Demographic Research a free, expedited, online journal
of peer-reviewed research and commentary
in the population sciences published by the
Max Planck Institute for Demographic Research
Konrad-Zuse Str. 1, D-18057 Rostock · GERMANY
www.demographic-research.org

DEMOGRAPHIC RESEARCH

VOLUME 14, ARTICLE 9, PAGES 157-178
PUBLISHED 07 MARCH 2006
http://www.demographic-research.org/Volumes/Vol14/9/
DOI: 10.4054/DemRes.2006.14.9

Research Article

A model for geographical variation in health and total life expectancy

Peter Congdon

Table of Contents

Demographic Research – Volume 14, Article 9
a research article

A model for geographical variation in health and total life expectancy

Peter Congdon[1]

Abstract

This paper develops a joint approach to life and health expectancy based on 2001 UK Census data for limiting long term illness and general health status, and on registered death occurrences in 2001. The model takes account of the interdependence of different outcomes (e.g. ill health and mortality) as well as spatial correlation in their patterns. A particular focus is on the proportionality assumption or 'multiplicative model' whereby separate age and area effects multiply to produce age-area mortality rates. Alternative non-proportional models are developed and shown to be more parsimonious as well as more appropriate to actual area-age interdependence. The application involves mortality and health status in the 33 London Boroughs.

[1]Department of Geography, Queen Mary London University, Mile End Rd, London E1 4NS. E-mail: p.congdon@qmul.ac.uk

1. Introduction

Health expectancy is increasingly emphasised as an indicator for population health that takes account of both mortality and morbidity or disability. While morbidity and disability data are often only obtainable from surveys, the recent UK 2001 Census includes questions on both limiting long term illness and general health status. Thus, in England and Wales 63% of adults (aged 16 and over) said they had good health, 26% reported they had fairly good health and 11% said their health was not good. A variety of measures of health expectancy are available that may be based on limited function or self-perceived health status; these include disability free life expectancy and healthy life expectancy (Bebbington et al, 1993; Robine and Ritchie, 1991).

While both total life expectancy and health expectancy have improved in the UK, there are wide variations between geographic areas and socio-economic groups. Analyses of such contrasts, especially of spatial variations, have typically used standard life table calculations. These do not take account of features such as interdependence of different outcomes (e.g. ill health and mortality), or of spatial correlation in their patterns, or of sampling variations in deaths or other outcomes. Where statistical modelling techniques are adopted, simplifying assumptions about the impacts of demographic variables and area are often made; for example, the proportionality assumption or 'multiplicative model' (Hoem, 1987) whereby separate age and area effects multiply to produce age-area mortality rates.

The present paper considers how evidence from mortality, limiting illness and self-rated health for sets of areas may be integrated in life tables for sets of contiguous areas. It includes consideration of the validity of the multiplicative model, and considers how interactions between age and area effects may be parsimoniously modelled. The application involves the 33 London Boroughs (Figure 1) and combines information from the two 2001 Census questions on disabling illness and self assessed health with recorded deaths in 2001 for the same areas. The result is a joint life table model for life and health expectancies by area.

2. The proportionality assumption (multiplicative model)

Let populations in area $a(a = 1, \ldots, A)$, and age band $x(x = 1, \ldots, X)$ be denoted P_{ax}. Then deaths D_{ax} by area and age band will be binomial

$$D_{ax}|\mu_{ax} \sim \text{Bin}(P_{ax}, \mu_{ax})$$

In line with many spatial epidemiology studies (e.g. Wakefield et al, 2000), the proportionality assumption is that

$$\mu_{ax} = \rho_a r_x \tag{1}$$

Figure 1: The London boroughs

where ρ_a are unknown relative risks for area a and r_x are death rates at age x. For rare outcomes, the binomial distribution may be approximated by a Poisson distribution for the D_{ax} (e.g. Sun et al, 2000, p. 2108)

$$D_{ax} \sim \text{Poi}(P_{ax}\mu_{ax})$$

where μ_{ax} is often taken as proportional as in (1). A relevant model (e.g. with log link) for μ_{ax} would then take $\log(\rho_a)$ and $\log(r_x)$ as independent effects. Alternatively under the proportionality assumption one may collapse over the age groups to obtain a model where the area death totals $D_a = \sum_x D_{ax}$ are Poisson with means $E_a\rho_a$ where $E_a = \sum_x P_{ax}r_x$ are expected deaths. If an internal demographic standardisation is used then $\sum_a D_a = \sum_a E_a$ and so the ρ_a will have average 1, and posterior densities for ρ_a concentrated on values over 1 (e.g. with 95% credible interval all above 1) then indicate excess relative risk in area a.

In models with more classifiers (e.g. time as well as age and area) a common assump-

tion is that age effects are independent of area (e.g. McNab and Dean, 2001), though changes over time in the age profile of mortality may be included (Sun et al, 2000).

The present analysis uses age and area classifiers only and considers either total deaths (males and females combined) or deaths for one sex only. Extensions to include more classifiers (e.g. time) or to bivariate life table analysis (male and female life tables in one overall model) are, however, possible. Life and health expectancy may be jointly modelled for a set of areas using data on health status and long term illness as well as mortality data. An initial analysis using the proportionality assumption for these outcomes is contrasted in terms of fit and substantive implications with an analysis allowing for age-area interactions. The age-area interaction model draws on the principles in the Carter and Lee (1992) model for age-time interactions in mortality, and the related log-linear model of Goodman (1979). More heavily parameterised models that use random effects for each age-area interaction are also considered.

3. A model based on proportionality

The relevant data are deaths D_{ax} for the year 2001, numbers of long term ill in area a at age x, G_{ax}, and the numbers H_{axj} in area a and age x in the $j = 1, \ldots, 3$ categories of the general health (good, fairly good, not good). There are $a = 1, \ldots, 33$ areas and $x = 1, \ldots, 19$ age bands (namely $0-4, 5-9, \ldots, 85-89$, over 90).

Let s_a denote spatially correlated area effects, u_a be random errors without any spatial structure, and δ_x denote age effects. To reflect correlated outcomes one may include a common spatial effect across the responses, since it is plausible that a common structure between excess mortality and morbidity exists and that it follows a spatial structure. Then coefficients θ_j may be introduced to express the differential impact of s_a on each outcome j. Hence the s_a can be seen as a spatially correlated factor scores, and $\theta = (\theta_1, \theta_2, ..)$ as factor loadings, that account for correlations between the outcomes. If $var(s_a)$ is taken as a free parameter then for identifiability one θ coefficient is assigned a set value (e.g. $\theta_1 = 1$), while if $var(s_a)$ is set, e.g. $var(s_a) = 1$, all θ coefficients may be free.

