Imperial College London

Spatial epidemiology

Dr Mireille B. Toledano

Dept of Epidemiology & Biostatistics

6th October 2010

Learning outcomes

- **At the end of this lecture, you should be:**
  - **able to distinguish the main types of studies employed in spatial epidemiology**
  - **able to give an example of each**
  - **aware of their main limitations**
  - **able to appreciate the relevance of spatial epidemiology in addressing important public health questions**

Small-area studies

- **Four types of small-area studies of environment and health:**

1. **Cluster detection**
2. **Disease mapping**
3. **Geographical correlation and semi-ecologic studies**
4. **Risk in relation to a point source**

1. Disease clusters

What is a Cluster? (Knox 1989)

A geographically and / or temporally bounded group of occurrences of sufficient size and concentration to be unlikely to have occurred by chance

MYSTERY OF BABIES WITH NO EYES

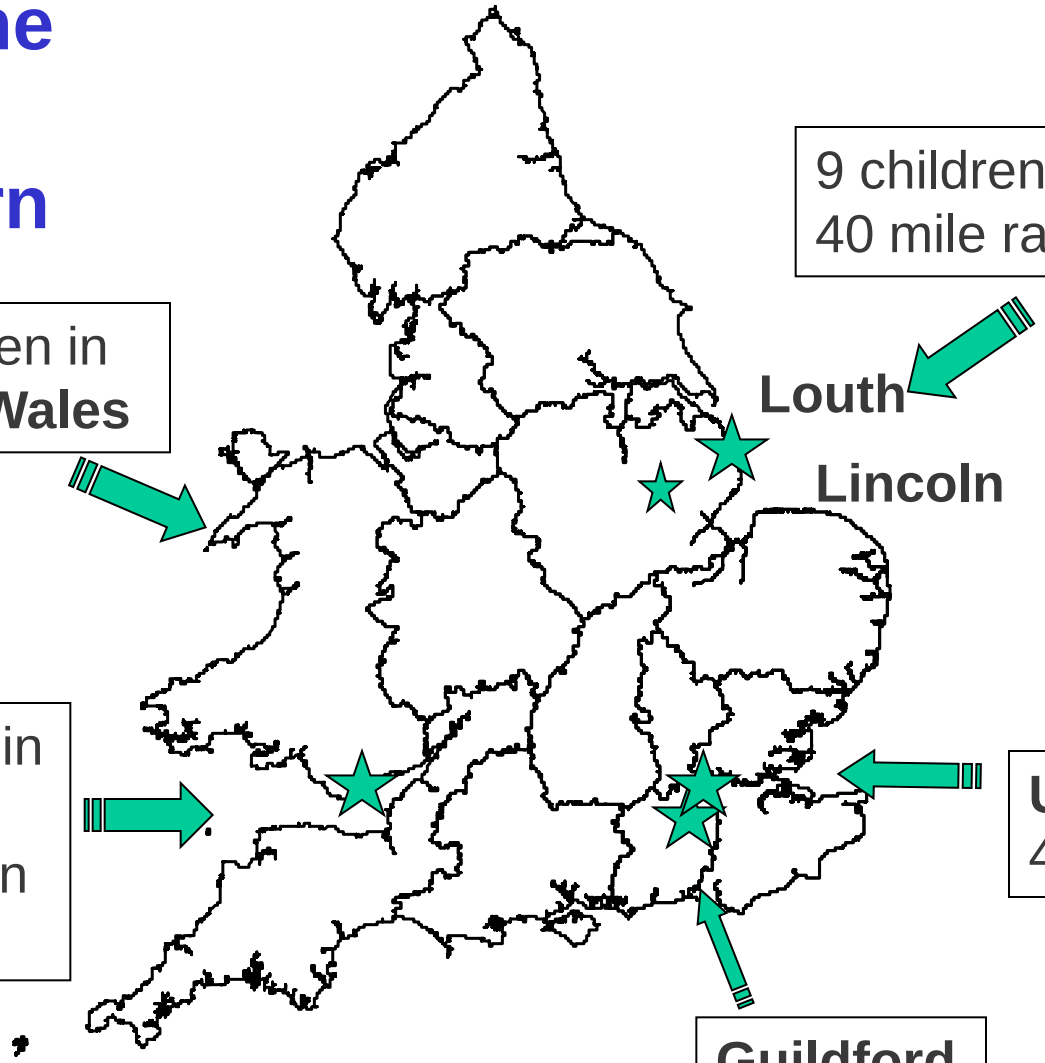
"Clusters" raise fears of link with pesticide

Observer, 17 January 1993, page 3

PESTICIDE FEARS GROW AS NUMBER OF BABIES BORN BLIND DOUBLES

Observer, 31 January 1993, page 3

Where the victims were born



9 children within 40 mile radius

3 children in **North Wales**

Louth

Lincoln

6 children within **Cardiff** area. Plus 5 others in **South Wales**

Uxbridge
4 children

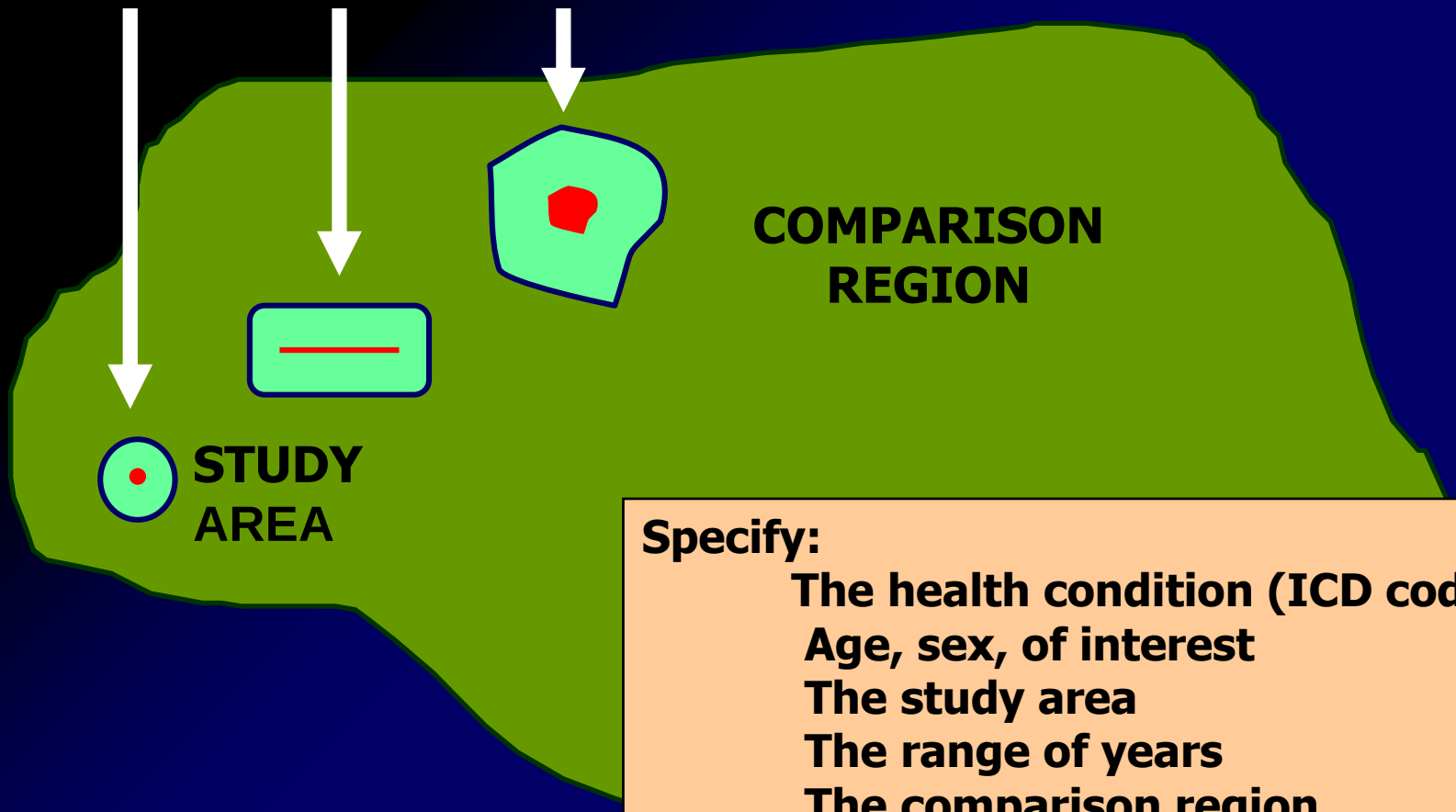
Guildford
4 children

Observer, 17 January 1993, page 3

Cluster investigations

Is the risk different here (compared to the comparison region)?

$$\frac{\text{Observed}}{\text{Expected}} = 1\ ?$$

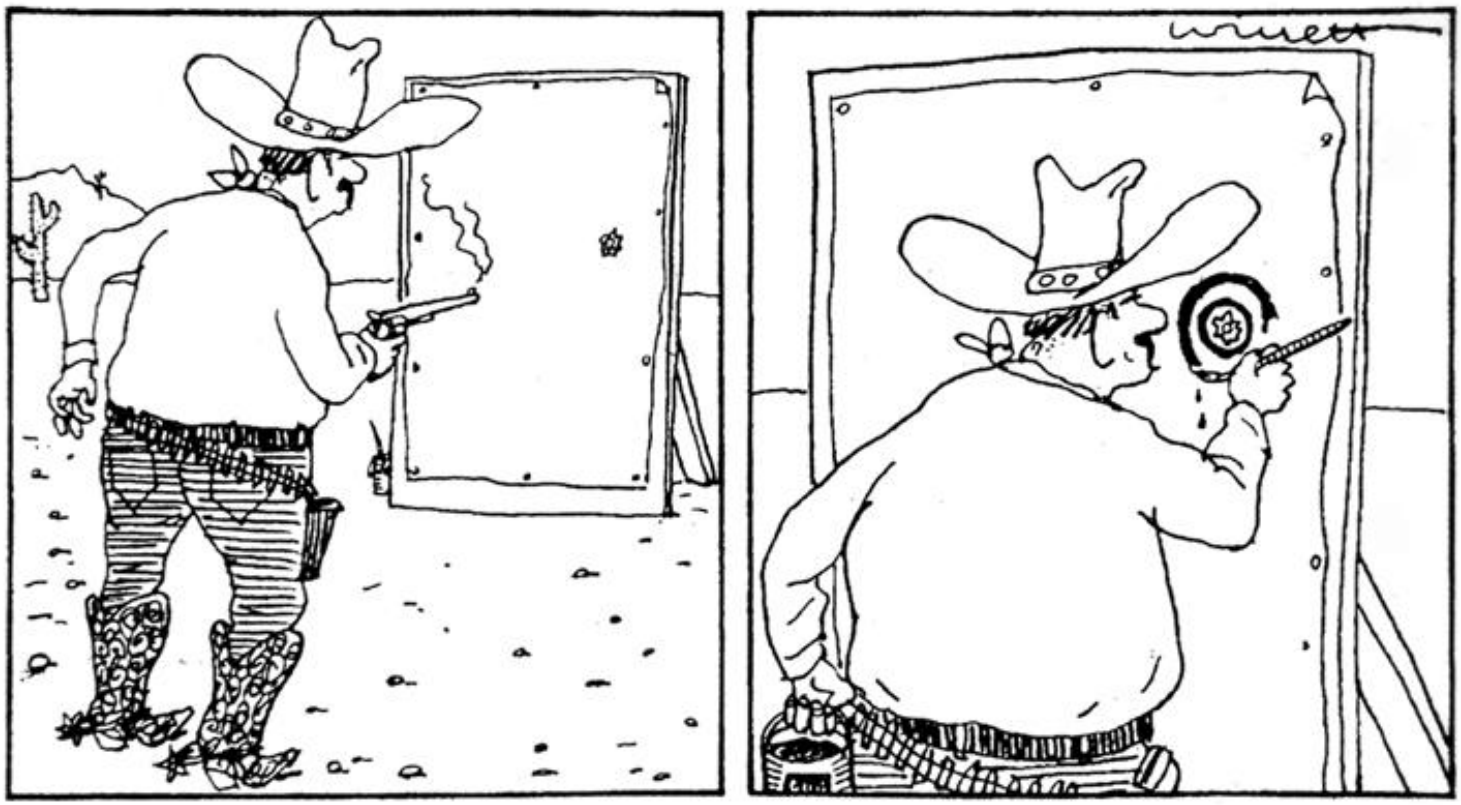


Specify:

- **The health condition (ICD codes)**
- **Age, sex, of interest**
- **The study area**
- **The range of years**
- **The comparison region**

"We found little or no evidence of large scale geographical variation or localised clustering in anophthalmia and microphthalmia during 1988-94"

Dolk et al. BMJ 1998; 317: 905-910



The identification of a cluster and its source population is all too often like the Texas sharp-shooter who first fires his bullet and only then draws the target round it.

