Assessing Univariate and Bivariate Spatial Clustering in Modelled Disease Risks

Peter Congdon*
Department of Geography, Queen Mary, University of London, Mile End Rd, London E1 4NS, UK

Abstract

Models for spatial variation in relative disease risk often consider posterior probabilities of elevated disease risk in each area, but for health prioritisation, the interest may also be in the broader clustering pattern across neighbouring areas. The classification of a particular area as high risk may or may not be consistent with risk levels in the surrounding areas. Local join-count statistics are used here in conjunction with Bayesian models of area disease risk to detect different forms of disease clustering over groups of neighbouring areas. A particular interest is in spatial clustering of high risk, which can be assessed by high probabilities of elevated risk across both a focus area and its surrounding locality. An application considers univariate spatial clustering in suicide deaths in 922 small areas in the North West of England, extending to an analysis of bivariate spatial clustering in suicide deaths and hospital admissions for intentional self-harm in these areas.

Keywords: Relative risk; Spatial clustering; Bayesian; Bivariate clustering; Join-count

Introduction

Spatial analyses of disease incidence or mortality in small areas are often used to identify elevated risk. For example, posterior probabilities of elevated disease risk in each area may be obtained from Bayesian models of area disease counts [1-3]. However, elevated risk identified in a particular area may not extend to nearby areas, whereas spatial clustering of high risk across several adjacent or nearby areas may be of particular importance for health policy prioritisation. Identification of high risk localities may be more precise than identification of elevated risk in individual areas, especially for less frequent outcomes. Interest in spatial clustering of high risk may extend to the geographic pattern of interrelated disease outcomes (e.g., different forms of mental illness or different cardiovascular diseases).

For identifying such clustering in conjunction with a disease model, local indices of spatial association, and particular adaptations of them, become relevant. The present analysis considers use of local join-count statistics to detect high risk locality clustering for disease count data aggregated into small areas. These statistics are used in conjunction with an area disease model, and Bayesian inferences based on updating of prior assumptions using Markov chain Monte Carlo (MCMC) estimation.

The proposed join-count statistics methodology is initially applied to clustering of suicide deaths in 922 small areas in NW England. The relative mortality risks are modelled using a Bayesian spatial convolution prior [4] with alternative forms of spatial interaction (e.g., binary adjacency vs distance decay) between neighbouring areas considered. Results obtained using the join count method in conjunction with a spatial model are compared with the widely used (albeit non-Bayesian) spatial scan method, as implemented in the FlexScan and SaTScan packages. The application is then extended to analyze bivariate spatial clustering in suicide deaths and self-harm hospitalisations.

Methods: Join-Count Measures for Local Clustering in Risk

In Bayesian small area disease applications, the classification of an area as high risk typically depends on unknown parameters. Consider disease counts \( y_i, i=1,...,n \) with expected values \( \epsilon_i \) obtained by multiplying area populations by the region-wide disease rate, and with \( \Sigma y_i = n \epsilon \). Then the \( y_i \) may be taken as Poisson,

\[ y_i \sim \text{Poi}(e_i r_i), \]

where the \( r_i \) are relative disease risks in area \( i \) with average 1. Such risks often show spatial correlation, and ignoring such correlation can lead to biased and inefficient inference, as the observations are not independent [5]. A widely applied model (known as the convolution model) involves two sets of random effects: (a) spatially structured effects \( s_i \) to represent spatially correlated risks and following a conditional autoregressive (CAR) scheme [4],

\[ s_i | \xi[i] \sim N \left( \sum_{j \neq i} w_{ij} s_j / S_{II}, \sigma^2 / S_{ii} \right), \]

where \( w_i \) are symmetric spatial interactions (with \( w_i=0 \)), \( \xi[i] \) represents the collection of \( S \) effects excluding \( s_i \), and \( S_{II} = \sum w_{ij}^2 \); and (b) iid random effects \( u_i \), typically normal with \( u_i \sim N(0,\tau) \), to represent possible over dispersion, or excess variability in relation to the Poisson assumption. Then if there are covariates \( X_i \) relevant to explaining variations in area disease risk, one has:

\[ \log(r_i) = X_i \beta + s_i + u_i, \]  

(1)

If there are no covariates to model spatial patterning of risks, the spatial random effects represent spatial pattern in the disease outcome, whereas otherwise they capture spatial structure in the residuals. To assess actual clustering, one may obtain measures such as Moran’s \( I \) for the \( s_i \); this entails deriving the index at each MCMC iteration, with posterior inferences (e.g. credible intervals) based on the values

*Corresponding author: Peter Congdon, Department of Geography, Queen Mary, University of London, Mile End Rd, London E1 4NS, UK, E-mail: p.congdon@qmul.ac.uk

Received January 24, 2012; Accepted February 19, 2013; Published February 23, 2013


Copyright: © 2013 Congdon P. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
accumulated over iterations.

Consider binary measures $b$ of disease risk for areas $i=1,...,n$, with $b=1$ for elevated risk, $b=0$ otherwise. For example, one may define $b=1$ for $r>r_τ$, where $r_τ$ is a relative risk threshold (e.g., $r_τ=1$ or $r_τ=1.25$, and $b=0$ otherwise. If relative risks have average 1, and $r_τ=1$ then the region-wide proportion of areas with elevated risk, $E(b)=π$, will be approximately 0.5.

