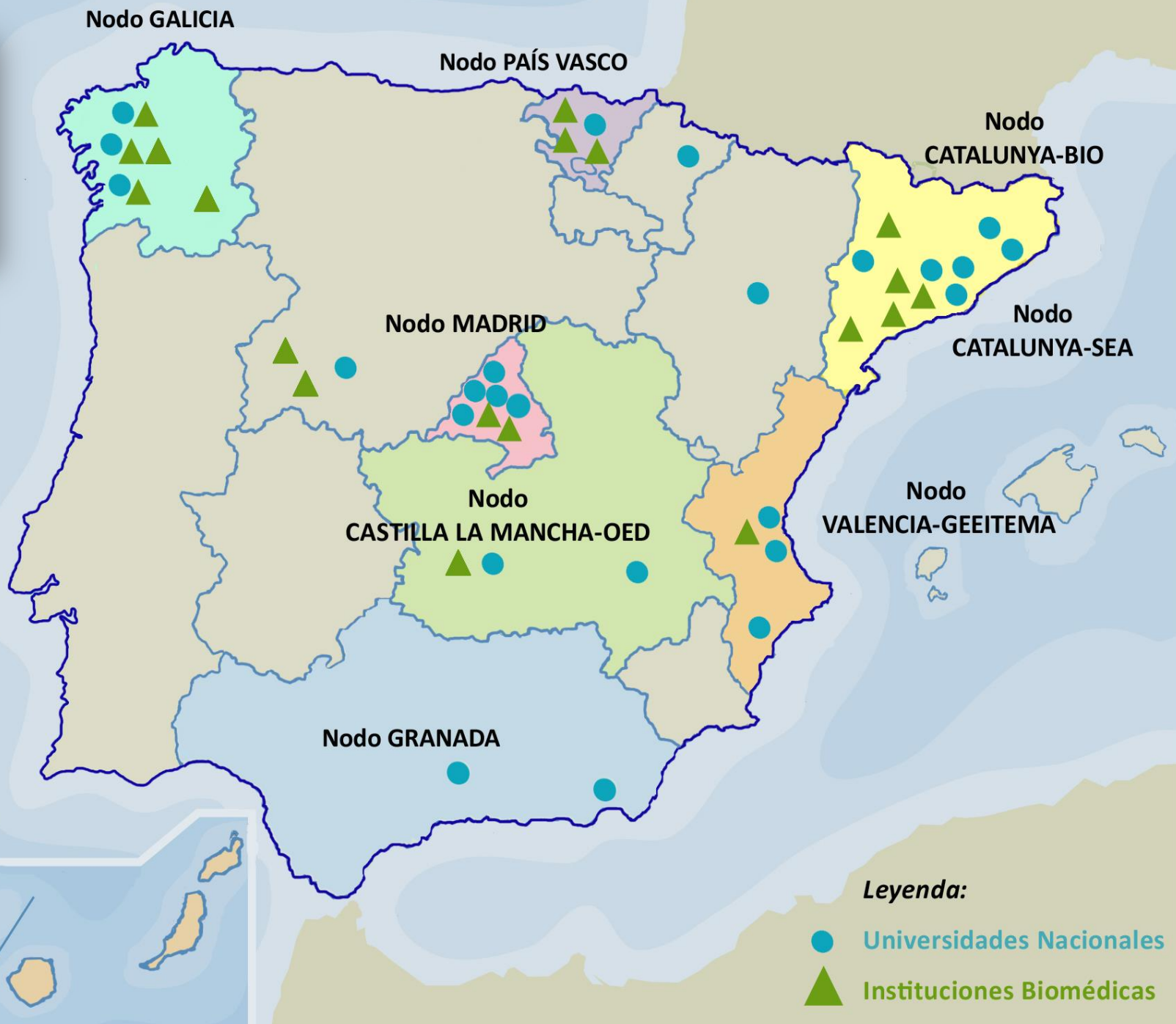
GRID[ECMB]#
Interdisciplinary Group of Statistics, Computation, Medicine and Biology

BIO
STA
TECH

SPIN-OFF académica

Interdisciplinary Center of Biostatistics of the USC (ICBUSC)







BIOSTATNET: Rede Nacional de Bioestatística

- **Membros:** Arredor de 190 **estatísticos e investigadores biomédicos**, pertencentes a:
 - De 25 Universidades e 23 Institucións Biomédicas españolas.
 - De 20 Universidades estranxeiras
- **Páxina web:** www.biostatnet.org