Cluster Investigation: Problems

- Often post hoc - no *a priori* hypothesis
- ‘Highs’ investigated - never ‘lows’
- Small numbers
- ‘Boundary shrinkage’ & multiple testing:
  - time/space/disease classification/age/sex

Sellafield nuclear reprocessing plant, Cumbria

The Black Enquiry
Recommendation 5

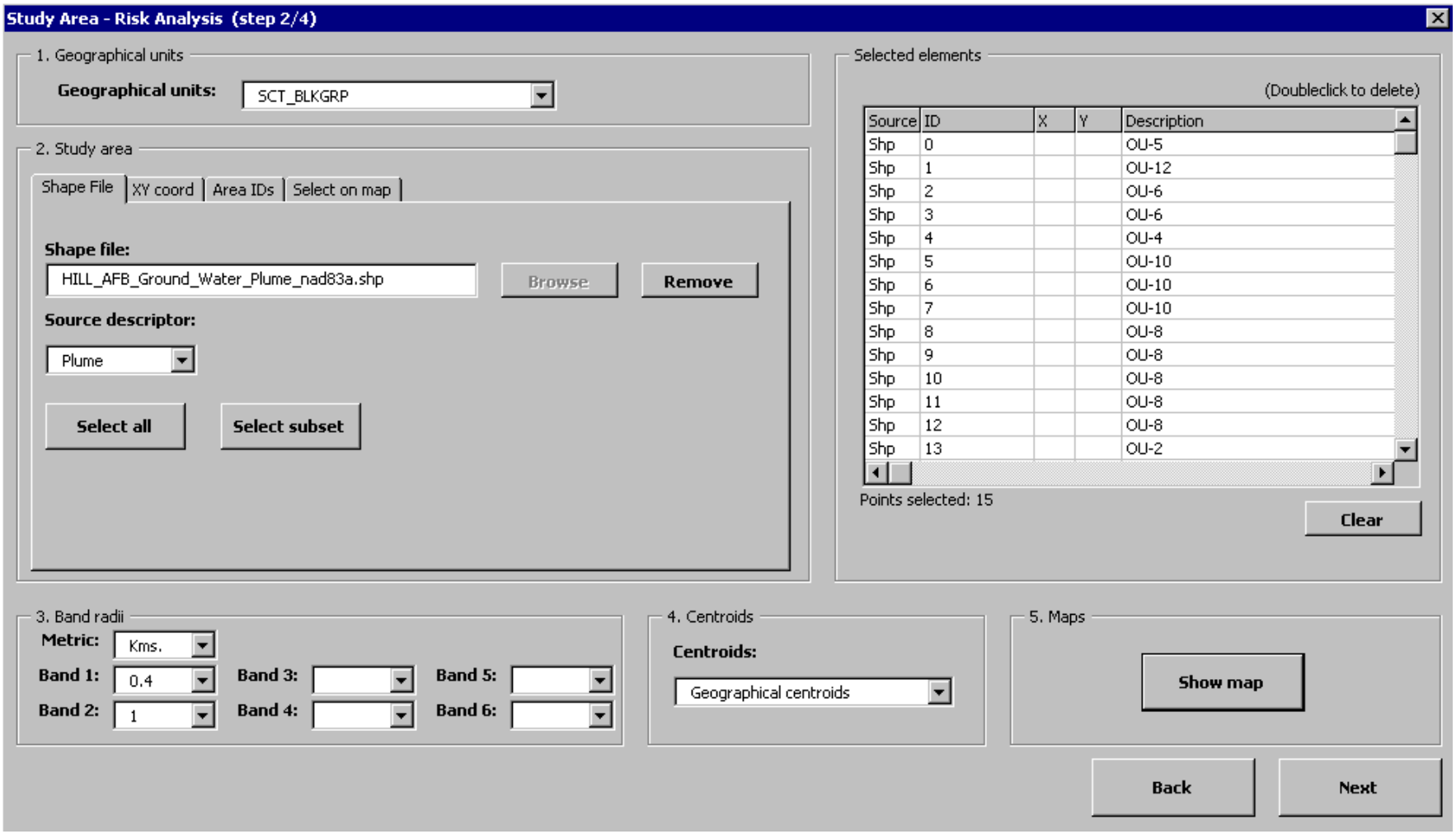
"… encouragement should be given to an organisation … to co-ordinate centrally the monitoring of small area statistics around major installations producing discharges that might present a carcinogenic or mutagenic hazard to the public. In this way, early warning of any untoward health effect could be obtained."

London, HMSO, 1984

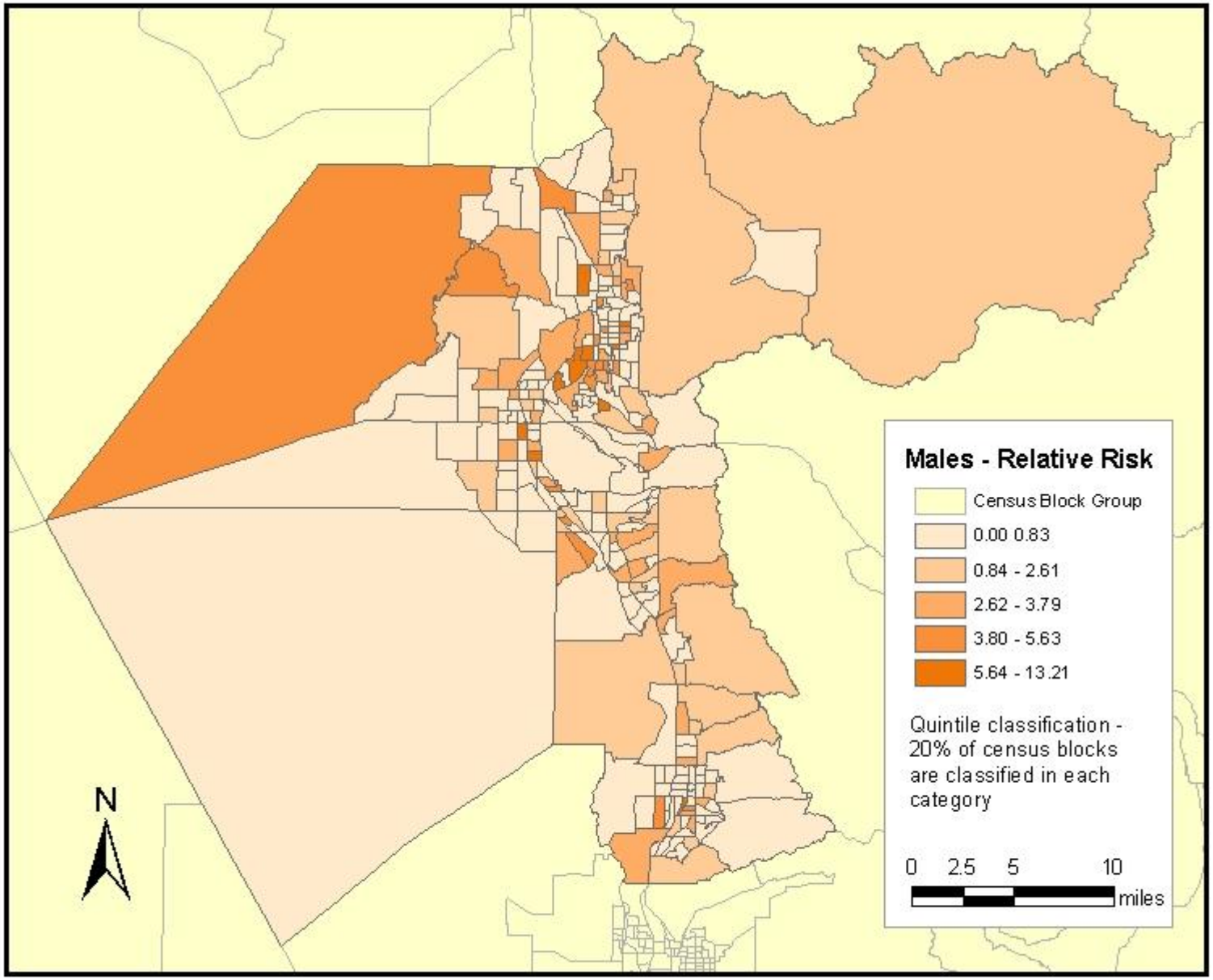
SAHSU - scope of work
www.sahsu.org

- To develop and maintain a comprehensive database of postcoded health data & environmental exposures and social confounding factors at the small-area level
- To develop small-area statistical methodology
- To respond rapidly to *ad hoc* queries about unusual clusters of disease, particularly in the neighbourhood of industrial installations
- To develop the RIF – a software tool developed by SAHSU to facilitate disease cluster investigation

