

□

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

EPIDEMIOLOGÍA Y EPIDEMIOLOGÍA ESPACIAL

Epidemiología

- La **epidemiología** se dedica al estudio de la distribución, frecuencia, causas y control de los factores relacionados con la salud y la enfermedad en poblaciones humanas bien definidas y a la aplicación de este estudio para defender y mejorar la salud de la población.
- La **epidemiología** estudia, sobre todo, la relación entre exposición y enfermedad.
- La **epidemiología** se considera la ciencia básica para la medicina preventiva y una fuente de información para la formulación de políticas de salud pública.

EPIDEMIOLOGÍA Y EPIDEMIOLOGÍA ESPACIAL

- Las enfermedades no se producen de forma aleatoria. Tienen causas, muchas de ellas de origen humano, que pueden evitarse.
- Los métodos epidemiológicos han sido cruciales para la identificación de numerosos factores etiológicos que, en su momento, han justificado la formulación de políticas sanitarias encaminadas a la prevención de enfermedades.

EPIDEMIOLOGÍA Y EPIDEMIOLOGÍA ESPACIAL

- La epidemiología surgió del estudio de las epidemias de las enfermedades infecciosas.
- En la actualidad, la epidemiología se interesa por el estudio demográfico de cualquier enfermedad con la ayuda de la estadística.

EPIDEMIOLOGÍA Y EPIDEMIOLOGÍA ESPACIAL

Epidemiología espacial

“Spatial epidemiology is the description and analysis of geographic variations in disease with respect to demographic, environmental, behavioural, socioeconomic, genetic, and infectious risk factors”.

Elliot P, Wartenberg D. Spatial epidemiology: current approaches and futures challenges. *Environ. Health Perspect.* 2004; 112(9):998-1006. doi: 10.1289/ehp.6735.

EPIDEMIOLOGÍA Y EPIDEMIOLOGÍA ESPACIAL

- La **epidemiología espacial** está relacionada con la descripción de las enfermedades y el estudio de las causas y prevención de las mismas, utilizando diferentes perspectivas de análisis donde la localización de los sucesos es una componente fundamental, ya que estudia también las variaciones geográficas de las enfermedades.
- La **epidemiología espacial** es un subcampo de la epidemiología centrado en el estudio de la distribución espacial de los resultados en salud.
- La **epidemiología espacial** se basa en un concepto de salud, en el cual los individuos son vistos en su contexto sociocultural-ambiental.

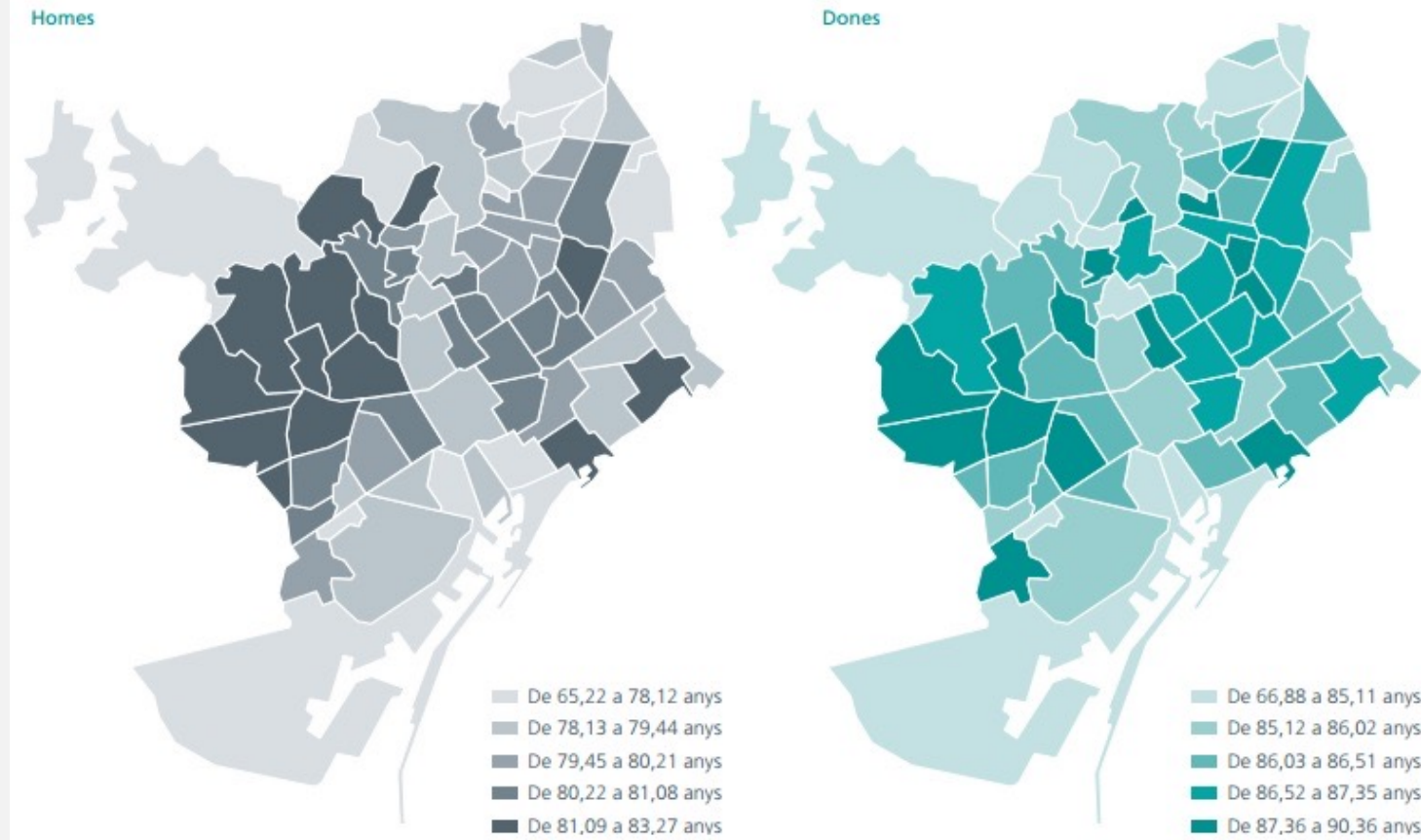
EPIDEMIOLOGÍA Y EPIDEMIOLOGÍA ESPACIAL

Principal objetivo de la epidemiología espacial:

- Mostrar qué parte de la variación espacial de la distribución de la ocurrencia de un suceso de salud no está explicada ni por la distribución espacial de factores explicativos conocidos, ni por una variación aleatoria.
- De hecho, muy a menudo nos interesa encontrar pistas sobre algún factor de riesgo desconocido de una determinada enfermedad.

EPIDEMIOLOGÍA Y EPIDEMIOLOGÍA ESPACIAL

Figura 1. Esperança de vida en homes i dones als barris. Barcelona, 2008-2012.



Font: Registre de Mortalitat de Barcelona. Agència de Salut Pública de Barcelona.

