

□

ANÁLISIS DE DATOS ESPACIALES EN EL ÁMBITO DE LA EPIDEMIOLOGÍA

Prof. Dr. Maria A Barceló y Prof. Dr. Marc Saez

8, 10, 14 y 16 de septiembre de 2021

Grupo de Investigación en Estadística, Econometría y Salud (GRECS), Universidad de Girona
CIBER de Epidemiología y Salud Pública (CIBERESP)

INTRODUCCIÓN AL CURSO

1. Introducción al curso
2. Introducción a la epidemiología y la estadística espacial
3. Panorámica de los modelos mixtos
4. Panorámica de los modelos mixtos - Prácticas
5. Introducción a INLA y R INLA
6. R INLA - Prácticas

Miércoles 8

Viernes 10

INTRODUCCIÓN AL CURSO

- 7. Mapas de enfermedades. Estandarización de razones de incidencia y mortalidad
- 8. **Mapas de enfermedades. Suavización de razones de incidencia y de mortalidad estandarizadas**
- 9. Mapas de enfermedades – Prácticas
- 10. Estudios de asociación geográfica. Regresión ecológica espacial
- 11. Regresión ecológica espacial - Prácticas

Martes 14

INTRODUCCIÓN AL CURSO

- 12. Agrupación de casos
- 13. Extensiones: BYM2, procesos puntuales, leaflet, pc priors
- 14. Extensiones – Prácticas

} Jueves 16

SUAVIZACIÓN DE RAZONES DE INCIDENCIA Y MORTALIDAD ESTANDARIZADAS

Recordemos que dijimos que nos interesaría tener en cuenta la componente espacial:

- Porque nos interesa explícitamente el patrón espacial del factor de riesgo: **mapas de enfermedades**
- Porque este recoge gran parte de la confusión no observada: **regresión espacial**
- Porque observamos aglomeraciones en el espacio: **detección de agrupaciones**
- Porque nos interesan los efectos de un foco contaminante en la salud de los habitantes residentes en sus alrededores: **identificación de focos**

SUAVIZACIÓN DE RAZONES DE INCIDENCIA Y MORTALIDAD ESTANDARIZADAS

Recordemos:

- los **mapas de enfermedades** proporcionan una primera visión de la distribución espacial de la enfermedad, del suceso de salud o bien de los factores de riesgo de los mismos. Son un resumen visual del riesgo geográfico.
- los indicadores de ocurrencia de la enfermedad más habitualmente representados en estos mapas son los indicadores de morbilidad (prevalencia e incidencia) y los indicadores de mortalidad.

SUAVIZACIÓN DE RAZONES DE INCIDENCIA Y MORTALIDAD ESTANDARIZADAS

- no es correcto utilizar números absolutos, ni tasas crudas. Como mínimo, deberían utilizarse tasas estandarizadas, las RIE (SIR) o las RME (SMR).
- para **estimar el riesgo relativo**, se utilizaría una **regresión de Poisson** o bien una **regresión binomial** en el caso de enfermedades más frecuentes.
- aun cuando utilicemos una regresión para estimar el riesgo relativo, no olvidemos que los modelos de regresión continúan estimando las RIE (o las RME).

SUAVIZACIÓN DE RAZONES DE INCIDENCIA Y MORTALIDAD ESTANDARIZADAS

Sin embargo, hemos visto que las tasas estandarizadas tienen una serie de problemas:

- **En el caso de enfermedades raras y/o áreas pequeñas, las RME (o RIE) son muy imprecisas** porque la varianza es proporcional al cuadrado del denominador (es decir, es proporcional al número de casos esperados al cuadrado)
- La **varianza** asociada a **áreas con casos esperados pequeños será muy alta**
- Las RME (o RIE) **se estiman independientemente en cada área**
- Las RME (o RIE) **no tienen en cuenta la muy probable dependencia espacial**

SUAVIZACIÓN DE RAZONES DE INCIDENCIA Y MORTALIDAD ESTANDARIZADAS

Posible solución:

- Utilizar la razón de incidencia suavizada (RIE suavizada) (en inglés, *smoothed standardized incidence ratio*, smoothed SIR) y/o la razón de mortalidad suavizada (RME suavizada) (en inglés, *smoothed standardized mortality ratio*, smoothed SMR).

SUAVIZACIÓN DE RAZONES DE INCIDENCIA Y MORTALIDAD ESTANDARIZADAS

Resumiendo:

- los ***problemas*** que encontramos al estimar razones estandarizadas (tanto de mortalidad como de incidencia) son la ***inestabilidad de los estimadores*** y los ***efectos de no tener en cuenta la muy probable dependencia espacial***.
- Estos problemas pueden solucionarse en parte mediante el ***suavizado espacial de las razones***.
- Existen diversos métodos para suavizar estas razones.