RIF menu example – study area definition



CDC EPHT pilot study for Utah – lung cancer risk



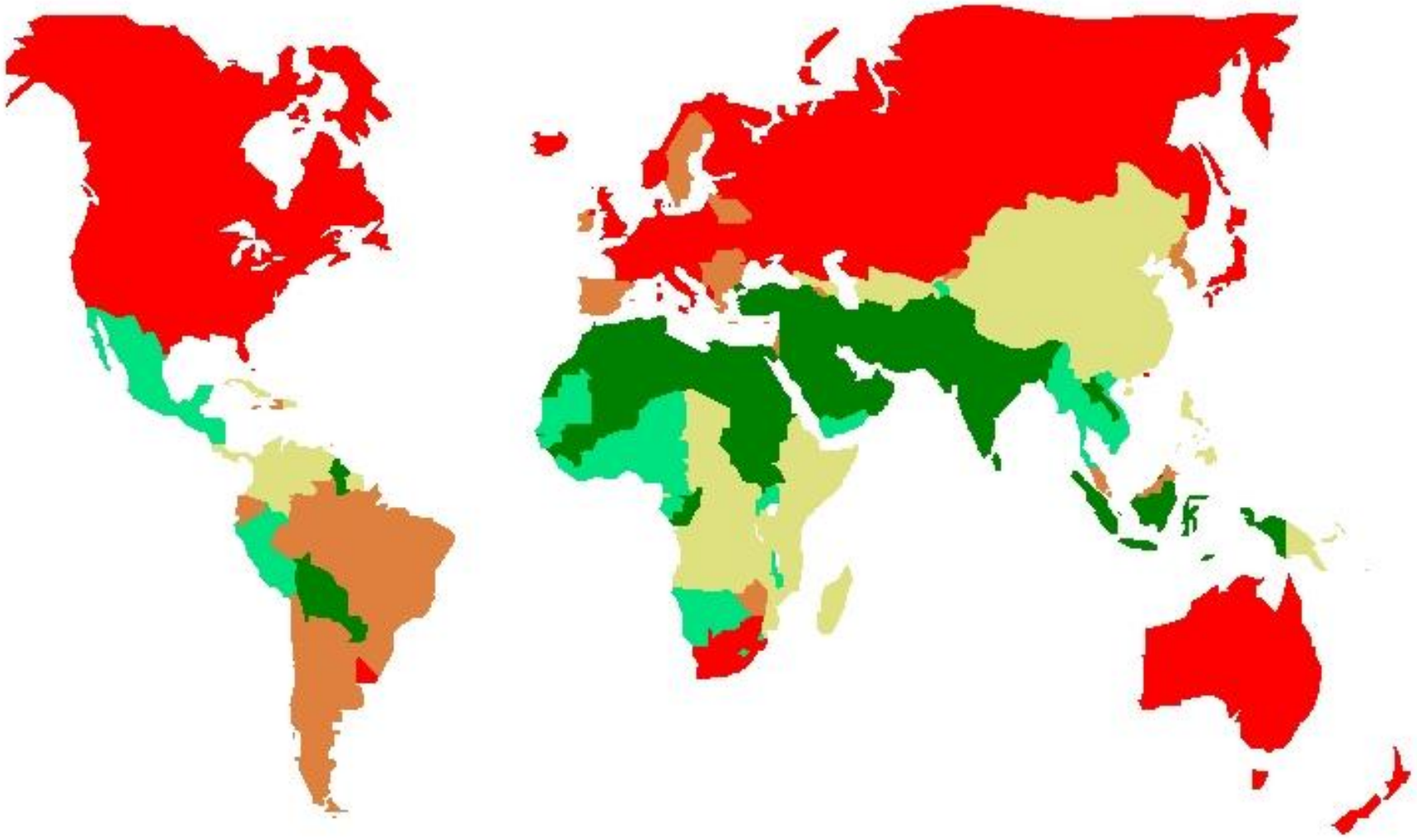
2. Disease mapping

- Usually information on disease only (sometimes exposure only)
- Geographical differences in disease risk
- Can address aetiological questions even in the absence of risk factor data

2. Disease mapping

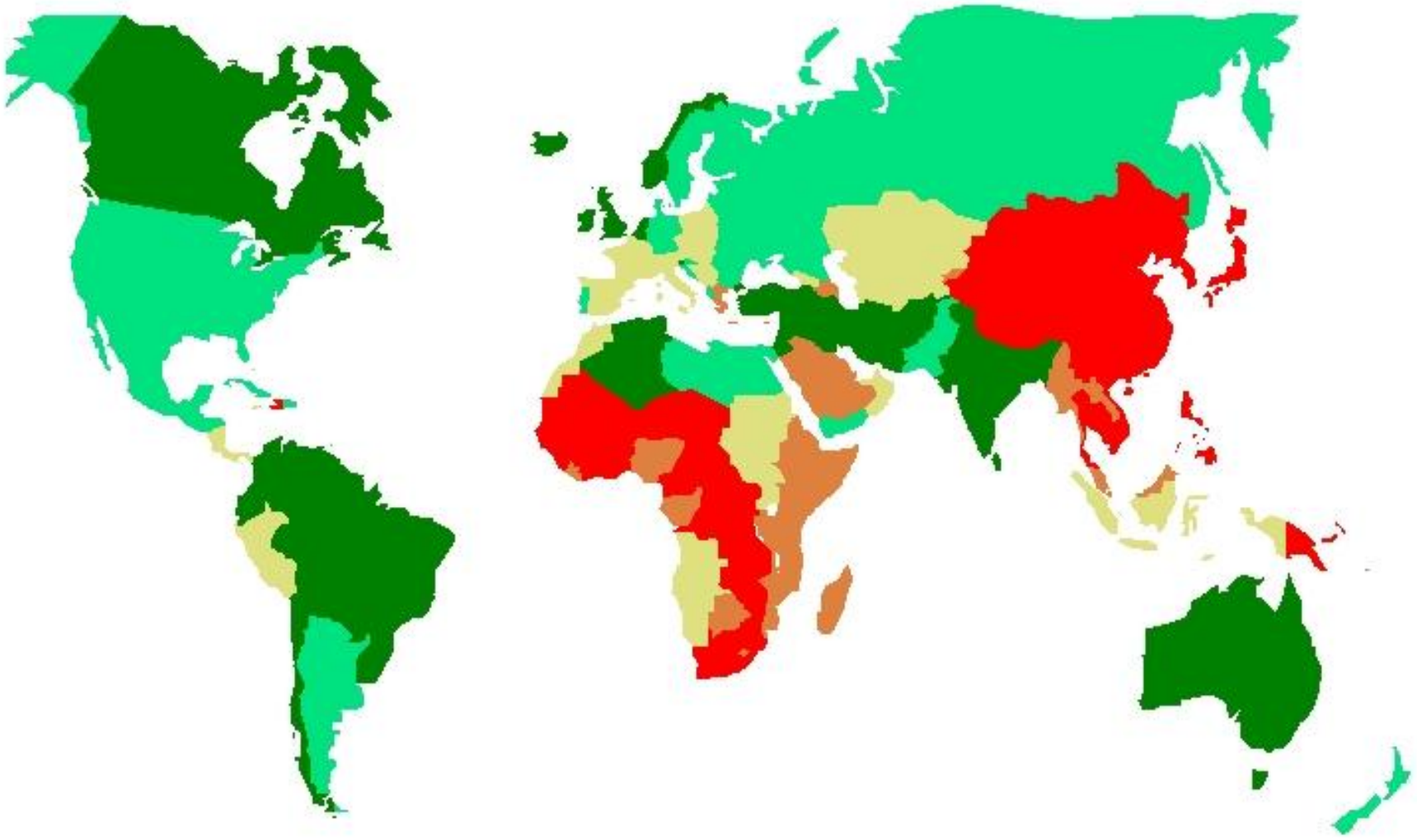
- **Scale:**
  - Global
  - National
  - ‘small area’

Incidence of All cancers: ASR (World)-Male (All ages)



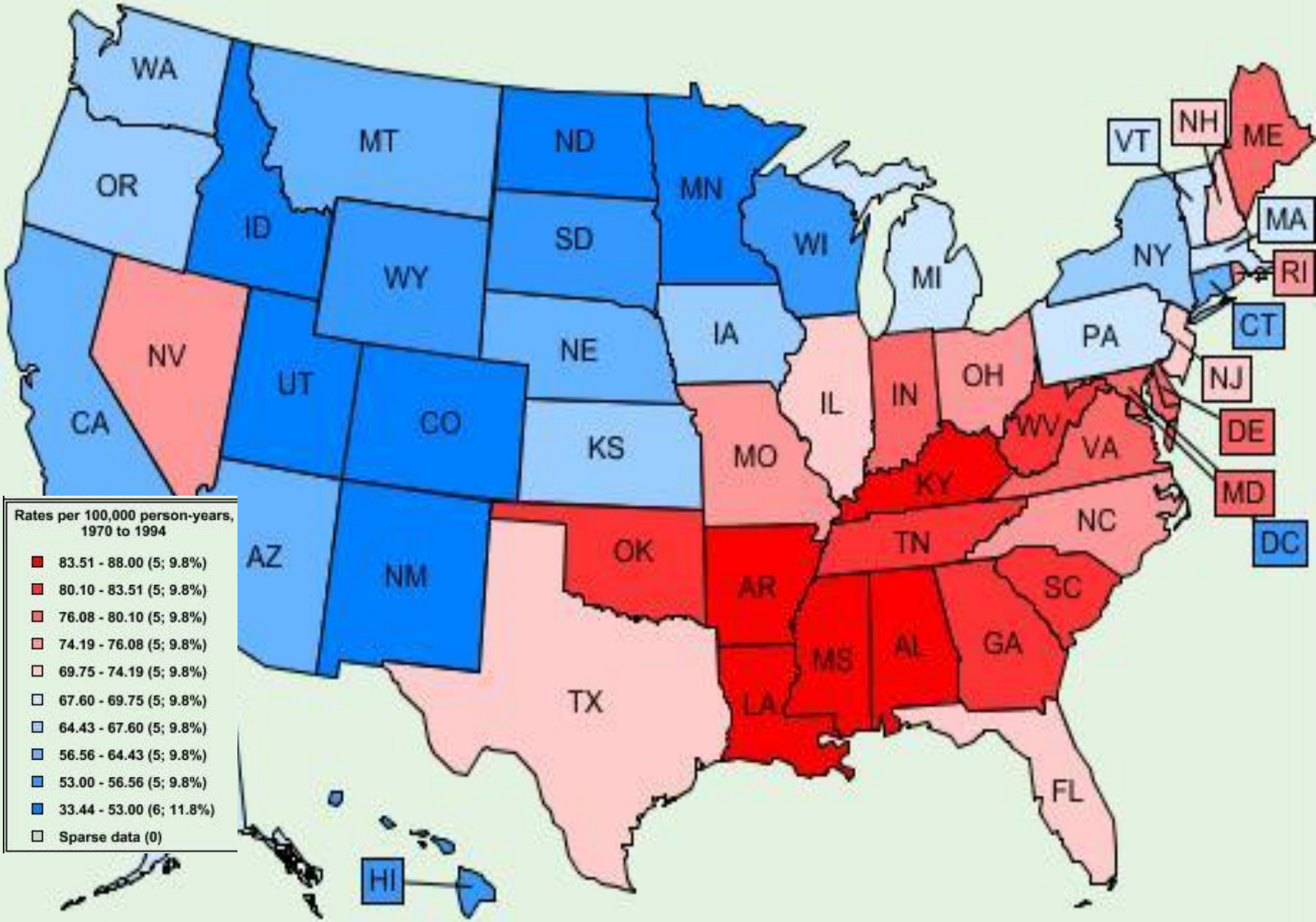
GLOBOCAN (IARC 1998)

Incidence of Liver cancer: ASR (World)-Male (All ages)



