

□

ANÀlisi DE DADES ESPACIALS EN L'ÀMBIT DE L'EPIDEMIOLOGIA

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

8, 10, 14 i 16 de setembre de 2021

Grup de Recerca en Estadística, Econometria i Salut (GRECS), Universitat de Girona
CIBER d'Epidemiologia i Salut Pública(CIBERESP)

INTRODUCCIÓ AL CURS

1. Introducció al curs
2. Introducció a l'epidemiologia i l'estadística espacial
3. Panoràmica del models mixtos
4. Panoràmica del models mixtos - Pràctiques
5. Introducció a INLA i R INLA
6. R INLA - Pràctiques

Dimecres 8

Divendres 10

INTRODUCCIÓ AL CURS

- 7. Mapes de malalties. Estandardització de raons d'incidència i mortalitat
- 8. Mapes de malalties. Suavització de raons d'incidència i de mortalitat estandarditzades**
- 9. Mapes de malalties – Pràctiques
- 10. Estudis d'associació geogràfica. Regressió ecològica espacial
- 11. Regressió ecològica espacial - Pràctiques

Dimarts 14

INTRODUCCIÓ AL CURS

- 12. Agrupació de casos
- 13. Extensions: BYM2, processos puntuals, leaflet, pc priors
- 14. Extensions – Pràctiques

} Dijous 16

SUAVITZACIÓ DE RAONS D'INCIDÈNCIA I MORTALITAT ESTANDARDITZADES

Recordem que vam dir que ens interessaria tenir en compte el component espacial:

- Perquè ens interessa explícitament el patró espacial del factor de risc: **mapes de malalties**
- Perquè aquest recull gran part de la confusió no observada: **regressió espacial**
- Perquè observem aglomeracions en l'espai: **detecció d'agrupacions**
- Perquè ens interessin els efectes d'un focus contaminant en la salut dels habitants residents en els seus voltants: **identificació de focus**

SUAVITZACIÓ DE RAONS D'INCIDÈNCIA I MORTALITAT ESTANDARDITZADES

Recordem:

- els **mapes de malalties** proporcionen una primera visió de la distribució espacial de la malaltia, de l'esdeveniment de salut o bé dels factors de risc dels mateixos. Són un resumen visual del risc geogràfic.
- els indicadors d'ocurrència de la malaltia més habitualment representats en aquests mapes són els indicadors de morbiditat (prevalença i incidència) i els indicadors de mortalitat.

SUAVITZACIÓ DE RAONS D'INCIDÈNCIA I MORTALITAT ESTANDARDITZADES

- no és correcte utilitzar nombres absoluts, ni taxes crues. Com a mínim, s'haurien d'utilitzar taxes estandarditzades, les RIE (SIR) o les RME (SMR).
- ***per estimar el risc relatiu***, s'utilitzaria una ***regressió de Poisson*** o bé una ***regressió binomial*** en el cas de malalties més freqüents.
- encara que fem servir una regressió per estimar el risc relatiu, no oblidem que els models de regressió continúen estimant les RIE (o les RME).

SUAVITZACIÓ DE RAONS D'INCIDÈNCIA I MORTALITAT ESTANDARDITZADES

Ara bé, hem vist que les taxes estandarditzades tenen una sèrie de problemes:

- **En el cas de malalties rares i/o àrees petites, les RME (o RIE) són molt imprecises** perquè la variància és proporcional al quadrat del denominador (és a dir, és proporcional al nombre de casos esperats al quadrat)
- La **variància** associada a les **àrees que tinguin casos esperats petits serà molt alta**
- Les RME (o RIE) **s'estimen independentment en cada àrea**
- Les RME (o RIE) **no tenen en compte la molt probable dependència espacial**

SUAVITZACIÓ DE RAONS D'INCIDÈNCIA I MORTALITAT ESTANDARDITZADES

Possible solució:

- Utilitzar la raó d'incidència suavitzada (RIE suavitzada) (amb anglès *smoothed standardized incidence ratio*, smoothed SIR) i/o la raó de mortalitat suavitzada (RME suavitzada) (amb anglès *smoothed standardized mortality ratio*, smoothed SMR).

SUAVITZACIÓ DE RAONS D'INCIDÈNCIA I MORTALITAT ESTANDARDITZADES

- Resumint, els **problemes** que ens trobem a l'estimar raons estandarditzades (tant de mortalitat com d'incidència) són la **inestabilitat dels estimadors** i els **efectes de no tenir en compte la molt probable dependència espacial**.
- Aquests problemes poden solucionar-se en part mitjançant el **suavitzat espacial de les raons**.
- Hi ha diversos mètodes per suavitzar aquestes raons.

