



R COMO HERRAMIENTA DE TRABAJO EN UNA UNIDAD DE INVESTIGACIÓN BIOMÉDICA

Teresa Seoane Pillado

Unidad de Epidemiología Clínica y Bioestadística.

Complejo Hospitalario Universitario de A Coruña-INIBIC

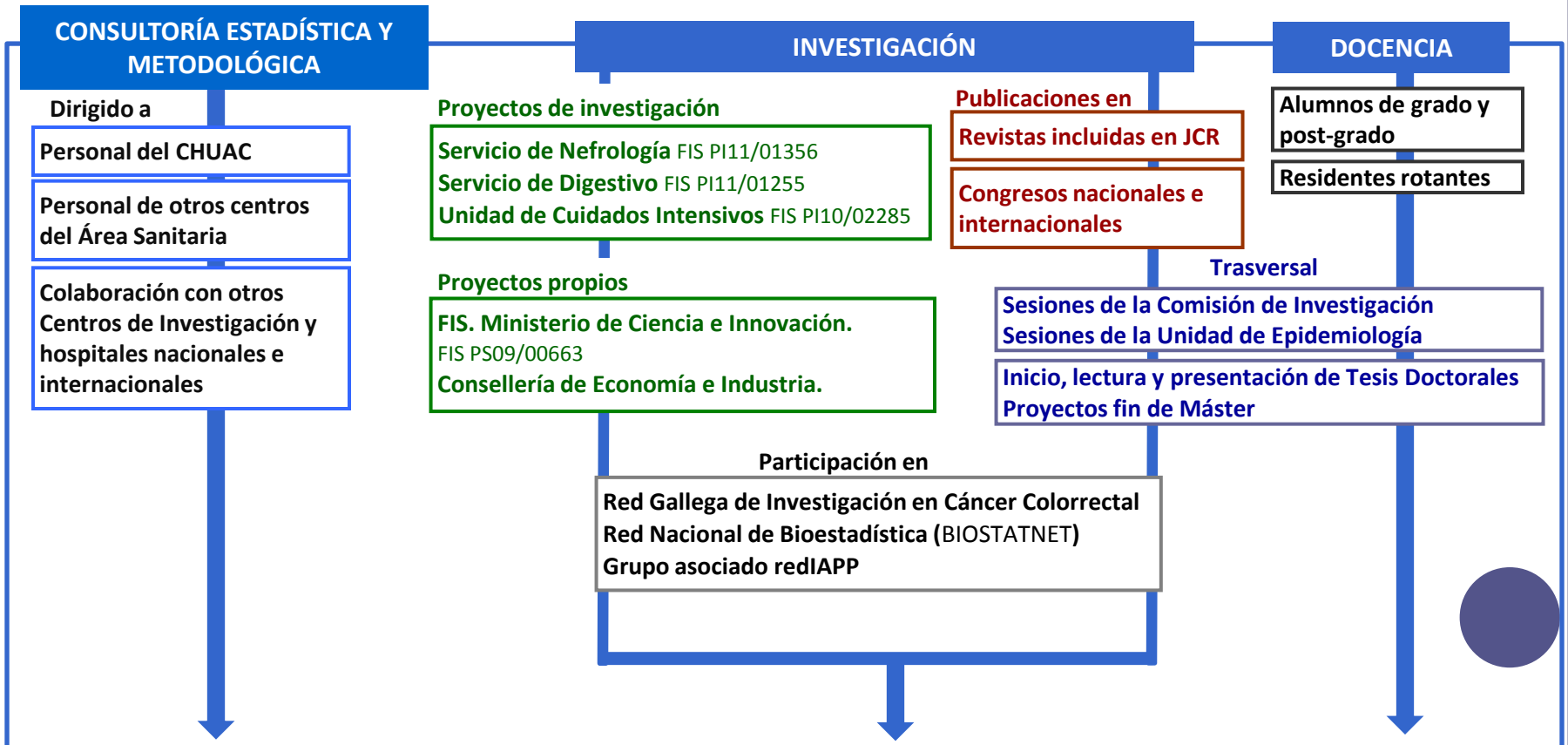
Quienes somos?