In spatial disease applications with correlated relative risks $r$, binary indicators such as $b=I(r>1)$ will also tend to be spatially correlated. Region-wide spatial clustering in the $b$ can be measured by join-count statistics, based on concordance in risk status between area pairs. Thus, a join-count measuring clustering in high risk across a region is

$$J_{11} = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} b_i b_j,$$

with 0.5$I_i$, known as the BB statistic [6]. Differing health status in neighbouring area-pairs is measured by a weighted total of joins with $b_i$ and $b_j$ discordant, which can be denoted

$$J_{10} = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} (b_i - b_j)^2.$$

Observed join-count totals can be compared with totals expected under a null hypothesis of spatial independence [7]. The expected total of discordant joins under the hypothesis of no spatial dependence is $E(J_{10})=S_\pi^2 n^2$, where $S_\pi = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}$, and $J_{11}$ will exceed $E(J_{11})$ when there is spatial patterning in the disease outcome [6]. Similarly, the observed join-counts $J_{ii}$ can be compared with $E(J_{ii})=2S_\pi(1-π)$, and will be less than this expected total when there is disease clustering. It may be noted that in a modelling application with $r$ unknown, the indicators $b_i$ (and related parameters such as $π$) are also unknown, and sampled at each MCMC iteration.

A localised set of join-count statistics (with area $i$ as the focus) can be used to decide whether area $i$ and its neighbouring areas form a high risk cluster, or demonstrate an alternative risk pattern in the locality. For measuring joint high risk, with both area $i$ and its neighbouring areas being high risk, one has

$$J_{11i} = I(b_i = 1) \sum_{j=1}^{n} w_{ij} I(b_j = 1) = h_i \sum_{j=1}^{n} w_{ij} b_j,$$

where $I(A)=1$ if condition $A$ holds, and $I(A)=0$ otherwise. When the focus is on area $i$, it is relevant to distinguish discordant high-low risk pairings ($b_i=1, b_j=0$) from low-high risk pairings ($b_i=0, b_j=1$). The relevant local join-count statistics in these cases are then

$$J_{10i} = I(b_i = 1) \sum_{j=1}^{n} w_{ij} I(b_j = 0) = h_i \sum_{j=1}^{n} w_{ij} (1-b_j),$$

and

$$J_{01i} = I(b_i = 0) \sum_{j=1}^{n} w_{ij} I(b_j = 1) = (1-h_i) \sum_{j=1}^{n} w_{ij} b_j.$$

The count $I_i$ captures situations where area $i$ is high risk, but nearby areas are mostly low risk, so that area $i$ may be termed a high risk local outlier. The count $H_i$ would be elevated when area $i$ itself does not have high risk, but neighbouring areas are mostly high risk. Finally,

$$J_{00i} = I(b_i = 0) \sum_{j=1}^{n} w_{ij} I(b_j = 0) = (1-h_i) \sum_{j=1}^{n} w_{ij} (1-b_j)$$

represents localities where both the focus and surrounding areas are low risk. The expected number of common high risk joins with area $i$ as the focus (i.e. area $i$ is a high risk cluster member) is $E(J_{11i})=S_\pi π^2$, while $E(J_{10i})=S_\pi(1-π)$ and $E(J_{00i})=S_\pi(1-π)^2$.

Consider a sequence $t=1,...,T$ of MCMC samples. From the indicators $b_i(t)$ of elevated risk at each MCMC iteration, one may estimate probabilities of elevated risk in area $i$ specifically (without regard to the broader locality), namely

$$\hat{H}_i = \frac{\sum_{t=1}^{T} I(b_i(t) = 1)}{T}.$$

One may also monitor join-counts indicating locality-wide elevated risk $J_{11i} = b_i(t) \sum_{j=1}^{n} w_{ij}(t)$, with posterior estimates

$$\hat{J}_{11i} = \frac{\sum_{t=1}^{T} J_{11i}(t)}{T}.$$

The estimated proportion of joins in the locality centred on area $i$ that are joint high risk, namely

$$\hat{π}_{11i} = \frac{\hat{J}_{11i}}{S_0},$$

provides a summary index of high risk across that locality. By contrast, the proportion of joins centred on area $i$ that are (1,0) pairs

$$\hat{π}_{10i} = \frac{\hat{J}_{10i}}{S_0},$$

provides an index that area $i$ is a high risk outlier relative to the broader locality.

The join-counts $J_{11i}$ and $J_{10i}$ can be written as $I(b_i = 1)\sum_{j=1}^{n} w_{ij} b_j$ and $I(b_i = 0)\sum_{j=1}^{n} w_{ij}(1-b_j)$ respectively, from which it follows that

$$J_{11i} + J_{10i} = I(b_i = 1)S_0,$$

and hence that

$$\hat{H}_i = \hat{π}_{11i} + \hat{π}_{10i}.$$

Hence, $\hat{π}_{11i}$ will be elevated when both $\hat{H}_i$ is elevated, and risk in the surrounding locality is elevated also. By contrast, $\hat{π}_{10i}$ will be elevated when $\hat{H}_i$ is elevated, but risk in the surrounding locality is relatively low. Similarly, $J_{00i} + J_{01i} = I(b_i = 0)\sum_{j=1}^{n} w_{ij} b_j + I(b_i = 0)\sum_{j=1}^{n} w_{ij}(1-b_j) = I(b_i = 0)S_0$, and defining $\hat{π}_{01i} = \frac{J_{01i}}{S_0}$ and $\hat{π}_{00i} = \frac{J_{00i}}{S_0}$, one has