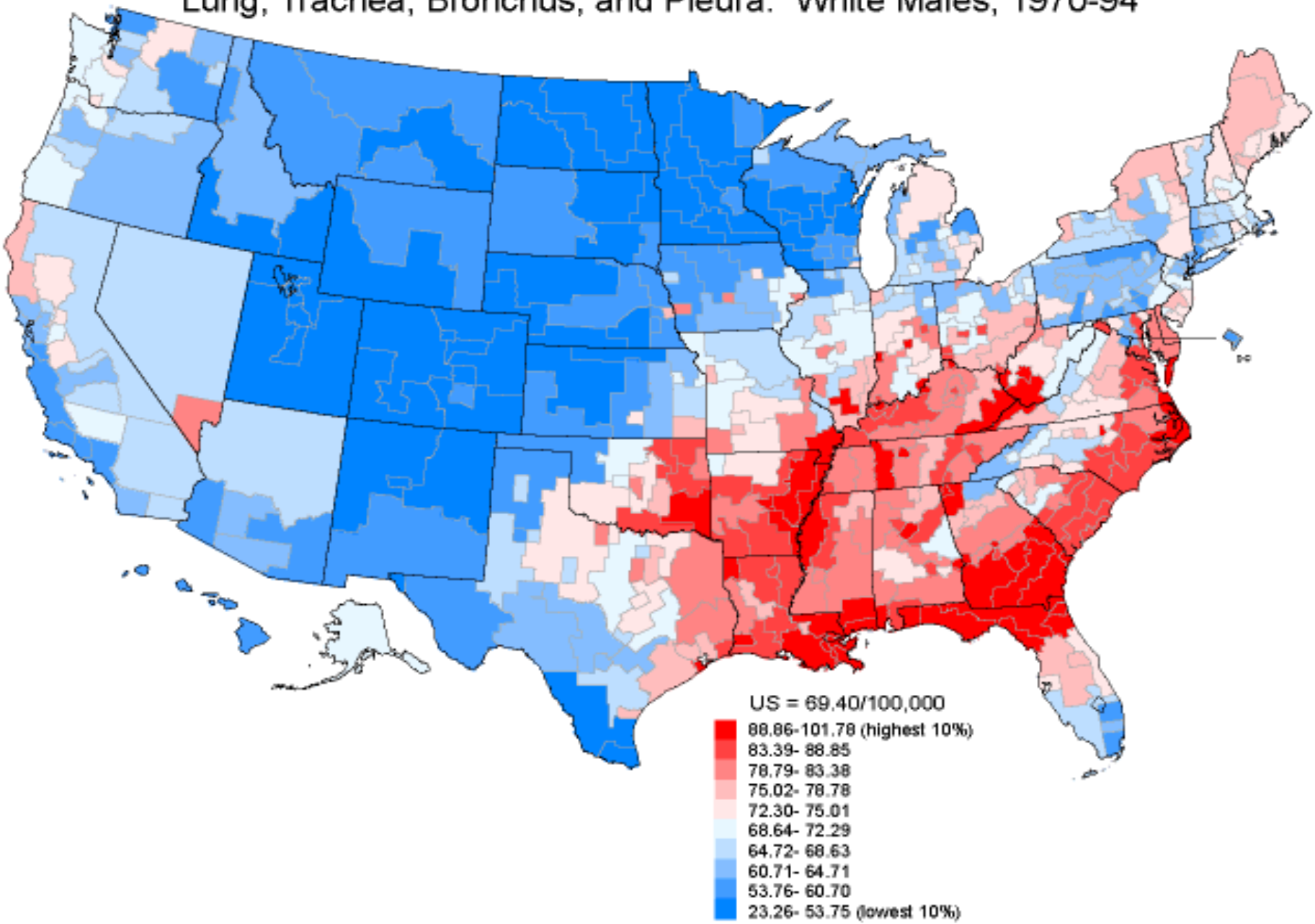
GLOBOCAN (IARC 1998)

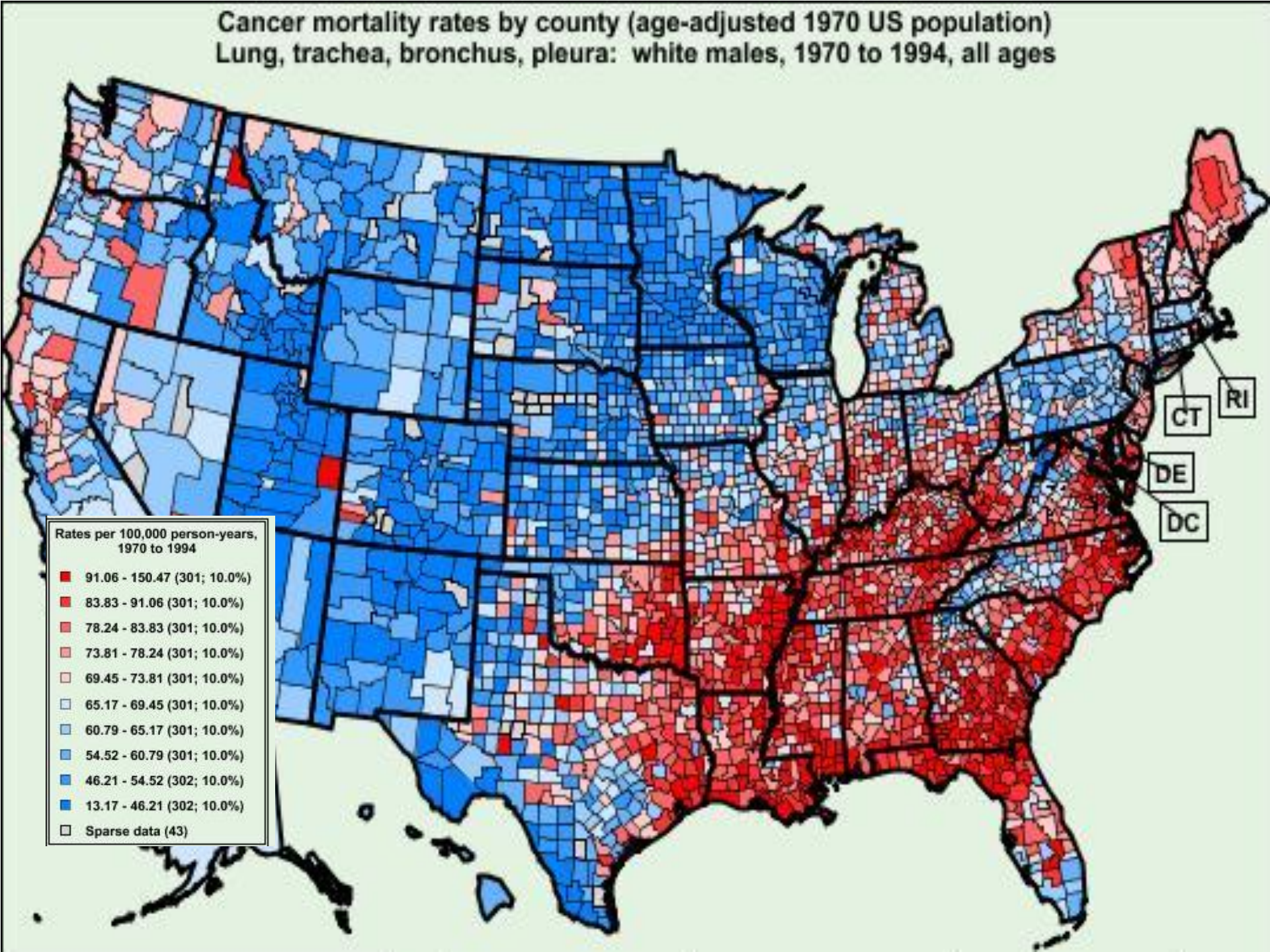
Cancer mortality rates by state (age-adjusted 1970 US population)
Lung, trachea, bronchus, pleura: white males, 1970 to 1994, all ages

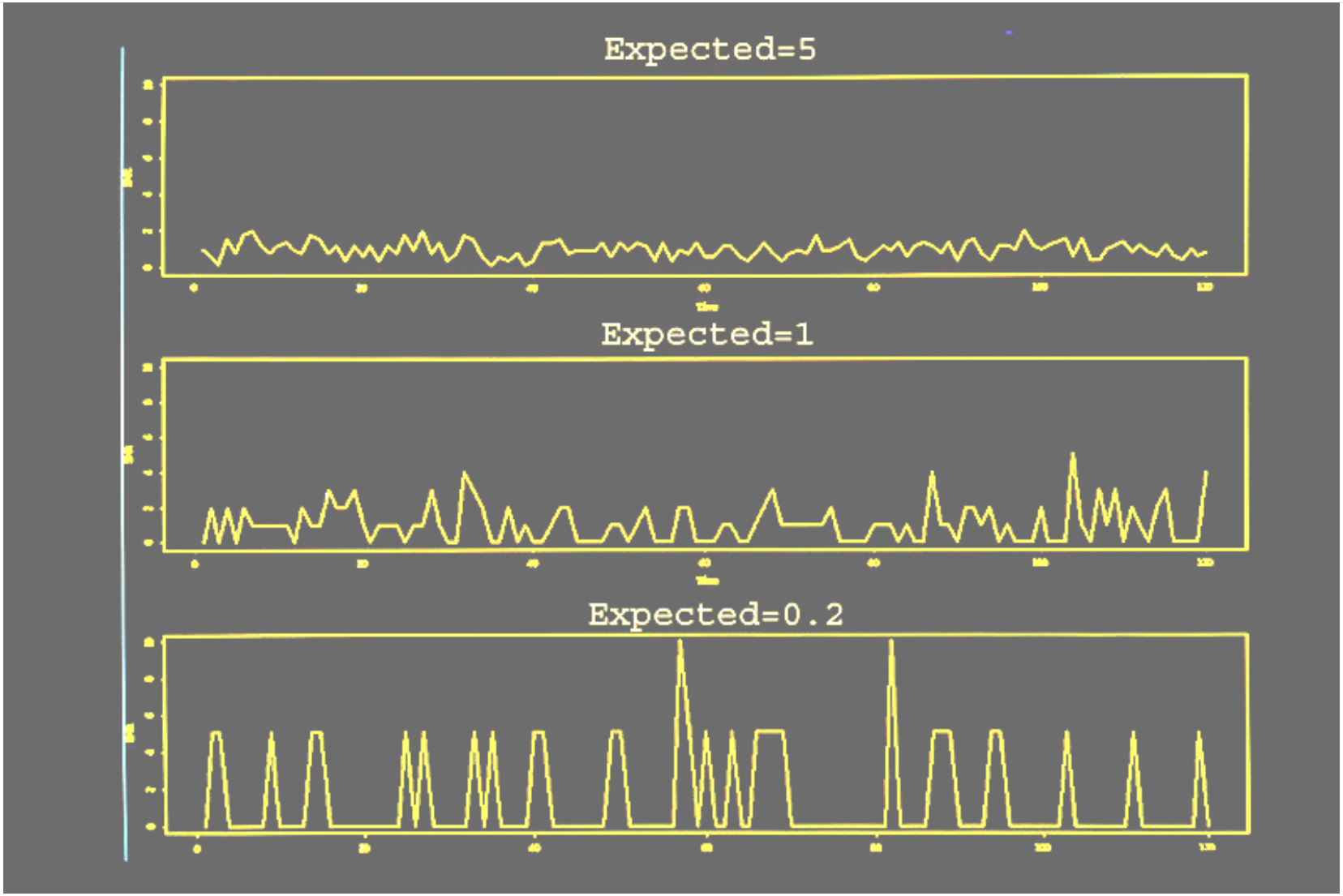


Cancer Mortality Rates by State Economic Area (Age-adjusted 1970 US Population)

Lung, Trachea, Bronchus, and Pleura: White Males, 1970-94




Cancer mortality rates by county (age-adjusted 1970 US population)
Lung, trachea, bronchus, pleura: white males, 1970 to 1994, all ages
Rates per 100,000 person-years, 1970 to 1994
91.06 - 150.47 (301; 10.0%)
83.83 - 91.06 (301; 10.0%)
78.24 - 83.83 (301; 10.0%)
73.81 - 78.24 (301; 10.0%)
69.45 - 73.81 (301; 10.0%)
65.17 - 69.45 (301; 10.0%)
60.79 - 65.17 (301; 10.0%)
54.52 - 60.79 (301; 10.0%)
46.21 - 54.52 (302; 10.0%)
13.17 - 46.21 (302; 10.0%)
Sparse data (43)
RI
CT
DE
DC


Expected=5
Expected=1
Expected=0.2

Small-area disease mapping studies

- Sparse data
- Standard statistical methods that estimate the risks for each area independently suffer from lack of precision
- Bayesian hierarchical models - developed to facilitate analysis of sparse data by better separating the true “signal” from random noise

Bayesian models in disease mapping

- Provide shrinkage or **smoothing** (over space or time or both) of the raw risk estimates (e.g. SMR or O/E) to give more stable estimates of the underlying pattern of risk.