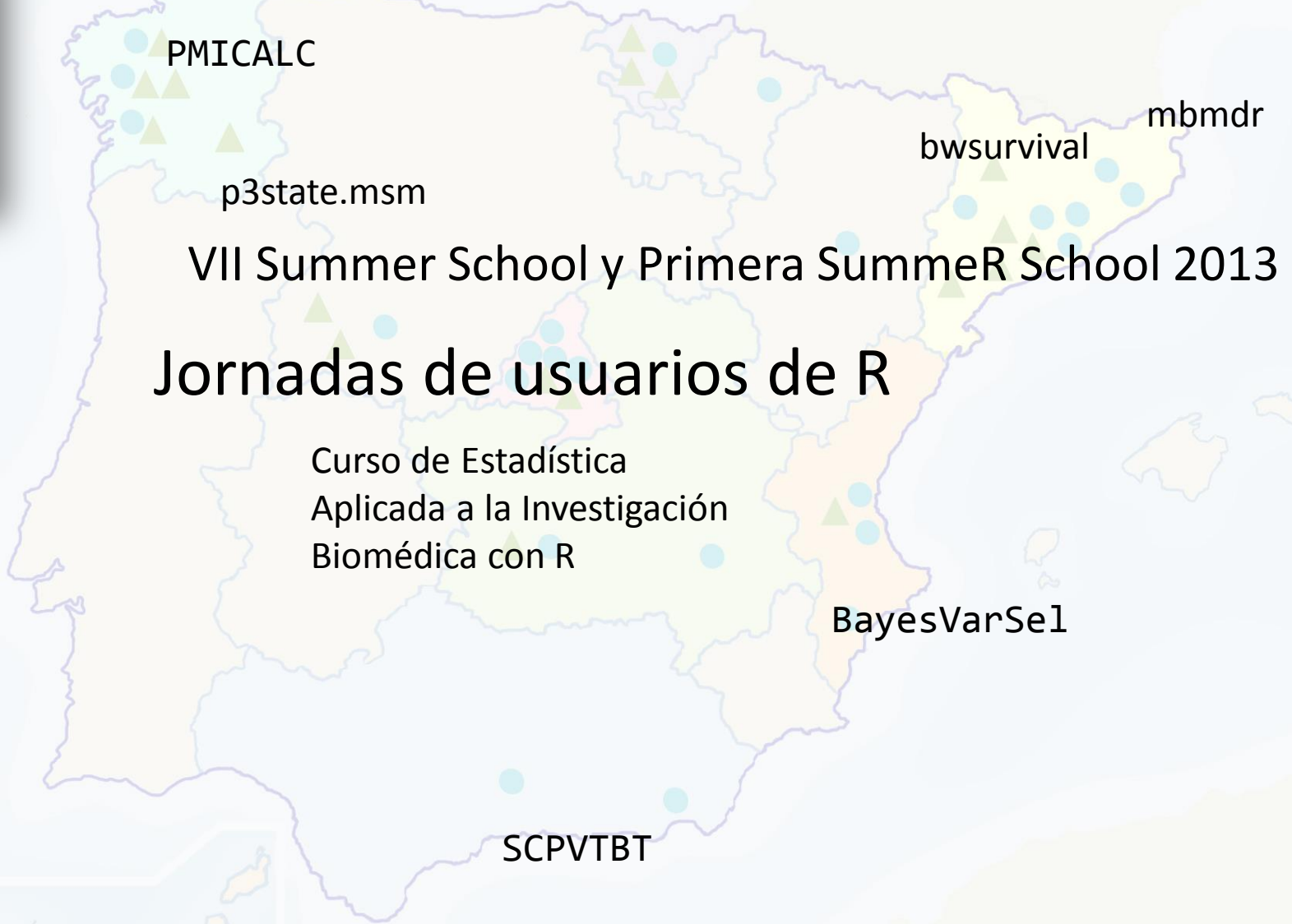
BIostatNET: Actividades con R



Jornadas de usuarios de R



BIostatNET: Actividades con R





BIostatNET: Actividades con R

PMICALC
ROCRegression

Curso de introducción a R

bwsurvival

mbmdr

p3state.msm

VII Summer School y Primera SummerR School 2013

Jornadas de usuarios de R

OptimalCutpoints

Curso de Estadística
