Small Area Analyses

Graham Moon
Institute for the Geography of Health
University of Portsmouth

University of Portsmouth

Aim

- To provide an oversight of the outcomes of a programme of integrated projects focussed on the derivation of small area indicators of health need

University of Portsmouth

Small Area Indicators

- People live in areas
- Agencies have area-based responsibilities
- BUT
- Home and work
- Changing boundaries
- Non-bounded processes
- No routine data available

University of Portsmouth

Method: Synthetic Estimation

- Method to estimate unknown area-level measures based on other known variables
- Unknown measures: health themes
- Known variables: census data, marketing data, sample data, administrative data…typically large national surveys

University of Portsmouth

Method: Synthetic Estimation

- Why needed?
  - No alternative
  - Survey sample sizes too small for direct estimates
  - Survey samples clustered so no data for large % of areas

Method: Synthetic Estimation

- **Crosstabulation approach**
  - Step 1: derive crosstabulation from national survey data
  - Step 2: apply cell percents from step 1 to small area crosstabulations of (usually) census data
- Limitations
  - limited number of limited cross-tabulations available
  - Erases geography
- Advantages:
  - computationally simple
  - transparent

University of Portsmouth

Method: Synthetic Estimation

- Modelling approach
  - Step 1: model relationship between outcome and predictor variables in a survey
  - Step 2: apply coefficients from step 1 to known small area counts of predictors from (usually) census data
- Individual predictors
  - computationally equivalent to the crosstabulation approach
  - same problems
- Area predictors
  - aggregate survey results to lower scale
  - same problems

University of Portsmouth

Method: Synthetic Estimation

- Multilevel approach
  - Both individual and area level predictors
  - Area predictors in a multilevel framework
  - Apply coefficients to known small area counts of predictors from (usually) census data
- Advantages:
  - Both individual and area-level factors affect health outcomes
  - Hierarchical nature of survey data lends itself to a multilevel model
- Disadvantages:
  - complexity

University of Portsmouth

Method: Synthetic Estimation

- Or... just use a proxy indicator of poor-health, such as a deprivation index
- Choice of model: f(complexity, data, objectives)
- Note: ***estimates*** : 'risk' measures, 'probable prevalence given...'; ***synthetic*** : educated best guess

University of Portsmouth

Two examples

- Environmental health outcome
  - Exposure to ETS
- Environmental health predictor
  - Air pollution and chronic respiratory disease

Exposure to ETS

- NHS Health Scotland
- Scottish Health Survey
- Sample: 16,915
- Non-smokers aware of being exposed to environmental tobacco smoke
  - Covers range of settings
- Objective: small area estimates

University of Portsmouth

Exposure to ETS

$$
\left.\begin{aligned}
&\text{nsetsaware}_{ijk} \sim \text{Binomial}(\text{denom}_{ijk}, \pi_{ijk}) \\
&\text{nsetsaware}_{ijk} = \pi_{ijk} + e_{0ijk}\text{bcons}^{*}_{ijk}
\end{aligned}\right\}
$$

$$
\begin{aligned}
\text{logit}(\pi_{ijk}) = {} & \beta_{1jk}\text{cons} + 0.23172(0.04452)\text{male}_{ijk} + -0.23673(0.04601)\text{single}_{ijk} + 0.11505(0.17051)\text{a1}_{ijk} + -0.14914(0.04810)\text{a3}_{ijk} + \\
& -0.28453(0.05057)\text{a4}_{ijk} + -0.37483(0.05118)\text{a5}_{ijk} + -0.63531(0.06766)\text{a6}_{ijk} + -0.01644(0.06536)\text{sm}_{ijk} + \\
& 0.35047(0.17697)\text{sa1}_{ijk} + -0.00481(0.00406)\text{rm6cr}_{ijk} + 0.00156(0.00511)\text{s.rm6cr}_{ijk} + -0.01048(0.00486)\text{m.rm6cr}_{ijk} + \\
& 0.01591(0.00729)\text{sm.rm6cr}_{ijk} + -0.01766(0.00542)\text{a6.rm6cr}_{ijk} + -0.00490(0.00591)\text{tenprcr}_{ijk} + \\
& -0.04803(0.00956)\text{a1.tenprcr}_{ijk} + -0.00416(0.00218)\text{tenlahacr}_{ijk} + -0.00707(0.00303)\text{a1.tenlahacr}_{ijk} + \\
& -0.00418(0.00254)\text{sc3m45cr}_{ijk} + 0.00741(0.00328)\text{a5.sc3m45cr}_{ijk} + -0.04615(0.01030)\text{scnotcr}_{ijk} + \\
& 0.04333(0.01304)\text{s.scnotcr}_{ijk} + 0.03958(0.01307)\text{m.scnotcr}_{ijk} + -0.04399(0.01901)\text{sm.scnotcr}_{ijk} + \\
& 0.02948(0.00911)\text{a3.scnotcr}_{ijk} + 0.02258(0.00887)\text{munempcr}_{ijk} + -0.03360(0.01034)\text{s.munempcr}_{ijk} + \\
& -0.03872(0.01062)\text{m.munempcr}_{ijk} + 0.04339(0.01509)\text{sm.munempcr}_{ijk} + -0.01201(0.00528)\text{a34.munempcr}_{ijk} + \\
& 0.01071(0.00447)\text{kidscr}_{ijk} + -0.01851(0.00488)\text{s.kidscr}_{ijk}
\end{aligned}
$$