SUAVITZACIÓ DE RAONS D'INCIDÈNCIA I MORTALITAT ESTANDARDITZADES

Mètodes per suavitzar mapes

- Algorismes de suavitzat local, ad hoc, per exemple, mitjanes mòbils espacials, algorisme de headbanging
- Anàlisi de superfícies de tendència, per exemple, kriging, suavitzat mitjançant polinomis/splines
- Models d'efectes aleatoris

SUAVITZACIÓ DE RAONS D'INCIDÈNCIA I MORTALITAT ESTANDARDITZADES

Mètodes per suavitzar mapes

- **Algorismes de suavitzat local, ad hoc,** per exemple, mitjanes mòbils espacials, algorisme de headbanging
 - Avantatges: ràpids i senzills d'implementar
 - Inconvenients: poden ser molt sensibles a l'elecció ad hoc de ponderacions, etc. i no hi ha estimacions per a la incertesa (errors estàndard)

SUAVITZACIÓ DE RAONS D'INCIDÈNCIA I MORTALITAT ESTANDARDITZADES

Mètodes per suavitzar mapes

- **Anàlisi de superfícies de tendència**, per exemple, kriging, suavitzat mitjançant polinomis/splines
 - Avantatges: l'estimació dels “paràmetres de suavitzat” està basada en l'equilibri entre l'ajustament i la suavització i els errors estàndard estan generalment disponibles
 - Inconvenients: poden ser sensibles a l'elecció del paràmetre de penalització per a l'equilibri

SUAVITZACIÓ DE RAONS D'INCIDÈNCIA I MORTALITAT ESTANDARDITZADES

Mètodes per suavitzar mapes

- **Models d'efectes aleatoris:** les alternatives més populars als mètodes vistos anteriorment consisteixen en els models d'efectes aleatoris
- Dos tipus de models aleatoris:
 - ***Empirical Bayes:*** inclouen com efecte aleatori únicament l'heterogeneïtat espacial
 - ***Models jeràrquics Bayesianos (models mixtos):*** inclouen dos efectes aleatoris, l'heterogeneïtat espacial i la dependència espacial

SUAVITZACIÓ DE RAONS D'INCIDÈNCIA I MORTALITAT ESTANDARDITZADES

Mètodes per suavitzar mapes

➤ Models d'efectes aleatoris

Abans del gran desenvolupament informàtic que es va produir a partir de l'any 2000, s'utilitzaven bàsicament els models empíric Bayes, ja que els altres eren computacionalment complexos. **Actualment**, aquests han quedat obsolets i s'utilitzen els **models jeràrquics Bayesianes**.

SUAVITZACIÓ DE RAONS D'INCIDÈNCIA I MORTALITAT ESTANDARDITZADES

Models d'efectes aleatoris

Amb l'objectiu de solucionar els problemes que hem vist que presenten les RME i/o RIE, s'utilitzen estimadors suavitzats bayesians de les RME i/o RIE obtinguts mitjançant:

- ***Model de Poisson-lognormal-espacial*** (model jeràrquic Bayesià amb un efecte aleatori)
- ***Model de Besag, York i Mollié, BYM*** (model jeràrquic Bayesià amb dos efectes aleatoris)

SUAVITZACIÓ DE RAONS D'INCIDÈNCIA I MORTALITAT ESTANDARDITZADES

Models d'efectes aleatoris: Model 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: $\alpha \sim N(0, 100)$

SUAVITZACIÓ DE RAONS D'INCIDÈNCIA I MORTALITAT ESTANDARDITZADES

Models d'efectes aleatoris: Model de Poisson-lognormal-espacial

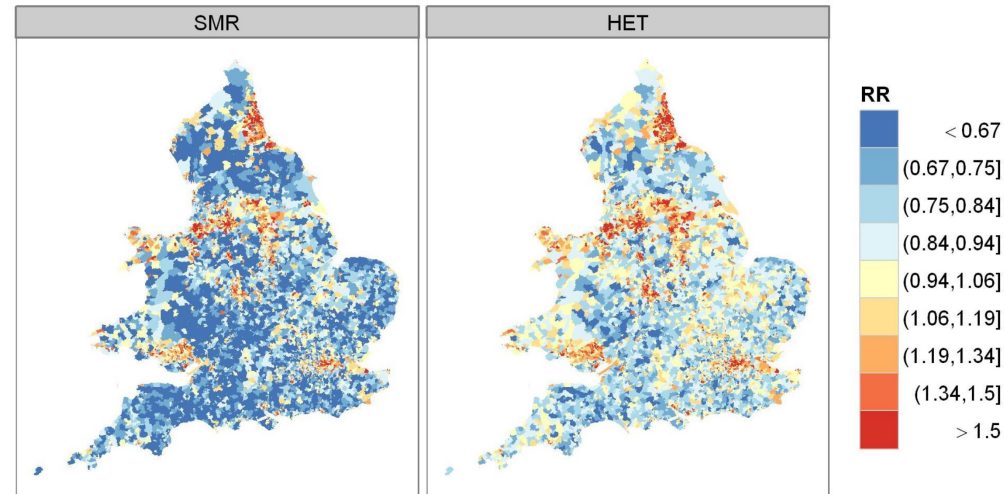
- O_i , E_i són el nombre de casos observats i esperats en l'àrea i , respectivament
- $\lambda_i = \exp(\alpha + V_i)$ és el risc relatiu (RR) en l'àrea i comparat amb el risc esperat en base a l'edat i sexe de la població
- Paràmetres V_i són efectes aleatoris específics per cada àrea per controlar l'heterogeneïtat (sobredispersió)