Several models are possible for age effects. Sun et al (2000) treat them as fixed effects; McNab and Dean (2001) and Nandram et al (1999) use spline models; Ibrahim et al (2001) suggest random walks, while demographic applications (Anson, 1991) may use polynomials in age. Here the first two models use a random effects approach combining a structured random walk prior with unstructured age effects. Thus with j denoting mortality/morbidity responses $(j = 1, ..K)$

$$\delta_{jx} = v_{jx} + w_{jx}$$

where the v_{jx} are structured (follow a state-space form) and the w_{jx} are unstructured

effects with $w_{jx} \sim N(0, \varphi_j)$. To reflect the correlation among the outcomes j it is assumed that rather than separate state space models for each of the K series of effects v_{jx}, the processes are interlinked according to

$$v_{jx} = \phi_j V_x$$

where the ϕ_j are loadings on a shared structured age effect

$$V_x \sim N(V_{x-1}, \xi).$$

One possible model (model 1) for deaths and long term illness totals based on age-area proportionality is then

$$\begin{aligned} D_{ax} &\sim \text{Po}(P_{ax}\mu_{ax}) \\ \log(\mu_{ax}) &= \alpha_1 + \phi_1 V_x + w_{1x} + \theta_1 s_a + u_{1a} \qquad \text{(1a)} \\ G_{ax} &\sim \text{Bin}(P_{ax}, \lambda_{ax}) \\ \text{logit}(\lambda_{ax}) &= \alpha_2 + \phi_2 V_x + w_{2x} + \theta_2 s_a + u_{2a} \qquad \text{(1b)} \end{aligned}$$

where α_j are intercepts. For numbers H_{axk} in the health status groups, a cumulative logit model involving

$$\upsilon_{axj} = \pi_{ax1} + \cdots + \pi_{axj} = \Pr(H_{axk} \leq j), j = 1, J-1$$

($J = 3$ here) is often assumed. Other links allowing for asymmetric departures from the cumulative logit might also be considered such as the cumulative log-log link or links involving a transformation parameter (Zayeri et al, 2005; Agresti, 2002). A proportional cumulative logit model would require common age gradients and area effects across j. In the current application considerable gains in fit were made if age gradients and area effects were allowed to differ between levels of health status, leading to a non-proportional model (Peterson and Harrell, 1990). While $J-1$ non-parallel regression lines may cross when explanatory variables are continuous, this problem does not occur for explanatory variables that are categorical, as here (Gibbons and Hedeker, 2000). The cumulative logit model is then

$$\text{logit}(\upsilon_{ax1}) = \kappa_1 - (\phi_3 V_x + w_{3x} + \theta_3 s_a + u_{3a}) \qquad \text{(1c)}$$

$$\text{logit}(\upsilon_{ax2}) = \kappa_2 - (\phi_4 V_x + w_{4x} + \theta_4 s_a + u_{4a}) \qquad \text{(1d)}$$

where the prior on the cutpoints κ_j has an order constraint. The form of the regression in (1c) and (1d) means that δ_{3x}, $\theta_3 s_a$, and u_{3a} will rise in line with increases in sub-optimal health (fair or not good) while δ_{4x}, $\theta_4 s_a$, and u_{4a} will be positive measures of poor health.

The probabilities π_{axj} of the health status distribution ($H_{ax1}, H_{ax2}, H_{ax3}$) in area a and age x are obtained as

$$\begin{aligned}\pi_{ax1} &= \upsilon_{ax1} \\ \pi_{ax2} &= \upsilon_{ax2} - \upsilon_{ax1} \\ \pi_{ax3} &= 1 - \upsilon_{ax2}\end{aligned}$$

The ICAR prior of Besag et al (1991) is used for the shared spatial effects s_a. Define the $A \times A$ contiguity matrix C with elements $c_{ab} = c_{ba} = 1$ if areas a and b are adjacent and zero otherwise, let L_a be the neighbourhood of areas adjacent to a (excluding area a itself) and let N_a be the number of areas in the neighbourhood. Then the Normal version of the ICAR prior (with variance τ) assumes

$$f(s_a|s_{[-a]}) = \left(\frac{N_a}{2\pi\tau}\right)^{0.5} \exp\{-0.5\frac{N_a}{\tau}(s_a - S_a)^2\}$$

where $s_{[-a]}$ denotes all $\{s_1, s_2, ..s_A\}$ except s_a, and S_a is the average of s_b for the areas b in the locality L_a of area a. Equivalently

$$s_a|s_{[-a]} \sim N(\sum_b c_{ab}s_b, \tau/N_a)$$

To ensure identification the s_a are recentred at each iteration to have mean zero. The u_{ja} are taken to be unstructured Normal random effects with mean zero.

Note that a close fit to the data may be attained by effectively modelling each observation, namely adding random age-area effects $\{e_{1ax}, e_{2ax}, e_{3ax}, e_{4ax}\}$ in (1a)-(1d). However, this approach is heavily parameterised, and leads to complex interpretation issues of model results in substantive terms. Instead the goal is relatively parsimonious and interpretable models that clearly improve fit as an alternative to introducing age-area interaction effects. This objective is pursued in subsequent model elaboration.

4. Estimation

The estimation of the above model, namely model 1 as set out in (1a)-(1d), for the London borough data for males was based on two parallel chains of $10,000$ iterations with dispersed starting values based on a pilot run. Convergence from $5,000$ iterations was obtained under Gelman-Rubin criteria (Gelman et al, 1995). In this and subsequent models $N(0, 1000)$ priors are used for fixed effects and Gamma priors with index 1 and scale 0.001 are used for precisions.

To assess model fit, one criterion used is the DIC of Spiegelhalter et al (2002), under which the number of effective parameters p_e is derived as the difference between the averaged sampled deviance $\overline{\text{Dev}}$ and the deviance at $\overline{\Phi}$, the posterior mean of the full parameter set Φ. The DIC is then the average deviance plus the effective parameter total (see Table 1 for fit statistics for model 1 and subsequent models). Another is the pseudo marginal likelihood based on the Monte Carlo estimate of the conditional predictive ordinate, as proposed by Gelfand and Dey (1994).