Unidad de apoyo metodológico y de soporte en el diseño, realización y análisis de proyectos de investigación en el ámbito de la salud y la biomedicina.



Equipo multidisciplinar formado por:

- Epidemiólogos
- Matemáticos
- Enfermeras
- Personal de apoyo administrativo



Qué hacemos en la parte de consultoría estadística y metodológica?

Atendemos a profesionales sanitarios proporcionando:

Asesoramiento metodológico:

- Diseño de protocolos de estudios de investigación.
- Presentación de proyectos a convocatorias I+D+i de financiación pública o privada
- Ejecución de proyectos académicos

Asesoramiento estadístico:

- Realizamos análisis estadísticos
- Ayudamos en la interpretación y discusión de resultados con la intención de aportar resultados concretos para la toma de decisiones y la obtención de conocimiento.
- Asesoramos en el uso de software de análisis estadístico para que el investigador pueda realizar el análisis de sus datos.

Formación:

- Realizamos actividades formativas sobre metodología de la investigación, bioestadística y manejo de software estadístico



Nuestra actividad

Consultoría:

Diseño del estudio
Análisis estadístico de datos
Interpretación de resultados

| CONSULTAS (1997-2012) | HOSPITALARIAS | EXTRA HOSPITALARIAS | TOTAL |
|--------------------------|---------------|---------------------|----------------|
| Media anual(±DT) | 333.0 (±67.1) | 103.9 (±54.4) | 436.9 (±103.6) |
| Datos brutos | 4995 (76.2%) | 1559 (23.8%) | 6554 |

Investigación:

Desarrollo de proyectos propios (5 proyectos financiados en convocatorias públicas en el 2012)
Participación en proyectos de investigación
Participación en Redes de Investigación Cooperativa

Formación:

Realización de actividades de formación sobre metodología de la investigación y bioestadística.
Cursos sobre metodología de la investigación
Cursos básicos de bioestadística
Cursos avanzados de bioestadística
Cursos de manejo de software estadístico (**R**, SPSS, Stata, Epidat...)

<http://www.fisterra.com/> → 35 trabajos sobre estadística y epidemiología clínica



Trabajando con R...

algunos ejemplos

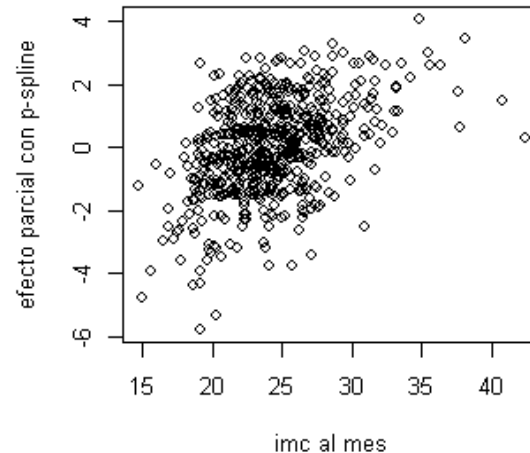
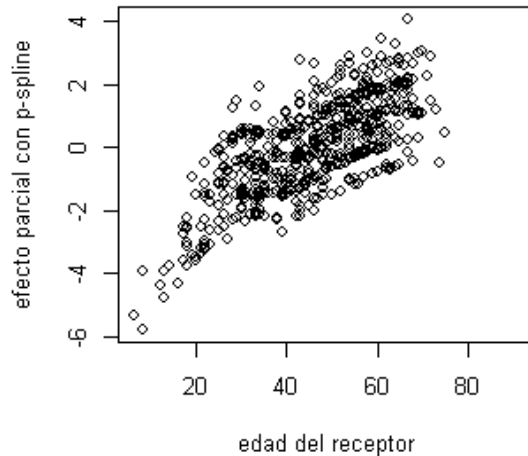


Ejemplo 1:

Ajuste de modelos flexibles mediante p-splines en modelos de regresión para la identificación de posibles factores de riesgo de eventos coronarios en pacientes sometidos a un trasplante renal.