SUAVITZACIÓ DE RAONS D'INCIDÈNCIA I MORTALITAT ESTANDARDITZADES

Models d'efectes aleatoris: Model de Poisson-lognormal-espacial

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

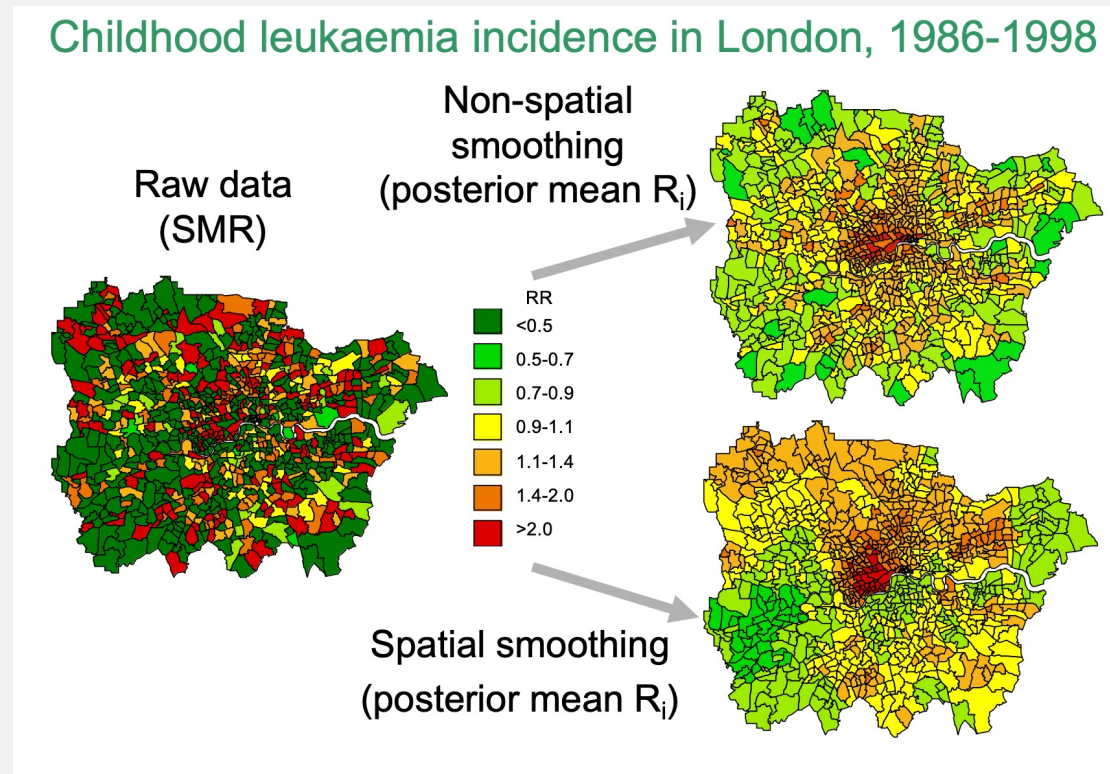
RR estimates using 2 methods



SMRs and non-spatially smoothed RRs

SUAVITZACIÓ DE RAONS D'INCIDÈNCIA I MORTALITAT ESTANDARDITZADES

Models d'efectes aleatoris: Model de Poisson-lognormal-espacial



SUAVITZACIÓ DE RAONS D'INCIDÈNCIA I MORTALITAT ESTANDARDITZADES

Models d'efectes aleatoris: Model BYM

- *L'efecte aleatori utilitzat per recollir l'heterogeneïtat espacial* és el mateix en tots els models (iid normal).
- *L'efecte aleatori per recollir la dependència espacial* es pot aproximar mitjançant un CAR (Conditional Autoregressive) (utilitzat quan es tenen dades d'àrea)

Actualment, la dependència espacial s'aproxima mitjançant una Màtern (**model denominat log-Gaussian-Cox**, utilitzat per tot tipus de dades espacials).