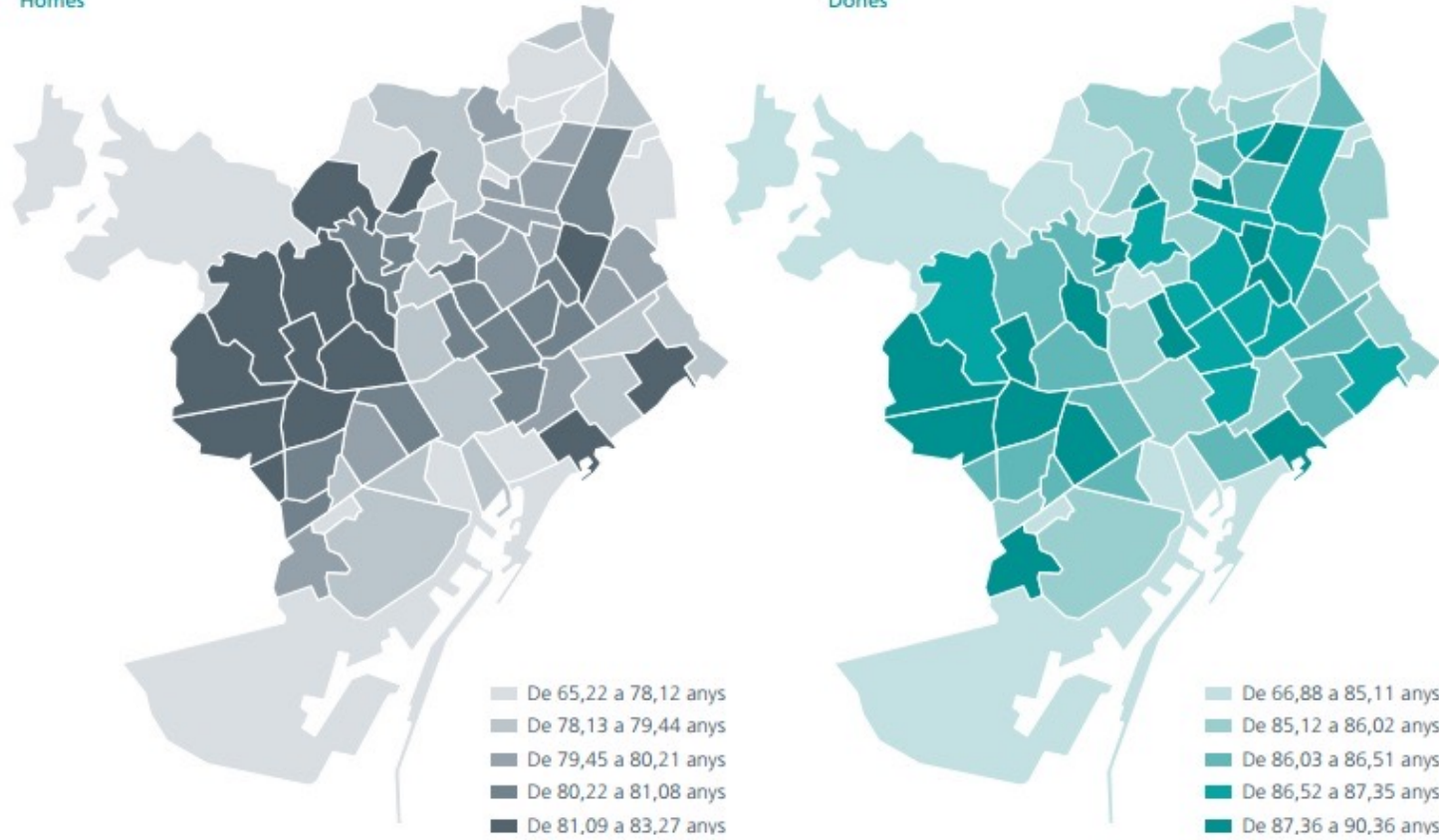
2. Introducción a la epidemiología y la estadística espacial

EPIDEMIOLOGÍA Y EPIDEMIOLOGÍA ESPACIAL

Figura 1. Esperança de vida en homes i dones als barris. Barcelona, 2008-2012.

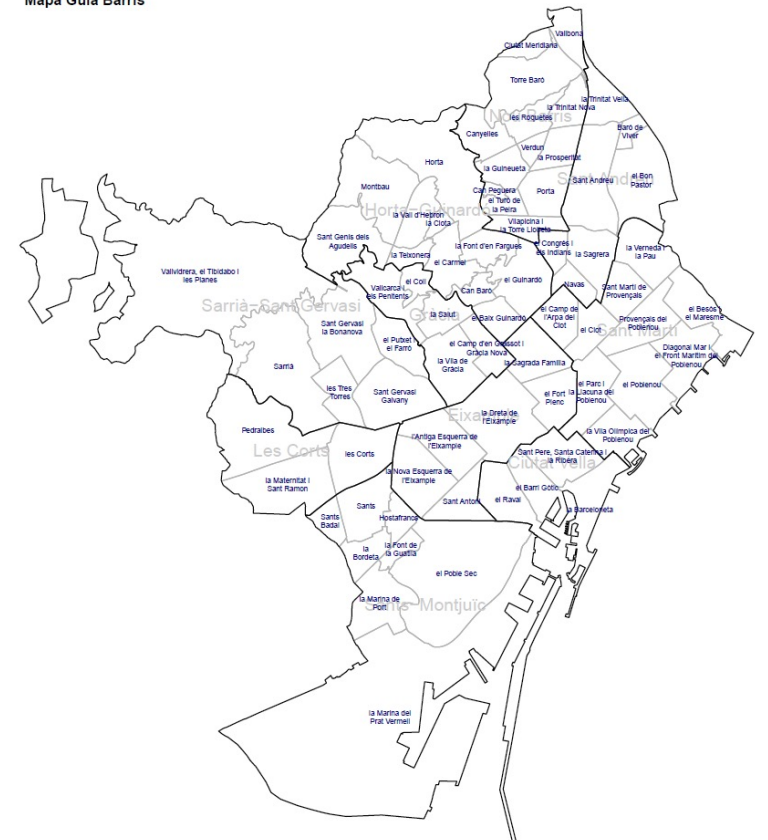
Homes

Dones



Font: Registre de Mortalitat de Barcelona. Agència de Salut Pública de Barcelona.

Mapa Guia Barris



EPIDEMIOLOGÍA Y EPIDEMIOLOGÍA ESPACIAL

- El estudio de la distribución geográfica de los sucesos de salud ha sido de gran interés entre los epidemiólogos durante las últimas décadas, aunque como veremos a continuación, tiene una historia de más de 200 años.
- Los **primeros mapas de enfermedades** se realizaron para representar la ubicación de casos de enfermedades infecciosas:
 - Fiebre amarilla en Nueva York (Seaman, 1798)
 - Cólera en Londres (Snow, 1854)

EPIDEMIOLOGÍA Y EPIDEMIOLOGÍA ESPACIAL

Mapa de Filippo Arrieta

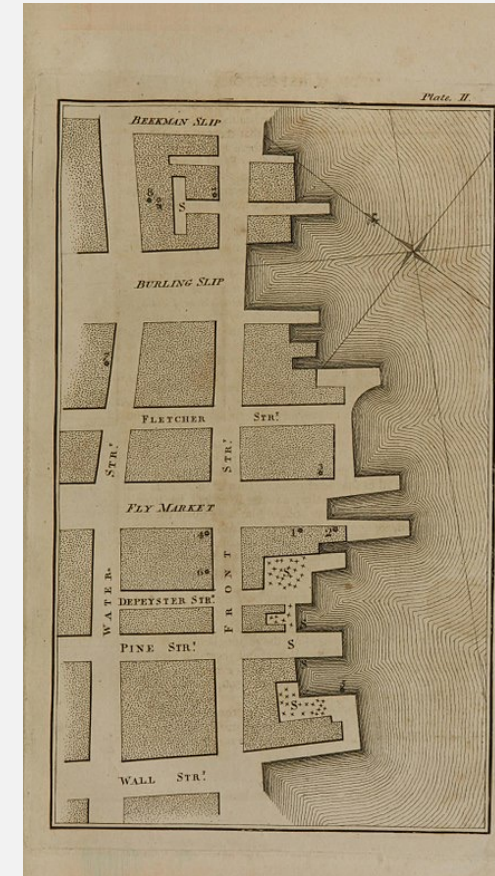
Visualización de la estrategia para contener la propagación de la peste en la región de Bari, Italia 1690-92.



EPIDEMIOLOGÍA Y EPIDEMIOLOGÍA ESPACIAL

Mapa de Seaman de casos de fiebre amarilla

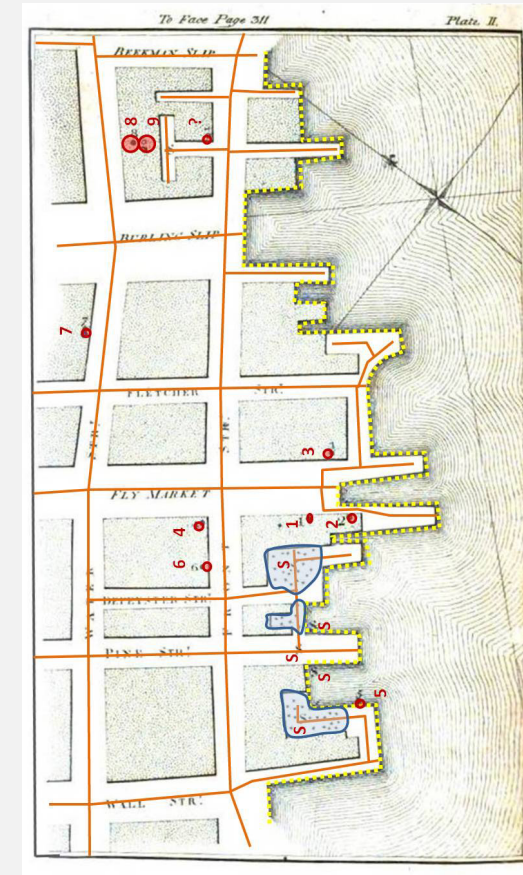
Los puntos representan las muertes por fiebre amarilla y las "S", los vertederos.