Aplicada a la Investigación
Biomédica con R

BayesVarSel

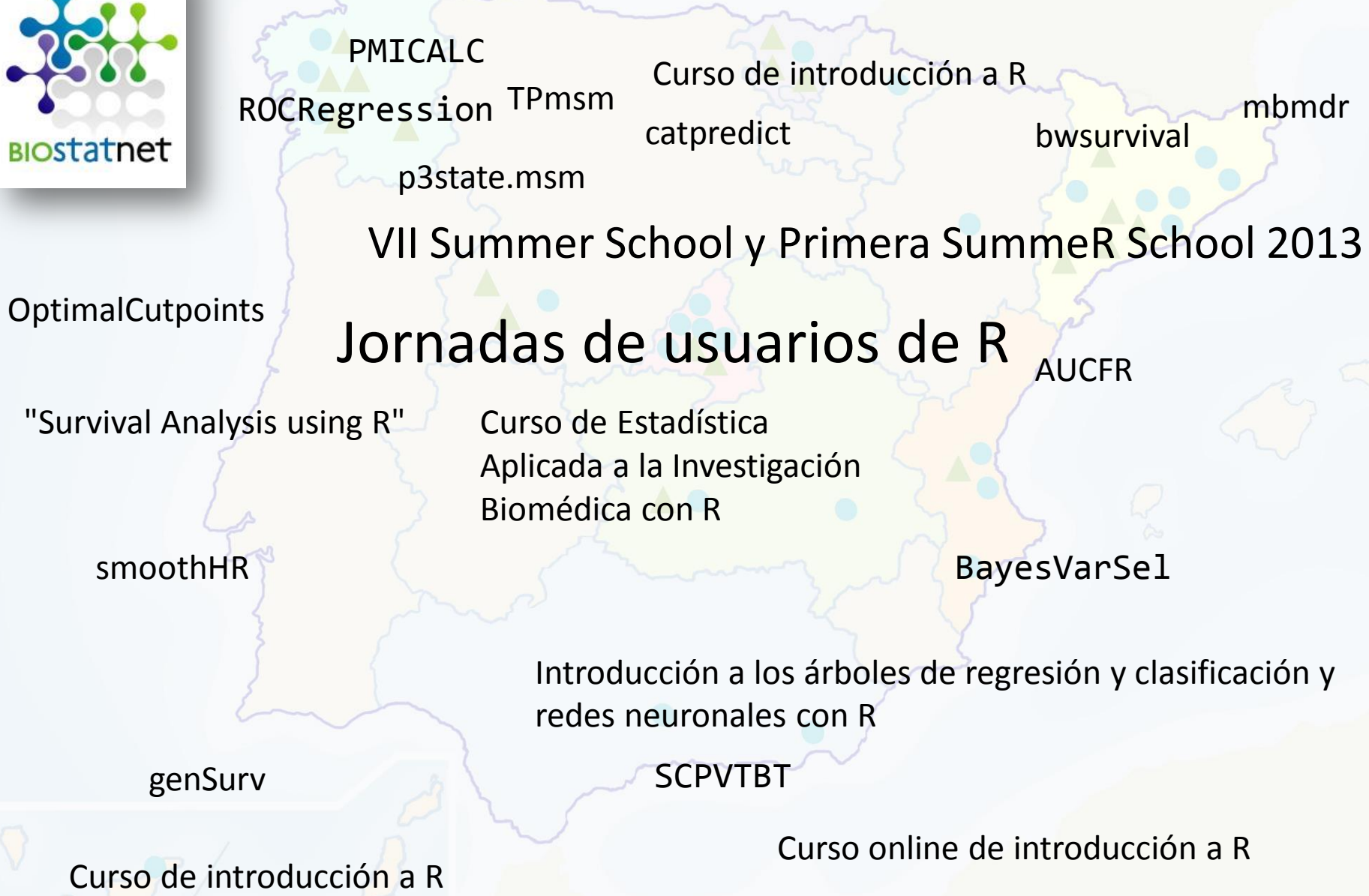
SCPVTBT

Curso de introducción a R

Curso online de introducción a R

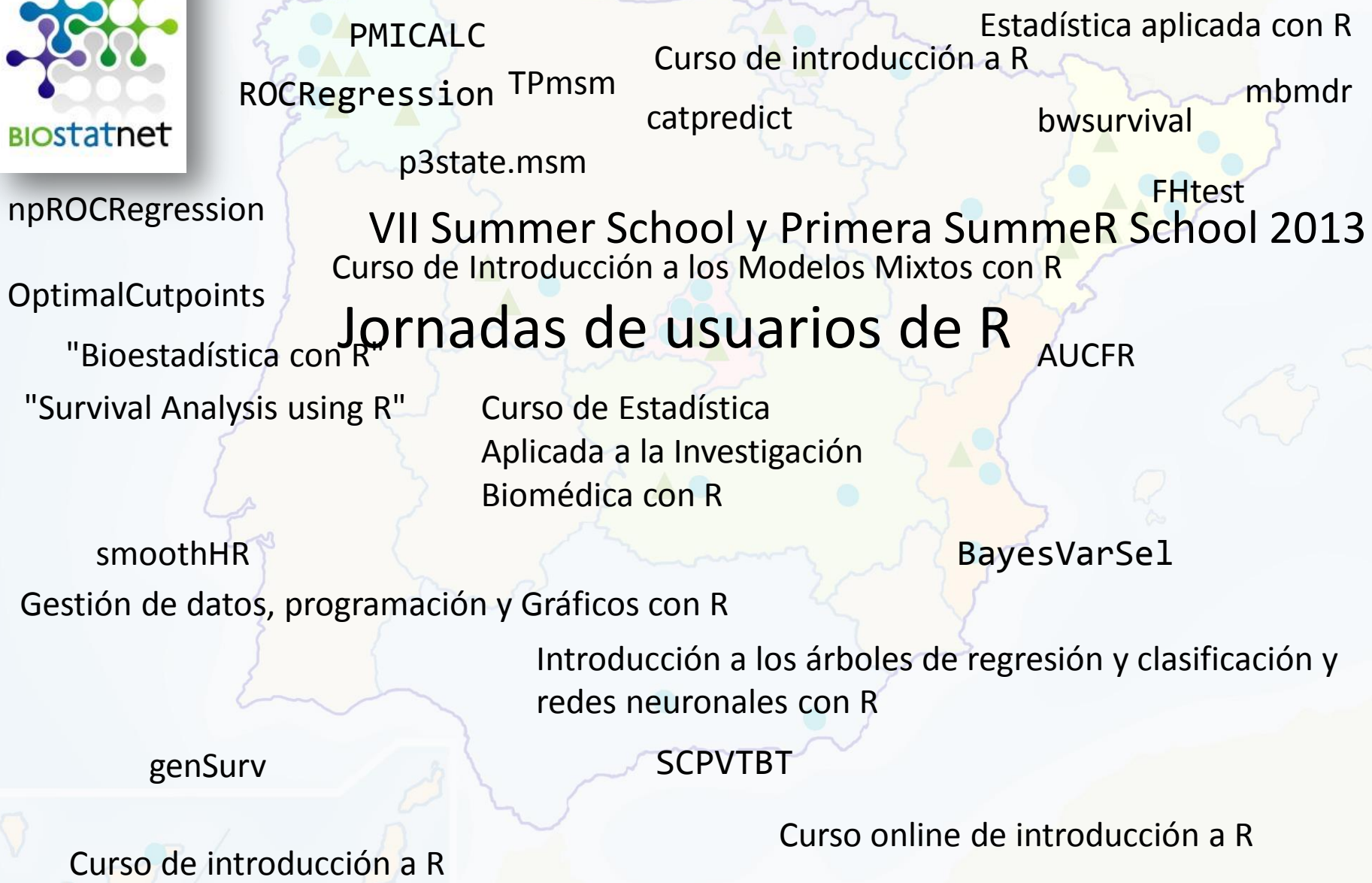


BIostatNET: Actividades con R

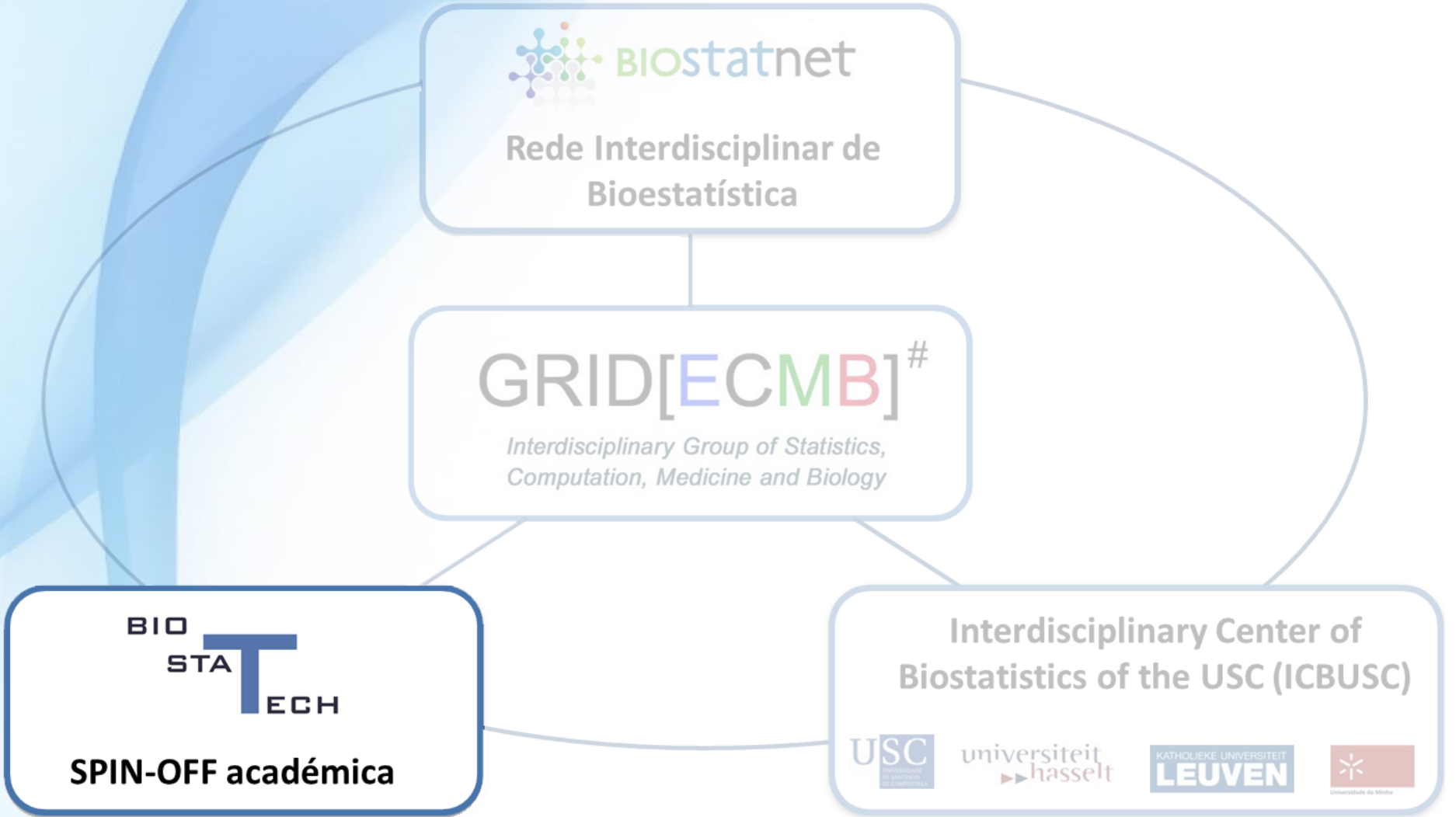




BIostatNET: Actividades con R



Estruturas a través das cales damos soporte



BIOSTATECH: SPIN-OFF académica da USC



Creada en Marzo do 2012

Promotores: Carmen Cadarso e varios membros do GRID[ECMB].

Páxina web: www.biostatech.com



Edificio Emprendia
Campus Vida
USC

BIOSTATECH: SPIN-OFF académica da USC



Actividades principais:

- **Consultoría e Asesoramento Biostatístico**
- **Formación** en diversas ferramentas e metodoloxías estatísticas
- **Investigación** en colaboración co GRID[ECMB]

Esta spin-off aposta por R, sendo o software máis utilizado para resolver os traballos propostos polos seus clientes.



Edificio Emprendia
Campus Vida
USC

O futuro...

O futuro...

Aprender...

Mellorar...

Compartir...

O futuro...

Aprender...

Mellorar...

Compartir...



GRAZAS POLA SÚA ATENCIÓN