$$
\beta_{1jk} = -0.17209(0.05059) + v_{1k} + u_{1jk}
$$

Prevalence decreases with age (the youngest married males are an exception). Since prevalence of Current Smoking does *not* increase overall with age, this decrease is largely a measure awareness or exposure, not of a decreasing pool of non-smokers.

University of Portsmouth

Exposure to ETS

$$\left[v_{1k}\right] \sim N(0,\ \Omega_v) \ :\ \Omega_v = \left[0.01416(0.00711)\right]$$

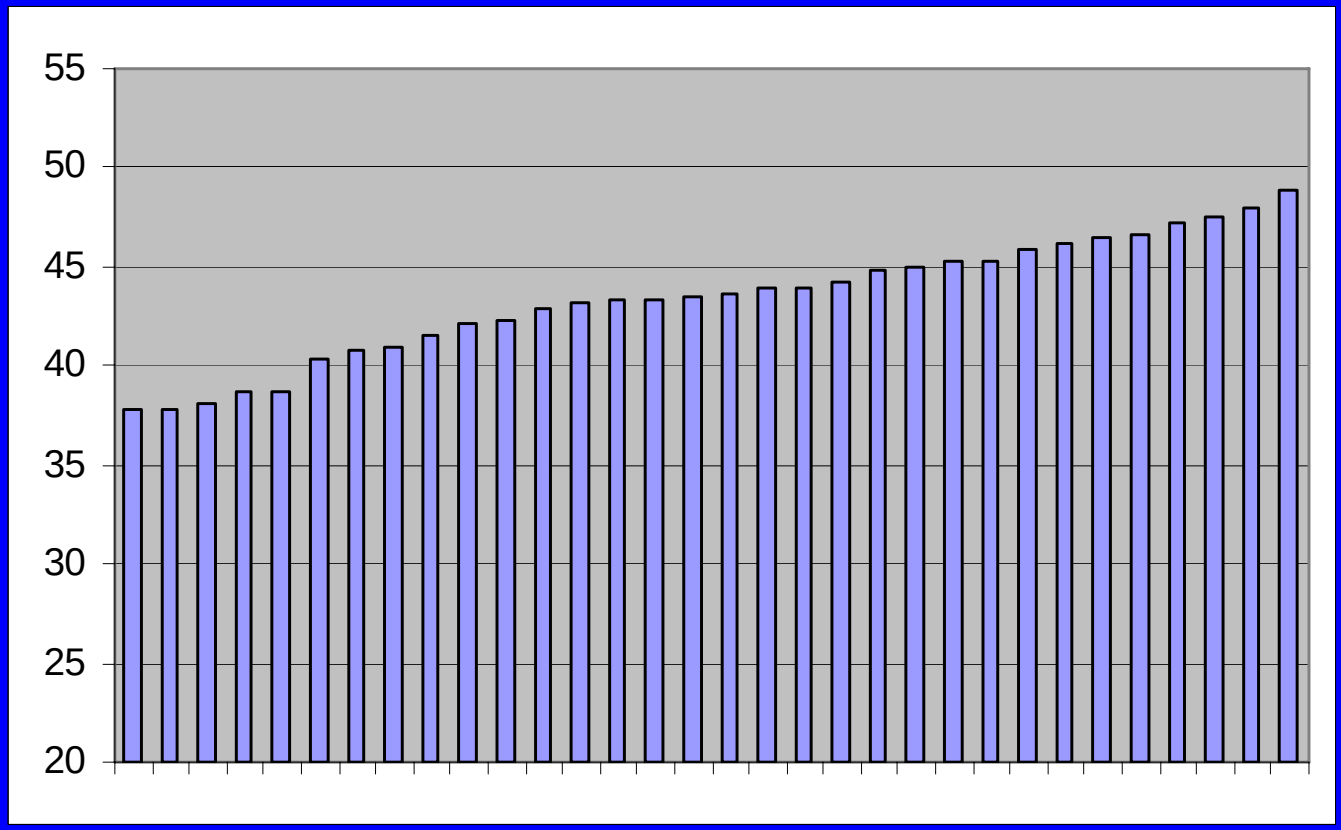
$$\left[u_{1jk}\right] \sim N(0,\ \Omega_u) \ :\ \Omega_u = \left[0.05390(0.01162)\right]$$