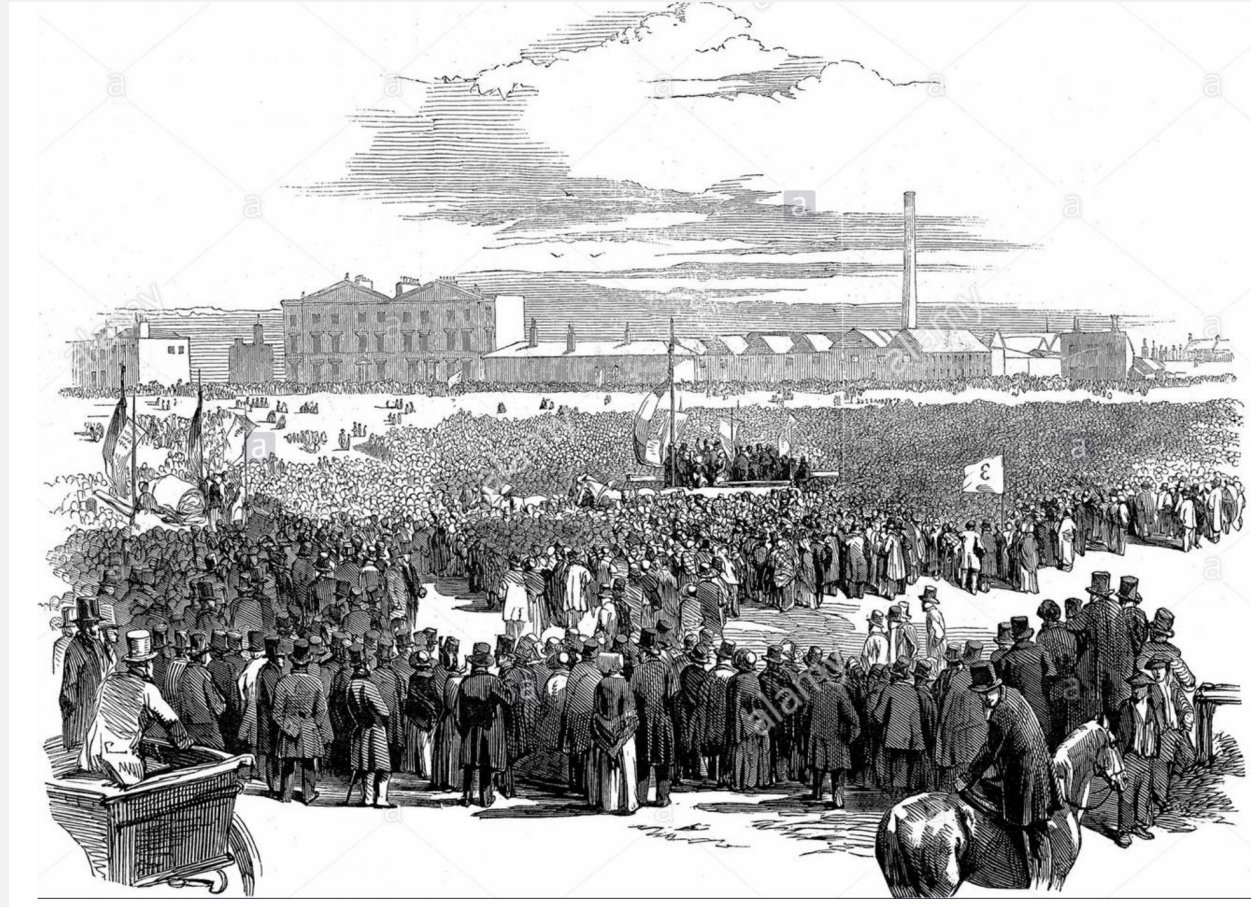
EPIDEMIOLOGÍA Y EPIDEMIOLOGÍA ESPACIAL

Mapa de Seaman de casos de fiebre amarilla

Los puntos representan las muertes por fiebre amarilla y las cruces o las "S", los vertederos.



EPIDEMIOLOGÍA Y EPIDEMIOLOGÍA ESPACIAL



2. Introducción a la epidemiología y la estadística espacial

EPIDEMIOLOGÍA Y EPIDEMIOLOGÍA ESPACIAL

- Durante el otoño de 1848, se produjo una segunda epidemia de cólera en Inglaterra, causando una gran mortalidad.
- No se conocía con certeza la etiología, ni la manera como la enfermedad se transmitía.
- Dos corrientes teóricas:
 - Contagio por contacto
 - Miasmas

EPIDEMIOLOGÍA Y EPIDEMIOLOGÍA ESPACIAL

Snow observó:

- **Miasmas no podían causar la enfermedad:** los pacientes deberían presentar síntomas respiratorios producto de la inhalación de las “miasmas” y no el síndrome diarreico agudo presente en el cólera.
- **Defunciones por cólera** ocurridas entre 1848-49 (segunda epidemia) estaban **concentradas** en los distritos de la **zona sur de Londres**. Las tasas de mortalidad observadas en esta zona eran muy superiores a las del resto de la ciudad (8,0 y 2,4 defunciones por 1.000 habitantes, respectivamente).

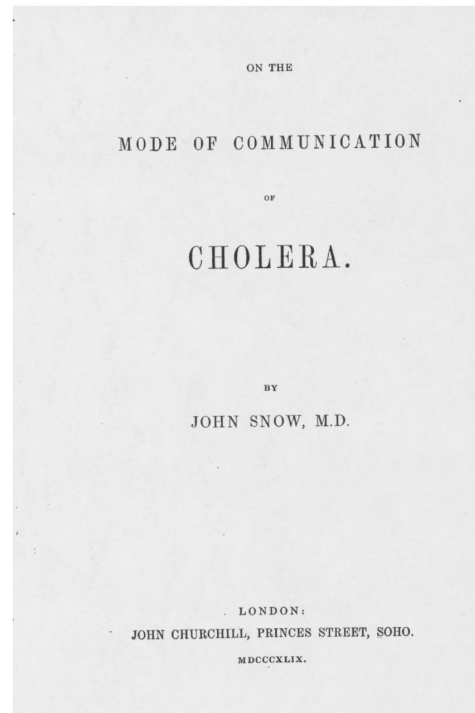
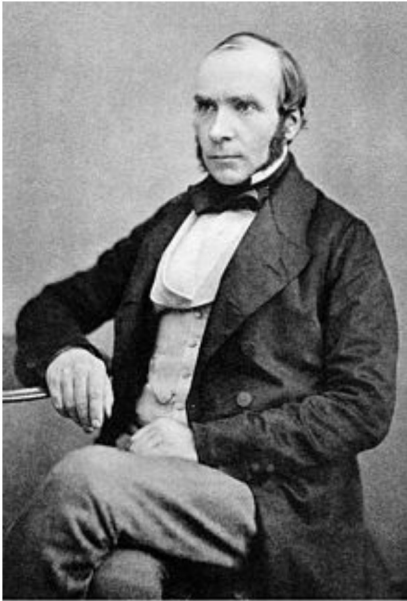
EPIDEMIOLOGÍA Y EPIDEMIOLOGÍA ESPACIAL

Además, Snow observó que:

- Los habitantes de la zona sur de Londres obtenían agua para beber río abajo del Támesis (aguas altamente contaminadas).
- Las otras zonas de Londres obtenían el agua para beber de sectores menos contaminados del río (aguas arriba del Támesis o de afluentes).

EPIDEMIOLOGÍA Y EPIDEMIOLOGÍA ESPACIAL

John Snow
1813-1858



*Deaths from Cholera in London, registered from
September 23d, 1848, to August 25th, 1849.*

Districts of London.	Population in 1841.	Deaths from Cholera.	Deaths to each 1,000 inhabitants.
West . .	300,711	533	1.77
North . .	375,971	415	1.10
Central . .	373,605	920	2.48
East . . .	392,444	1,597	4.06
South . .	502,548	4,001	7.95
Total . .	1,948,369	7,466	3.83

On the Mode of Communication of Cholera. John Snow. 1849

EPIDEMIOLOGÍA Y EPIDEMIOLOGÍA ESPACIAL

- En 1853-1854, tuvo lugar la tercera epidemia de cólera en Londres.
- Los habitantes de ciertos distritos del sur de la ciudad extraían el agua de pequeños afluentes del río Támesis o bien a partir de numerosas bombas de agua de uso público.
- Los desechos humanos eran lanzados en improvisadas alcantarillas o directamente al río.