SUAVIZACIÓN DE RAZONES DE INCIDENCIA Y MORTALIDAD ESTANDARIZADAS

Métodos para suavizar mapas

- Algoritmos de suavizado local, ad hoc, por ejemplo, medias móviles espaciales, algoritmos de headbanging
- Análisis de superficies de tendencia, por ejemplo, kriging, suavizado mediante polinomios/splines
- Modelos de efectos aleatorios

SUAVIZACIÓN DE RAZONES DE INCIDENCIA Y MORTALIDAD ESTANDARIZADAS

Métodos para suavizar mapas

- **Algoritmos de suavizado local, ad hoc**, por ejemplo, medias móviles espaciales, algoritmo de headbanging
 - Ventajas: rápidos y sencillos de implementar
 - Inconvenientes: pueden ser muy sensibles a la elección ad hoc de ponderaciones, etc. y no hay estimaciones para la incertidumbre (errores estándar)

SUAVIZACIÓN DE RAZONES DE INCIDENCIA Y MORTALIDAD ESTANDARIZADAS

Métodos per suavizar mapas

- **Análisis de superficies de tendencia**, por ejemplo, kriging, suavizado mediante polinomios/splines
 - Ventajas: la estimación de los “parámetros de suavizado” está basada en el equilibrio entre el ajuste y la suavización y los errores standard están generalmente disponibles
 - Inconvenientes: pueden ser sensibles a la elección del parámetro de penalización para el equilibrio

SUAVIZACIÓN DE RAZONES DE INCIDENCIA Y MORTALIDAD ESTANDARIZADAS

Métodos para suavizar mapas

- **Modelos de efectos aleatorios:** las alternativas más populares a los métodos vistos anteriormente consisten en los modelos de efectos aleatorios
- Dos tipos de modelos aleatorios:
 - ***Empirical Bayes:*** incluyen como efecto aleatorio sólo la heterogeneidad espacial
 - ***Modelos jerárquicos Bayesianos (modelos mixtos):*** incluyen dos efectos aleatorios, la heterogeneidad espacial y la dependencia espacial

SUAVIZACIÓN DE RAZONES DE INCIDENCIA Y MORTALIDAD ESTANDARIZADAS

Métodos para suavizar mapas

➤ Modelos de efectos aleatorios

Anteriormente al gran desarrollo informático que se produjo a partir del año 2000, se utilizaban básicamente los modelos empíricos Bayes, ya que los otros modelos eran computacionalmente complejos. **Actualmente**, estos han quedado obsoletos y se utilizan los **modelos jerárquicos Bayesianos**.

SUAVIZACIÓN DE RAZONES DE INCIDENCIA Y MORTALIDAD ESTANDARIZADAS

Modelos de efectos aleatorios

Con el objetivo de solucionar los problemas que hemos visto que presentan las RME y/o RIE, se utilizan estimadores suavizados bayesianos de las RME y/o RIE obtenidos mediante:

- ***Modelo de Poisson-lognormal-espacial*** (modelo jerárquico Bayesiano con un efecto aleatorio)
- ***Modelo de Besag, York y Mollié, BYM*** (modelo jerárquico Bayesiano con dos efectos aleatorios)

SUAVIZACIÓN DE RAZONES DE INCIDENCIA Y MORTALIDAD ESTANDARIZADAS

Modelos de efectos aleatorios: Modelo de Poisson-lognormal-espacial

Poisson-logNormal model

$$\begin{aligned}O_i &\sim \text{Poisson}(\lambda_i E_i) \\ \log \lambda_i &= \alpha + V_i \\ V_i &\sim \text{Normal}(0, \sigma_v^2)\end{aligned}$$

Priors (vague, non informative):

- between-area variance σ_v^2 :
 $\sigma_v^2 \sim \text{Inverse Gamma}(0.5, 0.0005) \Leftrightarrow \tau_v \sim \text{Gamma}(0.5, 0.0005)$
 $\sigma_v \sim \text{Truncated Normal}(0, 100)_{[0, \text{Inf})}$
- mean log relative risk: $\alpha \sim N(0, 100)$

SUAVIZACIÓN DE RAZONES DE INCIDENCIA Y MORTALIDAD ESTANDARIZADAS

Modelos de efectos aleatorios: Modelo de Poisson-lognormal-espacial

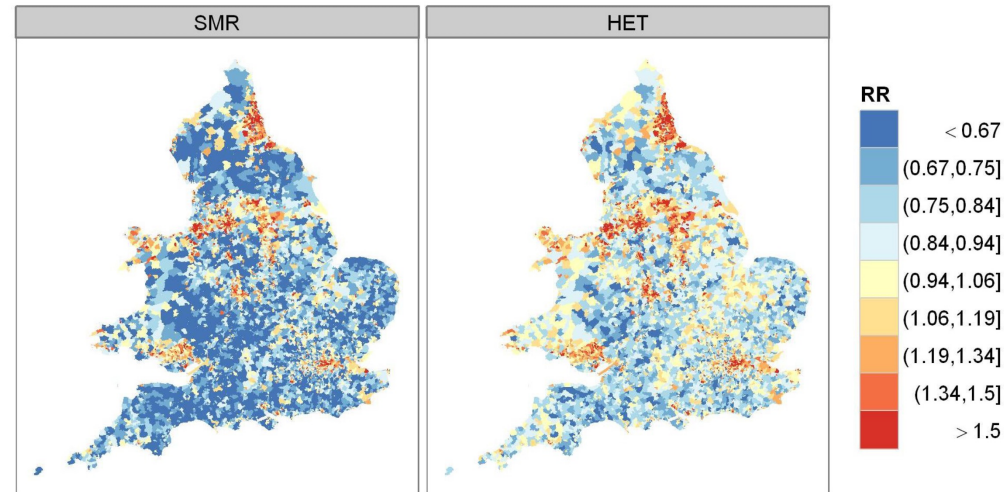
- O_i , E_i son el número de casos observados y esperados en el área i , respectivamente
- $\lambda_i = \exp(\alpha + V_i)$ es el riesgo relativo (RR) en el área i comparado con el riesgo esperado en base a la edad y sexo de la población
- Parámetros V_i son efectos aleatorios específicos para cada área para controlar la heterogeneidad (sobredispersión)