Paquetes R:

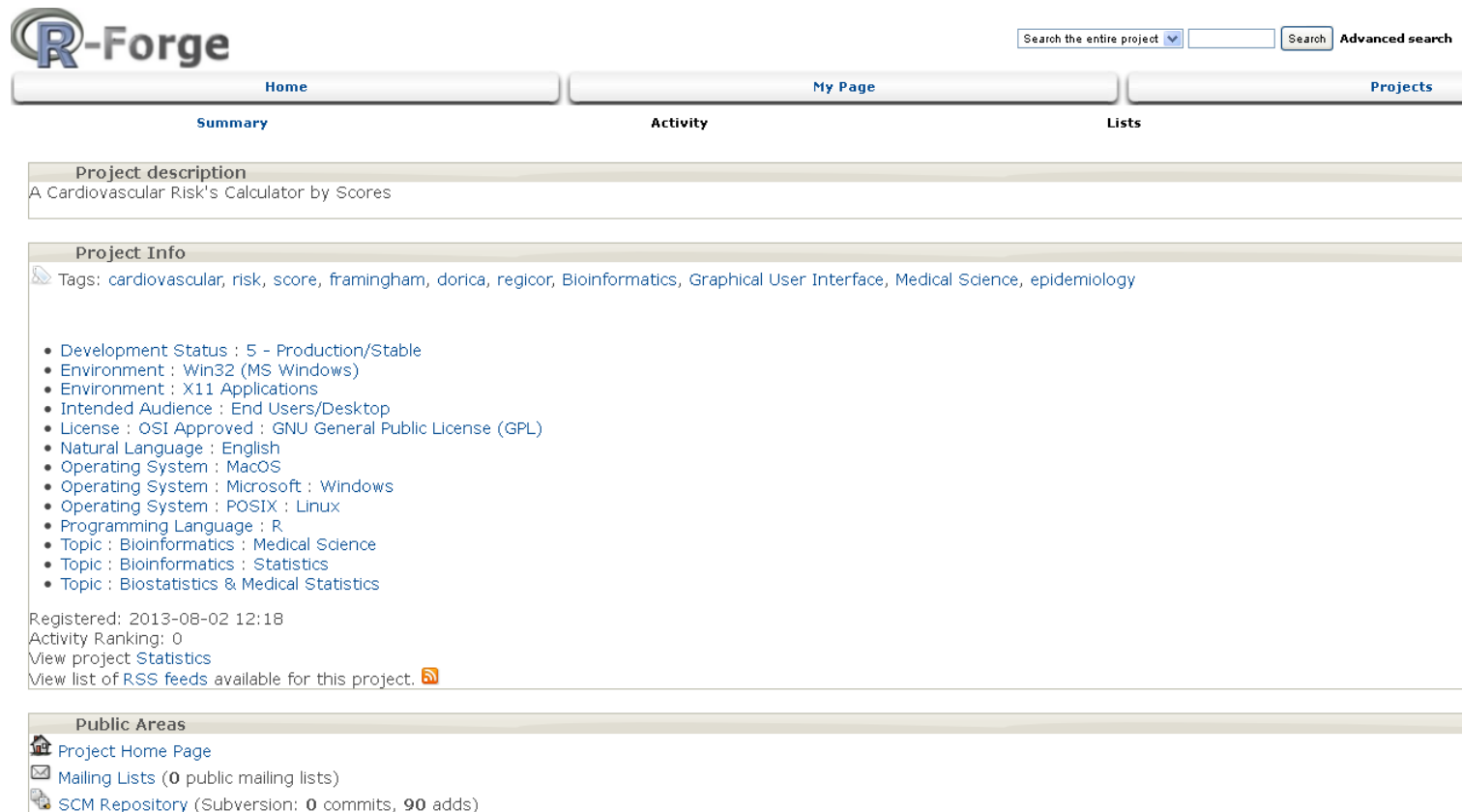
- **survival:** Survival analysis. descriptive statistics, two-sample tests, parametric accelerated failure models, Cox model. Delayed entry (truncation) allowed for all models; interval censoring for parametric models. Case-cohort designs
- **splines:** Regression spline functions and classes



Ejemplo 2:

Implementación en R de una calculadora “**cvrcalc**” de los scores de riesgo cardiovascular Framingham, Regicor, Dorica y Score.