- Risk estimates for any particular area "borrow strength" from relevant data about the disease risks in neighbouring areas

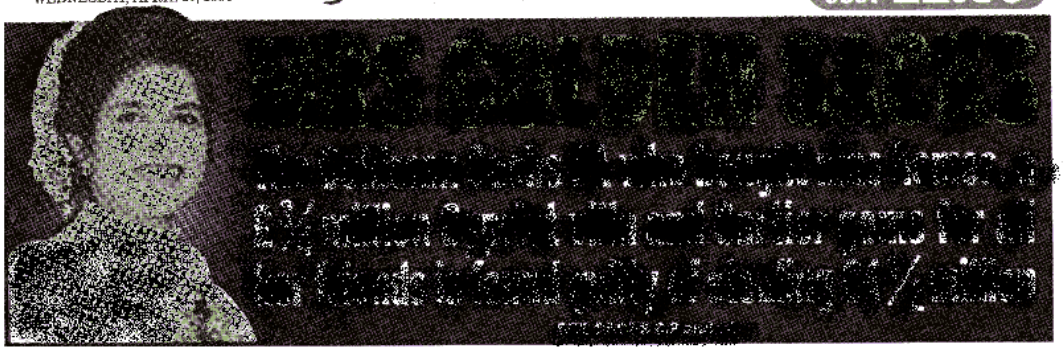
Beware!

- Over-smoothing? “SAHSU white washing”
- Do not over-interpret lack of spatial variation!
- “A good map of bad data looks better than a bad map of good data”

3. Semi-ecologic studies

- Information on exposure and disease
- Taking ‘geographical correlation’ studies one step further
- Often at varying geographical levels

Daily Mail
WEDNESDAY, APRIL 21, 2004
40p



Experts' fears over chlorine...

DOES TAP WATER HARM UNBORN BABIES?

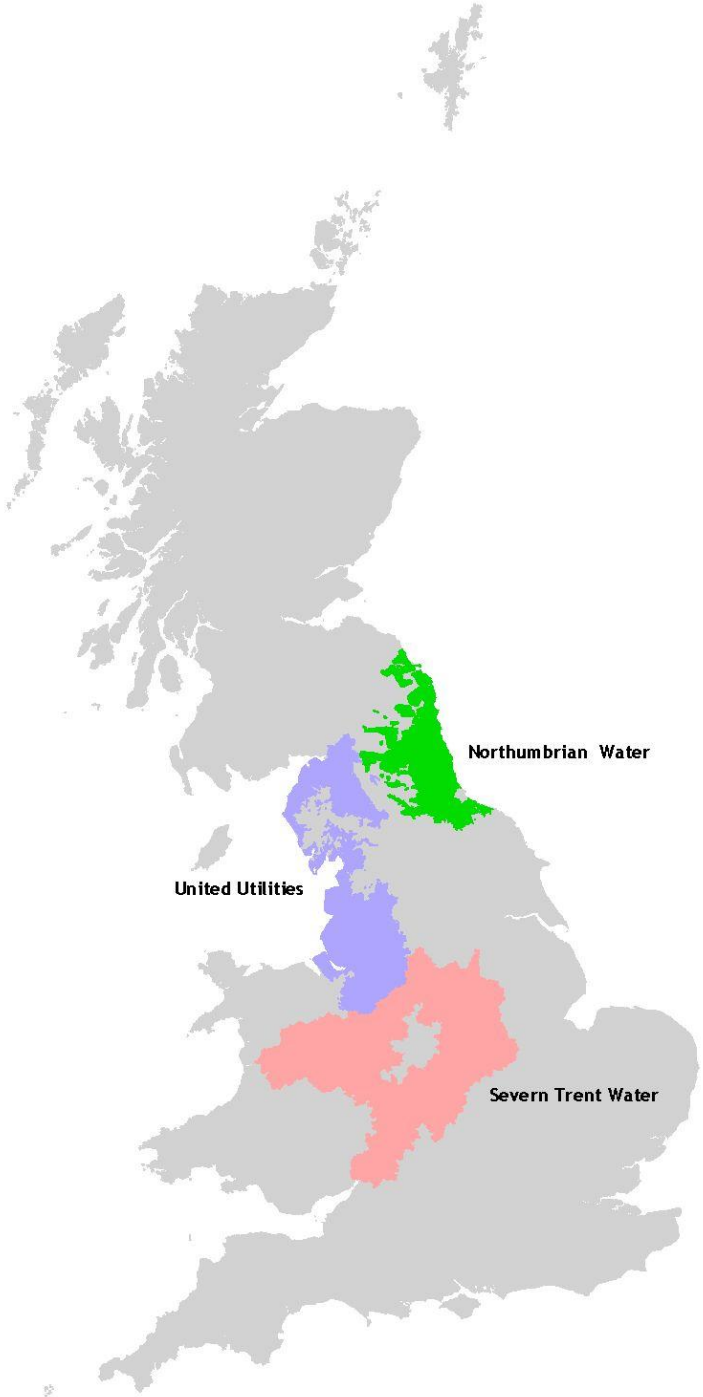
WATER companies have been warned to limit levels of chlorine used to disinfect tap supplies amid fears of a link to birth defects and miscarriage.

By **Sean Poulter**
Consumer Affairs Correspondent

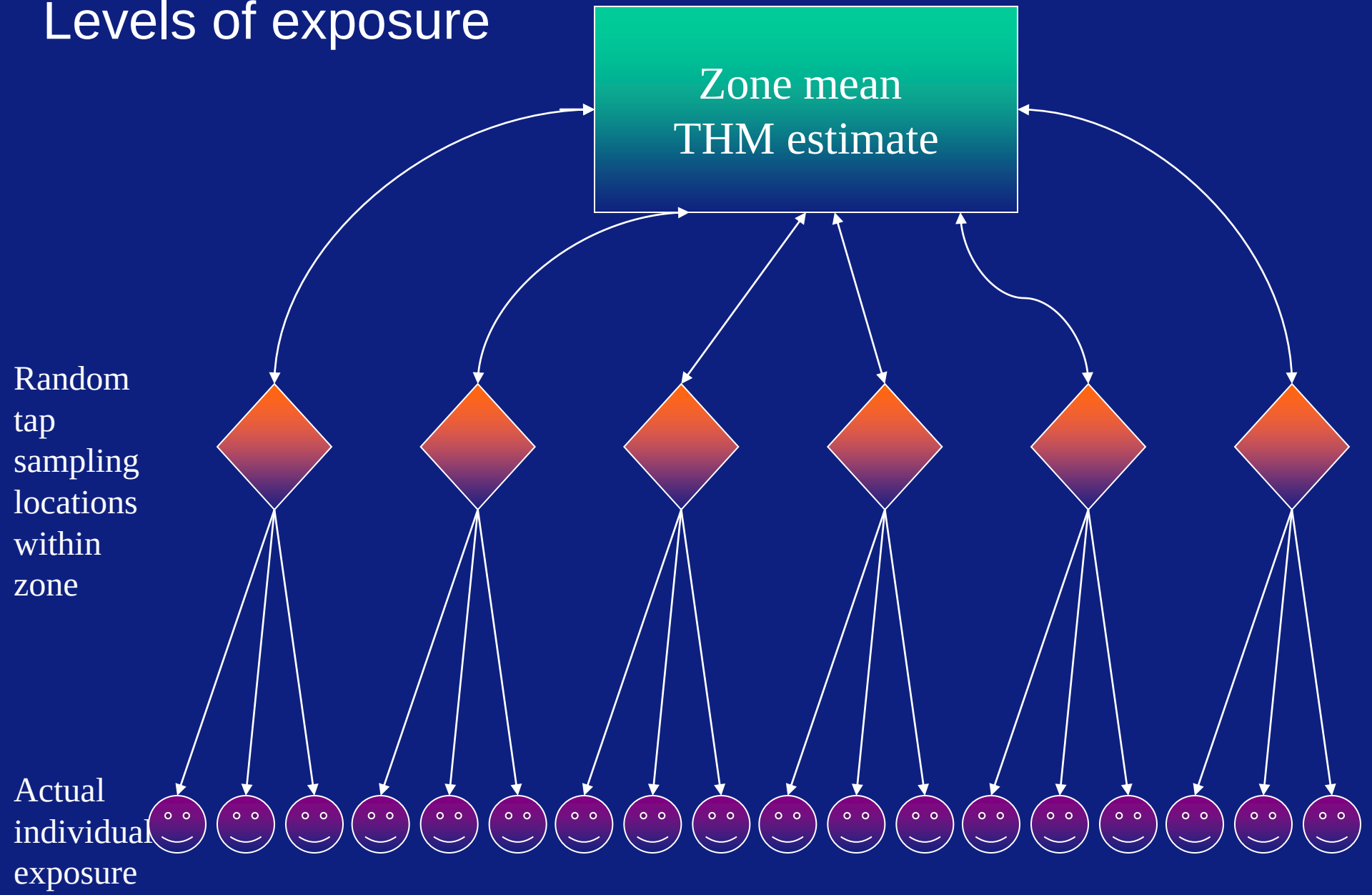
Government experts said research suggests chlorine by-products may harm the developing baby in the womb.