EPIDEMIOLOGÍA Y EPIDEMIOLOGÍA ESPACIAL

- Había dos compañías responsables del abastecimiento: Sothwark and Vauxhall Company y Lambeth Water Company.
- Durante la segunda epidemia de cólera de 1848-49, ambas compañías extraían el agua de sectores contaminados del Támesis, presentando, los distritos abastecidos por una y otra compañía, similar número de muertos.

EPIDEMIOLOGÍA Y EPIDEMIOLOGÍA ESPACIAL

- En 1853:
 - Lambeth Water Company había trasladado sus instalaciones río arriba (aguas no contaminadas).
 - Southwark and Vauxhall Water Company seguía aguas abajo.
- Snow observó que la tasa de mortalidad por cólera en hogares abastecidos por la segunda compañía era 8,5 veces mayor que la de los hogares abastecidos por la primera.

EPIDEMIOLOGÍA Y EPIDEMIOLOGÍA ESPACIAL

- A principios de septiembre de 1854:
 - En el sector de Londres “Golden Square” (en el Soho), se produjo un brote epidémico de cólera de inusual intensidad (500 muertos en tan solo 10 días).
 - La mayoría de los residentes del sector extraían el agua a partir de una bomba de uso público situada en Broad Street.

EPIDEMIOLOGÍA Y EPIDEMIOLOGÍA ESPACIAL



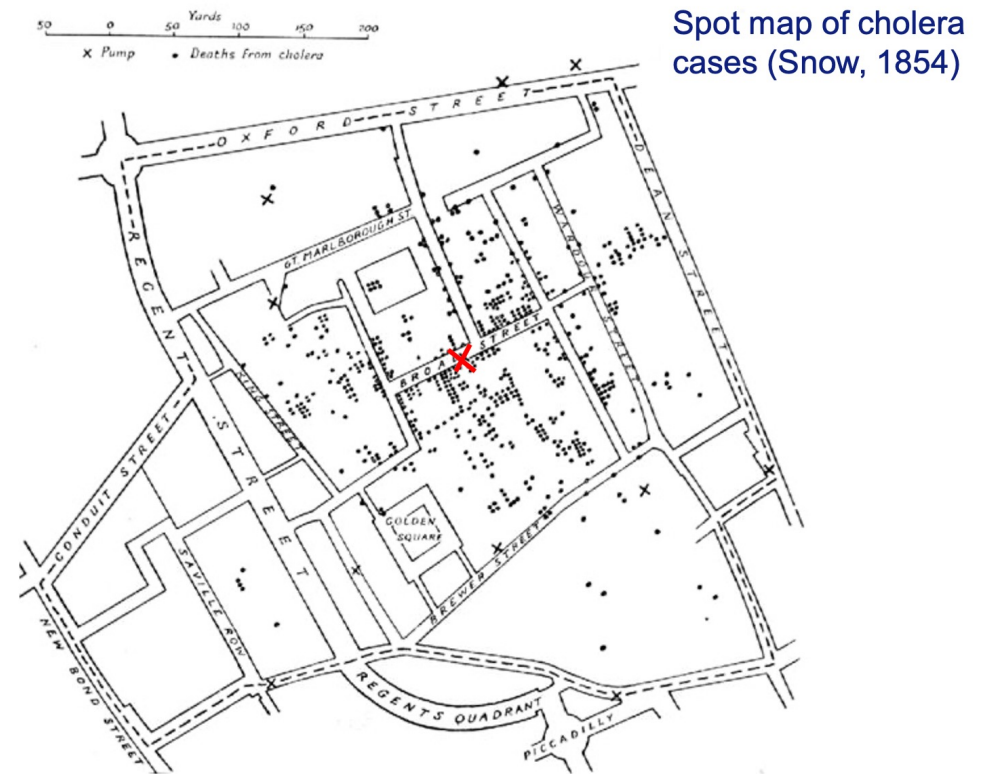
2. Introducción a la epidemiología y la estadística espacial

EPIDEMIOLOGÍA Y EPIDEMIOLOGÍA ESPACIAL

- Realizó un **mapa del sector**, marcando los puntos correspondientes a defunciones por cólera y las diferentes bombas de agua potable existentes.
- Planteó que este brote se debía a la ingestión de aguas contaminadas procedentes de la bomba de Broad Street.

EPIDEMIOLOGÍA Y EPIDEMIOLOGÍA ESPACIAL

Snow identificó la agregación espacial de los casos de cólera en Londres en 1857.



EPIDEMIOLOGÍA Y EPIDEMIOLOGÍA ESPACIAL

- Recogió muestras de la bomba de Broad Street y de otras 4 bombas cercanas (había diferencias en relación a la claridad del agua).
- Calculó la distancia entre la residencia de cada difunto y la bomba de agua más cercana.
- Observó que, en 73 de los 83 casos, la bomba de Broad Street era la más cercana y comunicó esto a las autoridades sanitarias y estas cerraron la bomba.

EPIDEMIOLOGÍA Y EPIDEMIOLOGÍA ESPACIAL

Evolución de los mapas en epidemiología espacial

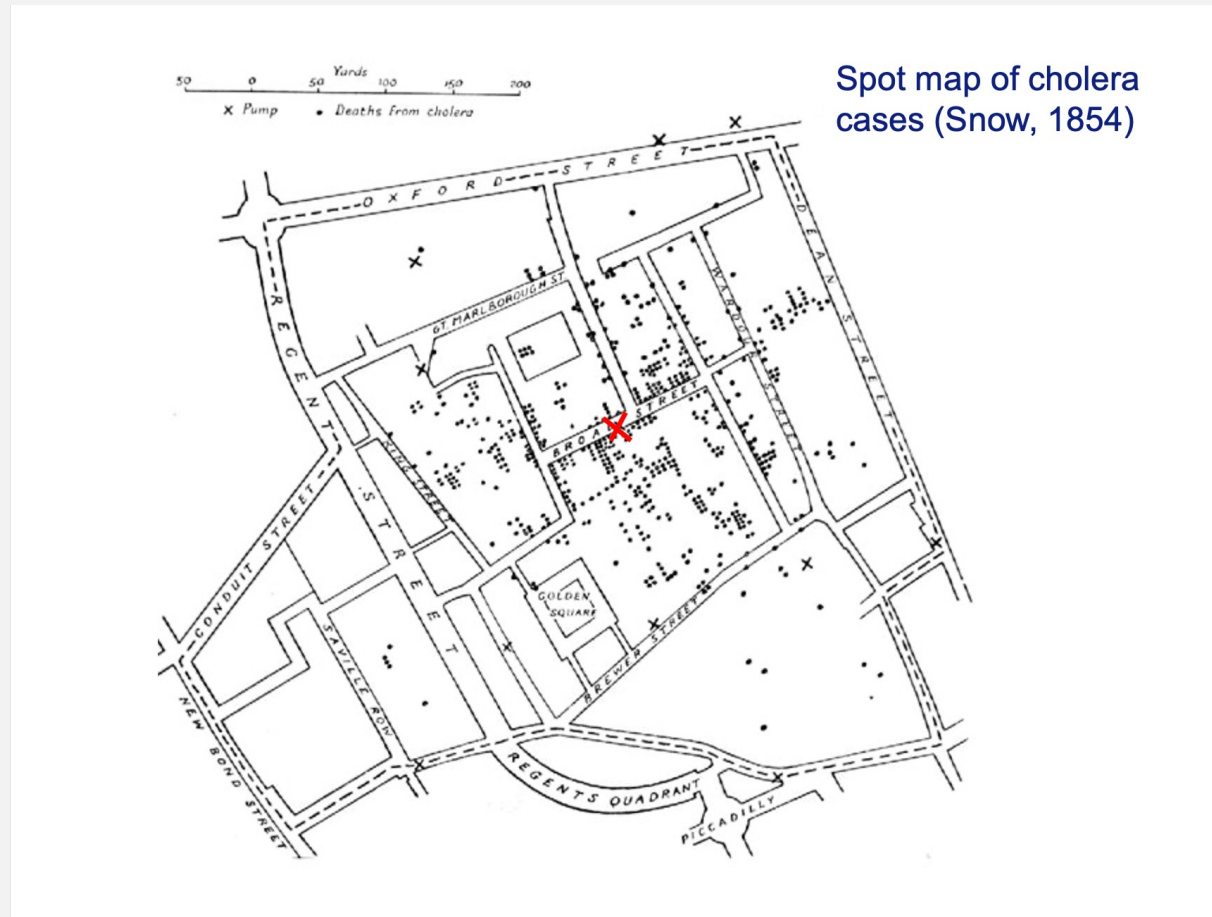
- Mapas de puntos
- Mapas de coropletas
- Atlas de enfermedades, a nivel nacional e internacional