SUAVITZACIÓ DE RAONS D'INCIDÈNCIA I MORTALITAT ESTANDARDITZADES

Models d'efectes aleatoris: Model BYM

Els λ_i solen estar correlacionats espacialment perquè reflecteixen, en part, factors de risc que varien espacialment

Per corregir això hauríem de:

- incorporar la dependència espacial en la distribució dels λ_i
- en el model BYM, aquesta dependència es modelitza com un CAR (model condicional autoregressiu)

SUAVITZACIÓ DE RAONS D'INCIDÈNCIA I MORTALITAT ESTANDARDITZADES

➤ Model BYM. Model 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

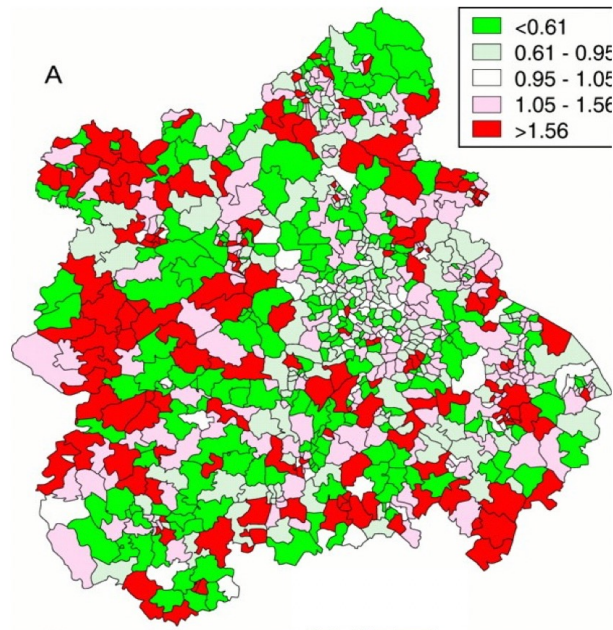
Besag, York, Mollie (1991) *Annals of the Institute of Statistics and Mathematics*, 43: 1-59

SUAVITZACIÓ DE RAONS D'INCIDÈNCIA I MORTALITAT ESTANDARDITZADES

Exemple:

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 ?

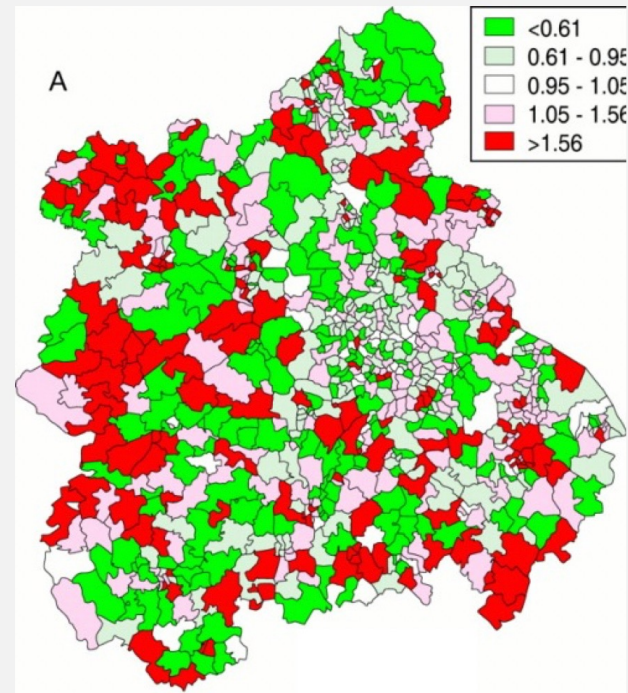
Have the red
highlighted areas
truly got a raised
relative risk?

SUAVITZACIÓ DE RAONS D'INCIDÈNCIA I MORTALITAT ESTANDARDITZADES

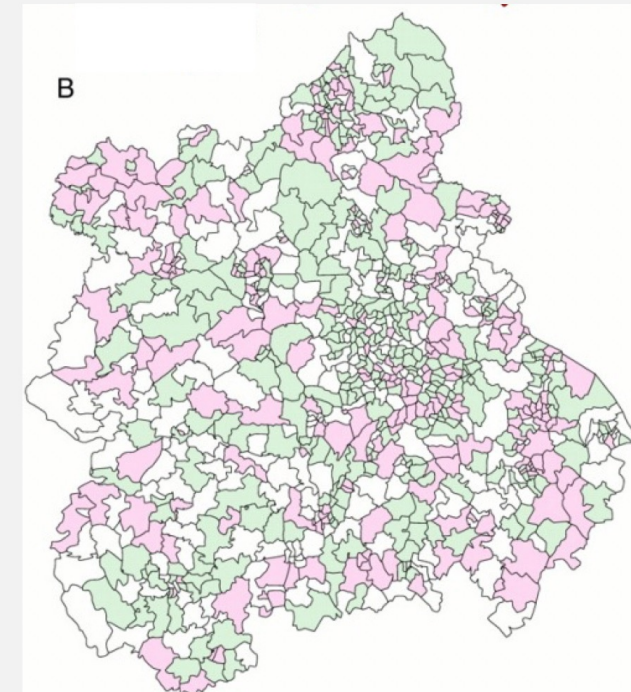
Exemple:

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

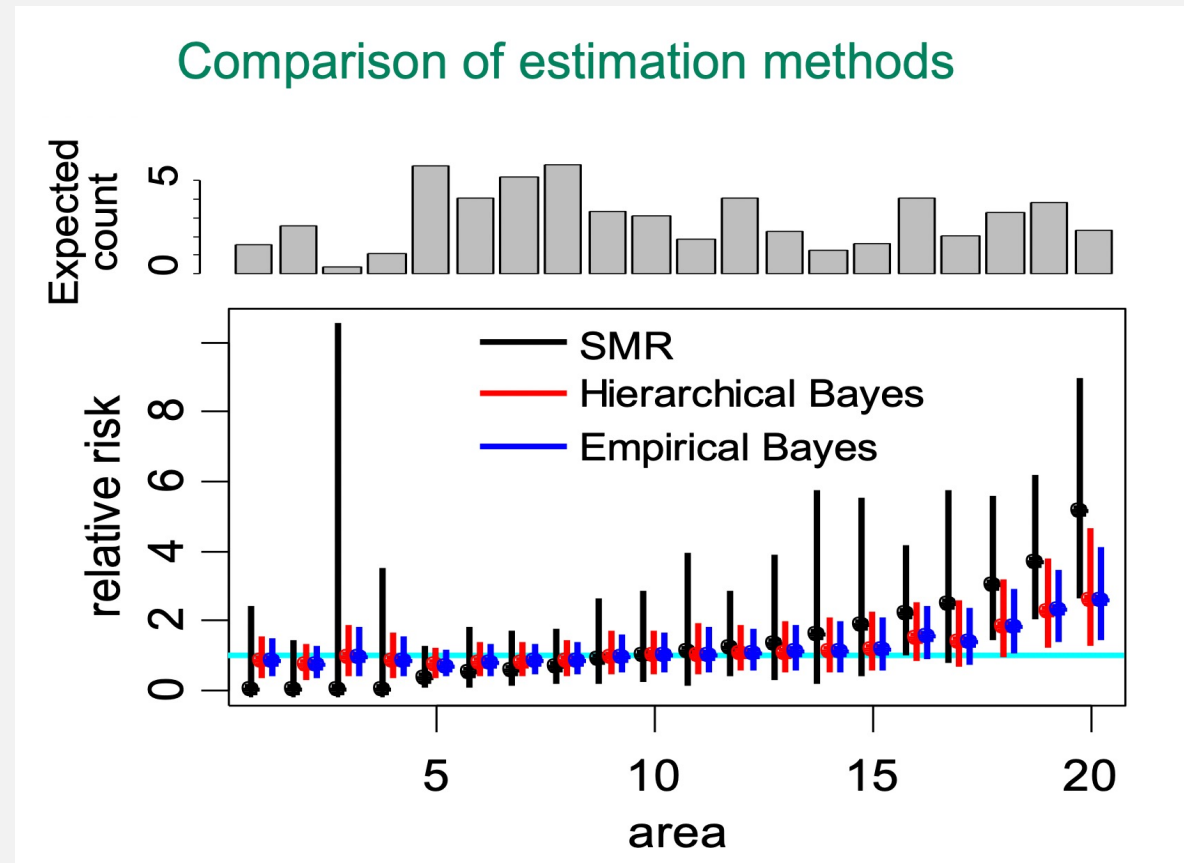
(A) unsmoothed SMR



(B) smoothed by Bayesian methods

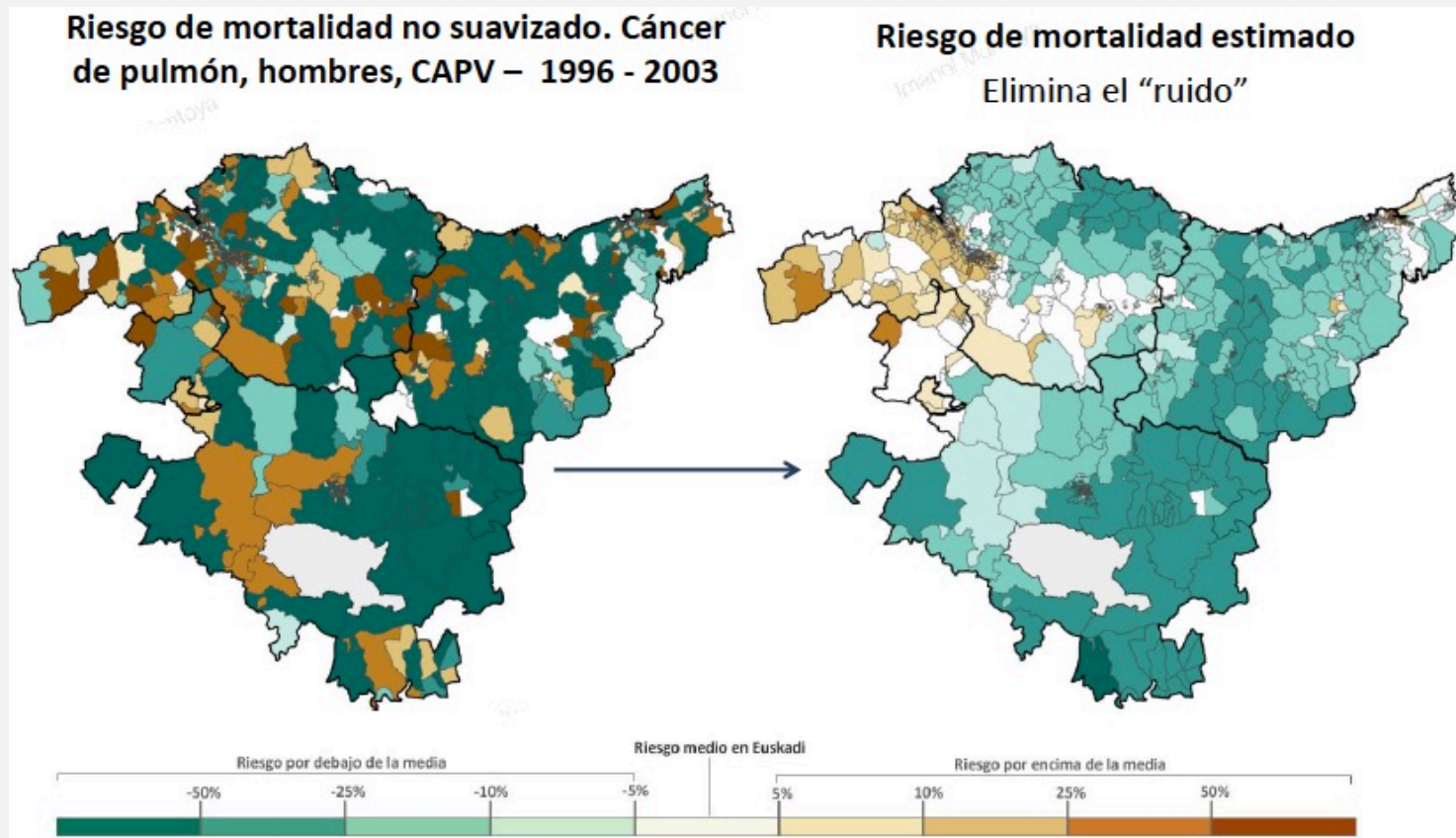


SUAVITZACIÓ DE RAONS D'INCIDÈNCIA I MORTALITAT ESTANDARDITZADES



SUAVITZACIÓ DE RAONS D'INCIDÈNCIA I MORTALITAT ESTANDARDITZADES

Exemple:



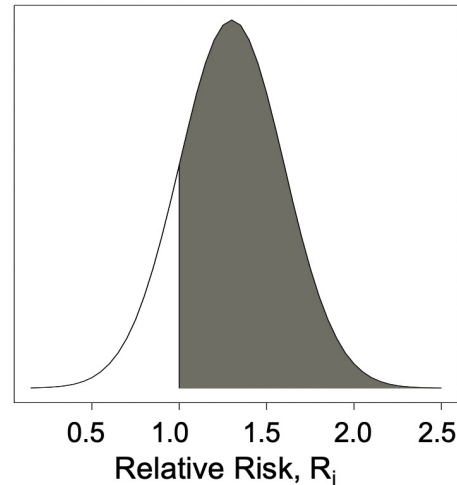
8. Mapes de malalties. Suavització de raons d'incidència i mortalitat estandarditzades

SUAVITZACIÓ DE RAONS D'INCIDÈNCIA I MORTALITAT ESTANDARDITZADES

Models d'efectes aleatoris: Model 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 ($R_i > 1$)**