$$1 - \hat{H}_i = \hat{π}_{01i} + \hat{π}_{00i}.$$

Areas can be ranked in terms of $\hat{π}_{11i}$ to indicate which likely high risk cluster centres are. Alternative tests regarding high risk clustering in the locality around area $i$ might be envisaged. One involves expectation weighted averages

$$R_i = \sum_{j=1}^{n} \sum_{j=1}^{n} r_{ij} e_j / \sum_{j=1}^{n} e_j$$

of modelled relative risks across localities $I_i$ that include both the focus area $i$ and areas adjacent to it. These weighted averages can be monitored during the MCMC updating and the probabilities that $R_i$ exceed 1 obtained. However, this test may be affected by unusually high relative risks in one or two areas within the locality, or by situations where a low risk area is surrounded by high risk areas.

Another option is to compare the sampled $I_i(t)$ at each MCMC iteration to the expected count $S_\pi π^2$ under a no clustering hypothesis, and obtain estimates of the probabilities $\hat{π}_{11i} = P(J_{11i} > S_\pi π^2 | y)$.

If $b_i=1$, $J_{11i} - b_i\sum_{j=1}^{n} w_{ij} b_j - \sum_{j=1}^{n} w_{ij}$, and the comparison $J_{11i} > S_\pi π^2$
reduces to \( \sum_j w_{ij} > S_{0i}^2 \), a condition very likely to be met in high risk localities (where risk is elevated in both the focus area and surrounding areas, so that \( R_{ij} \) is high). Hence, the comparison \( I(A_{ij} > S_{0i}^2) \) will tend to have a similar probability of holding as that for \( I(b_i = 1) \) in such localities.

**Methods: Clustering in Bivariate Risk**

An extension of the proposed join-count statistics is in the detection of elevated bivariate risk across localities, based on local join-counts for the joint high risk binary event. Methods for bivariate spatial association have been proposed [8], and bivariate LISA methods indicate association between the value for one variable at a given location and the average of another variable at neighbouring locations. However, there is no widely applied cluster detection method (e.g. spatial scan technique) for bivariate outcomes. Let \( A \) and \( B \) denote two health outcomes \( y_{Ai} \sim \text{Poi}(e_{Ai} \beta_A) \), \( y_{Bi} \sim \text{Poi}(e_{Bi} \beta_B) \) and consider join-counts corresponding to the joint high risk classification: \( b_{AB} = I(r_{Ai} > T_{ij}, r_{Bi} > T_{ij}) \). The event risks \( (r_{Ai}, r_{Bi}) \) can be obtained via the models

\[
\log(r_{Ai}) = X_i \beta_A + u_{Ai} + s_{Ai},
\]

\[
\log(r_{Bi}) = X_i \beta_B + u_{Bi} + s_{Bi},
\]

with one option for priors on the random effects being

\[
u_{Ai} \sim N(0, \sigma_A^2), \quad u_{Bi} \sim N(0, \sigma_B^2)
\]

To assess high risk clustering in both events jointly, the bivariate local join-counts

\[
J_{AB11} = b_{AB} \sum_{j=1}^{n} w_{ij} b_{ij} = b_{AB} w_{ij} \sum_{j=1}^{n} w_{ij} b_{ij}
\]

can be monitored. The estimated probability of elevated bivariate risk in area \( i \) specifically is

\[
\hat{H}_{AB} = \sum_{j=1}^{T} b_{ij} / T,
\]

but this elevated bivariate risk may not apply across the broader locality. However, the estimated proportion of bivariate joins in the locality centred on area \( i \) that are joint high risk, namely

\[
\hat{\pi}_{AB11} = J_{AB11} / S_{0i}
\]

provides a summary index of high bivariate risk across that locality. For detecting isolated elevated bivariate risk (high risk in the focus area but not extending to the broader locality), the relevant join count is

\[
J_{AB10} = b_{AB} \sum_{j=1}^{n} w_{ij} (1 - b_{ij}) = b_{AB} w_{ij} \sum_{j=1}^{n} w_{ij} (1 - b_{ij})
\]

Just as implications about smoothed relative risks may depend on the form of spatial interaction assumed [5], so may the inferences about clustering patterns. Implications about risk patterns for interdependent events, especially when one event is less frequent than another, may also be influenced by the form of random effects assumption (and the extent to which there is borrowing of strength). For example, clustering inferences in the less common outcome may be affected if a bivariate spatial prior (allowing correlation in spatial risks between outcomes with elevated bivariate risk) is adopted instead of separate univariate spatial priors as in equation (3).

**Results: Locality Risk Patterns in Suicide Deaths in NW England**

While relatively rare, suicide is a major reason for premature mortality. To assess risk patterns in individual areas as compared to their broader localities, we consider suicide deaths \( y \) over the period 2006 to 2010 in 922 small areas (Middle Level Super Output Areas or MSOAs) across the North West of England (Table 1). These areas are designed to be of similar size in population terms, with an average population of 7500. Expected deaths \( \hat{y}_i \) are based on applying an England wide schedule of age specific suicide rates to MSOA populations, with scaling applied to ensure \( \sum_i \hat{y}_i = \sum_i y_i \).