EPIDEMIOLOGÍA Y EPIDEMIOLOGÍA ESPACIAL

Evolución de los mapas en epidemiología espacial

- Mapas de puntos
 - Fiebre amarilla en Nueva York (Seaman, 1798)
 - Cólera en Londres (Snow, 1854)

EPIDEMIOLOGÍA Y EPIDEMIOLOGÍA ESPACIAL



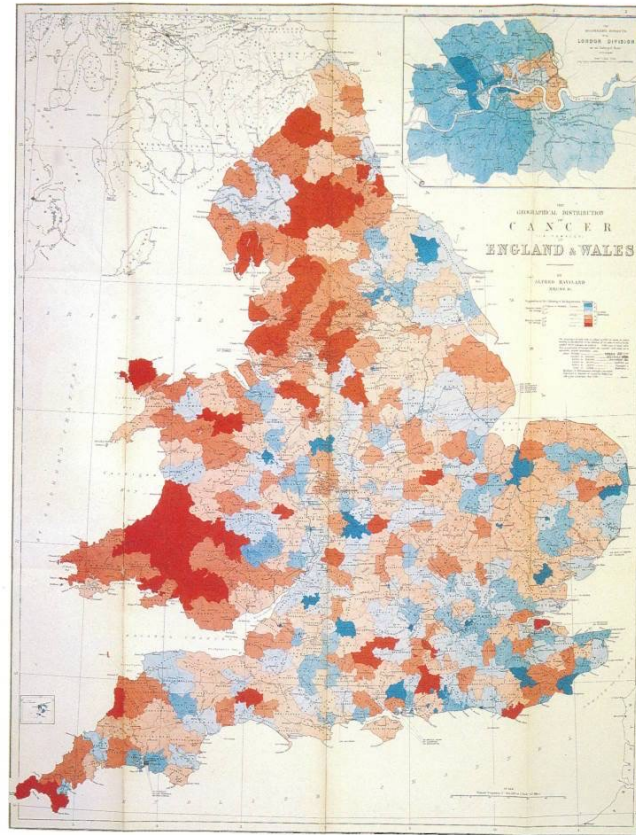
EPIDEMIOLOGÍA Y EPIDEMIOLOGÍA ESPACIAL

Evolución de los mapas en epidemiología espacial

- Mapas de coropleta
 - Distribución geográfica de la mortalidad por enfermedad cardíaca, cáncer y tuberculosis en Inglaterra y Gales (Haviland, 1878)
 - Mortalidad por cáncer por condado en Inglaterra y Gales, ajustada por edad y sexo (Stocks, 1936,1937,1939)

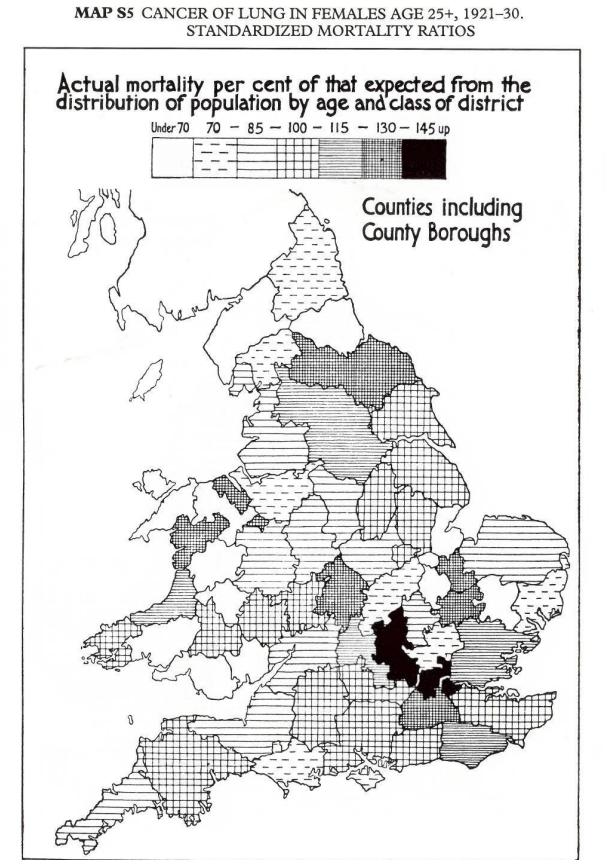
EPIDEMIOLOGÍA Y EPIDEMIOLOGÍA ESPACIAL

Cáncer en mujeres,
1851-1860
(Haviland, 1878)



The geographical distribution of cancer in females in England and Wales, 1851-60. From Haviland (1878).

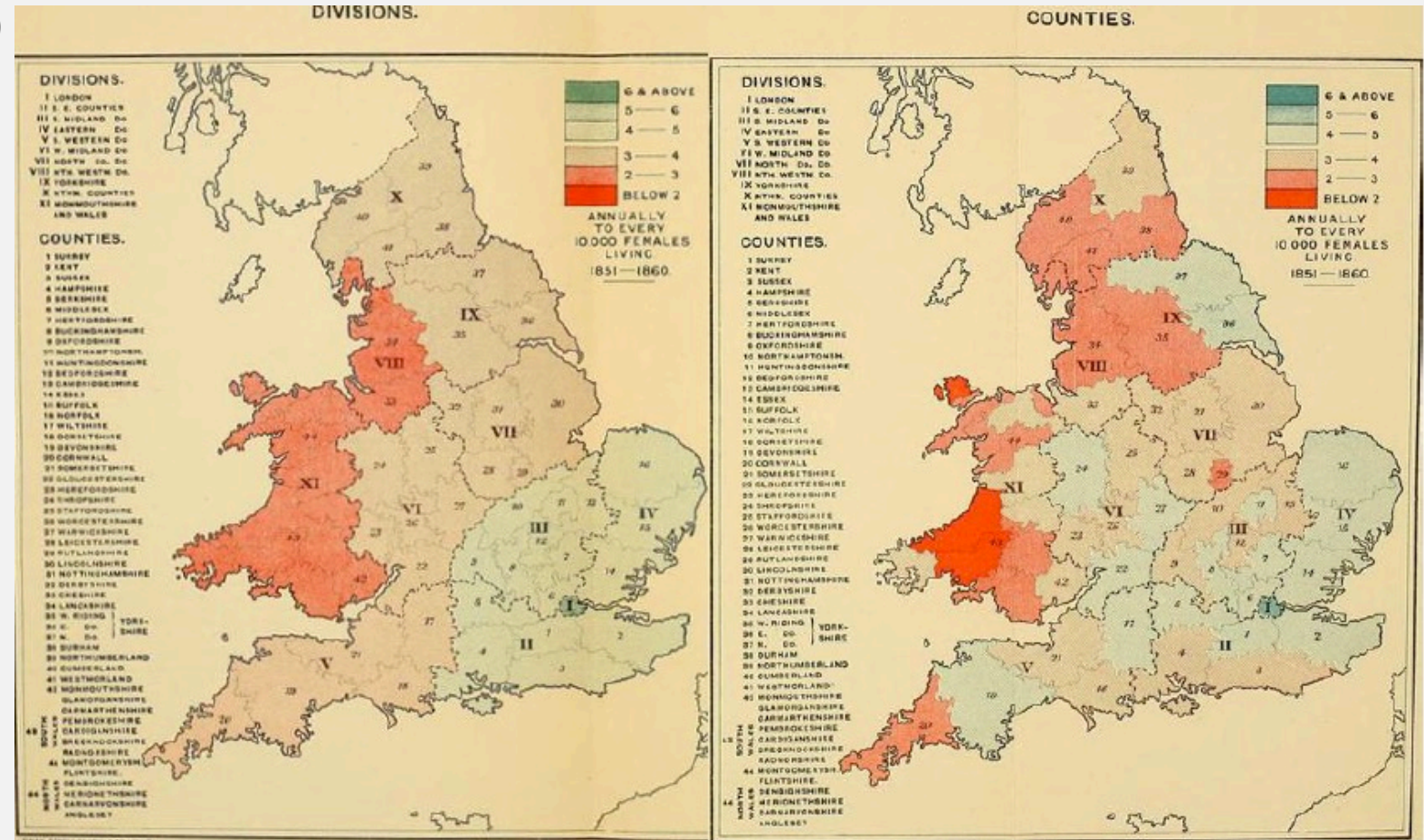
Cáncer de pulmón en
mujeres, 1921-1930
(Stocks, 1939)



From Stocks (1939).