SUAVIZACIÓN DE RAZONES DE INCIDENCIA Y MORTALIDAD ESTANDARIZADAS

Modelos de efectos aleatorios: Modelo de Poisson-lognormal-espacial

Lung cancer incidence in males, 1985-2009, England and Wales (I)

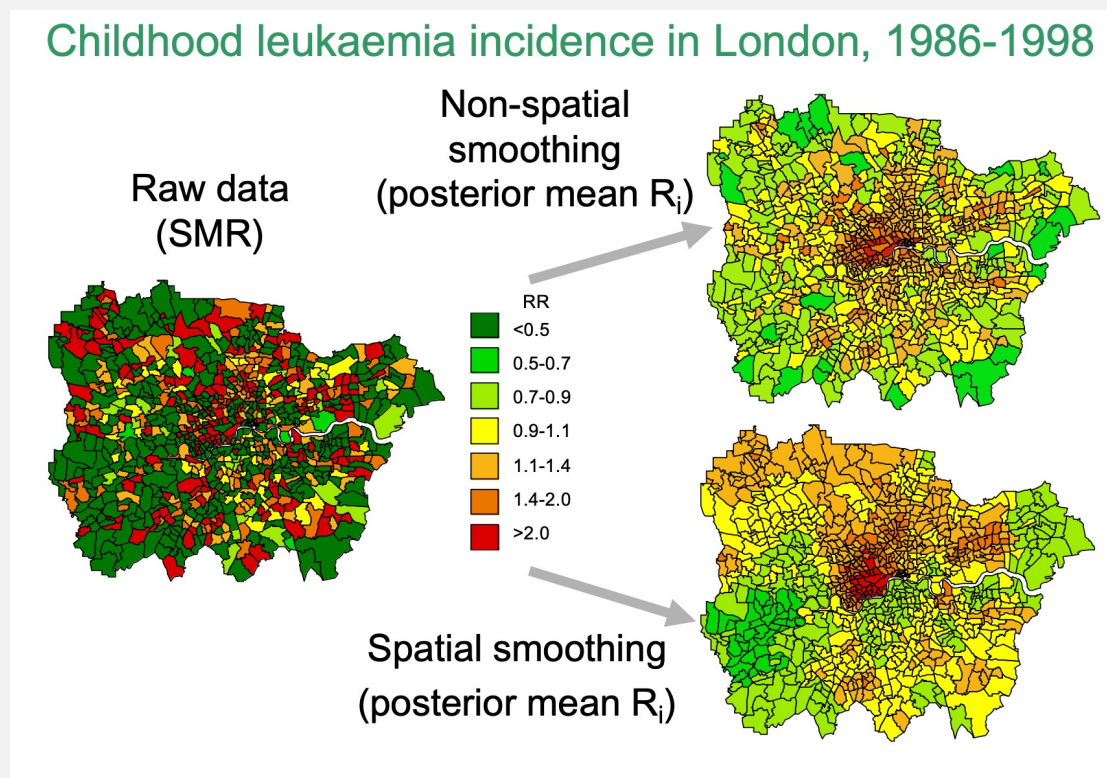
RR estimates using 2 methods



SMRs and non-spatially smoothed RRs

SUAVIZACIÓN DE RAZONES DE INCIDENCIA Y MORTALIDAD ESTANDARIZADAS

Modelos de efectos aleatorios: Modelo de Poisson-lognormal-espacial



SUAVIZACIÓN DE RAZONES DE INCIDENCIA Y MORTALIDAD ESTANDARIZADAS

Modelos de efectos aleatorios: Model BYM

- *El efecto aleatorio utilizado para recoger la heterogeneidad espacial* es el mismo en todos los modelos (iid normal).
- *El efecto aleatorio para recoger la dependencia espacial* se puede aproximar mediante un CAR (Conditional Autoregressive) (utilizado al tener datos de área)

Actualmente, la dependencia espacial se aproxima mediante una Mátern (**modelo denominado log-Gaussian-Cox**, utilizado para todo tipo de datos espaciales).

SUAVIZACIÓN DE RAZONES DE INCIDENCIA Y MORTALIDAD ESTANDARIZADAS

Modelos de efectos aleatorios: Modelo BYM

Los λ_i suelen estar correlacionados espacialmente porque reflejan, en parte, factores de riesgo que varían espacialmente.