Permite el cálculo masivo de los scores de riesgo cardiovascular a partir de una base de datos. Está publicado en R-Forge (plataforma central para el desarrollo de paquetes de software de R).



The screenshot shows the R-Forge project page for 'cvrcalc'. The page has a header with the R-Forge logo, a search bar, and navigation tabs for Home, My Page, and Projects. Below the tabs are sections for Summary, Activity, and Lists. The main content area is divided into three sections: Project description, Project Info, and Public Areas.

R-Forge Search the entire project Search [Advanced search](#)

[Home](#) [My Page](#) [Projects](#)

Summary **Activity** **Lists**

Project description
A Cardiovascular Risk's Calculator by Scores

Project Info
Tags: cardiovascular, risk, score, framingham, dorica, regicor, Bioinformatics, Graphical User Interface, Medical Science, epidemiology

- Development Status : 5 - Production/Stable
- Environment : Win32 (MS Windows)
- Environment : X11 Applications
- Intended Audience : End Users/Desktop
- License : OSI Approved : GNU General Public License (GPL)
- Natural Language : English
- Operating System : MacOS
- Operating System : Microsoft : Windows
- Operating System : POSIX : Linux
- Programming Language : R
- Topic : Bioinformatics : Medical Science
- Topic : Bioinformatics : Statistics
- Topic : Biostatistics & Medical Statistics

Registered: 2013-08-02 12:18
Activity Ranking: 0
[View project Statistics](#)
[View list of RSS feeds available for this project.](#)

Public Areas

- [Project Home Page](#)
- [Mailing Lists](#) (0 public mailing lists)
- [SCM Repository](#) (Subversion: 0 commits, 90 adds)

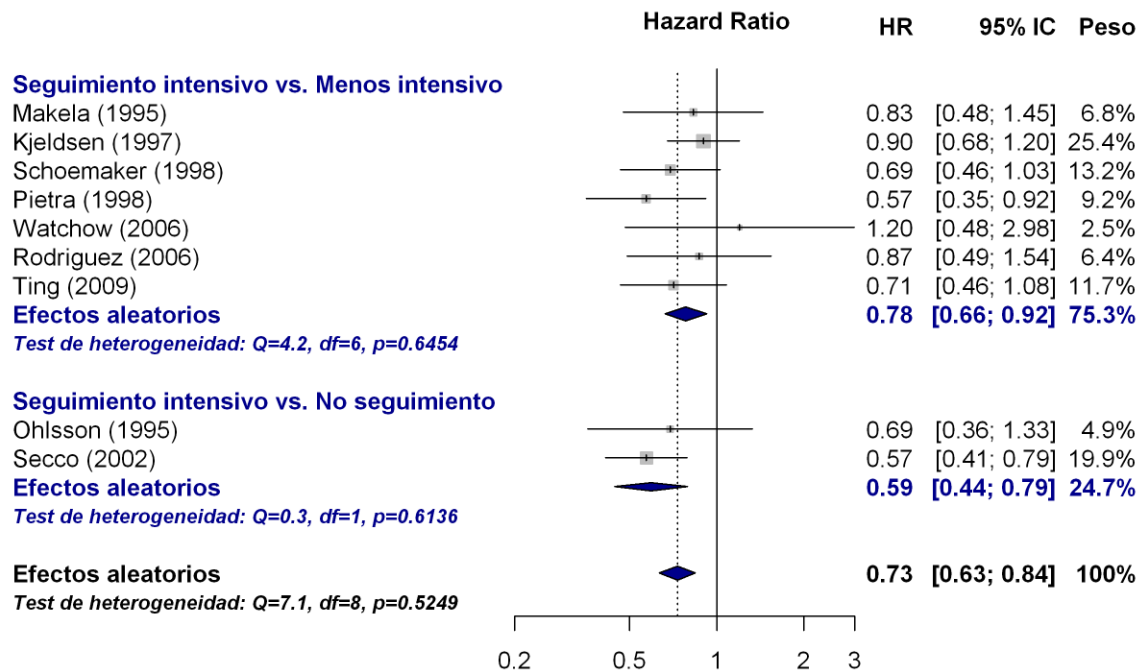


Ejemplo 3:

Metaanálisis para evaluar el impacto de estrategias de seguimiento intensivo en el pronóstico de pacientes con cáncer de colon y recto no metastásico sometidos a cirugía con intención curativa

Paquetes R:

- **meta:** Meta-Analysis with R. Fixed and random effects meta-analysis. Functions for tests of bias, forest and funnel plot.

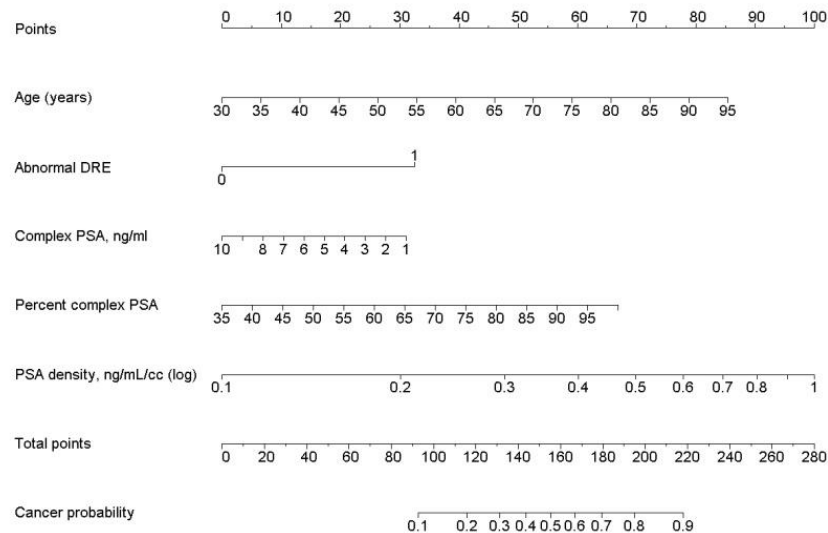


Ejemplo 4:

Construcción y validación de un nomograma para la toma de decisiones en pacientes con sospecha de cáncer de próstata candidatos a biopsia prostática.

Paquetes R:

- **design**
- **Hmisc:** The Hmisc package contains many functions useful for data analysis, high-level graphics, utility operations, functions for computing sample size and power, importing datasets, imputing missing values, advanced table making, variable clustering, character string manipulation, conversion of S objects to LaTeX code, and recoding variables.
- **ROCR:** Visualizing the performance of scoring classifiers. ROC graphs, sensitivity/specificity curves, lift charts, and precision/recall plots are popular examples of trade-off visualizations for specific pairs of performance measures.