Table 1: Model Criteria

		Deaths	Long Term Ill	Health Status	Total
Model 1	Effective Parameters	48	45	72	165
	DIC	1084	5150	7950	14184
	BIC	1297	5350	8319	14916
	Pseudo Marginal Likelihood	-2031	-5249	-9178	-16458
Model 2	Effective Parameters	63	76	154	293
	DIC	902	2651	4611	8164
	BIC	1181	2990	5404	9467
	Pseudo Marginal Likelihood	-1939	-3964	-7464	-13367
Model 3	Effective Parameters	61	100	194	355
	DIC	904	2278	3987	7169
	BIC	1175	2722	4983	8746
	Pseudo Marginal Likelihood	-1928	-3744	-7114	-12786
Model 4	Effective Parameters	159	467	905	1531
	DIC	868	1154	2291	4313
	BIC	1574	3228	6937	11112
	Pseudo Marginal Likelihood	-1920	-3126	-6181	-11227

The estimated age parameters from model 1 show a typical mortality 'bathtub' profile for males (Figure 2), with an accident hump in the late 20s and a virtually linear ascent after age 35 in the log death rate. Figure 3 contrasts the parameters δ_{2x}, δ_{3x} and δ_{4x}, namely the age effects for long term limiting illness/disability, for fair or poor health combined, and for poor health only. The variation in s_a (Figure 4) closely reproduces dimensions of mortality and ill health based on socio-economic structure and inner vs. outer city contrasts. The highest values are in inner east London in deprived boroughs such as Hackney and Tower Hamlets whereas the lowest are in affluent suburban boroughs in south west London (Kingston, Richmond). The posterior means of s_a have a correlation of 0.92 with area deprivation scores developed by Noble et al (2000).

Figure 2: Mortality Effects by Age

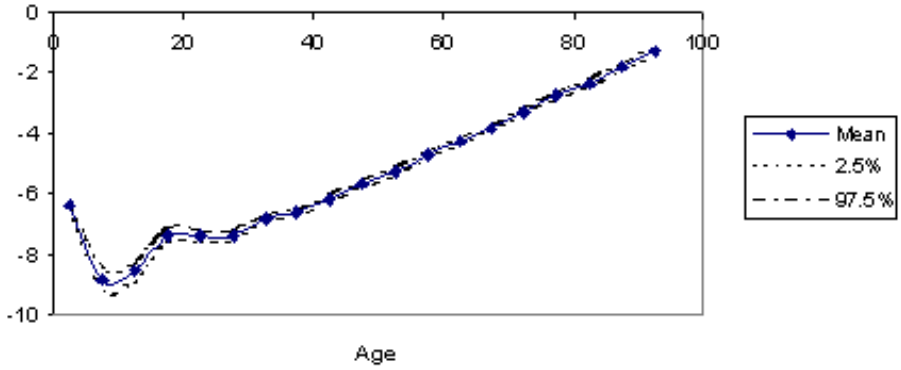


Figure 3: Age Effects for Health and Limiting Illness

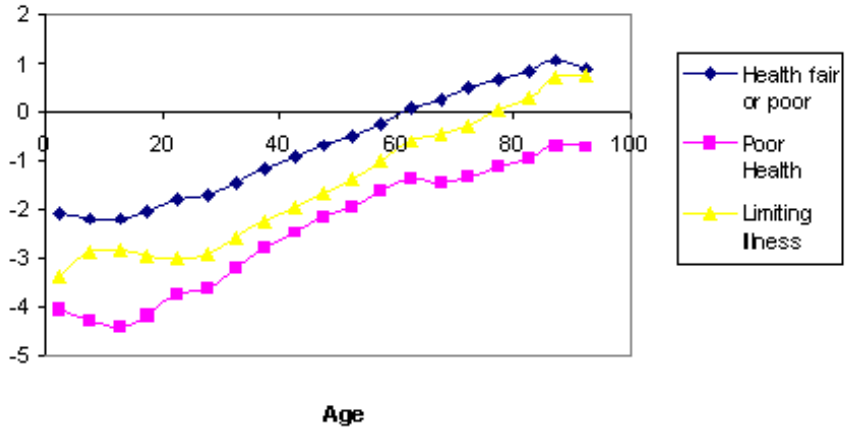
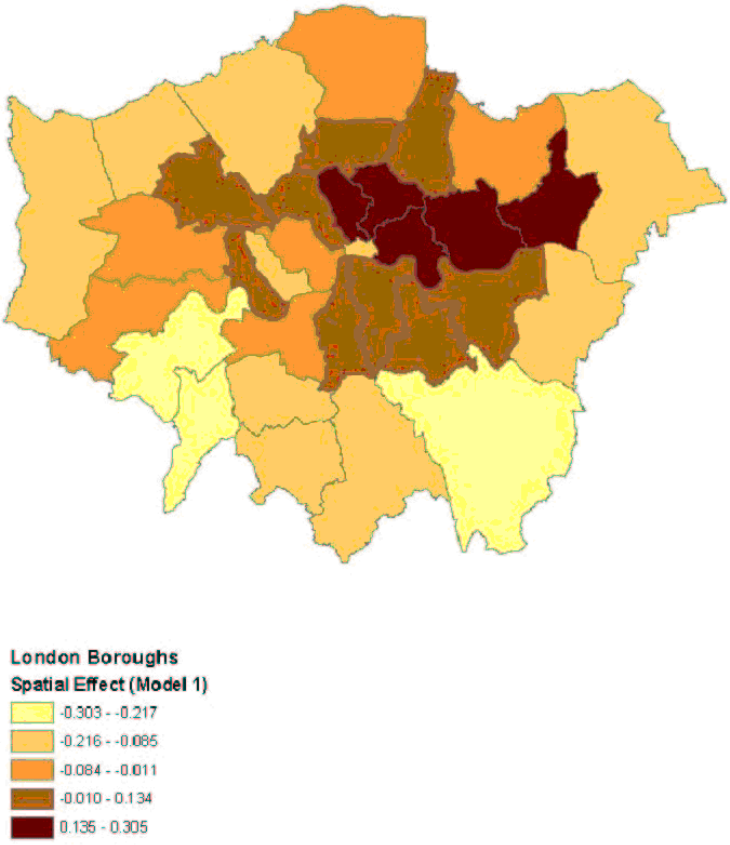


Figure 4: Spatial Effect by Borough (Model 1)



5. Modelling non-proportional age and area effects

To allow for non-proportional impacts means age effects and area effects will interact. As noted above, the most heavily parameterised models allowing this are random effects models with terms e_{jax} specific to area, age and outcome. These would usually be assumed spatially unstructured, though for a small number of age groups x one might assume $e_{ja1}, e_{ja2}, \ldots$, etc to have distinct spatially correlated densities.