Para corregir esto deberíamos:

- incorporar la dependencia espacial en la distribución de los λ_i
- en el modelo BYM, esta dependencia se modeliza como un CAR (modelo condicional autoregresivo)

SUAVIZACIÓN DE RAZONES DE INCIDENCIA Y MORTALIDAD ESTANDARIZADAS

➤ Modelo BYM. Modelo CAR

$$\log(R_i) \sim \text{Normal}(m_i, v_i)$$

$m_i = \sum_k R_k / n_i$ = average risk in neighbouring areas

$v_i = v / n_i \rightarrow$ variance inversely proportional to number of neighbours

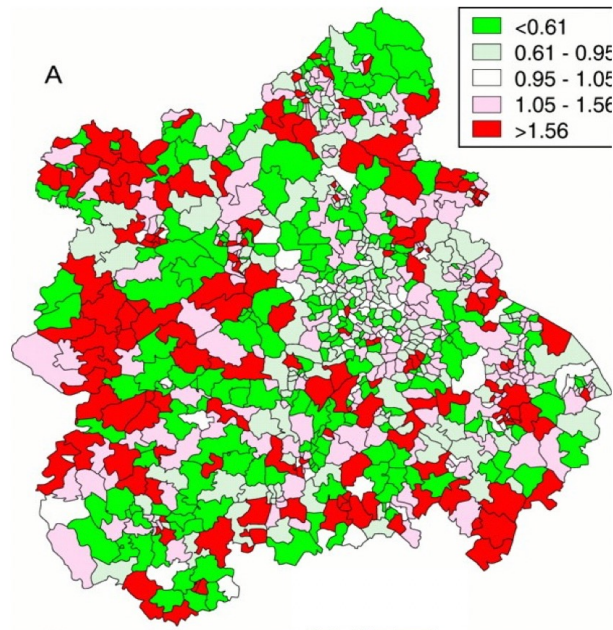
Besag, York, Mollie (1991) *Annals of the Institute of Statistics and Mathematics*, 43: 1-59

SUAVIZACIÓN DE RAZONES DE INCIDENCIA Y MORTALIDAD ESTANDARIZADAS

Ejemplo:

Map of SMR of adult leukaemia in West Midlands Region, England 1974-86

(Olsen, Martuzzi and Elliott, *BMJ* 1996;313:863-866).



Is the variability real
or simply reflecting
unequal expected
counts ?

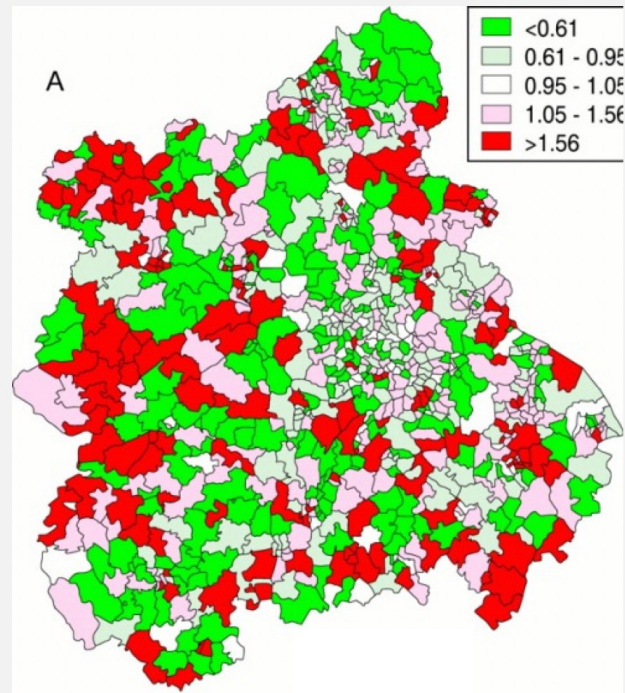
Have the red
highlighted areas
truly got a raised
relative risk?

SUAVIZACIÓN DE RAZONES DE INCIDENCIA Y MORTALIDAD ESTANDARIZADAS

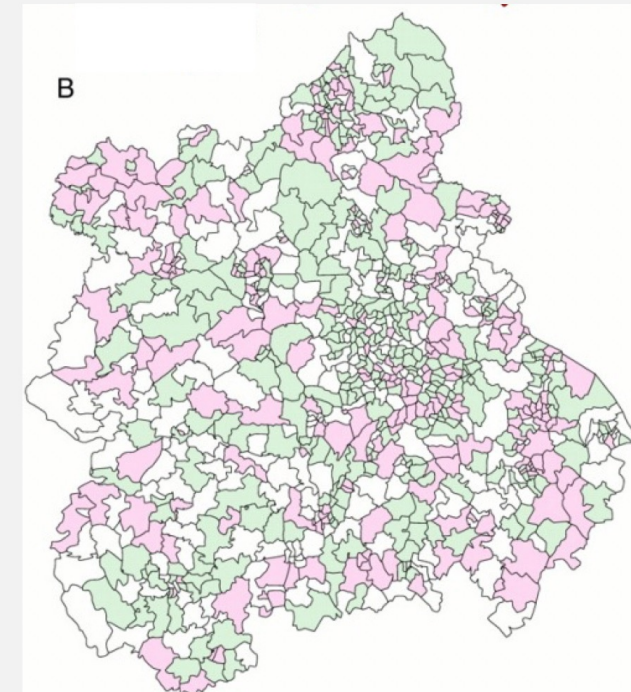
Ejemplo:

Map of SMR of adult leukaemia in West Midlands Region,
England 1974-86
(Olsen, Martuzzi and Elliott, *BMJ* 1996;313:863-866).