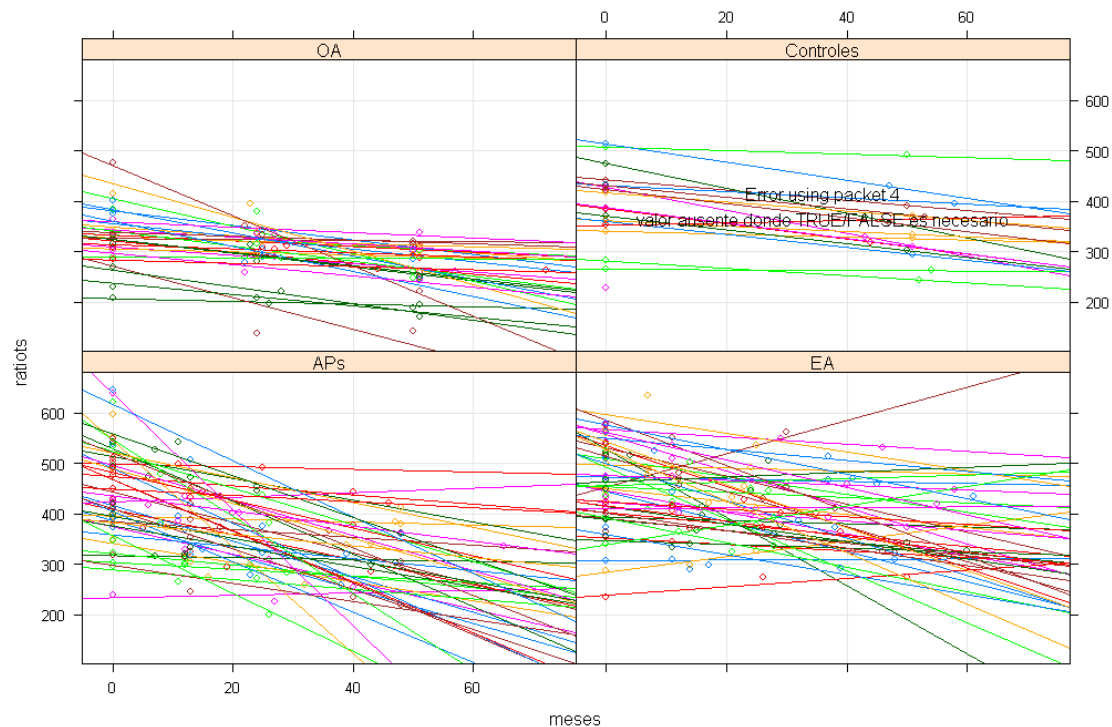
Ejemplo 6:

Modelos mixtos para:

1. Determinar la influencia de los haplogrupos de ADN mitocondrial humano en la progresión de la artrosis.
2. Analizar la variación en la longitud de los telómeros en pacientes con enfermedades reumatológicas.
3. Estudiar el efecto que la modalidad de diálisis peritoneal puede tener sobre las pérdidas proteicas peritoneales en una cohorte de pacientes incidentes en diálisis peritoneal.

Paquetes R:

- **plyr**: Tools for splitting, applying and combining data
- **nlme**: Linear and Nonlinear Mixed Effects Models



Ejemplo 7:

Determinar la incidencia de cáncer en pacientes sometidos a un trasplante renal, utilizando un análisis de riesgos competitivos.

Paquetes R:

- **cmprsk:** Subdistribution Analysis of Competing Risks
- **Survival:** Survival analysis.

Ejemplo 8:

Modelo de regresión logística basados en ecuaciones de estimación generalizada para determinar variables sociodemográficas y psicopatológicas asociadas a los intentos de suicidio atendidos en un hospital general.

Paquetes R:

- **geepack:** Generalized estimating equations solver for parameters in mean, scale, and correlation structures, through mean link, scale link, and correlation link. Can also handle clustered categorical responses

Ejemplo 9:

Análisis de ratios de coste-efectividad para comparar diferentes programas de fisioterapia utilizando técnicas bootstrap.

Paquetes R:

- **boot:** Bootstrap Functions (originally by Angelo Canty for S).