Alternatively for a relatively parsimonious model with substantive interpretability one may adapt the Carter-Lee model for forecasting mortality (Carter and Lee, 1992) to the present spatial application. The Carter-Lee model for mortality rates in age and time μ_{tx} (without an area dimension) takes the form

$$\log(\mu_{tx}) = \alpha_1 + \delta_x + \beta_x \kappa_t$$

with constraints on the multiplicative function $\beta_x \kappa_t$ to ensure identifiability. Lee (2000) assumes the β_x to be positive and sum to 1 over all x, and constrains the κ_t to sum to zero. The β_x parameters express variations between ages in the adherence to the overall mortality trend represented by the κ_t parameters. If the κ_t were declining as mortality fell then larger β_x indicate for which age groups the rates are declining more rapidly.

In the present spatial mortality application one may incorporate this form of non-proportionality (leading to model 2). This involves first re-defining the mortality model as

$$\begin{aligned} D_{ax} &\sim \text{Po}(P_{ax}\mu_{ax}) \\ \log(\mu_{ax}) &= \alpha_1 + \delta_{1x} + \beta_{1x}\gamma_a + u_{1a} \end{aligned} \tag{2a}$$

where the β_{1x} are assumed positive and sum to 1 and the γ_a are centred to sum to zero. The mixed random effects model for the age effects δ_{jx} used in model (1) is retained in model 2. The remaining components of the model are redefined as

$$\text{logit}(\lambda_{ax}) = \alpha_2 + \delta_{2x} + \beta_{2x}\gamma_a + u_{2a} \tag{2b}$$

$$\text{logit}(\gamma_{ax1}) = \kappa_1 - (\delta_{3x} + \beta_{3x}\gamma_a + u_{3a}) \tag{2c}$$

$$\text{logit}(\gamma_{ax2}) = \kappa_2 - (\delta_{4x} + \beta_{4x}\gamma_a + u_{4a}) \tag{2d}$$

The β_{1x} in (2a) represent differences between age groups in adherence to the spatial mortality regime defined by the γ_a. For example, if the γ_a are higher in deprived areas, then β_{1x} would peak at ages where deprivation has most impact on mortality. Mortality in childhood and at middle ages is most enhanced in deprived areas, while area contrasts are less pronounced at older ages (Eames et al, 1993), so the β_{1x} would be highest at

childhood and middle age bands, but low at old ages. Disability or poor health in middle age also tends to be elevated in deprived areas.

Extensions of the basic non-proportional model with generic form

$$\log(\mu_{ax}) = \alpha + \delta_x + u_a + \beta_x \gamma_a$$

may be envisaged. For example, it may be that there are discordant spatial effects or that the interaction between age and spatial effects is less clearly defined in some areas than others. The generic model reduces to the proportional model

$$\log(\mu_{ax}) = \alpha + \delta_x + u_a + \gamma_a$$

when all the β_x are equal, so one might propose a two group discrete mixture whereby in one group the β_x vary less than in another group. Thus

$$\log(\mu_{ax}) = \alpha + \delta_x + u_a + \beta_{xG_a} \gamma_a$$

where $G_a \in (1, 2)$. One possible prior for the β_x involves a multiple logit link, namely

$$\beta_x = \exp(a_x) / \left[1 + \sum_{x=1}^{X-1} \exp(a_x)\right] \qquad x = 1, \ldots, X-1$$

$$\beta_X = 1 / \left[1 + \sum_{x=1}^{X-1} \exp(a_x)\right]$$

where a_x are random effects, e.g. $a_x \sim N(0, \tau_a)$. So the discrete mixture would involve constraining τ_a to be lower in one group than the other. Another possible model allowing for spatial outliers would mix over a normal ICAR spatial effect γ_{1a} and a heavy tailed (e.g. Laplace) spatial effect γ_{2a}. This can be done using continuous mixing using beta weights $h_a \sim \text{Beta}(g_1, g_2)$ where g_1 and g_2 are known (Lawson and Clark, 2002). So

$$\log(\mu_{ax}) = \alpha + \delta_x + u_a + \beta_x \left[h_a \gamma_{1a} + (1 - h_a)\gamma_{2a}\right].$$

Here we consider only the basic non-proportional form as in model 2 above (equations 2a to 2d). Estimation again involves a two chain run to 10000 iterations. Figure 5 shows the spatial pattern of the γ_a effects common to all outcomes. They are, like the s_a, highest in deprived boroughs in inner London; the correlation between the means of γ_a and the deprivation scores of Noble et al (2000) is 0.88. The 'adherence' parameters β_{jx} (Figure 6) show that the spatial effect particularly impacts on mortality and poor health in childhood and at middle age. Table 1 shows the considerable improvement in fit by adopting model 2 as compared to model 1.