The average mortality count is 3.5, but event totals \( y_i \) in individual

<table>
<thead>
<tr>
<th>Locality</th>
<th>Index of focus area</th>
<th>( \pi_{0i} ) posterior estimate</th>
<th>( H_i ) posterior estimate</th>
<th>Relative risk ( r_i ) in focus area (posterior mean)</th>
<th>Total areas in locality (including focus)</th>
<th>Indices of areas in locality (other than focus)</th>
<th>Modelled relative risk ( R_i ) across locality (expectation weighted average)</th>
<th>( Pr(R_i &gt; 1) ) (elevated locality risk)</th>
<th>SMR across locality</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>594</td>
<td>0.854</td>
<td>0.979</td>
<td>1.717</td>
<td>7</td>
<td>590, 592, 593, 596, 597, 599</td>
<td>1.547</td>
<td>0.999</td>
<td>2.151</td>
</tr>
<tr>
<td>2</td>
<td>16</td>
<td>0.834</td>
<td>0.989</td>
<td>1.845</td>
<td>9</td>
<td>5, 10, 11, 15, 17, 21, 22, 25</td>
<td>1.433</td>
<td>1.000</td>
<td>1.872</td>
</tr>
<tr>
<td>3</td>
<td>11</td>
<td>0.820</td>
<td>0.971</td>
<td>1.729</td>
<td>5</td>
<td>5, 18, 15, 16</td>
<td>1.478</td>
<td>0.997</td>
<td>2.011</td>
</tr>
<tr>
<td>4</td>
<td>249</td>
<td>0.761</td>
<td>0.969</td>
<td>1.692</td>
<td>966</td>
<td>247, 248, 251, 252</td>
<td>1.385</td>
<td>0.991</td>
<td>1.782</td>
</tr>
<tr>
<td>5</td>
<td>856</td>
<td>0.753</td>
<td>0.977</td>
<td>1.742</td>
<td>95</td>
<td>849, 851, 853, 857, 858, 859</td>
<td>1.375</td>
<td>0.995</td>
<td>1.957</td>
</tr>
<tr>
<td>6</td>
<td>595</td>
<td>0.743</td>
<td>0.889</td>
<td>1.460</td>
<td>889</td>
<td>593, 594, 596, 599</td>
<td>1.427</td>
<td>0.992</td>
<td>1.844</td>
</tr>
<tr>
<td>7</td>
<td>251</td>
<td>0.741</td>
<td>0.885</td>
<td>1.420</td>
<td>885</td>
<td>247, 248, 250, 252, 258</td>
<td>1.534</td>
<td>0.999</td>
<td>2.181</td>
</tr>
<tr>
<td>8</td>
<td>597</td>
<td>0.735</td>
<td>0.966</td>
<td>1.755</td>
<td>963</td>
<td>594, 599, 601</td>
<td>1.452</td>
<td>0.989</td>
<td>1.901</td>
</tr>
<tr>
<td>9</td>
<td>590</td>
<td>0.726</td>
<td>0.986</td>
<td>1.882</td>
<td>985</td>
<td>587, 589, 592, 594</td>
<td>1.465</td>
<td>0.993</td>
<td>1.741</td>
</tr>
<tr>
<td>10</td>
<td>710</td>
<td>0.723</td>
<td>0.966</td>
<td>1.768</td>
<td>943</td>
<td>709, 711, 712</td>
<td>1.405</td>
<td>0.986</td>
<td>1.580</td>
</tr>
<tr>
<td>11</td>
<td>10</td>
<td>0.714</td>
<td>0.895</td>
<td>1.452</td>
<td>894</td>
<td>2, 5, 6, 13, 16, 17</td>
<td>1.353</td>
<td>0.998</td>
<td>1.757</td>
</tr>
<tr>
<td>12</td>
<td>258</td>
<td>0.713</td>
<td>0.999</td>
<td>2.207</td>
<td>994</td>
<td>250, 251, 252, 255, 259, 262, 264</td>
<td>1.369</td>
<td>0.995</td>
<td>1.660</td>
</tr>
<tr>
<td>13</td>
<td>15</td>
<td>0.713</td>
<td>0.870</td>
<td>1.397</td>
<td>870</td>
<td>8, 11, 16, 18, 21</td>
<td>1.419</td>
<td>0.998</td>
<td>1.795</td>
</tr>
</tbody>
</table>

Table 1: Suicide mortality, areas with highest estimates (\( \pi_{0i} \)) for elevated locality risk.
areas vary widely, and moment estimates of relative risk $\gamma_j / e_j$
(sometimes called standard mortality ratios or SMRs), also vary widely. Such
moment estimates are unreliable with variance instability when there are
small numbers of suicide deaths, as in many MSOAs [9,10].

To provide stabilised estimates of relative risk including spatial borrowing of
strength, a convolution model is applied with $y_i \sim \text{Poi}(\phi_i)$, where
\[ \log(\phi_i) = \beta_0 + u_i + s_i, \]
where $s_i \sim N \left( \sum w_{ij} e_j / S_0, \sigma^2 / S_0 \right)$, and $u_i \sim N(0, \tau^2)$. A flat prior
on $\theta_0$ is assumed, and a gamma prior with index 1 and shape 0.001 on
the inverse spatial variance $1 / \sigma^2$ [11,12]. Convergence is improved by
linking the variance parameters; thus $\tau^2 = \sigma^2 / \rho$, where $\rho$ is
assigned an exponential prior with rate 1. Inferences in this and subsequent
models are based on the second halves of two chain runs of 10,000
iterations, with convergence assessed according to BGR statistics [13].