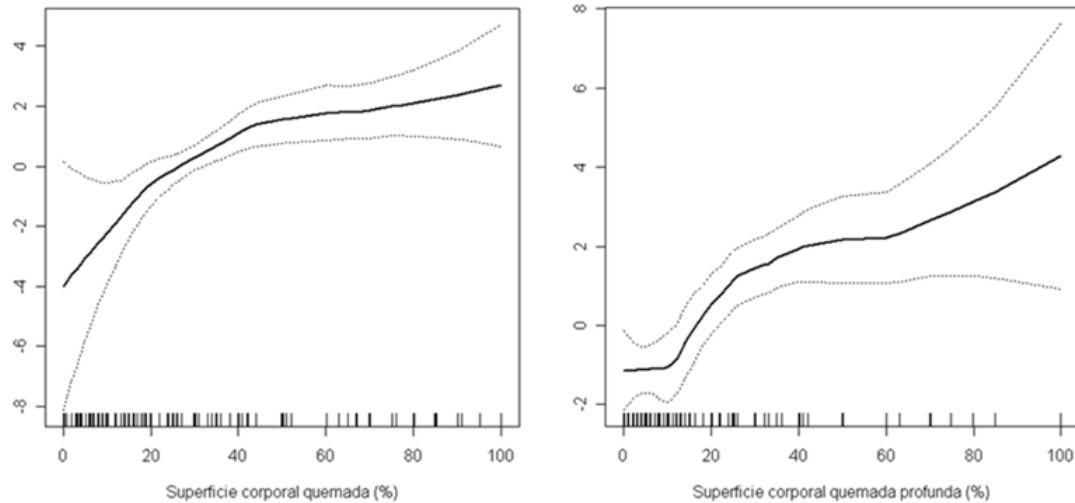


Ejemplo 10:

Modelos aditivos generalizados (GAM) para estudiar la relación de la superficie quemada total y profunda con la mortalidad en pacientes ingresados en una Unidad de Quemados.

Paquetes R:

- **mgcv:** Mixed GAM Computation Vehicle with GCV/AIC/REML smoothness estimation

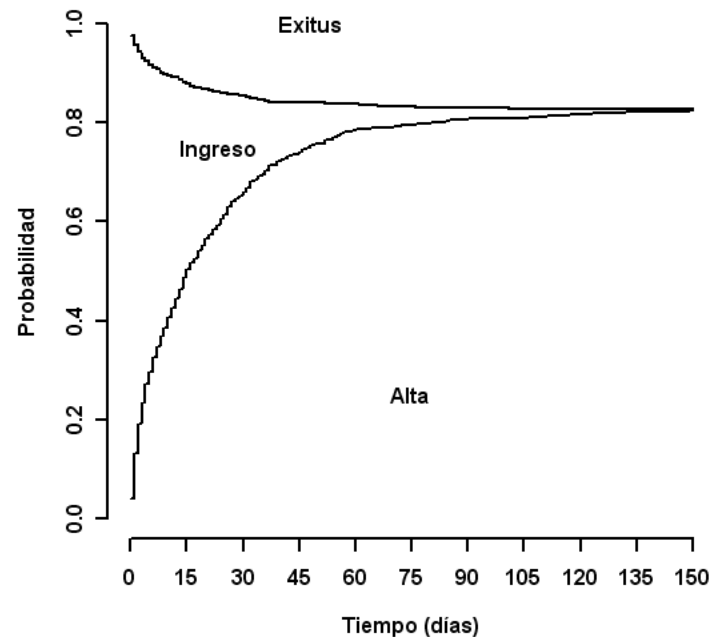


Ejemplo 11:

Análisis del pronóstico de pacientes quemados críticos ingresados en una Unidad de Quemados, mediante modelos de riesgos competitivos.

Paquetes R:

- **cmprsk:** Subdistribution Analysis of Competing Risks



Ejemplo 12:

Estimación de la función de supervivencia con indicadores de censura en datos perdidos aleatoriamente para estudiar la supervivencia específica de cáncer colorrectal.



Formación:

Curso de Metodología de la Investigación

Programa:

...

Introducción al análisis estadístico de datos con R

Introducción a R

Software estadístico y tipo de licencias

Instalación de R y paquetes

Interfaces gráficas: R-commander y Deducer

Manejo de datos con R: organización, carga y edición

Análisis descriptivo y exploratorio de datos

Medidas características (centralización, dispersión y forma, tablas de frecuencias)

Representaciones gráficas (barras, histogramas, sectores, diagramas de cajas)

Introducción a la inferencia estadística: contrastes de hipótesis e intervalos de confianza.

...

Dirigido a:

Especialistas en Formación (R2, R3, R4 ó R5) de la Xerencia de Xestión Integrada A Coruña



Entornos de trabajo:

Hemos colaborado con dos proyectos de la Dirección Xeral de Salud Pública (uno bajo Linux y otro bajo Windows) que facilitan el uso de software bioestadístico y epidemiológico. Incluyen R con entornos gráficos y paquetes específicos.