Figure 5: Spatial Effect from Age-Area Interaction Model

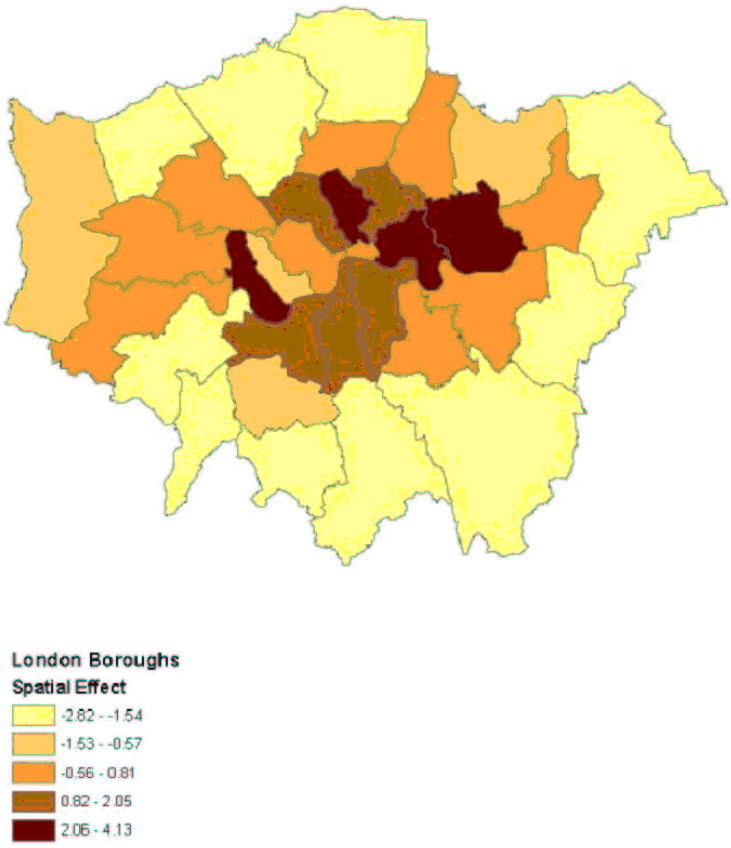
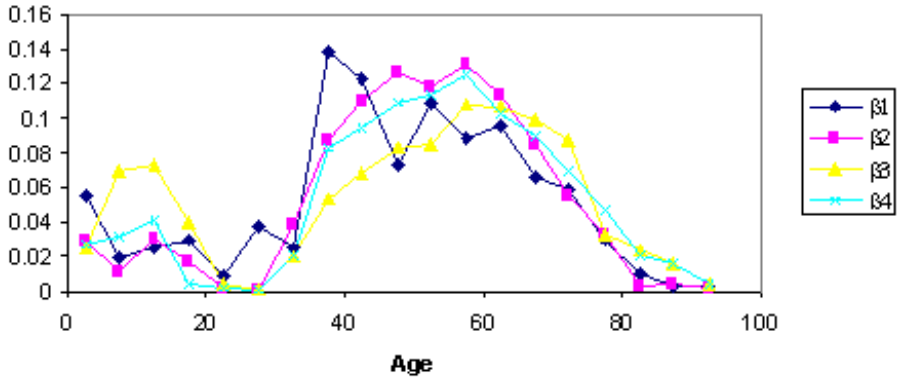


Figure 6: Adherence Parameters by Age



6. Other models for age effects and spatially varying age effects

Instead of assuming a random walk prior for the structured component in the age mixture model $\delta_{jx} = v_{jx} + w_{jx}$, one might represent v_{jx} by a basis function (e.g. a polynomial spline or B spline), and then assume spatially varying coefficients applied to certain components in each function. This may be combined with predictor selection on other components of the function, leading to averaging over a number of different models. As noted by Smith and Kohn (1996) this implies a nonparametric regression model for age effects in which several predictor variables may be redundant. Here a cubic spline in age with terms $\{x, x^2, x^3, (x-t_1)^3_+, \ldots, (x-t_M)^3_+\}$ is assumed in a third model. Define $B_1(x) = x, B_2(x) = x^2, \ldots, B_{M+3}(x) = (x-t_M)^3_+$ then the smooth in age has the form

$$v_{jx} = \sum_{k=1}^{M+3} g_{jk}\eta_{jk}B_k(x)$$

where g_{jk} are binary selection indicators with Bernoulli(0.5) priors, and η_{jk} are coefficients applied to $B_k(x)$ only when $g_{jk} = 1$. The linear coefficient in age η_{j1} is taken as necessary by default so that $g_{j1} = 1$ (e.g. see Figures 2 and 3). All other terms are subject to predictor selection. The M potential knots are taken as the mid-points of each five year age band, excluding the first and last so there are seventeen potential knots (at ages 7.5, 12.5, etc. to 87.5). An unstructured random age term w_{jx} is retained to model remaining residual age impacts for outcome j.

The third model allows for spatially varying linear impacts of age on the mortality

and illness outcomes, so that linear coefficient in age for outcome j is area specific, η_{j1a}. There is evidence at a higher geographical scale, for example, that high and low mortality regimes in developed societies may differ in their age slopes (Gakidou et al, 2000). To additionally reflect the correlation between outcomes (death, long term limiting illness, etc) the area linear effect on age is modelled as

$$\eta_{j1a} = \omega_{ja} + \psi_j \xi_a$$

where ξ_a is a shared spatially correlated error, ω_{ja} is an outcome specific unstructured error with mean η_{j1} and ψ_j are outcome specific loadings. The remainder of the model is an in model 2.

Then model 3 for mortality is

$$\begin{aligned}\log(\mu_{ax}) = \alpha_1 + w_{1x} + (\omega_{1a} + \psi_1\xi_a)x + [g_{12}\eta_{12}x^2 + g_{13}\eta_{13}x^3 + \\ g_{14}\eta_{14}(x-t_1)_+^3 + \cdots + g_{1,M+3}\eta_{1,M+3}(x-t_M)_+^3] + \\ \beta_{1x}\gamma_a + u_{1a}\end{aligned} \tag{3a}$$

The models for illness and health status are accordingly

$$\begin{aligned}\text{logit}(\lambda_{ax}) = \alpha_2 + w_{2x} + (\omega_{2a} + \psi_2\xi_a)x + [g_{22}\eta_{22}x^2 + g_{23}\eta_{23}x^3 + \\ g_{24}\eta_{24}(x-t_1)_+^3 + \cdots + g_{2,M+3}\eta_{2,M+3}(x-t_M)_+^3] + \\ \beta_{2x}\gamma_a + u_{2a}\end{aligned} \tag{3b}$$

$$\begin{aligned}\text{logit}(\gamma_{ax1}) = \kappa_1 - (w_{3x} + [\omega_{3a} + \psi_3\xi_a]x + \{g_{32}\eta_{32}x^2 + g_{33}\eta_{33}x^3 + \\ g_{34}\eta_{34}(x-t_1)_+^3 + \cdots + g_{3,M+3}\eta_{3,M+3}(x-t_M)_+^3\} + \\ \beta_{3x}\gamma_a + u_{3a})\end{aligned} \tag{3c}$$

$$\begin{aligned}\text{logit}(\gamma_{ax2}) = \kappa_2 - (w_{4x} + [\omega_{4a} + \psi_4\xi_a]x + \{g_{42}\eta_{42}x^2 + g_{43}\eta_{43}x^3 + \\ g_{44}\eta_{44}(x-t_1)_+^3 + \cdots + g_{4,M+3}\eta_{4,M+3}(x-t_M)_+^3\} + \\ \beta_{4x}\gamma_a + u_{4a}\end{aligned} \tag{3d}$$

As compared to model 2 this representation produces a further gain in fit at the expense of a relatively small increase in the effective parameter total. The spatially varying linear age effects η_{j1a} tend to be higher in deprived boroughs, but the correlation with deprivation is higher for health and illness outcomes than for mortality.