EPIDEMIOLOGÍA Y EPIDEMIOLOGÍA ESPACIAL

Cáncer en mujeres, 1851-1860
(Haviland, 1878)



EPIDEMIOLOGÍA Y EPIDEMIOLOGÍA ESPACIAL

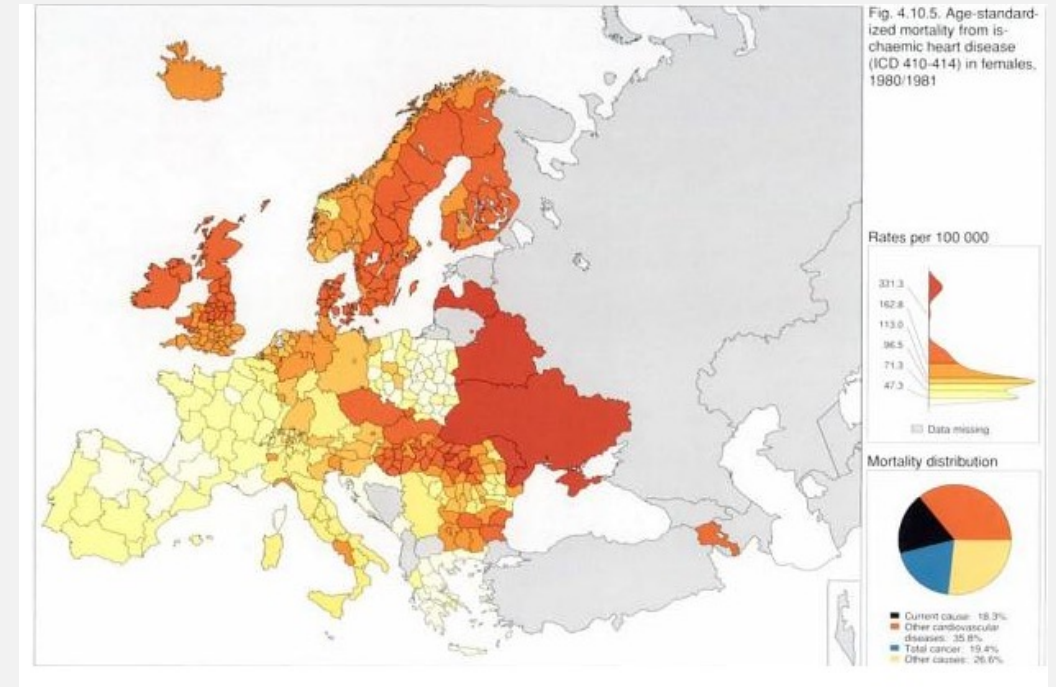
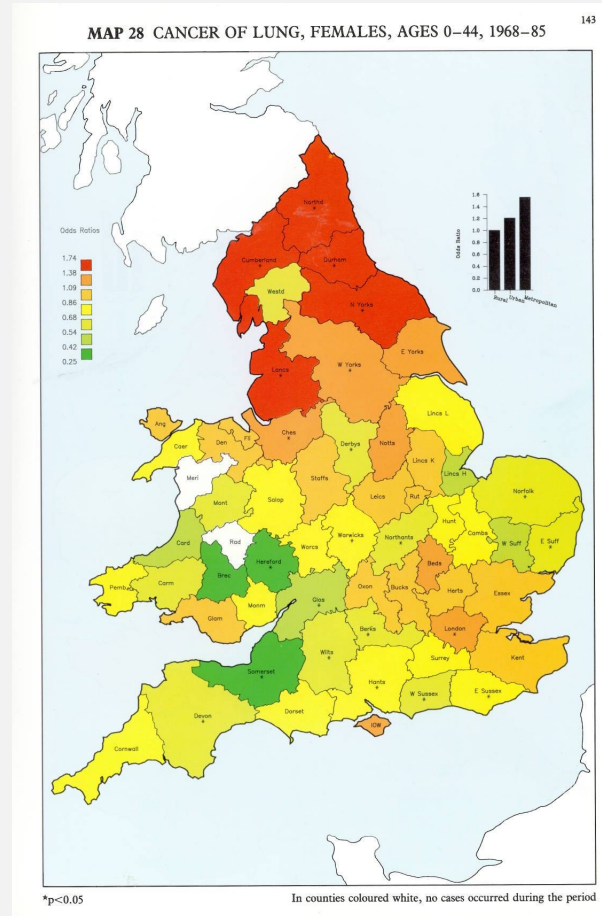
Evolución de los mapas en epidemiología espacial

- Atlas de enfermedades, a nivel nacional e internacional
 - Atlas de incidencia del cáncer en Inglaterra y Gales 1968-85 (Swerdlow y dos Santos Silva, 1993)
 - Atlas de mortalidad en Europa 1980-81 y 1990-91 (OMS, 1997)

EPIDEMIOLOGÍA Y EPIDEMIOLOGÍA ESPACIAL

Incidencia de cáncer de pulmón
en mujeres 1968-1985

(Swerdlow and dos Santos Silva,
1993)



Mortalidad por Infarto de miocardio
estandarizada por edad, 1980-81

(OMS)

2. Introducción a la epidemiología y la estadística espacial

EPIDEMIOLOGÍA Y EPIDEMIOLOGÍA ESPACIAL

- Como podemos ver, la epidemiología espacial surgió de la representación de la distribución espacial de sucesos en salud con el doble objetivo de caracterizar su extensión y establecer hipótesis de las posibles causas.
- Posteriormente, la epidemiología espacial creció mucho en complejidad (métodos de análisis, unidades pequeñas) y utilidad.
- La confluencia entre la epidemiología, la estadística y la informática, juntamente con el gran desarrollo tecnológico, han permitido que la epidemiología espacial se desarrolle a pasos agigantados.

EPIDEMIOLOGÍA Y EPIDEMIOLOGÍA ESPACIAL

Desarrollos a partir de 1996

- Sistemas de Información Geográfica (SIG)
- Software para representar mapas y analizar datos espaciales
- Mayor disponibilidad de datos georeferenciados (GPS, etc.)

EPIDEMIOLOGÍA Y EPIDEMIOLOGÍA ESPACIAL

Desarrollos a partir de 1996

- Desarrollo de métodos estadísticos especializados
 - Técnicas sofisticadas para separar la señal de ruido
 - Métodos para controlar la dependencia espacial y temporal
 - Métodos para detectar aglomeraciones de casos (clústeres)

EPIDEMIOLOGÍA Y EPIDEMIOLOGÍA ESPACIAL

Problemas en epidemiología espacial:

- Diferentes escalas (por ejemplo, comunidades autónomas, provincia, municipio, distrito, barrio, código postal, sección censal, ...)
- Cambios en los límites de algunas de estas unidades
- Pueden existir errores en la georeferenciación (a causa de direcciones erróneas o inexistentes, etc.)
- Desalineación (misalignment)

ESTADÍSTICA ESPACIAL

- La **estadística espacial** se ocupa de la exploración, descripción, visualización y análisis de los datos, considerando sus características de distribución en el espacio, que suelen expresarse a través del uso de coordenadas geográficas.
- La **estadística espacial** es la rama de la estadística que analiza datos georeferenciados, es decir, datos de los que se dispone de sus coordenadas espaciales (datos espaciales).

ESTADÍSTICA ESPACIAL

- Por **datos espaciales** se entienden aquellas medidas u observaciones realizadas en localizaciones o en áreas específicas. Además del valor de la medida, incorporan la localitzación/posición de los valores observados.

ESTADÍSTICA ESPACIAL

Características de los datos espaciales:

1. Heterogeneidad espacial

- Las observaciones no son homogéneas en el espacio

2. Dependencia espacial

- Las observaciones de una localización dependen de otras observaciones en otras localizaciones (generalmente próximas).