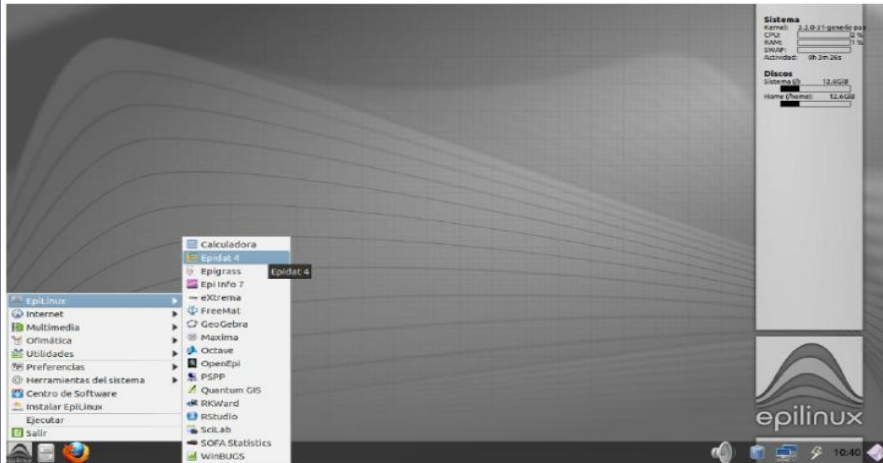
EPILINUX 3

www.galiciasaude.es/dxsp

Sistema Operativo completo y autónomo basado en GNU/Linux Ubuntu 12.04 LTS.

Es una recopilación de software bioestadístico (contiene entre otros R y EpiDat) y programas de uso general.

Se actualiza desde los repositorios originales. Se puede ejecutar desde un DVD o un USB en modo "live", evitando su instalación.



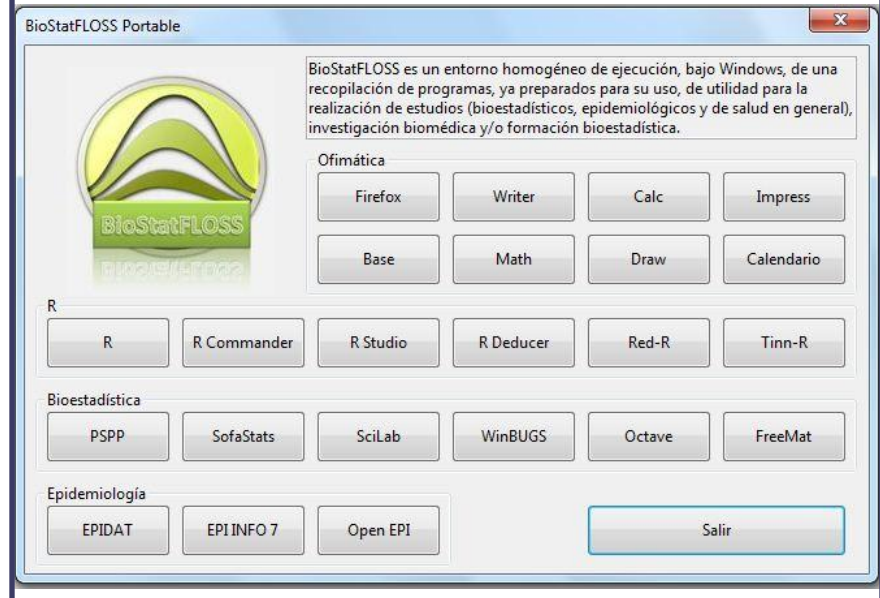
BIOSTATFLOSS

www.galiciasaude.es/dxsp

Entorno unificado y homogéneo de ejecución bajo el sistema operativo Microsoft Windows

Recopilación de software de fuentes abiertas (p.ej. R)

No necesita instalación. Se puede ejecutar desde una carpeta, USB, CDROM, etc.



Gracias...



<http://www.inibic.es>

Edei@sergas.es