There is some remaining overdispersion in relation to the $N_c = 627 (= 19 \times 33)$ categories in the mortality and illness analysis and the $N_c = 1254$ categories in the health

status analysis. For the mortality analysis this is only slight with $\text{Dev}(\overline{\Phi}) = 782$ but for illness the same quantity is 2078, while for health status it is 3599. As mentioned above one generalisation of model 1 or subsequent models is to include unstructured age-area random effects (Dean et al, 2001). So let effects e_{jax} replace the unstructured are effects u_{ja} in model 2.

This leads to model 4

$$\log(\mu_{ax}) = \alpha_1 + \delta_{1x} + \beta_1 s_a + e_{1ax} \tag{4a}$$

$$\text{logit}(\lambda_{ax}) = \alpha_2 + \delta_{2x} + \beta_2 s_a + e_{2ax} \tag{4b}$$

$$\text{logit}(\gamma_{ax1}) = \kappa_1 - (\delta_{3x} + \beta_3 s_a + e_{3ax}) \tag{4c}$$

$$\text{logit}(\gamma_{ax2}) = \kappa_2 - (\delta_{4x} + \beta_4 s_a + e_{4ax}) \tag{4d}$$

where the e_{jax} are assumed to be unstructured with outcome specific variances

$$e_{jax} \sim N(0, \tau_{ej}).$$

While producing a clear reduction in the average deviance, this approach also has a cost in model complexity. The effective parameter total of around 1500 compares to the number of categories being modelled, namely $N_c = (19\times33)+(19\times33)+(2\times19\times33) = 2508$.

Alternative measures of fit such as the BIC that penalise complexity more heavily than the DIC (or its classical equivalent the AIC) are available. There is evidence that the AIC tends to select complex models, i.e. is prone to overfitting (Geweke and Meese, 1981). An informal definition of the BIC that uses the effective parameter estimate for each of the three outcomes is contained in Table 1. This is based on the average deviance plus the product of the effective parameters by the log of the number of categories N_c being modelled:

$$\text{BIC} = \text{Dev}(\overline{\Phi}) + p_e \log(N_c).$$

Although model 4 has a relatively low DIC and the highest pseudo marginal likelihood, its BIC exceeds those for the less complex models. Model 3 has the lowest BIC.

7. Implications for life table parameters

One benefit of jointly modelling mortality and morbidity for areas is in providing measures of total and healthy life expectancy for areas and at particular ages and the resulting 'disease burden' measured by years lived in ill-health (Murray and Lopez, 1996). Estimation via repeated simulation has the benefit of providing posterior profiles on structural indices that combine data and parameters in their derivation. Of interest for mapping health need are the following

a) life expectancy by area at age x, E_{ax};
b) disability free life expectancy W^1_{ax}, namely years to be lived beyond age x before the onset of limiting long term illness;
c) healthy life expectancy W^2_{ax}, in terms of years to be lived in good health beyond age x
d) G^1_{ax}, average years lived with disability, namely the gap between E_{ax} and W^1_{ax};
e) and average years lived in poor health G^2_{ax}, the gap between E_{ax} and W^2_{ax}.

Table 2 shows posterior means and standard deviations by borough for total life expectancy and the two forms of health expectancy (at birth and age 65) under model 3. Table 2 also contains a deprivation index devised by the UK Department of Environment, Transport and Regions. For example, the disability free life expectancy at birth W^1_{a0} varies from 69 to 78.1 and correlates -0.85 with deprivation.

In terms of the disease burden at age 65, Table 2 shows that years lived in poor health $G^2_{a,65}$ after age 65 is typically around three years, or half of the years lived in disability $G^1_{a,65}$. Hence the worst category of the health status question is apparently identifying more severe morbidity than the long term illness (limiting disability) question. There is a 0.95 correlation between the disease burden measure $G^2_{a,65}$ and deprivation. For $G^1_{a,65}$ the correlation with deprivation is slightly lower, namely 0.92.

Of interest for health needs profiling is the disease burden at different ages and how this varies between geographic areas. As noted above the area gradients for illness on age η_{j1a} are more highly correlated with area deprivation than those for mortality. This implies that the age profile of the disease burden would be discrepant between affluent and deprived boroughs, and Figure 7 contrasts the burden-age profile for the deprived inner city borough of Tower Hamlets with that in the affluent suburban area of Bromley. The clear excess in morbidity in the inner city borough, especially in middle ages, can be seen.

8. Conclusion

This paper has sought to develop and investigate the fit of a set of models that depart from the often used proportionality assumption for mortality and morbidity data which are crossed by age and area. Instead relatively parsimonious models for age-area interactions in data on deaths and health in London have shown that the proportionality assumption is very much a simplification that does not match actuality for this city region.

The model variants developed have the intention of modelling area life tables that incorporate health status and survival, and to base parameterisation on central features of area contrasts in health. Thus an adaptation of the Lee-Carter model reflects how different

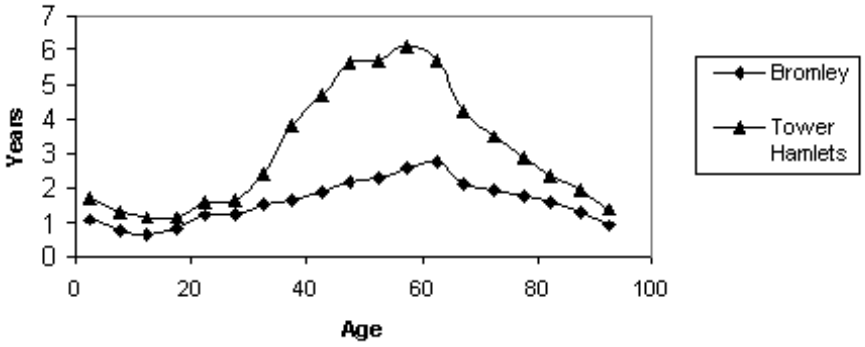
Table 2: Life Table Parameters, Mortality and Health, London Boroughs (Model 3)