Chlorine is used to kill bugs and to disinfect

Turn to Page 2, Col. 3

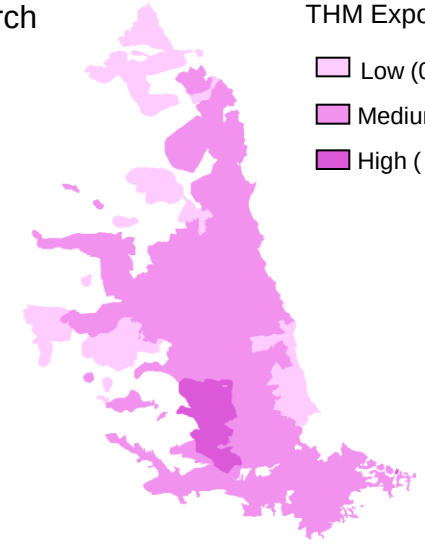

Northumbrian Water
United Utilities
Severn Trent Water

Levels of exposure

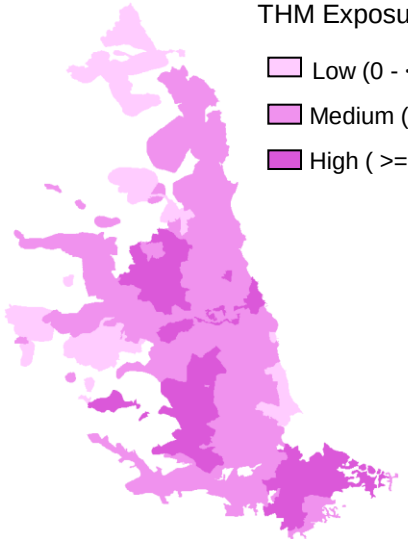


Northumbrian Water

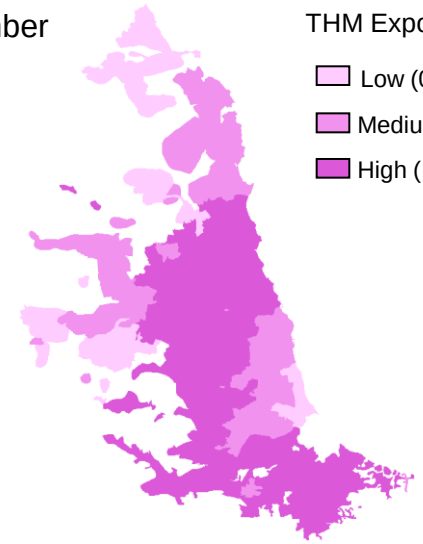
January - March



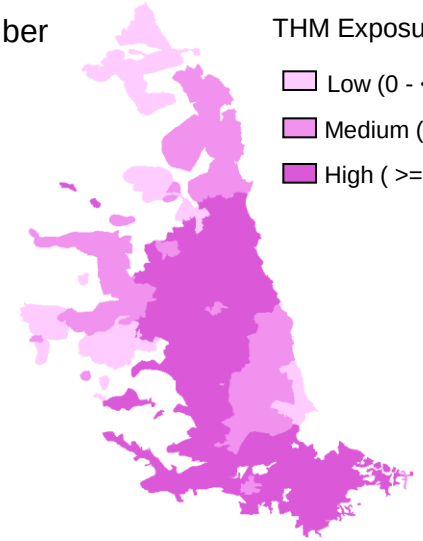
April - June

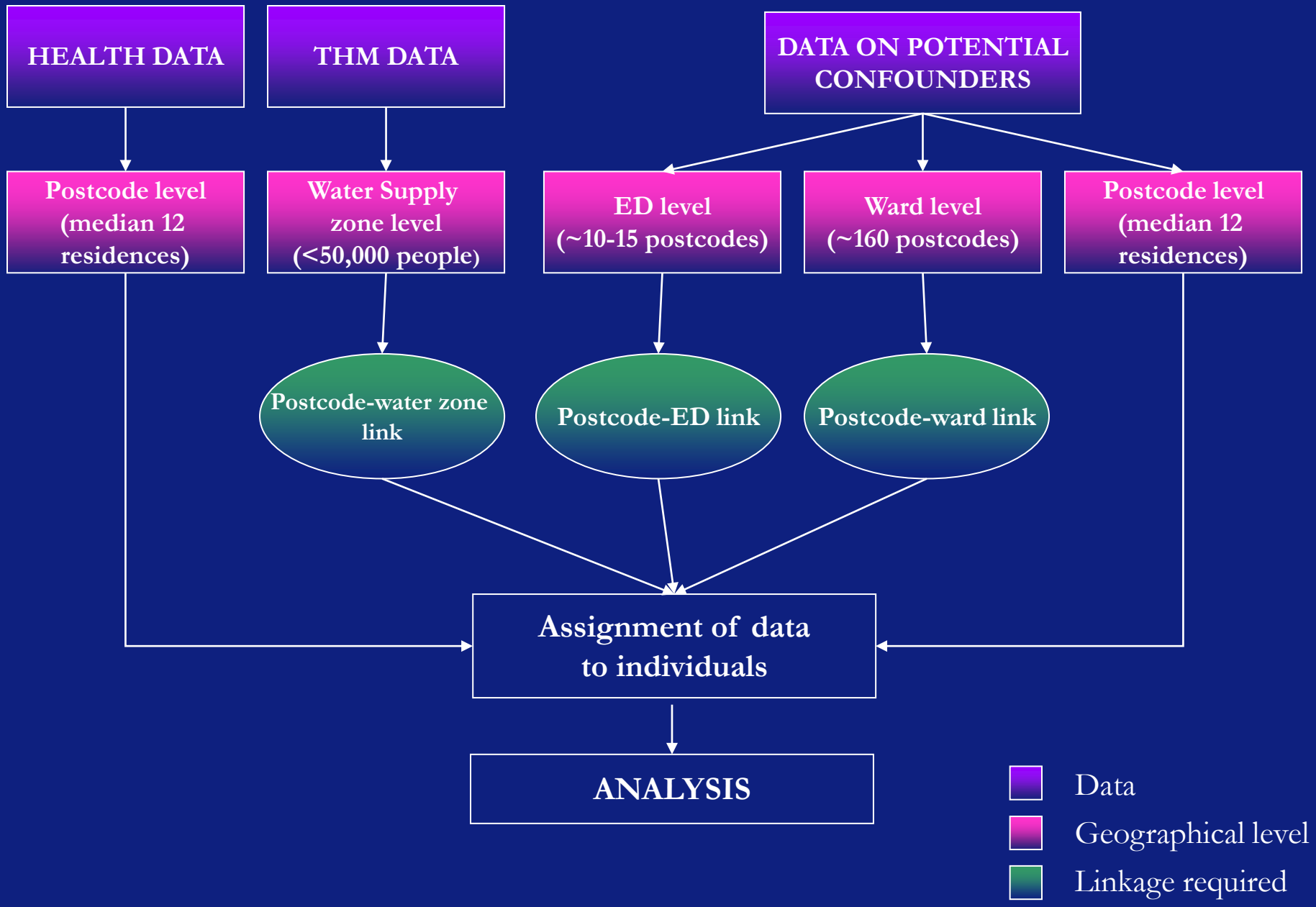


July - September



October - December




HEALTH DATA
THM DATA
DATA ON POTENTIAL CONFOUNDERS
Postcode level (median 12 residences)
Water Supply zone level (<50,000 people)
ED level (~10-15 postcodes)
Ward level (~160 postcodes)
Postcode level (median 12 residences)
Postcode-water zone link
Postcode-ED link
Postcode-ward link
Assignment of data to individuals
ANALYSIS
Data
Geographical level
Linkage required

Adjusted odds ratios (OR) for stillbirths and low and very low birth weight by TTHM (low <30µg/l, med 30-60µg/l, high>60µg/l) category and water region

	TTHM category	Overall	Northumbrian	United Utilities	Severn Trent
		OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Stillbirths	*Low*	**1.00**	**1.00**	**1.00**	**1.00**
	Medium	**1.06 (0.99,1.15)**	**1.19 (0.51, 2.75)**	**1.16 (1.00,1.35)**	**1.03 (0.95,1.13)**
	High	**1.11 (1.00, 1.23)**	**1.09 (0.46, 2.55)**	**1.21 (1.03,1.42)**	**1.04 (0.93,1.18)**
Low birth weight	*Low*	**1.00**	**1.00**	**1.00**	**1.00**
	Medium	**1.05 (0.96, 1.15)**	**1.02 (0.80, 1.30)**	**1.11 (1.07,1.16)**	**1.00 (0.98,1.03)**
	High	**1.09 (0.93, 1.27)**	**1.11 (0.87, 1.41)**	**1.19 (1.14,1.24)**	**0.98 (0.95,1.03)**
Very low birth weight	*Low*	**1.00**	**1.00**	**1.00**	**1.00**
	Medium	**1.03 (0.96, 1.10)**	**1.20 (0.66, 2.18)**	**1.09 (0.98,1.21)**	**1.00 (0.94,1.06)**
	High	**1.05 (0.82, 1.34)**	**1.11 (0.61, 2.03)**	**1.20 (1.07,1.34)**	**0.90 (0.82,0.99)**

Toledano *et al.* 2005

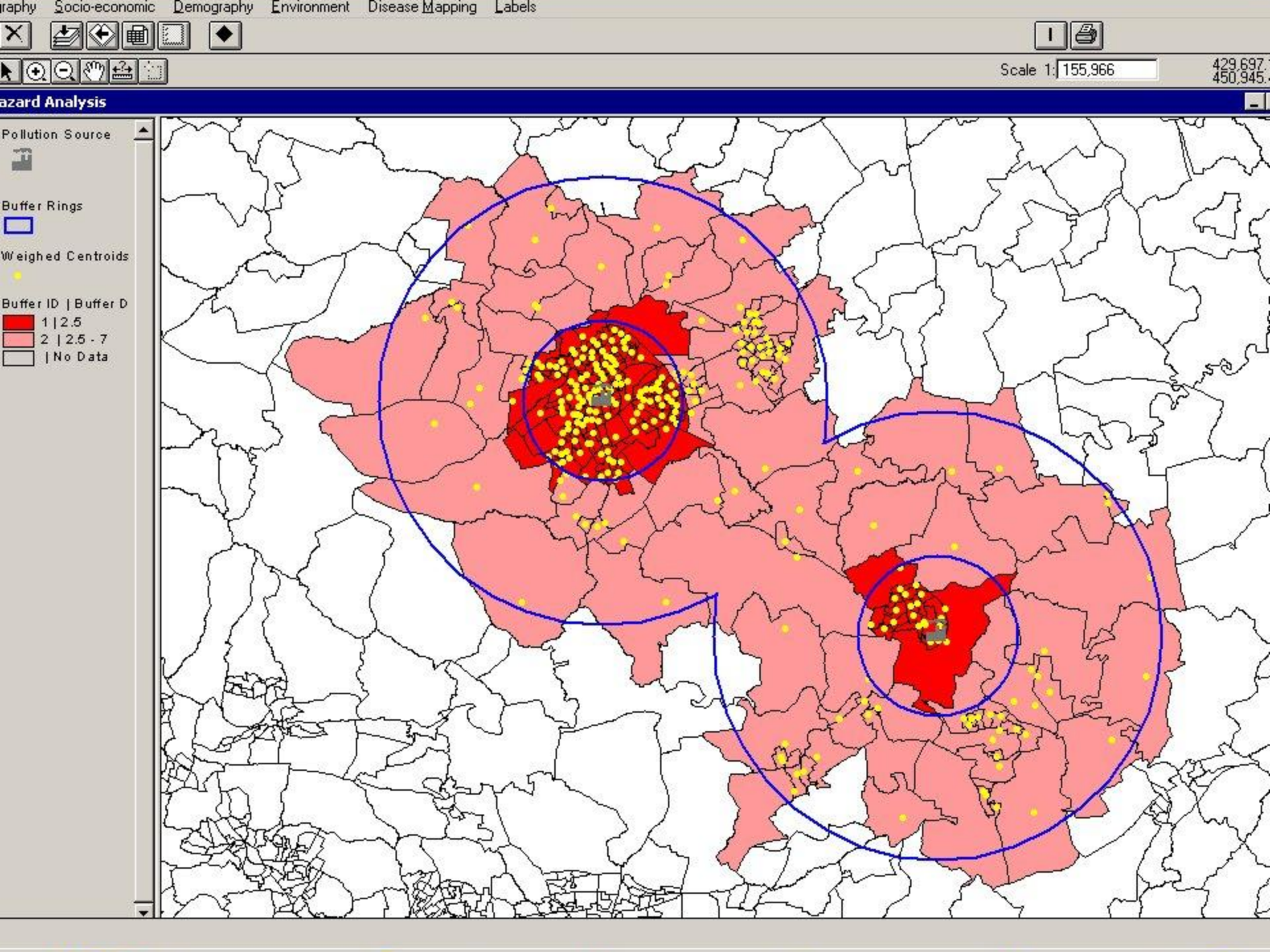
4. Point source studies

- Hypothesis testing of specific ‘point/linear’ exposures and disease associations