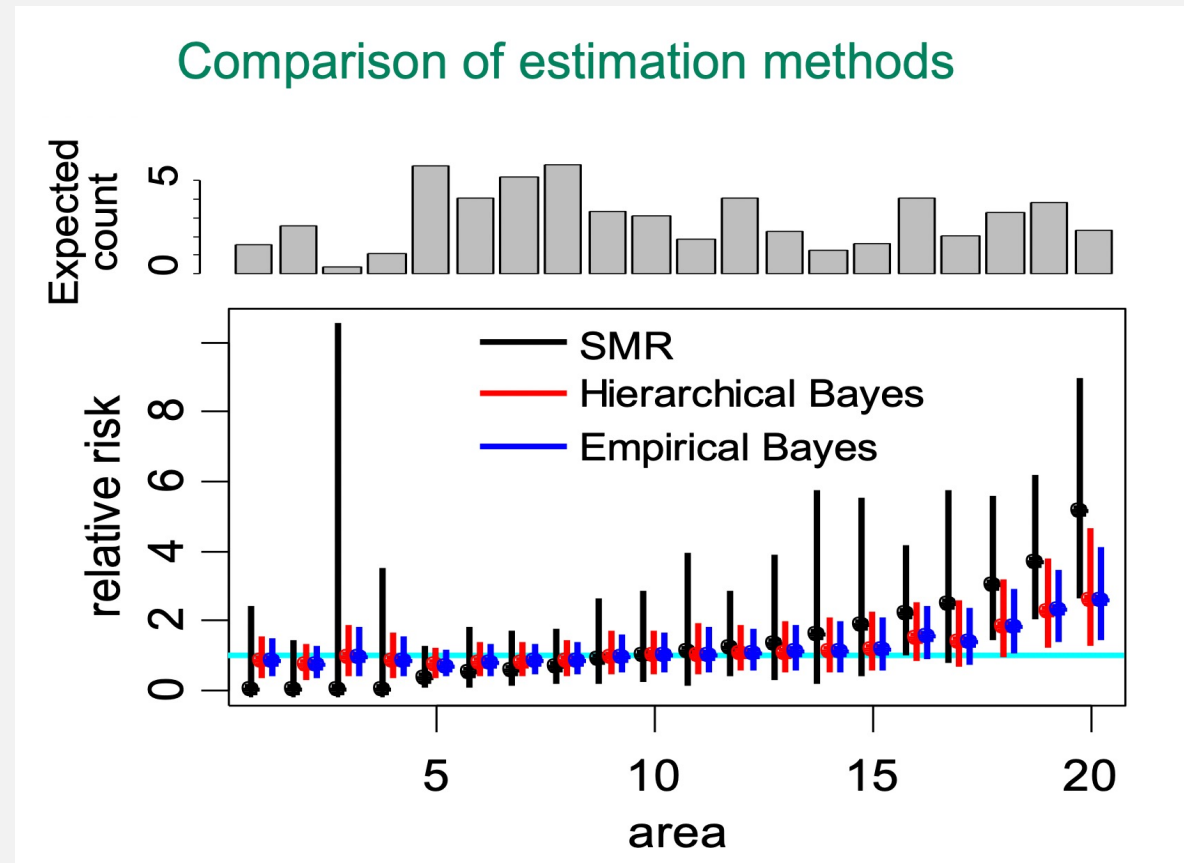
(A) unsmoothed SMR



(B) smoothed by Bayesian methods



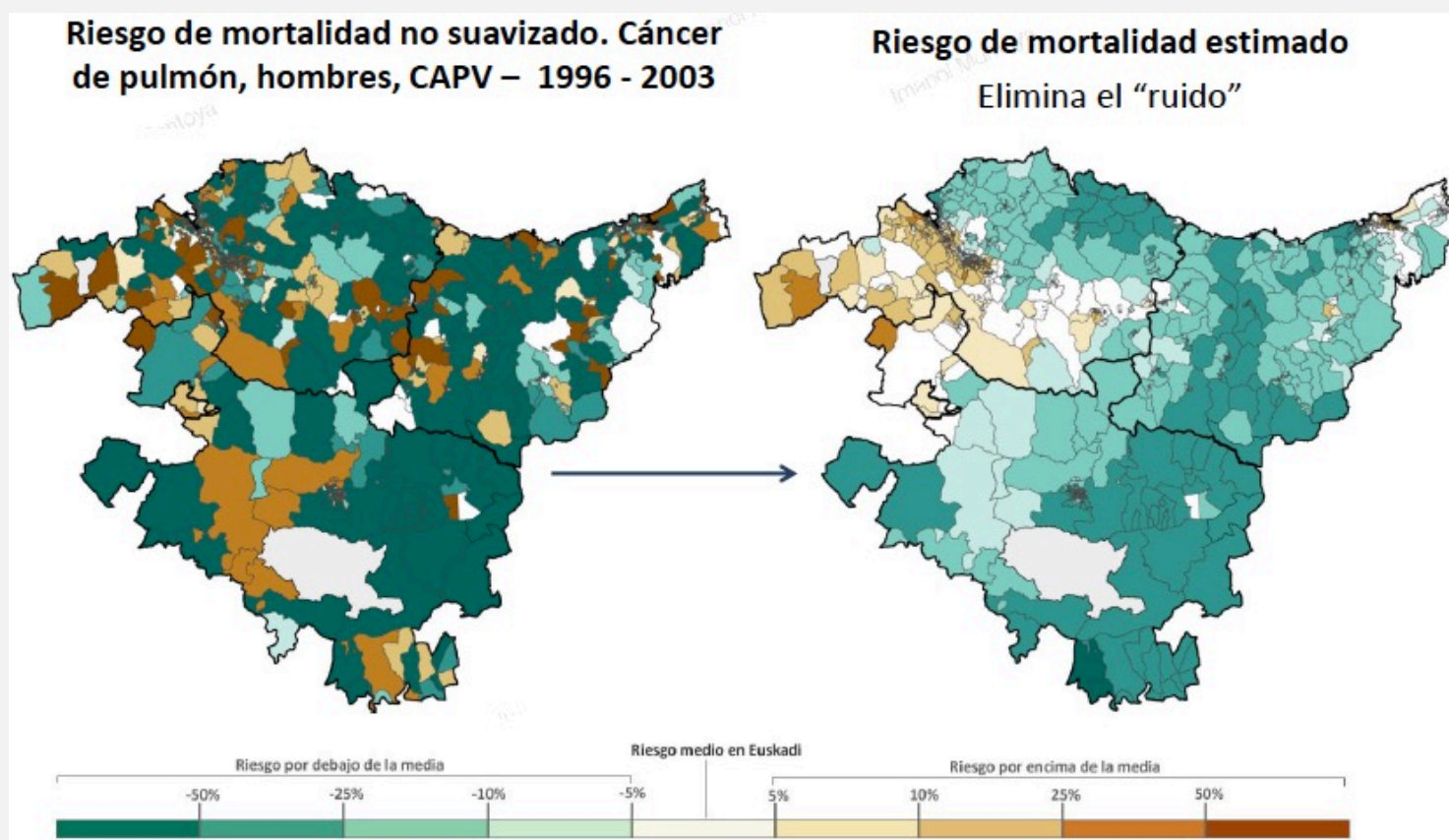
SUAVIZACIÓN DE RAZONES DE INCIDENCIA Y MORTALIDAD ESTANDARIZADAS



8. Mapas de enfermedades. Suavización de razones de incidencia y mortalidad estandarizadas

SUAVIZACIÓN DE RAZONES DE INCIDENCIA Y MORTALIDAD ESTANDARIZADAS

Ejemplo:



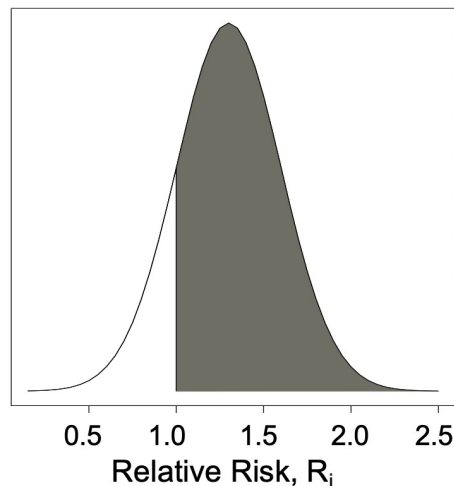
8. Mapas de enfermedades. Suavización de razones de incidencia y mortalidad estandarizadas

SUAVIZACIÓN DE RAZONES DE INCIDENCIA Y MORTALIDAD ESTANDARIZADAS

Modelos de efectos aleatorios: Modelo de Poisson-lognormal-espacial

Mapping uncertainty

- Mapping the mean posterior value of R_i does not make full use of the posterior distribution



