Science of the Total Environment 408 (2010) 4429–4439

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

# Spatial analysis of air pollution and cancer incidence rates in Haifa Bay, Israel

Ori Eitan [a], Yuval [a], Micha Barchana [b,d], Jonathan Dubnov [c,d], Shai Linn [d,e], Yohay Carmel [a], David M. Broday [a,*]

[a] Faculty of Civil and Environmental Engineering, Technion, Israel Institute of Technology, Haifa 32000, Israel
[b] Israel National Cancer Registry, Ministry of Health, Israel
[c] Haifa District Health Office, Ministry of Health, Israel
[d] School of Public Health, University of Haifa, Israel
[e] Epidemiology Unit, Rambam Medical Center, Haifa, Israel

ARTICLE INFO

*Article history:*
Received 13 October 2009
Received in revised form 14 May 2010
Accepted 17 June 2010
Available online 13 July 2010

*Keywords:*
Bayesian inference
Kriging
Lung cancer
Non-Hodgkin's lymphoma
$PM_{10}$
Population exposure
$SO_2$
Spatial randomness

ABSTRACT

The Israel National Cancer Registry reported in 2001 that cancer incidence rates in the Haifa area are roughly 20% above the national average. Since Haifa has been the major industrial center in Israel since 1930, concern has been raised that the elevated cancer rates may be associated with historically high air pollution levels. This work tests whether persistent spatial patterns of metrics of chronic exposure to air pollutants are associated with the observed patterns of cancer incidence rates. Risk metrics of chronic exposure to $PM_{10}$, emitted both by industry and traffic, and to $SO_2$, a marker of industrial emissions, was developed. Ward-based maps of standardized incidence rates of three prevalent cancers: Non-Hodgkin's lymphoma, lung cancer and bladder cancer were also produced. Global clustering tests were employed to filter out those cancers that show sufficiently random spatial distribution to have a nil probability of being related to the spatial non-random risk maps. A Bayesian method was employed to assess possible associations between the morbidity and risk patterns, accounting for the ward-based socioeconomic status ranking. Lung cancer in males and bladder cancer in both genders showed non-random spatial patterns. No significant associations between the $SO_2$-based risk maps and any of the cancers were found. Lung cancer in males was found to be associated with $PM_{10}$, with the relative risk associated with an increase of 1 $\mu g/m^3$ of $PM_{10}$ being 12%. Special consideration of wards with expected rates <1 improved the results by decreasing the variance of the spatially correlated residual log-relative risk.



## 1. Introduction

In 2001, the Israel National Cancer Registry presented data showing that incidence rates of several cancers in the Haifa district in the years 1984–1999 were above the national average rates (Barchana, 2001). Although the report indicated high morbidity rates, it did not suggest possible causes for these observations. Haifa Bay Area (HBA) has been Israel's primary industrial region since the 1930s. It hosts petroleum refineries (8 million tons/year), an oil fired power plant (425 MW), and several large petrochemical, chemical and agrochemical industries. It is also the home of many medium and small size factories, and of Israel's largest seaport (Fig. 1). These industries, as well as the dense population in the area, result in much traffic with a large proportion of heavy trucks. In general, emissions from motor vehicles increased sharply in the last 3–4 decades, mainly due to increased traffic volumes. Moreover, vehicle emissions in HBA are relatively high because of the steep city terrain and due to poor maintenance of part of the vehicle fleet. Consequently, public concerns were raised that the elevated cancer rates might be associated with environmental exposures of HBA residents to airborne pollutants originating from the industries and traffic in the area.

Evidence that air pollution is associated with various diseases and mortality causes has been accumulated since the 1970s. At first, temporal relationships between occurrence of adverse health effects and exposure to various air pollutants were sought, accounting for possible time lag between the supposedly triggering exposure event and the vital status data (Dockery et al., 1993; Nyberg et al., 2000; Schwartz, 2000; Sunyer et al., 2003; Pope et al., 2004; Brunekreef and Forsberg, 2005). In particular, associations between increased $PM_{10}$ concentrations and adverse health effects (including lung cancer) have been reported (cf. Dockery et al., 1993; Pope et al., 1995, 2002; Jerrett et al., 2005). High PM concentrations were found to promote lung cancer, possibly by chronic irritation of the lining of the respiratory tract or by decreasing the clearance rates in the lungs. Chronic exposure to substandard ambient PM concentrations was also found associated with increased morbidity and mortality. Similarly, $SO_2$ is a known respiratory irritant that may act as a promoter or co-carcinogen (Nisbet et al., 1984). It has been associated with respiratory outcomes and was suggested to have an indirect role in bioactivation and development of

* Corresponding author. Tel.: +972 4 8293468; fax: +972 4 8298898.
*E-mail address:* dbroday@tx.technion.ac.il (D.M. Broday).