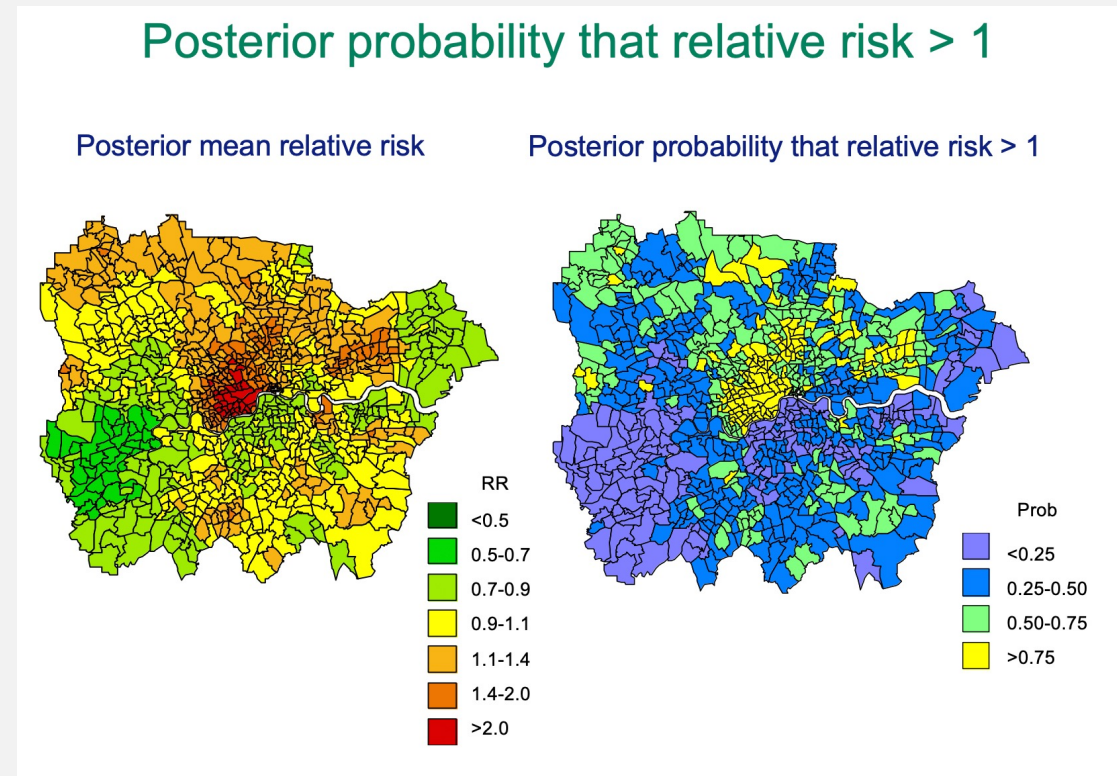
Note – this is not the same as a classical p-value

SUAVITZACIÓ DE RAONS D'INCIDÈNCIA I MORTALITAT ESTANDARDITZADES

- Richardson et al. (2004): estudi de simulació que investiga l'ús de probabilitats a posteriori en estudis enfocats en la representació de malalties en mapes
 - Es classificarà una àrea com d'elevat risc si $[Prob(R_i > 1)] > 0.8$
 - Es classificarà una àrea com d'elevada especificitat si *falsa detecció* < 10%
 - Sensibilitat 60%-95% per E_i de 5-20 i vertader R_i de 1.5-3.0