- Map **posterior SD**
- Map **Probability ($R_i > 1$)**

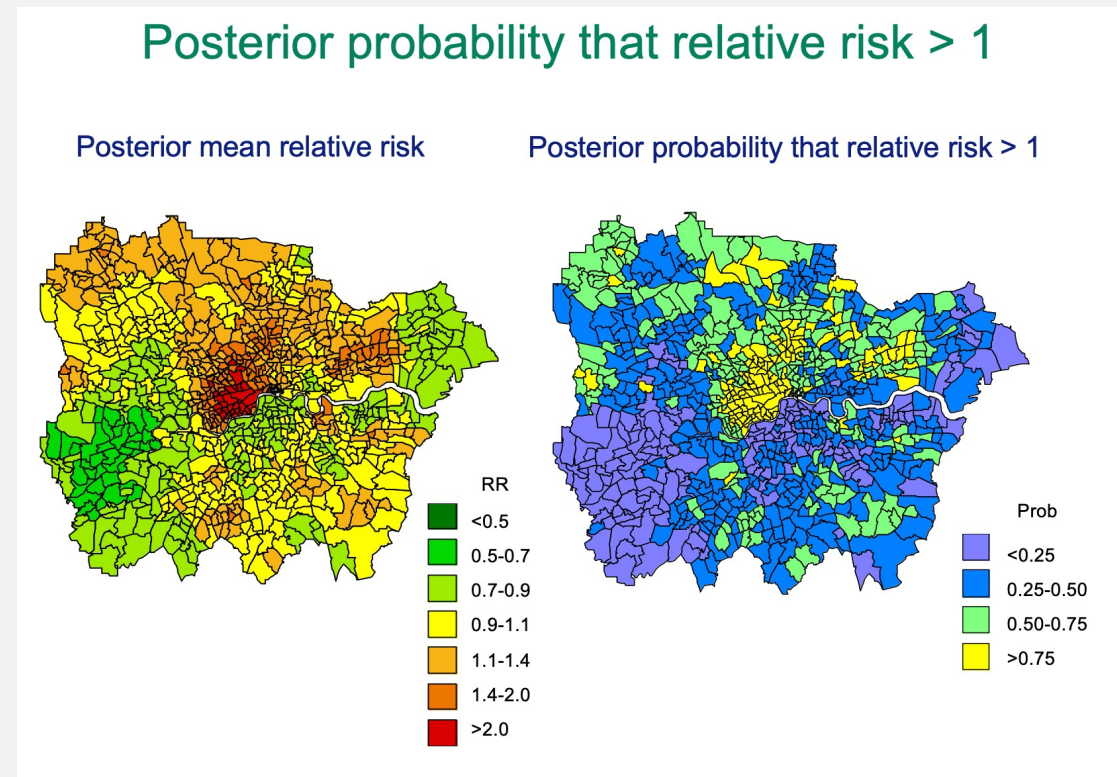
Note – this is not the same as a classical p-value

SUAVIZACIÓN DE RAZONES DE INCIDENCIA Y MORTALIDAD ESTANDARIZADAS

- Richardson et al. (2004): estudio de simulación que investiga el uso de probabilidades a posteriori **en estudios enfocados en la representación de enfermedades en mapas**
 - Se clasificará una área como de elevado riesgo si $[Prob(R_i > 1)] > 0.8$
 - Se clasificará una área como de elevada especificidad si $falsa\ detecció < 10\%$
 - Sensibilidad 60%-95% para E_i de 5-20 y verdadero R_i de 1.5-3.0