$$\text{bcons}^{*}_{ijk} = \text{bcons}_{ijk}\left[\pi_{ijk}(1-\pi_{ijk})/\text{denom}_{ijk}\right]^{0.5}$$

$$\left[e_{0ijk}\right] \sim (0,\ \Omega_e) \ :\ \Omega_e = \left[1.00000(0.00000)\right]$$

University of Portsmouth

Exposure to ETS



LA Scale

University of Portsmouth

Air Pollution and CRD

- Multilevel Synthetic Estimation
- Consolidated 1998-2004 database
- 95,000 respondents aged 20+
  - HSfE, SHS, NIH&SCS
  - census
  - CRD as a function of (1) age,sex, ethnicity; (2) socioeconomics, PM10

University of Portsmouth

Air Pollution and CRD

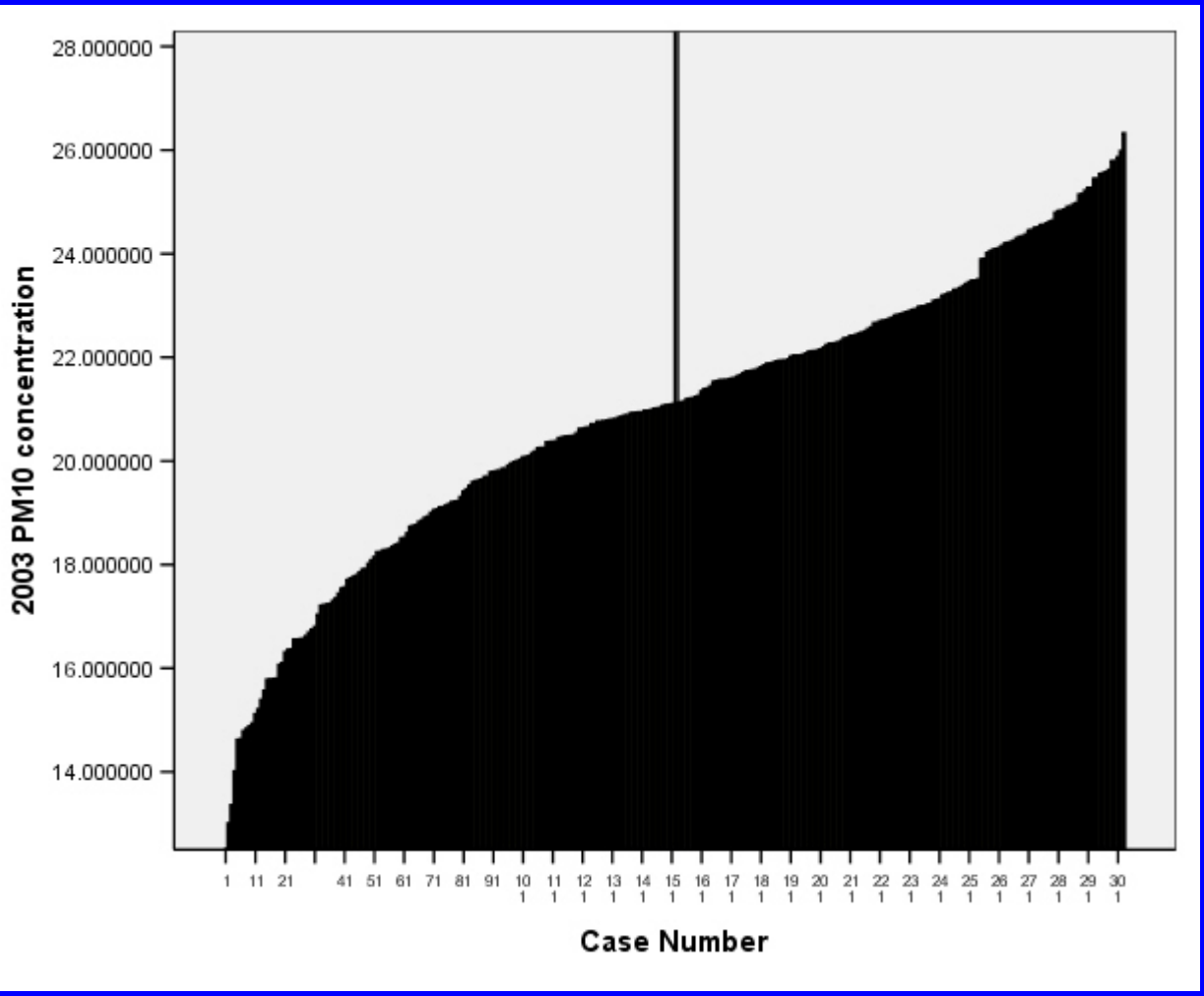
- PM10 data
  - modelled an overall PM10 profile across scattered empirical observation sites.
  - Estimated PM10 concentrations for 1km grid derived from modelled profile
  - These were aggregated to Lower Special Output Areas (in England)