Localities are defined as areas adjacent to area $i$, though the
weighting attached to different areas within such localities can be
varied. To assess possible sensitivity regarding inferences about locality
risk, alternative assumptions about $w$ are investigated: equal weighting
of all adjacent areas as compared to alternative forms of inverse distance
decay $w_{ij} = 1 / d_{ij}^{\alpha}$. The binary indicators $I_j = 1 (\phi_j > 1)$ and local
join-counts $j_{11}^{(i)}$ and $j_{11}^{(i)}$ are monitored to provide posterior estimated probabilities $\pi_{11}$
of high risk common to the focus and its locality, and estimated
marginal probabilities of elevated risk, namely $H_i$.

Inferences for Locality Risks

Consider first a binary adjacency assumption for the $w$ ($w_{ij} = 1$ if
areas are adjacent, $w_{ij} = 0$, otherwise), under which the Moran spatial
correlation index for the $s$ is obtained as 0.56 with 95% interval (0.47,
0.66). Figure 1 maps out the posterior mean relative risks $r_j$ across
the region, though this map tends to be dominated by low density rural
areas (such as in the Lake District in the northern part of the map).

Subsequently higher resolution maps are used to depict risk and
clustering patterns, since in the case study, high risk clustering tends
to be in densely populated urban areas. Maps of the administrative
geography of the region (including maps of MSOAs) are available at the
UK Map Collection page http://www.ons.gov.uk/ons/guide-method/
geography/beginner-s-guide/maps/index.html.

The estimated $\pi_{11}$ have an average of 0.221, with a 0.975 percentile
of 0.638, and with the maximum $\pi_{11}$ being 0.854. The estimated
marginal probabilities $H_i$ have an average of 0.426, with a 0.975
percentile of 0.923, and a maximum of 0.999. The $\pi_{11}$, which provide
indicators of isolated high risk not extending to the broader locality,
have an average of 0.205, with a maximum of 0.583. The estimated $h_{11}$
are also shown; as discussed above, these are similar to $H_i$ in localities
characterised by high risk clustering, but their ordering of potential
cluster centres is similar to that of the $\pi_{11}$. Of the 13 areas with highest
$h_{11}$ values, 10 are also among the 13 areas with highest $\pi_{11}$ Values.

There are 5 areas with $\pi_{11}$ over 0.75, and 13 areas with $\pi_{11}$ over
0.70. Table 1 summarises locality risk patterns for the 13 areas with
$\pi_{11}$ over 0.70, ranked by the size of $\pi_{11}$, and also including estimates
of $H_i$ and $r_i$ (posterior means). The relatively low values for both $H_i$
and $\pi_{11}$ reflect the rarity of the suicide outcome; more frequent
outcomes (such as self-harm hospitalisations considered in the bivariate
analysis) are more likely to have high $H_i$ and $\pi_{11}$ (e.g. close to 1).
Table 1 also shows posterior means of expectation weighted averages
$R_{ij} = \sum r_j e_j / \sum e_j$, of modelled relative risks across localities $L_j$
comprising both the focus area $i$ and areas adjacent to it. Also
shown are estimated probabilities that $R_{ij} > 1$, namely that the
entire locality has elevated risk, and unsmoothed suicide SMRs across
localities $L_j$ encompassing both the focus area $i$ and areas adjacent to it.

The $\pi_{11}$ identify focus areas with high probabilities of elevated
risk and of belonging to a high risk locality, rather than clusters per
se. So some areas are present in more than one locality in table 1; for
example, areas 594 and 599 appear twice. There are 47 distinct areas in
the localities in table 1, and their posterior mean $r_i$ range from 0.99 to
2.21 with average 1.36.

Probabilities that the average locality risk $R_{ij}$ exceeds 1 are all over
0.968. The average locality risks $R_{ij}$ may be used to confirm what the
join-count statistics indicate, in particular the $\pi_{11}$ statistics, but in
themselves are not conclusive about elevated risk common to both a
focus area and areas around it. Weighted averages such as $R_{ij}$ may be
affected by unusually high relative risks in a subset of areas within the
locality, whereas $\pi_{11}$ is specifically focussing on elevated risk status
across all areas in a locality. An example is provided by area 28 which has
1 suicide death against 4.4 expected, with an estimated exceedance
probability $H_{28} = 0.36$. However, the areas adjacent to area 28 have 34
deaths in relation to 20 expected, with a probability of 0.98 that the
locality wide $R_{1,28}$ exceed 1 (where the locality encompasses area 28). Note
that this type of pattern would be detected by the join-counts $I_{11}$
and corresponding probabilities $\pi_{01} = y_{01} / S_{01}$.

Delineation of high risk localities using local join-counts in
conjunction with a relative risk model such as equation (1) contrasts
with the spatial scan procedure which is applied to observed area
disease counts without any modelling preliminaries, for example,
smoothing or borrowing strength procedures to reduce unreliability in fixed effects relative risk estimates. Despite this fundamental difference, the localities of table 1 can be compared with clusters identified by the SaTScan and FleXScan packages developed by Kulldorff [14] and Tango and Takahashi [15], respectively. It also implies to the work done by Holowaty et al. [1] and Więckowska et al. [16]. SaTScan identifies five high rate clusters with Monte-Carlo p-values under 0.2. The average $\pi_{11i}$ is 0.666 for the 29 areas in these five clusters, and the overlap with the local join-count method is apparent in that only 2.2% of the 922 areas have $\pi_{11i}$ over 0.666. Similarly, the 59 areas identified by FleXScan (in 7 clusters with p-values under 0.2) have an average $\pi_{11i}$ of 0.618.