SUAVITZACIÓ DE RAONS D'INCIDÈNCIA I MORTALITAT ESTANDARDITZADES

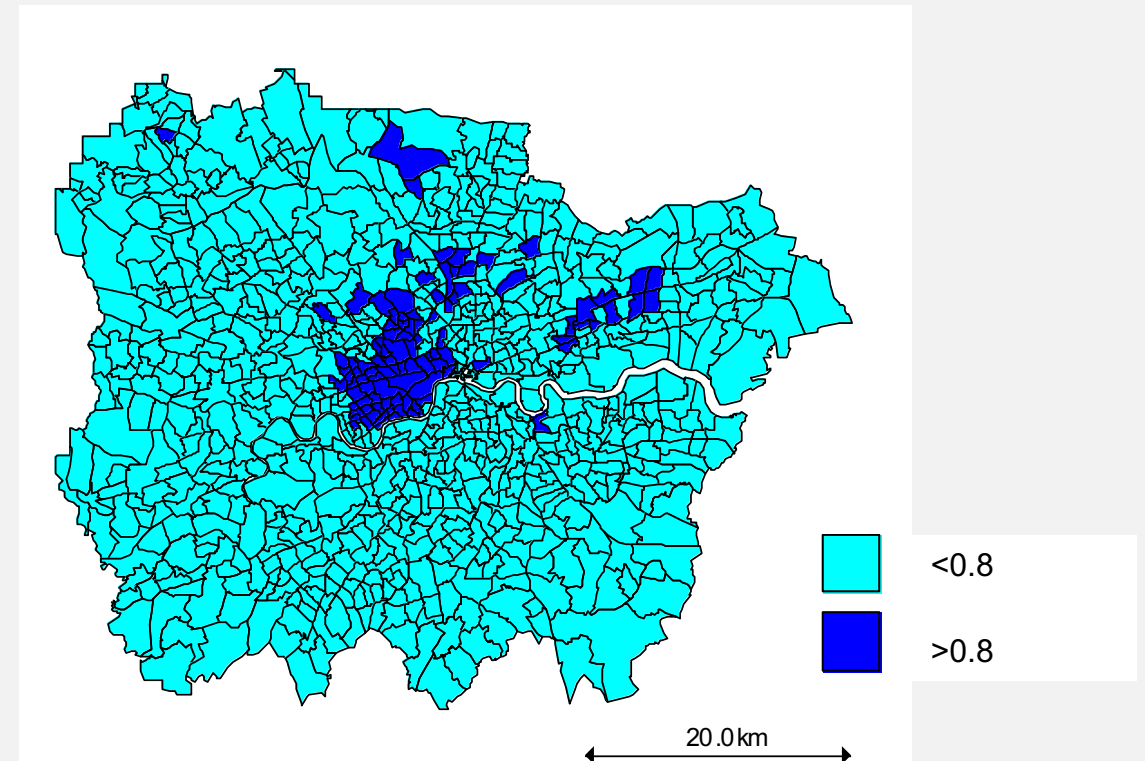
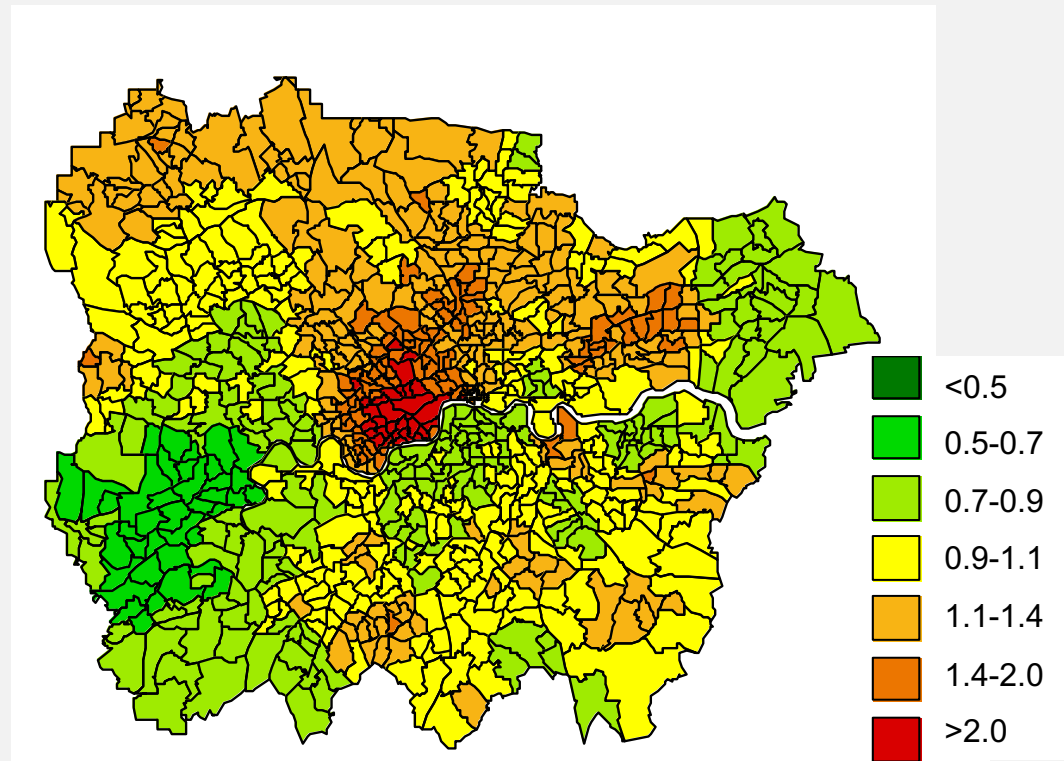
Exemple: Childhood leukaemia incidence in London, 1986-1998



8. Mapes de malalties. Suavitació de raons d'incidència i mortalitat estandarditzades

SUAVITZACIÓ DE RAONS D'INCIDÈNCIA I MORTALITAT ESTANDARDITZADES

Exemple: Childhood leukaemia incidence in London, 1986-1998



SUAVITZACIÓ DE RAONS D'INCIDÈNCIA I MORTALITAT ESTANDARDITZADES

Software

- Per realitzar el suavitzat mitjançant Empirical Bayes, es pot utilitzar la Rapid Inquiry Facility (RIF) de la SAHSU de l'Imperial College.
- L'estimació de models jeràrquics bayesians es feia mitjançant mètodes de simulació computacionalment complexos (MCMC) implementats en el programari gratuït WinBUGS i GeoBUGS
- Des de 2008, el programari gratuït INLA (Rue et al, 2008) implementa una aproximació ràpida i més eficient als MCMC per estimar aquests models.

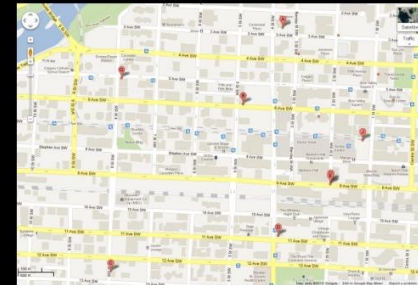
GIS



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