SUAVIZACIÓN DE RAZONES DE INCIDENCIA Y MORTALIDAD ESTANDARIZADAS

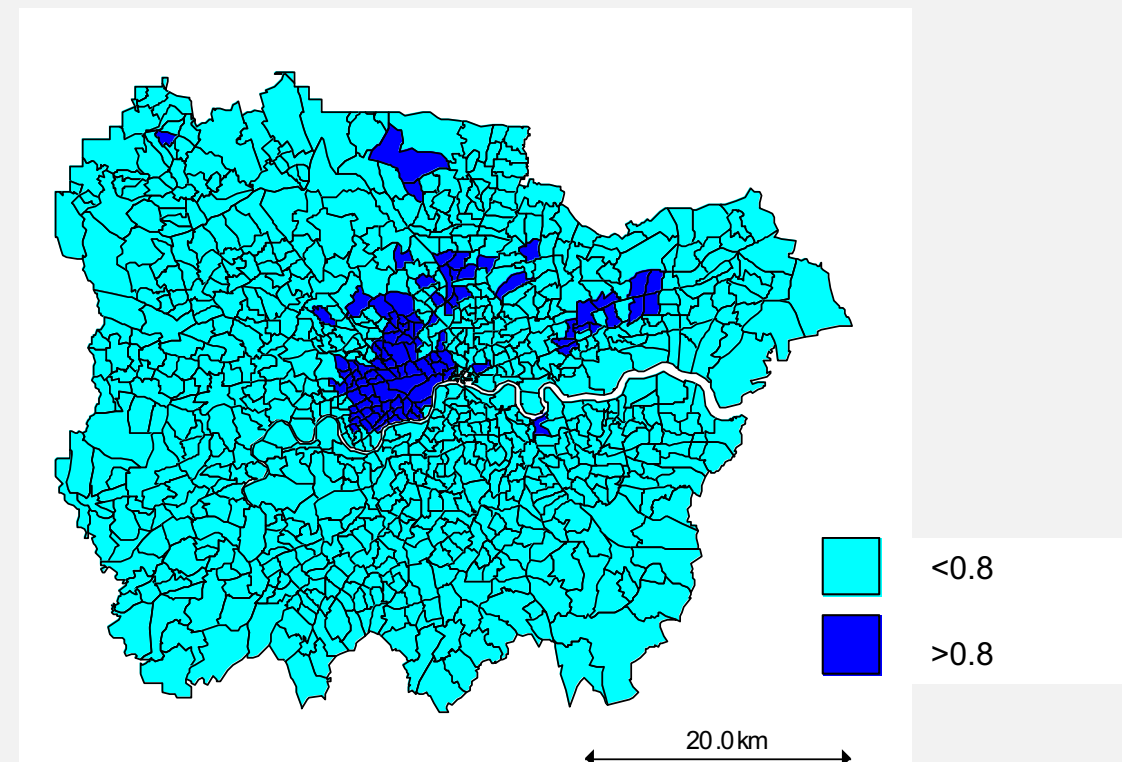
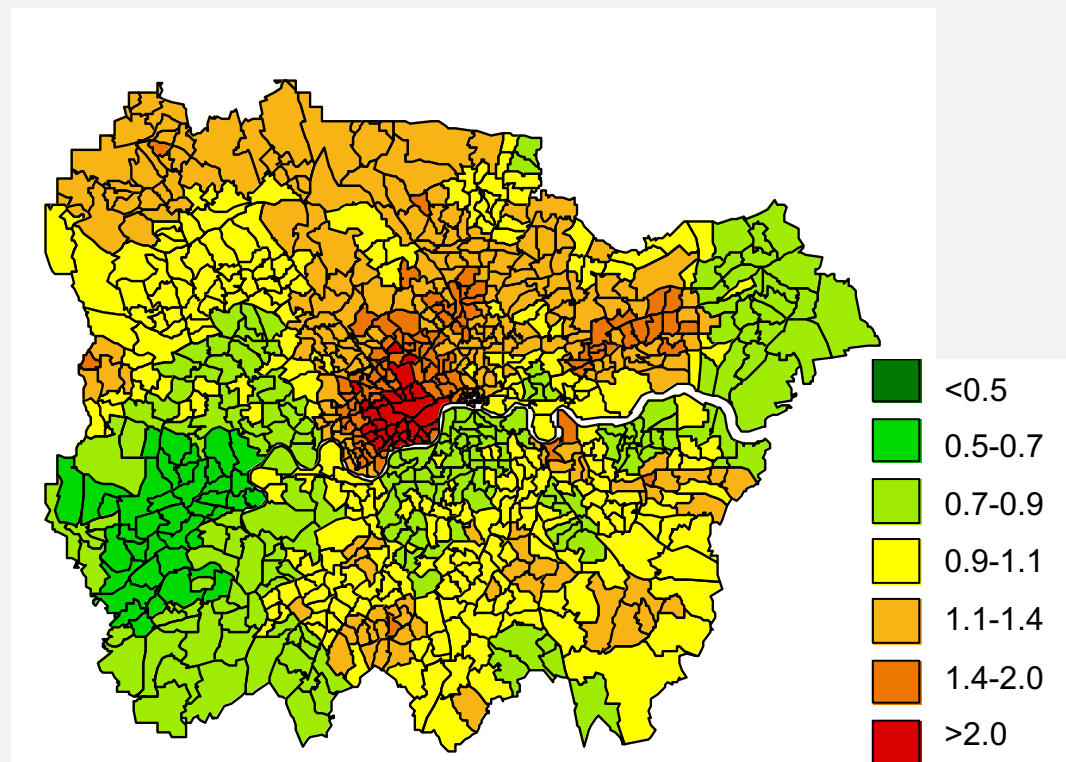
Ejemplo: Childhood leukaemia incidence in London, 1986-1998



8. Mapas de enfermedades. Suavización de razones de incidencia y mortalidad estandarizadas

SUAVIZACIÓN DE RAZONES DE INCIDENCIA Y MORTALIDAD ESTANDARIZADAS

Ejemplo: Childhood leukaemia incidence in London, 1986-1998



SUAVIZACIÓN DE RAZONES DE INCIDENCIA Y MORTALIDAD ESTANDARIZADAS

Software

- Para realizar el suavizado mediante Empirical Bayes, se puede utilizar la Rapid Inquiry Facility (RIF) de la SAHSU del Imperial College.
- La estimación de modelos jerárquicos bayesianos se hacía mediante métodos de simulación computacionalmente complejos (MCMC) implementados en el programario gratuito WinBUGS y GeoBUGS
- Desde 2008, el programario gratuito INLA (Rue et al, 2008) implementa una aproximación rápida y más eficiente a los MCMC para estimar estos modelos.

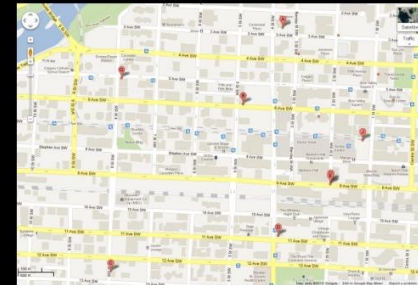
GIS



What my friends think I do



What my mom thinks I do



What society thinks I do



What my clients think I do



What I think I do



What I really do