ESTADÍSTICA ESPACIAL

Características de los datos espaciales:

1. Heterogeneidad espacial

- Las observaciones no son homogéneas en el espacio

ESTADÍSTICA ESPACIAL

Hay dos tipos de **heterogeneidad espacial**:

➤ ***Heteroscedasticidad:***

- *Causa estructural*: a causa de la utilización de datos procedentes de unidades espaciales arbitrarias
- *Causa mostral*: existencia de outliers, omisión de variables relevantes, errores de medida y otros errores de especificación.

➤ ***Cambio-inestabilidad estructural***: los individuos no son homogéneos en el espacio (por ejemplo, norte-sur; centro-periferia, ...).

ESTADÍSTICA ESPACIAL

Características de los datos espaciales:

2. Dependencia espacial

- Las observaciones de una localización dependen de otras observaciones en otras localizaciones (generalmente próximas).

ESTADÍSTICA ESPACIAL

Existen dos tipos de **dependencia espacial**:

- ***Interacción espacial (dependencia espacial substantiva)***: es un efecto spillover de un individuo sobre otro individuo.
- ***Autocorrelación en los residuos (dependencia espacial residual)***: causada por errores de medida espacialmente relacionados.

ESTADÍSTICA ESPACIAL

Tipos datos espaciales:

Los datos espaciales se han clasificado históricamente en tres grandes grupos (Cressie, 1993):

1. Datos en rejilla o **datos de área** ('lattice data' o 'areal data')
2. Datos referenciados puntualmente: **Procesos puntuales** ('point processes')
3. Datos referenciados puntualmente: **Datos geoestadísticos** ('geoestatistical data')

ESTADÍSTICA ESPACIAL

1. Datos en rejilla o datos de área ('lattice data' o 'areal data')

- Los **datos de área** o en rejilla se corresponden con variables aleatorias discretas (recuentos) ("count data").
- En los datos de área ***no se conoce la localización exacta del caso.***
- ***Las localizaciones son áreas*** con límites geográficos bien definidos, habitualmente unidades administrativas (ciudades, barrios, secciones censales, etc.).

ESTADÍSTICA ESPACIAL

1. Datos en rejilla o datos de área ('lattice data' o 'areal data')

- La ***variable respuesta es el número agregado de casos en esta área*** en un período de tiempo determinado.
- Este era el tipo de dato más comunmente utilizado al principio de la estadística espacial.

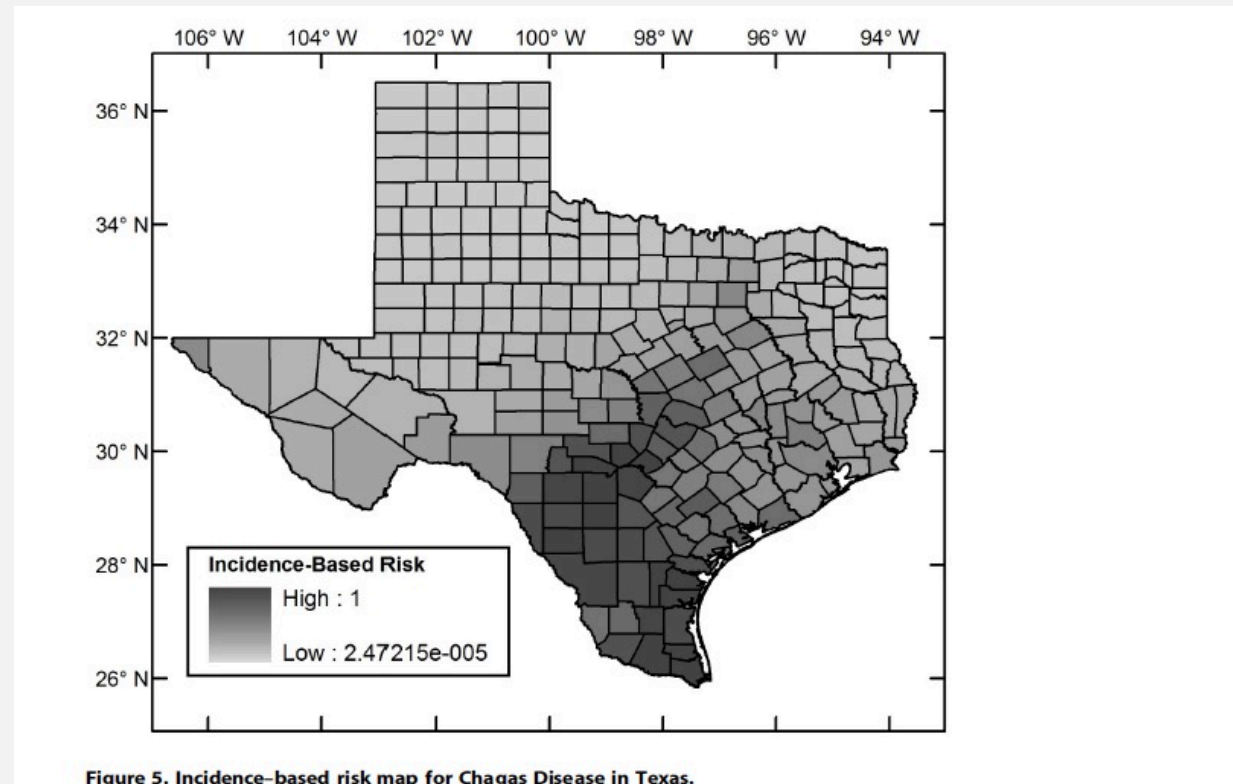
ESTADÍSTICA ESPACIAL

1. Datos en rejilla o datos de área ('lattice data' o 'areal data')

- Observaciones procedentes de un proceso aleatorio sobre un conjunto de regiones espaciales distribuidas espacialmente de forma regular o irregular.
- Quedan definidas matemáticamente como un conjunto de índices de localizaciones con un conjunto asociado de vecinos.
- **Vecinos:** áreas vecinas (proximidad, contiguidad, etc.) a una región concreta.

ESTADÍSTICA ESPACIAL

1. Datos en rejilla o datos de área ('lattice data' o 'areal data')



2. Introducción a la epidemiología y la estadística espacial

ESTADÍSTICA ESPACIAL

2. **Datos referenciados puntualmente: Procesos puntuales ('point processes')**
 - Los **procesos puntuales** se corresponden con variables aleatorias Bernouilli.
 - En los procesos puntuales ***se conoce la localización exacta del caso y esta es aleatoria.***
 - Las ***localizaciones*** del suceso de interés ***se observan en una región determinada.*** Por ejemplo, las coordenadas de los domicilios de los casos de ELA en Cataluña.

ESTADÍSTICA ESPACIAL

2. Datos referenciados puntualmente: Procesos puntuales ('point processes')

- La **información** sobre estos datos **no es pública**. Se recogen en estudios de casos y controles o en estudios de cohortes.
- Se pueden agregar por unidades espaciales, pudiéndose crear datos de área.