University of Portsmouth

Air Pollution and CRD

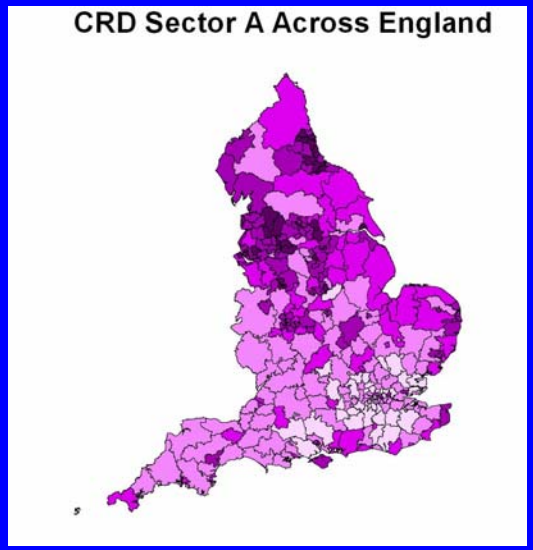
- PM10 data
  - We can only work with PCO data
  - Aggregate LSOA to PCO
  - 107 LSOAs/PCO in England
  - Population versus area weighting
  - Chose population weighting

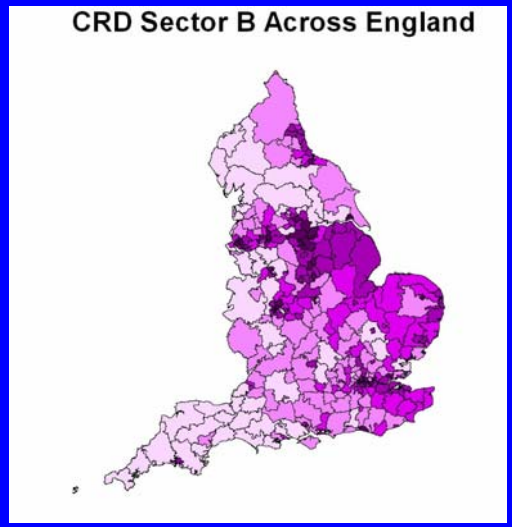
Air Pollution and CRD

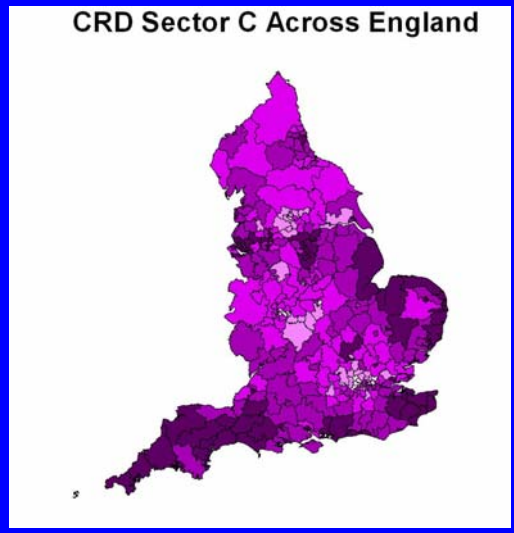


University of Portsmouth

Chronic Respiratory Disease







→A: diagnosed emphysema, asthma or bronchitis

→B: Symptoms: undiagnosed but smokes and wheezes or smokes and poor lung function

→C: Signs: Undiagnosed but smokes

University of Portsmouth

Conclusions

- Major utility in
  - health profiling
  - care planning
- Clear advantages in absence of other data

University of Portsmouth