The most likely cluster identified by SaTScan [15] contains areas {248, 249, 251, 252, 253, 258}, while FleXScan identifies the area set {249, 251, 252, 253, 258, 262, 263, 265, 271} as its leading secondary cluster (with lowest p-value after the most likely cluster). Areas {248, 249, 250, 251, 252, 255, 258, 259, 262, 264} are included in the localities identified using join-count statistics in table 1, and in fact consist of neighbouring areas in Tameside, a local authority district in the south east of the region, with the district of Oldham to the North and with Stockport to the South. Figure 2 (of MSOAs in the three local government districts of Tameside, Oldham and Stockport) shows a cluster of MSOAs in the centre of the mapped sub-region, mostly in Tameside, all having posterior mean relative risks above 1.10.

The most likely cluster identified by FleXScan consists of the areas [587, 588, 590, 591, 593, 594, 597, 599], and the similar area set [590, 592, 593, 594, 595, 597, 599] is also the leading secondary cluster identified by SatScan. Areas [587, 589, 590, 592, 593, 594, 595, 596, 597, 599 and 601] are included in the areas in table 1, and consist of a set of areas in the coastal town of Blackpool. Figure 3 (of MSOAs in the three local government districts of Blackpool, Wyre and Fylde) shows this cluster of adjacent MSOAs at the westernmost centre of the plot, all having posterior mean relative risks above 1.15 except for area 589 with modelled relative risk of 0.994, but encompassed within surrounding higher risk areas.

The local join-count procedure also provides estimates of $\pi_{10i}$, which will be elevated when $H_i$ is elevated, but risk in the surrounding locality is relatively low. These may be considered as local high risk outliers, discordant in terms of health status from their neighbours. To demonstrate the contrasting risk patterns between the focus area and surrounding areas, we define $A_i$, encompassing areas adjacent to the focus area but not including that area.

Thus, table 2 shows the 12 MSOAs with $\pi_{10i}$ over 0.5, the modelled

<table>
<thead>
<tr>
<th>Index of focus area</th>
<th>$H_i$ posterior estimate</th>
<th>$\pi_{10i}$ posterior estimate</th>
<th>Modelling relative risk, $r_i$, in focus area (posterior mean)</th>
<th>Number of areas in surrounding locality (excluding focus)</th>
<th>Modelling relative risk $R_2i$ across rest of locality (excluding focus)</th>
<th>$\text{Pr}(R_2i&gt;1)$ (elevated risk in adjacent areas)</th>
<th>SMR across rest of locality</th>
</tr>
</thead>
<tbody>
<tr>
<td>532</td>
<td>0.853</td>
<td>0.583</td>
<td>1.379</td>
<td>5</td>
<td>0.920</td>
<td>0.248</td>
<td>0.768</td>
</tr>
<tr>
<td>620</td>
<td>0.800</td>
<td>0.582</td>
<td>1.358</td>
<td>2</td>
<td>0.890</td>
<td>0.239</td>
<td>0.490</td>
</tr>
<tr>
<td>439</td>
<td>0.787</td>
<td>0.580</td>
<td>1.288</td>
<td>6</td>
<td>0.888</td>
<td>0.113</td>
<td>0.702</td>
</tr>
<tr>
<td>772</td>
<td>0.974</td>
<td>0.550</td>
<td>1.805</td>
<td>4</td>
<td>1.009</td>
<td>0.492</td>
<td>0.739</td>
</tr>
<tr>
<td>554</td>
<td>0.739</td>
<td>0.545</td>
<td>1.233</td>
<td>6</td>
<td>0.847</td>
<td>0.091</td>
<td>0.693</td>
</tr>
<tr>
<td>406</td>
<td>0.804</td>
<td>0.521</td>
<td>1.299</td>
<td>6</td>
<td>0.917</td>
<td>0.234</td>
<td>0.856</td>
</tr>
<tr>
<td>862</td>
<td>0.896</td>
<td>0.514</td>
<td>1.469</td>
<td>4</td>
<td>0.977</td>
<td>0.411</td>
<td>0.695</td>
</tr>
<tr>
<td>763</td>
<td>0.743</td>
<td>0.509</td>
<td>1.252</td>
<td>3</td>
<td>0.887</td>
<td>0.216</td>
<td>0.869</td>
</tr>
<tr>
<td>158</td>
<td>0.855</td>
<td>0.506</td>
<td>1.329</td>
<td>7</td>
<td>0.956</td>
<td>0.329</td>
<td>0.877</td>
</tr>
<tr>
<td>740</td>
<td>0.960</td>
<td>0.503</td>
<td>1.145</td>
<td>7</td>
<td>1.022</td>
<td>0.551</td>
<td>0.754</td>
</tr>
<tr>
<td>744</td>
<td>0.765</td>
<td>0.502</td>
<td>1.244</td>
<td>4</td>
<td>0.924</td>
<td>0.271</td>
<td>0.782</td>
</tr>
<tr>
<td>198</td>
<td>0.749</td>
<td>0.502</td>
<td>1.251</td>
<td>5</td>
<td>0.900</td>
<td>0.209</td>
<td>0.532</td>
</tr>
</tbody>
</table>

Table 2: Suicide mortality, areas with highest estimates ($\pi_{10i}$) for outlier high risk.
relative risk $r_j$ (posterior mean) in the focus area, and posterior mean relative risk in the surrounding area, namely

$$R_{2j} = \sum_{j \in A} r_j e_j / \sum_{j \in A} e_j.$$  

Also shown are unsmoothed suicide SMRs across adjacent areas $\sum_{j \in A} y_j / \sum_{j \in A} e_j$ For all but one area, the probabilities that $R_j$ (average risk in the locality excluding the focus area) exceed 1 are under 0.5, whereas the probabilities $H_i$ of elevated risk in the focus area itself all exceed 0.7.