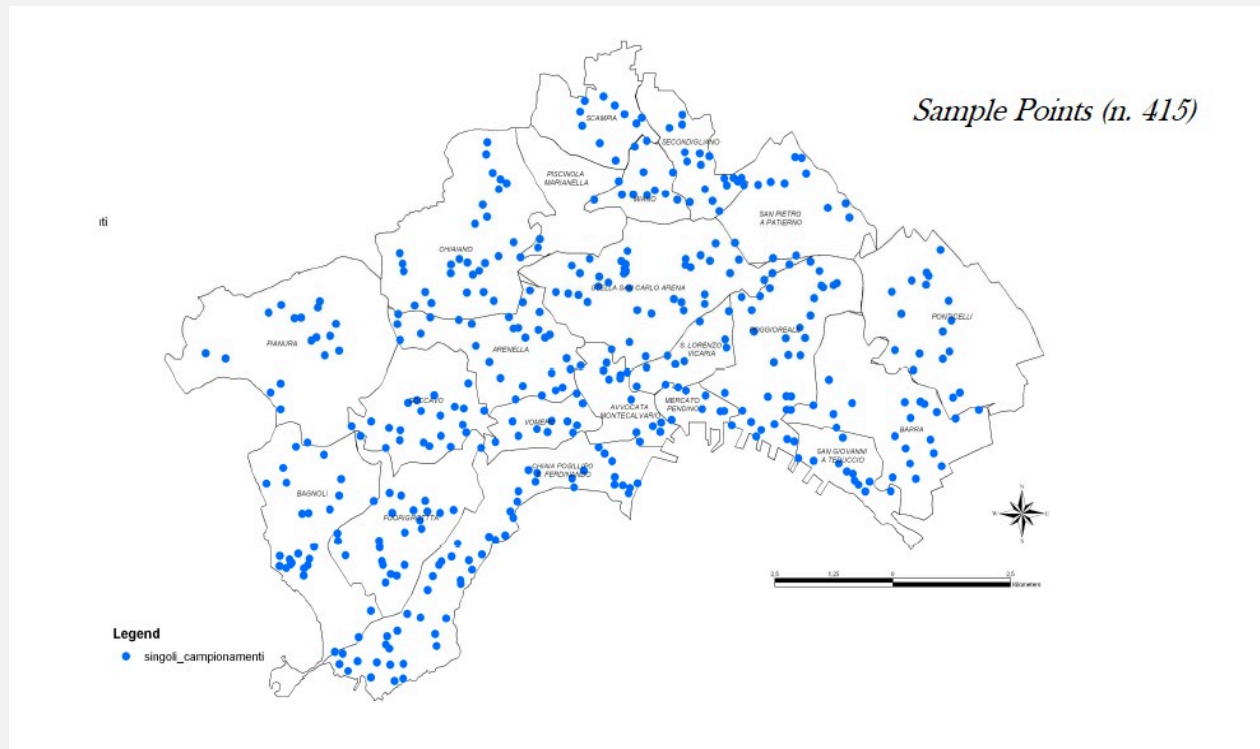
ESTADÍSTICA ESPACIAL

2. Datos referenciados puntualmente: Procesos puntuales ('point processes')

- En este caso, nos puede interesar:
 - Valorar si los sucesos siguen un patrón espacial determinado (agregación, forma regular, etc.)
 - Estudiar si un patrón observado se asocia a alguna variable (exposición a alguna variable ambiental, como la contaminación atmosférica; proximidad a focos contaminantes; contexto socioeconómico, etc.)

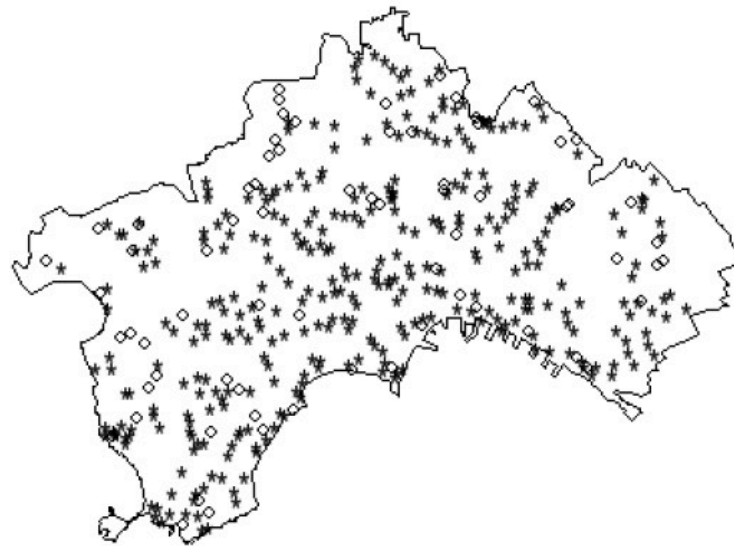
ESTADÍSTICA ESPACIAL

2. Datos referenciados puntualmente: Procesos puntuales ('point processes')



ESTADÍSTICA ESPACIAL

2. Datos referenciados puntualmente: Procesos puntuales ('point processes')



Distribution of cases (positives: circle) and controls (negatives: stars). Naples, February - May 2005

ESTADÍSTICA ESPACIAL

3. Datos referenciados puntualmente: Datos geoestadísticos ('geoestatistical data')

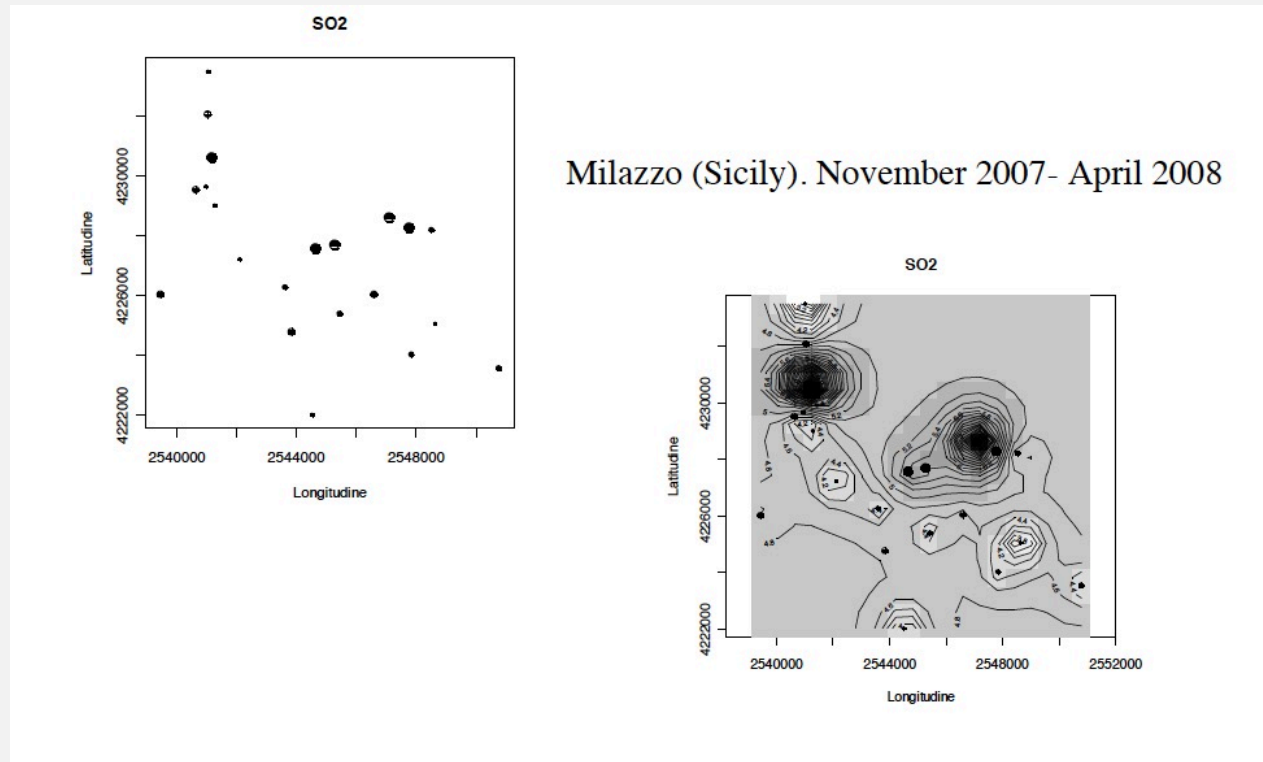
- Los **datos geoestadísticos** se corresponden con variables aleatorias continuas.
- En los datos geoestadísticos ***se conoce la localización exacta del caso y esta es fija*** (estaciones captadoras de contaminación atmosférica, vetas de un determinado metal, etc.)
- ***La variable respuesta se mide en cada localización*** (por ejemplo, medidas de contaminación atmosférica, medidas de composición química del suelo, temperatura, etc.).

ESTADÍSTICA ESPACIAL

3. Datos referenciados puntualmente: Datos geoestadísticos ('geoestatistical data')
- Es decir, disponemos de medidas tomadas en puntos fijos, definidas en cualquier lugar del espacio, de manera que sus **localizaciones definen una superficie espacialmente continua**.
 - Se suele **extender la distribución espacial** de los valores de un atributo **sobre la región de estudio** mediante modelos matemáticos (por ejemplo, kriging).

ESTADÍSTICA ESPACIAL

3. Datos referenciados puntualmente: Datos geoestadísticos ('geoestatistical data')



ESTADÍSTICA ESPACIAL

Pero, los tipos de diseño espacial, ¿qué son en realidad? ¿procesos? ¿modelos? ¿métodos?

- El 2012, Diggle propone un cambio de paradigma y re-define la estadística espacial como ***'un conjunto de modelos y métodos estadísticos que pretenden ayudar a los científicos a entender fenómenos espaciales, que no pueden ser observados directamente, sino de forma indirecta, con información incompleta'***, en forma de datos en rejilla, procesos puntuales y datos geoestadísticos. El model estadístico que propuso, como modelo único base, es el **model log Cox**.

ESTADÍSTICA ESPACIAL

- Es decir, Diggle (2012) unificó la estadística espacial, tal y como McCullagh y Nelder (1989) unificaron los modelos lineales generalizados.