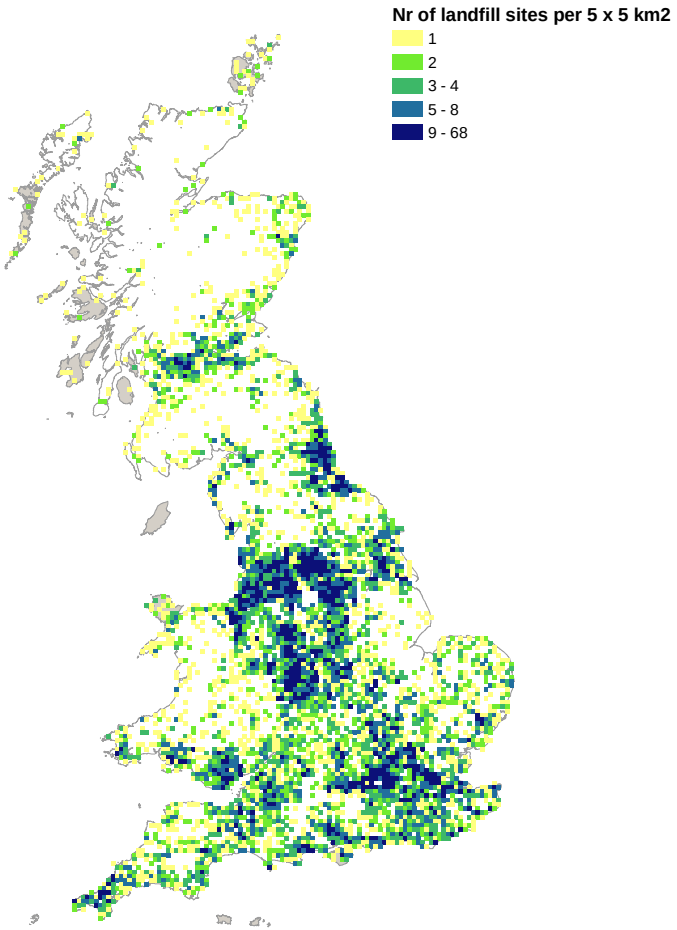
Public health/aetiological questions

- Am I at excess risk if I live near ‘X’ (polluting source or industry) and, if so, by how much?

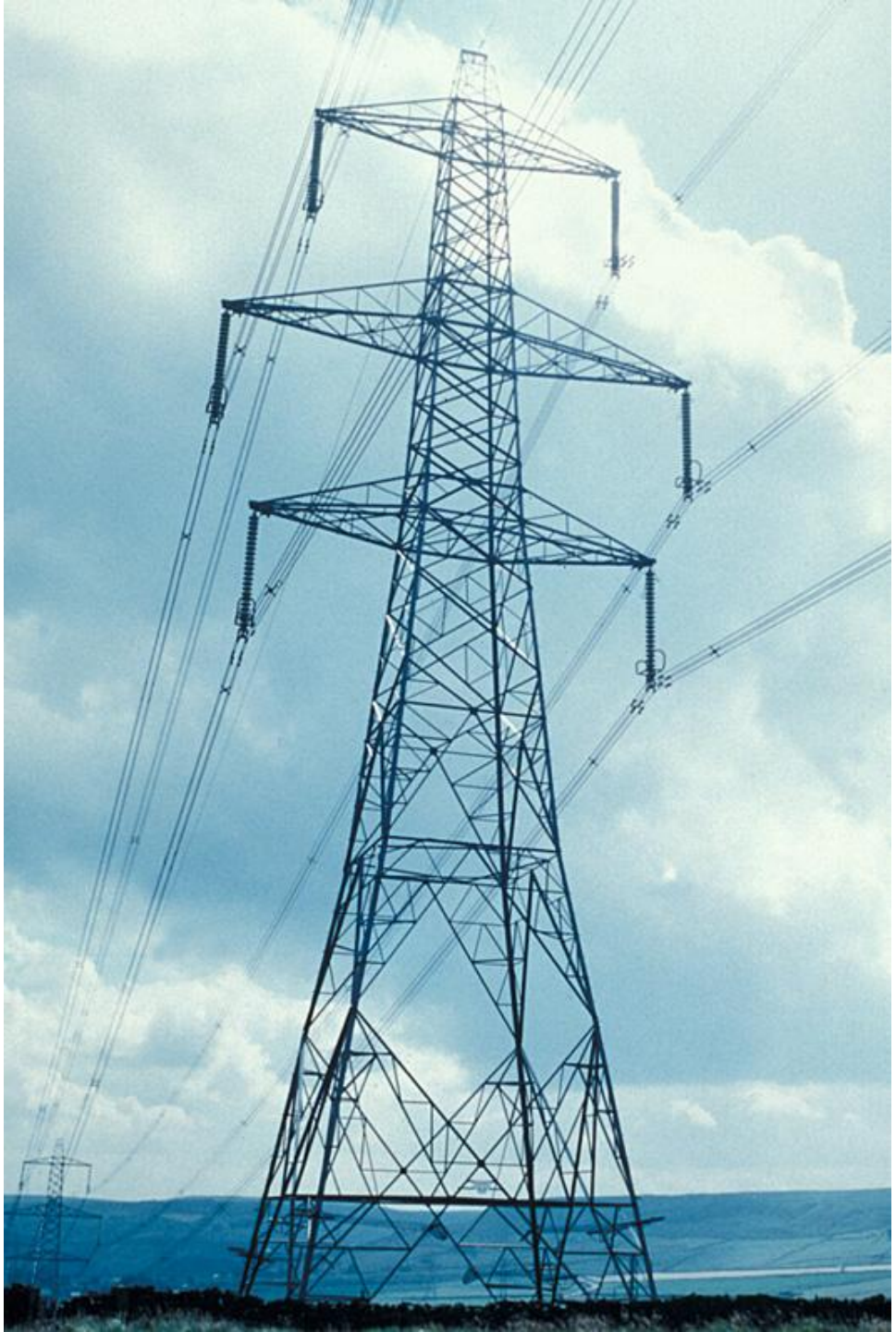
- Does ‘X’ cause the excess risk?


raphy Socio-economic Demography Environment Disease Mapping Labels
Scale 1: 155,966
429,697
450,945
azard Analysis
Pollution Source
Buffer Rings
Weighed Centroids
Buffer ID | Buffer D
1 | 2.5
2 | 2.5 - 7
| No Data

Landfill



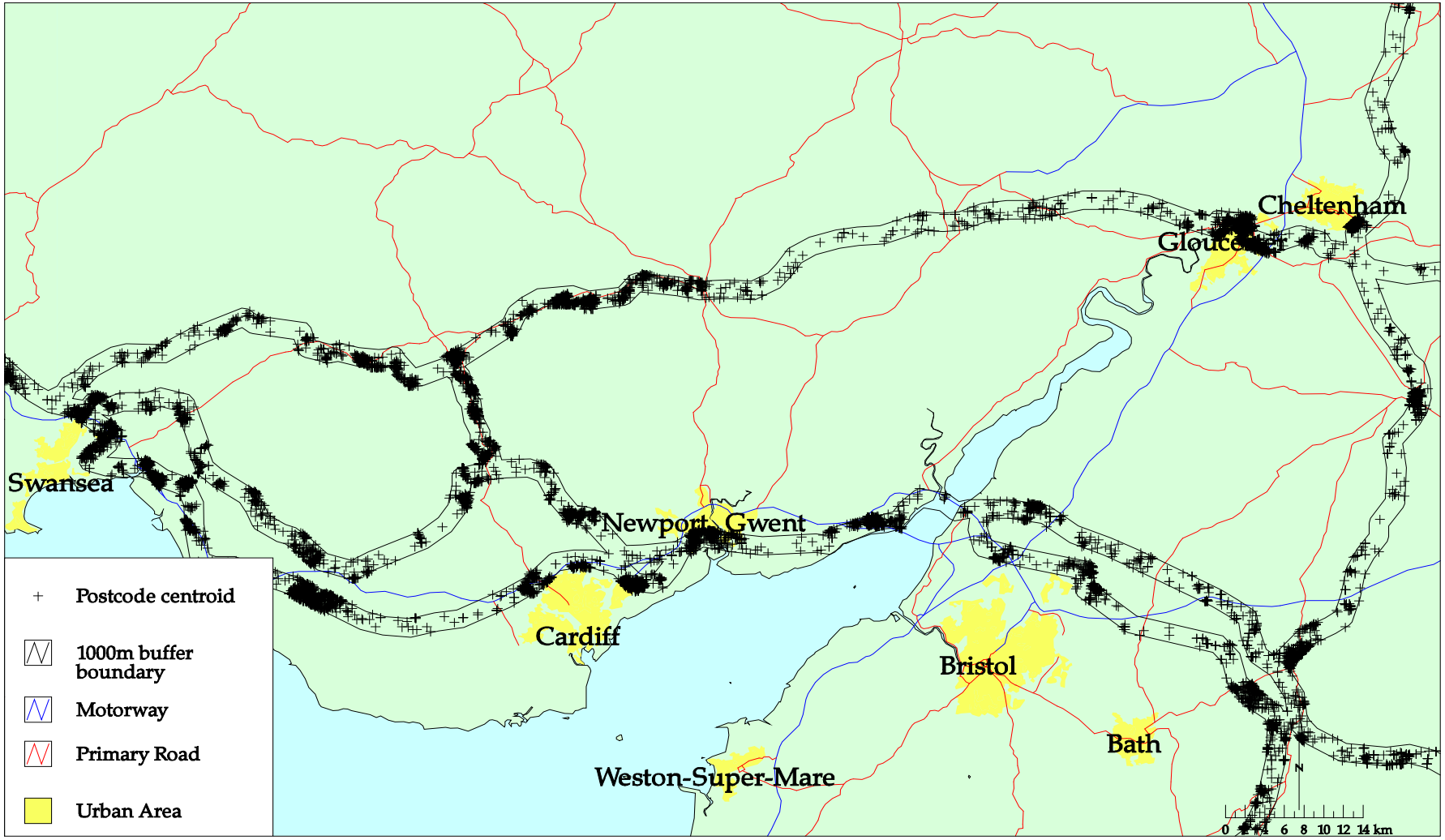
Risk of adverse birth outcomes in populations living near landfill sites
Elliott, Briggs, Morris et al *BMJ 2001*