**Locality Risk Patterns under Alternative Spatial Weights**

Inferences from convolution or other area disease count models may be affected by the form of spatial interaction assumed. Two alternatives to binary adjacency are considered, which involve down weighting areas at greater distance from the focus area (with inter-area distances based on population centroids). These assume distance decay according to $w_{ij} = 1 / d_{ij}^\gamma$ ($\gamma > 0$) with values of $\gamma = 0.5$ and $\gamma = 1$ considered. These values are based on a preliminary analysis using model (4) to find an optimal value for $\gamma$ using a discrete prior over values $0, 0.1, 0.2, ..., 1.5$, which produced a posterior mean for $\gamma$ of 0.69.

We focus on elevated locality risk in particular, and table 3 summarises locality risk patterns under the two distance decay options.

The table considers only areas with $\tilde{\gamma}_{1i}$ over 0.70, ranked by $\tilde{\gamma}_{1i}$. The weighted averages of modelled relative risks across localities $L_i$ (centred on and including area $i$) now adjust also for distance decay as well as expected deaths, namely

$$R_{3j} = \sum_{j \in L_i} r_j w_j e_j / \sum_{j \in L_i} w_j e_j.$$  

Table 3 shows posterior mean $R_{3j}$ and probabilities that $R_j$ exceed 1. Unsmoothed suicide SMRs across the locality $\sum_{j \in L_i} y_j / \sum_{j \in L_i} e_j$ are defined as before.

There is considerable overlap between table 3 and table 1 in those focus areas identified as having both elevated "own area" risk (high $H_i$) and elevated risk across the locality also. Thus of the 13 areas with high $\pi_{1i}$ identified in table 1, 11 also appear as focus areas in the top panel (high distance decay $w_j$) of table 3, and the other two (areas 251, 590) are included in the broader localities listed there. All 13 cluster-centre areas identified in table 1 appear as such areas in the lower panel of table 3 (less marked distance decay).

**Results: Bivariate Spatial Clustering under Alternative Spatial Priors**

We now consider local join-counts for detecting bivariate risks that are both significantly elevated and also spatially clustered. Consider...
from scaling applied to ensure count later actual suicide, but considerably more frequent with average event suicide deaths \(y_{Ai}\) for 2006-10 as discussed above, and self-harm hospitalisations \(y_{Bi}\) for 2006-7 to 2010-11 (five financial years, with IDCN10 X66--X84 codes) across the 922 MSOAs in NW England (the data can be obtained at http://wwwapho.org.uk/resource). Expected hospitalisations \(e_{Ai}\) are based on England wide age specific rates, with scaling applied to ensure \(\sum e_{Ai} = \sum y_{Ai}\). Self-harm is often a precursor to later actual suicide, but considerably more frequently with average event count \(\overline{y}_{Bi} = 93\).

A convolution model is applied with \(y_{Ai} \sim \text{Poi}(\lambda_{Ai} f_{Ai}(A))\), \(y_{Bi} \sim \text{Poi}(\lambda_{Bi} f_{Bi}(B))\), but comparing two alternative procedures to provide stabilised estimates of relative risk. The first includes spatial borrowing of strength within outcomes, but without such borrowing between outcomes, and unrelated CAR and iid priors for each event

\[
\log(r_{Ai}) = X_i \beta_{Ai} + u_{Ai} + s_{Ai},
\]

\[
\log(r_{Bi}) = X_i \beta_{Bi} + u_{Bi} + s_{Bi},
\]

with

\[
\begin{align*}
\lambda_{Ai} &\sim N(0, \tau_{Ai}^2), \\
\lambda_{Bi} &\sim N(0, \tau_{Bi}^2), \\
\mathbf{s}_{Ai} \mid \mathbf{s}_{Bi} &\sim N\left(\sum_{i,j} w_{ij} s_{Ai} / S_i, \sigma_j^2 / S_j\right).
\end{align*}
\]

and

\[
\mathbf{s}_{Bi} \mid \mathbf{s}_{Ai} \sim N\left(\sum_{i,j} w_{ij} s_{Bi} / S_i, \sigma_j^2 / S_j\right).
\]

Spatial interactions \(w_{ij}\) are binary based on adjacency. The second procedure assumes \(\mathbf{s}_{Ai} = \mathbf{s}_{Bi}\) follow a bivariate CAR prior [17] with unknown within area covariance matrix

\[
\Sigma = \begin{pmatrix}
\sigma_{\lambdaAi} & \sigma_{\lambdaAi \lambdaBi} \\
\sigma_{\lambdaBi \lambdaAi} & \sigma_{\lambdaBi}
\end{pmatrix}
\]

and

\[
\Sigma^{-1} = \begin{pmatrix}
\sigma_{\lambdaAi}^{-1} & -\sigma_{\lambdaAi \lambdaBi} \rho_{AB} \\
-\sigma_{\lambdaBi \lambdaAi} \rho_{BA} & \sigma_{\lambdaBi}^{-1}
\end{pmatrix}
\]

with \(\Sigma^{-1}\) taken to be Wishart with 2 degrees of freedom and identity scale matrix.