	Life Expectancy				*Disability Free Life Expectancy*			
	Age 0		Age 65		Age 0		Age 65	
Borough	Mean	St devn	Mean	St devn	Mean	St devn	Mean	St devn
City of London	80.4	2.3	19.8	1.9	78.1	2.3	13.6	1.3
Barking	73.9	0.4	14.9	0.2	70.3	0.4	8.1	0.1
Barnet	77.0	0.4	16.6	0.2	74.7	0.3	11.1	0.2
Bexley	76.5	0.3	16.2	0.2	74.3	0.3	10.5	0.2
Brent	75.6	0.4	16.2	0.3	73.1	0.4	9.7	0.2
Bromley	76.8	0.3	16.4	0.2	74.5	0.3	11.2	0.2
Camden	73.2	0.5	15.3	0.3	69.9	0.4	8.9	0.2
Croydon	76.8	0.3	16.6	0.2	74.3	0.3	10.8	0.2
Ealing	75.7	0.4	16.1	0.2	73.4	0.3	9.7	0.2
Enfield	76.5	0.3	16.2	0.2	73.8	0.3	10.1	0.2
Greenwich	73.8	0.4	15.2	0.3	70.7	0.4	8.8	0.2
Hackney	74.1	0.5	15.4	0.3	70.8	0.4	7.6	0.2
Hammersmith	75.5	0.5	16.3	0.3	73.1	0.5	9.6	0.2
Haringey	74.6	0.4	15.5	0.3	71.8	0.4	8.9	0.2
Harrow	78.1	0.4	17.5	0.3	75.8	0.4	11.6	0.2
Havering	76.3	0.4	15.9	0.2	73.9	0.4	10.1	0.2
Hillingdon	76.1	0.4	16.3	0.2	73.7	0.4	10.6	0.2
Hounslow	75	0.4	15.8	0.3	72.7	0.4	9.7	0.2
Islington	72.3	0.5	14.5	0.3	69	0.4	7.6	0.2
Kensington & Chelsea	78.3	0.5	17.9	0.4	75.9	0.5	12.0	0.3
Kingston	76.5	0.5	16.2	0.3	74.5	0.5	11.2	0.2
Lambeth	72.5	0.4	14.9	0.3	70.1	0.4	8.7	0.2
Lewisham	73.5	0.4	14.6	0.2	70.8	0.4	8.6	0.1
Merton	76.2	0.4	16.2	0.3	74.1	0.4	10.6	0.2
Newham	72.4	0.4	14.3	0.3	69.5	0.4	6.9	0.1
Redbridge	76.1	0.4	16.2	0.2	73.7	0.4	10.0	0.2
Richmond	77.4	0.4	16.8	0.3	75.8	0.4	11.9	0.2
Southwark	73.3	0.4	15.3	0.3	70.6	0.4	8.7	0.2
Sutton	76.1	0.4	15.8	0.3	73.7	0.4	10.5	0.2
Tower Hamlets	72.1	0.4	14.2	0.3	69.1	0.4	6.8	0.2
Waltham Forest	73.6	0.4	14.7	0.2	70.9	0.4	8.5	0.2
Wandsworth	74	0.4	14.7	0.2	71.9	0.4	9.2	0.2
Westminster	75.7	0.5	16.5	0.3	73.1	0.4	10.6	0.2

Table 2: (Continued)

	Healthy Life Expectancy (Years lived before entering 'poor health')				*Disease Burden (age 65), Years lived in Disability or Poor Health*				*DETR Deprivation Index*
	Age 0		Age 65		Disability Based		Health Status Based		
Borough	Mean	St devn	Mean	St devn	Mean	St devn	Mean	St devn	
City of London	79.0	2.3	16.8	1.6	6.2	0.62	3.1	0.34	-0.88
Barking	72.1	0.4	11.4	0.2	6.8	0.13	3.4	0.07	0.68
Barnet	75.9	0.3	14.1	0.2	5.5	0.08	2.4	0.05	-0.84
Bexley	75.5	0.3	13.7	0.2	5.7	0.10	2.5	0.05	-0.91
Brent	74.4	0.4	13.0	0.2	6.5	0.12	3.2	0.06	0.41
Bromley	75.7	0.3	14.3	0.2	5.3	0.08	2.1	0.05	-1.12
Camden	71.4	0.4	11.9	0.2	6.3	0.14	3.4	0.09	0.56
Croydon	75.5	0.3	14.0	0.2	5.8	0.09	2.6	0.05	-0.52
Ealing	74.6	0.4	13.0	0.2	6.3	0.10	3.1	0.06	-0.14
Enfield	75.1	0.3	13.4	0.2	6.1	0.10	2.8	0.05	-0.13
Greenwich	72.2	0.4	12.0	0.2	6.4	0.12	3.2	0.06	0.56
Hackney	72.5	0.5	11.0	0.2	7.8	0.17	4.4	0.11	2.02
Hammersmith	74.4	0.5	12.9	0.3	6.7	0.16	3.5	0.09	0.12
Haringey	73.1	0.4	12.0	0.2	6.6	0.14	3.4	0.08	0.92
Harrow	77.0	0.4	15.1	0.3	5.9	0.12	2.5	0.06	-0.89
Havering	75.2	0.4	13.4	0.2	5.8	0.10	2.5	0.05	-0.85
Hillingdon	74.9	0.4	13.8	0.2	5.7	0.10	2.5	0.05	-0.7
Hounslow	73.9	0.4	12.9	0.2	6.1	0.12	2.9	0.06	-0.22
Islington	70.5	0.4	10.7	0.2	6.9	0.16	3.8	0.10	1.21
Kensington & Chelsea	77.0	0.5	14.9	0.3	5.9	0.14	2.9	0.08	-0.62
Kingston	75.5	0.4	14.1	0.3	5.1	0.11	2.1	0.06	-1.32
Lambeth	71.3	0.4	11.6	0.2	6.2	0.12	3.3	0.07	0.66
Lewisham	72.2	0.4	11.6	0.2	6.0	0.10	3.0	0.06	0.56
Merton	75.2	0.4	13.7	0.2	5.6	0.11	2.5	0.06	-0.74
Newham	70.9	0.4	10.2	0.2	7.4	0.15	4.0	0.09	2.02
Redbridge	75.0	0.4	13.4	0.2	6.3	0.11	2.9	0.06	-0.49
Richmond	76.6	0.4	14.8	0.3	4.9	0.11	2.0	0.05	-1.47
Southwark	72.0	0.4	11.8	0.2	6.6	0.13	3.5	0.08	1.21
Sutton	74.9	0.4	13.6	0.2	5.3	0.10	2.2	0.05	-1.01
Tower Hamlets	70.4	0.4	9.9	0.2	7.4	0.16	4.3	0.11	2.36
Waltham Forest	72.3	0.4	11.6	0.2	6.2	0.12	3.1	0.07	0.34
Wandsworth	73.1	0.4	12.0	0.2	5.6	0.11	2.7	0.06	-0.37
Westminster	74.3	0.5	13.4	0.3	5.9	0.13	3.0	0.07	-0.43

Figure 7: Disease Burden by Age, Deprived and Affluent Boroughs Compared



age groups accord with a single spatial health index γ_a. Similarly the basic linear age effect on log death rates or logit illness/health rates may vary over areas.