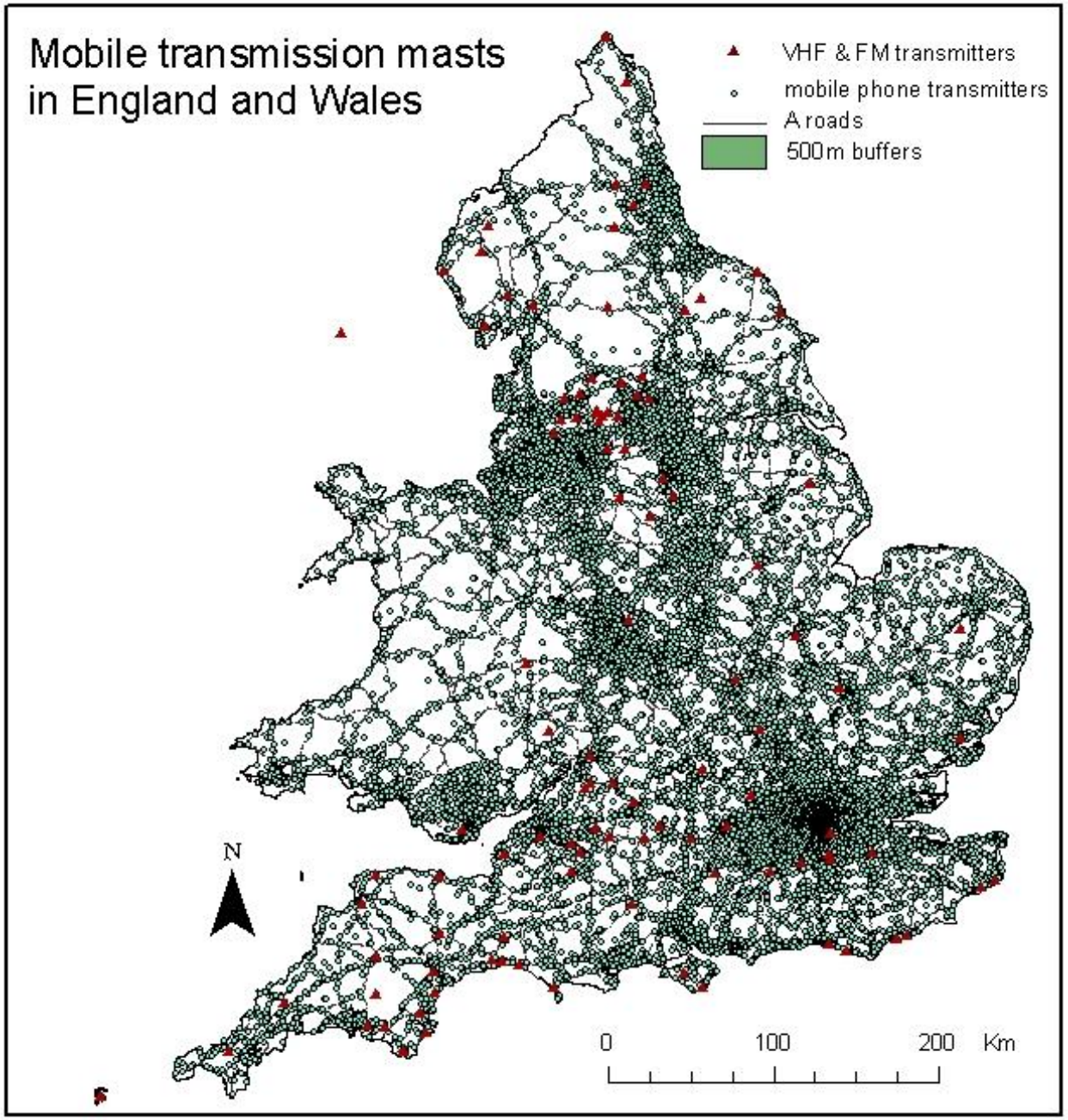
Non-ionising radiation

- radio/TV transmitter
- mobile masts
- Power lines

Adult cancers in proximity to overhead power lines (Elliott, Toledano et al)



Thanks to National Grid

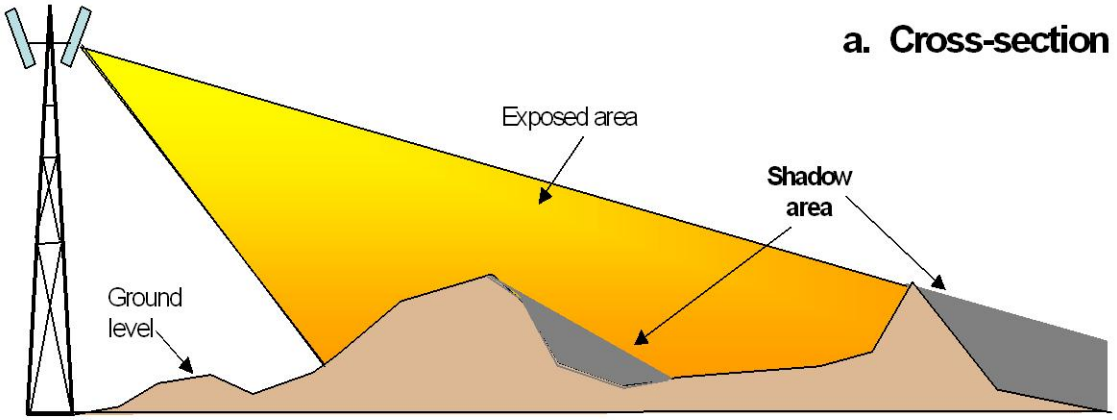


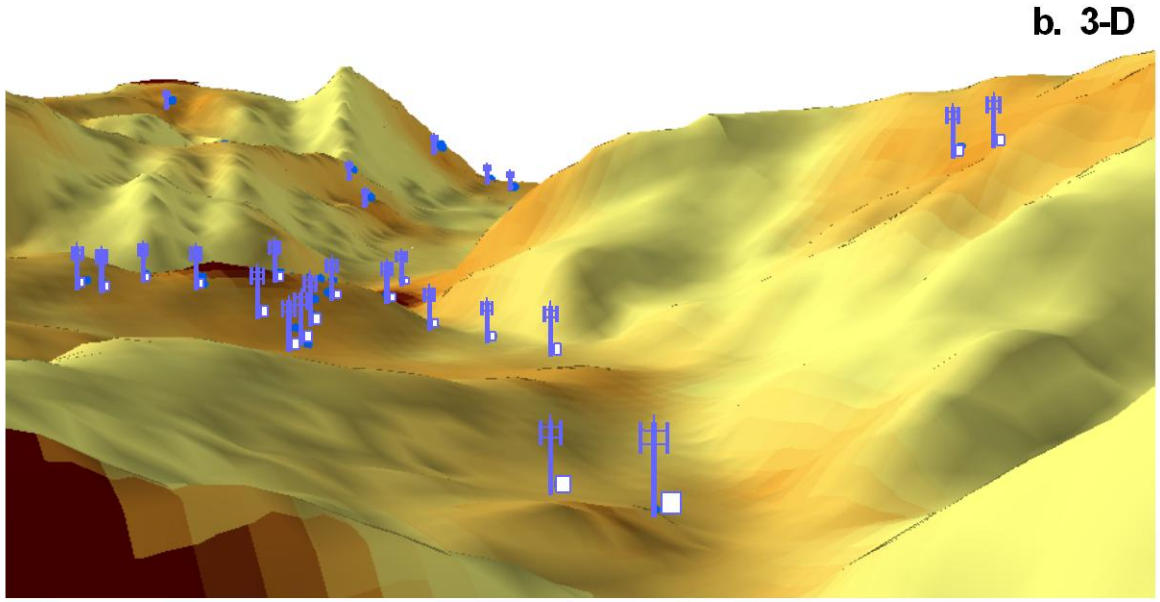
Mobile phone base stations and early childhood cancers

Elliott, Toledano, Bennett et al

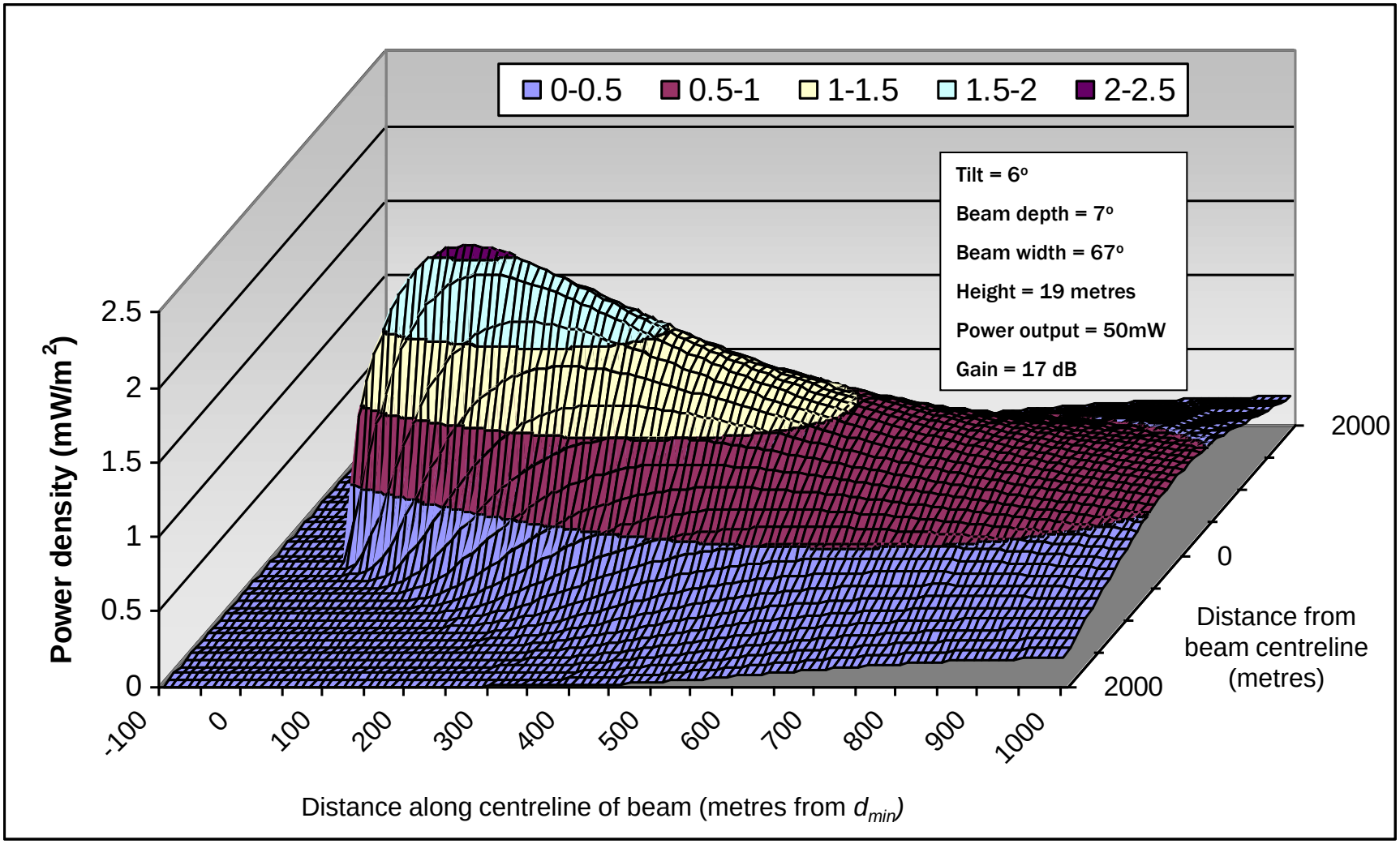
BMJ 2010

Gaussian Model





Gaussian Model



GEM3 simulates 3D distribution of power density around the focus of the beam using Gaussian functions.

Spatial epidemiological studies: some issues in interpretation

1. **Quality of routine health and population data**

   - **Data problems/errors (diagnostic accuracy and coding misclassification & case ascertainment)**
   - **Boundary changes over time**
   - **Latency periods - migration**

Spatial epidemiological studies: some issues in interpretation

2. **Environmental exposure data**
   - Level of analysis - individual vs ecological (ecological bias)
   - Proxy measures often used e.g. concentric circles, distance
   - Exposure model validation - environmental monitoring expensive & may only give partial picture of true individual exposure
3. **Potential confounding variables**

- **Interpret with caution!**

Spatial epidemiology: summary

- **Data/exposure model often fall short of ideal**

- **Near-ubiquitous low-level exposures, small elevations of risk, data problems ⟶ limit ability to detect causal relationships & differentiate them from effects of bias, confounding etc**

Spatial epidemiology: summary

- **Nevertheless, these types of studies serve a valuable public health function:**
  - **it is often not feasible to investigate their questions on a national scale at the individual level**
  - **they therefore help direct scarce resources towards problems/areas of greatest need**