The bivariate indicators

\[
b_{ABi} = I(r_{Ai} > 1, r_{Bi} > 1)
\]

are monitored in each case to provide bivariate join counts \(J_{ABi}\). From these one obtains indicators of elevated bivariate risk encompassing both the focus area and its surrounding locality.
The lower panel of table 4 shows two new cluster centres (197, 492), as compared to the upper panel, these being areas where self-harm risk (both observed and modelled) is high, and estimated suicide risk is pulled towards the risk for the more common outcome under the bivariate spatial prior. Using extra information about risk patterns provided by a more frequent outcome (or by intercorrelation between outcomes in general) is generally regarded as beneficial. This is a form of borrowing strength [19] enabling stronger inferences for an infrequent outcome. However, analysis such as that here, of potential impacts on inferences about clustering, may provide an additional facet for assessing sensitivity to alternative spatial priors.

**Criteria for Cut-off Points**

Considering the results of both the univariate and bivariate clustering analyses together, one may set out some criteria for choosing focus areas for high risk localities. The choice of cut-off for \( \pi_W \) (or \( \pi_{AB111} \) for bivariate outcomes) should be based on the profile of their ranked values, in conjunction with information about risk variation (e.g. the profile of \( H_i \) and \( R_i \)). The necessary interconnection with \( H_i \) follows from the relation \( H_i = \pi_W + \pi_{AB11} \).

Health outcome data for small areas vary considerably in the extent to which significant variations in area relative risk (and hence locality clustering) can be detected and this affects cut-off choice. For example, for the relatively rare suicide outcome, there are only 18 areas with \( \Pr(R_i > 1|y) = H_i \) exceeding 0.95, and a cut-off \( \pi_W > 0.7 \) was used, with a minimum \( H_i \) of 0.87 among the 13 areas above this cut-off. A slightly lower cut-off could be entertained, though the 17th ranked area in terms of \( \pi_W \) (with \( \pi_W = 0.69 \)) has a relatively low \( H_i \) of 0.79, below the threshold of \( H = 0.8 \) for elevated risk suggested by Richardson et al. [3]. The probabilities \( \Pr(R_i > 1|y) \) that the locality wide modelled SMRs exceed 1 are also relevant, provided the \( R_i \) are obtained for localities where both the focus and surrounding areas have elevated risk. All 13 areas with \( \pi_W > 0.7 \) have \( \Pr(R_i > 1|y) \) exceeding 0.95.

Whereas suicide is a rare outcome, self-harm is around 25 times more frequent. When a univariate clustering analysis (comparable to that carried out for completed suicide and reported on above) is carried out for self-harm, there are 275 MSOAs with \( H_i \) exceeding 0.95, and 16 MSOAs with \( \pi_W > 0.9 \), so a higher cut-off point could be used to detect high risk clusters for this outcome.

The locality wide modelled SMRs exceed 1 are also relevant, provided the \( R_i \) are obtained for localities where both the focus and surrounding areas have elevated risk. All 13 areas with \( \pi_W > 0.7 \) have \( \Pr(R_i > 1|y) \) exceeding 0.95.
For the bivariate outcome analysis (suicide and self-harm) without borrowing of strength between outcomes (e.g., Table 4 upper panel), there are 14 areas with $\Pr(\pi_{0i} = 1|y_i = 1,y_j,y_k) = H_{AB}i$, exceeding 0.95, and a relatively low cut-off of $\pi_{AB}i > 0.5$ was used. The implications of using a slightly lower cut-off point could be considered, since even in this analysis, the locality relative risks $R_{ai}$ and $R_{bi}$ are significantly elevated (above 1) at lower values of $\pi_{AB}i$ than the illustrative cut-off taken.

It follows from the above discussion that there are no simple rules for a low threshold $\pi_{i}$ or $\pi_{AB}i$ below which clustering is implausible. It depends on the profile of $H_i$ and $R_i$ as well as on the profile of $\pi_{i}$. Also relevant is the relative size of $\pi_{i}$ and $\pi_{AB}i$, the latter being the probability of a high risk area surrounded by low risk areas. Where an area has $\Pr(R_i > 1|y_i) < 0.9$, or $H_i < 0.75$, or $\pi_{AB}i > 0.5$ clearly exceeding $\pi_{1i}$ then high risk clustering becomes considerably less likely.

Conclusions

Small area disease models often use exceedance probabilities for each individual area to make inferences about risk patterns. However, elevated risk in an area may not necessarily extend to the surrounding locality. This paper has sought to identify areas where elevated risk extends to the broader locality using local join-count statistics. These statistics can identify local outliers as well as high risk cluster centres, and can be applied to assess high risk clustering in more than one health outcome.

The procedure here can be used in conjunction with a disease model where risk status is unknown, so enabling the clustering implications of contrasting likelihood and prior assumptions (e.g. regarding pooling between areas, and outcomes) to be assessed. In particular, inferences about clustering patterns in two outcomes considered jointly may well be influenced by alternative assumptions, particularly when a spatial prior borrows strength over outcomes as well as areas. Sensitivity of clustering inferences to alternative priors for spatial effects, such as the approach of Leroux et al. [20] in contrast to the convolution prior, also provides an additional area of research.

References