In a joint life table pooling over outcomes it is important to model the correlation between outcomes. The correlation over outcomes in both age and area impacts is reflected in

a) the pooled random walk effect V_x in δ_{jx} in models 1 and 2
b) the shared spatial effects $\theta_j s_a$ in model 1
c) the adherence by age $\beta_{jx}\gamma_a$ interacting with shared spatial effects in model 2, and
d) the common area effect multiplied by an outcome specific loading in the linear age effects in model 3, namely $\eta_{j1a} = \omega_{ja} + \psi_j \xi_a$.

Further stratifiers may be introduced into such a framework, for example deaths, illness and health may be specific for gender or ethnicity as well as for age and area. A time dimension could be added also.

A range of inferences is possible from this type of model in terms of contrasts in life expectancy, health or disability free expectancies, and resulting disease burdens. Variations in the disease burden are closely related to health need and use of health care (Murray and Lopez, 1996). Unlike deprivation proxies for need that are often used in health care resourcing the outputs from spatial life tables form a direct rather than proxy measure of morbidity (Newbold et al, 1988).

References

Agresti, A. (2002). Categorical data analysis. *Wiley, 2nd Edition*.

Anson, J. (1991). Model mortality schedules: a parametric evaluation. *Population Studies*, 45:137–153.

Bebbington, J. (1993). Regional and social variations in disability-free life expectancy in great britain. *In: Robine J-M, Mathers C, Bone I, Romieu, I, eds. Calculation of health expectancies: harmonization, consensus achieved and future perspectives. London: John Libbey*.

Besag, J., York, J., and Mollié, A. (1991). Bayesian image restoration with two applications in spatial statistics. *Annals of the Institute of Statistics and Mathematics*, 43:1–59.

Carter, L. and Lee, R. (1992). Modeling and forecasting us sex differentials in mortality. *International Journal of Forecasting*, 8:393–411.

Dean, C., Ugarte, M., and Militino, A. (2001). Detecting interaction between random region and fixed age effects in disease mapping. *Biometrics*, 57:197–202.

Eames, M., Ben-Shlomo, Y., and Marmot, M. (1993). Social deprivation and premature mortality: regional comparison across england. *British Medical Journal*, 307:1085–6.

Gakidou, E., Murray, C., and Frenk, J. (2000). Defining and measuring health inequality: an approach based on the distribution of health expectancy. *Bulletin of the World Health Organisation*, 78:42–54.

Gelfand, A. and Dey, D. (1994). Bayesian model choice: asymptotics and exact calculations. *J. Royal Statist. Soc.*, 56(B):501–514.

Gelman, A., Carlin, J., Stern, H., and Rubin, D. (1995). Bayesian data analysis. *London: Chapman and Hall*.

Geweke, J. and Meese, R. (1981). Estimating regression models of finite but unknown order. *International Economic Review*, 22:55–70.

Gibbons, R. and Hedeker, D. (2000). Applications of mixed-effect models in biostatistics. *Sankhya*, 62(B):70–103.

Goodman, L. (1979). Simple models for the analysis of association in cross-classifications having ordered categories. *Journal of the American Statistical Association*, 74:537–551.

Hoem, J. (1987). Statistical analysis of a multiplicative model and its application to the standardization of vital rates: a review. *International Statistical Review*, 55:119–152.

Ibrahim, J., Chen, M., and Sinha, D. (2001). Bayesian survival analysis. *Springer Verlag: New York*.

Lawson, A. and Clark, A. (2002). Spatial mixture relative risk models applied to disease mapping. *Statistics in Medicine*, 21:359–370.

Lee, R. (2000). The lee-carter method for forecasting mortality, with various extensions and applications. *North American Actuarial Journal*, 4:80–93.

MacNab, Y. and Dean, C. (2001). Autoregressive spatial smoothing and temporal spline smoothing for mapping rates. *Biometrics*, 57:949–56.

Murray, C. and Lopez, A. (1996). The global burden of disease. *Harvard University Press*.

Nandram, B., Sedrank, J., and Pickle, L. (1999). Bayesian analysis of mortality rates for us health service areas. *Sankhya*, 61(B):145–165.

Newbold, K., Eyles, J., Birch, S., and Spencer, A. (1998). Allocating resources in health care: alternative approaches to measuring needs in resource allocation formula in ontario. *Health and Place*, 4:79–89.

Noble, M., Penhale, B., Smith, G., Wright, G., Dibben, C., Owen, T., and Lloyd, M. (2000). Indices of deprivation 2000. *Regeneration Research Summary Number 31, Department of Transport, Environment and the Regions*.

Peterson, B. and Harrell, F. (1990). Partial proportional odds models for ordinal response variables. *Applied Statistics*, 39:205–217.

Robine, J.-M. and Ritchie, K. (1991). Healthy life expectancy : evaluation of a new global indicator for change in population health. *British Medical Journal*, 302:457–460.

Smith, M. and Kohn, R. (1996). Nonparametric regression using bayesian variable selection. *Journal of Econometrics*, 75:317–34.

Spiegelhalter, D., Best, N., Carlin, B., and van der Linde, A. (2001). Bayesian measures of model complexity and fit. *J. Royal Statist. Soc*, 64(B):583–639.

Sun, D., Tsutakawa, R., Kim, H., and He, Z. (2000). Spatio-temporal interaction with disease mapping. *Statistics in Medicine*, 19:2015–2035.

Wakefield, J., Best, N., and Waller, L. (2000). Bayesian approaches to disease mapping. *In Elliott P, Wakefield J, Best N, Briggs D (eds) Spatial Epidemiology; Methods and applications. Oxford University Press, Oxford*, pages 106–127.

Zayeri, F., Kazemnejad, A., Khanafshar, N., and Nayeri, F. (2005). Modeling repeated ordinal responses using a family of power transformations: application to neonatal hypothermia data. *BMC Medical Research (Methodology) 5:29*.