0048-9697/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.scitotenv.2010.06.031

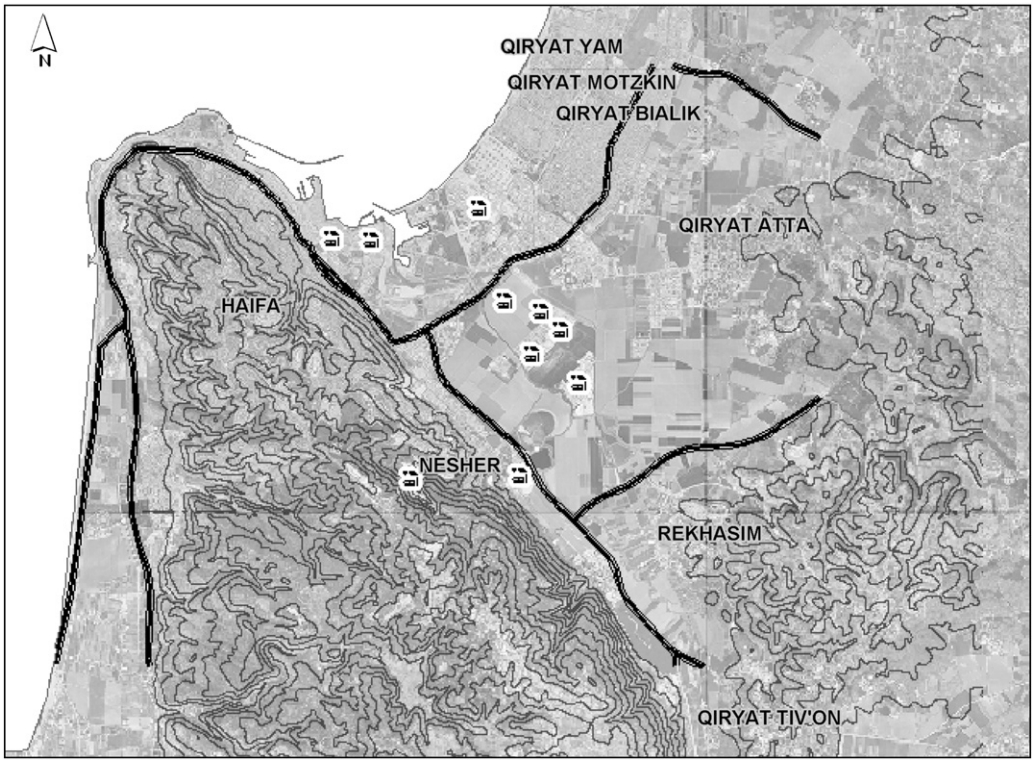


**Fig. 1.** The Haifa Bay study area with topographic contours every 50 m, main roads, the location of the major industrial plants, and the towns surrounding the industrial zone.

damage from exposure to co-pollutants, e.g. PM-bound carcinogenic polyaromatic hydrocarbons (Pott and Stöber, 1983; Bascom, 1996). Pope et al. (1995) suggested that lung cancer mortality was more strongly associated with sulfate particles than with simply fine PM. In HBA, like in many other regions worldwide, sulfate makes up the largest mass fraction of fine PM and is originated both from atmospheric oxidation of locally emitted $SO_2$ under conditions of high solar radiation and from long range trans-boundary transport. Spatial patterns of risk metrics can be represented by pollutant concentration maps combined with geo-referenced demographic data (Maslia et al., 1994; Georgopoulos and Lioy, 1994; Phillips et al., 1997). The spatial distribution of ambient pollutants can be obtained using atmospheric dispersion models or by spatial interpolation of monitoring data.

Due to the long latency period of most cancers, their numerous multi-factorial possible causes, and the common lack of detailed individual-level activity log files, the possibly small contribution of air pollution to cancer etiology makes the isolation of its effect extremely difficult (Doll and Peto, 1981). The influence of confounding factors such as genetic predisposition, smoking, life style, diet, occupation, etc., need to be carefully accounted for to unmask the role of air pollution on cancer if such a relationship exists (Samet et al., 2000; Pope et al., 2002; Jerrett et al., 2005; Vineis et al., 2007). However, normally there is a severe dearth of detailed long-term data on individual-level risk factors and on complete exposure history, especially in cases of chronic exposure to minute yet potent pollutants such as carcinogenic, mutagenic and genotoxic agents. In the absence of patient-specific data, composite level data are frequently used, with the underlying assumption that the aggregate approach provides exposure–response relationships that are reasonable surrogates of the individual's actual exposure–response relationships (e.g. Best et al., 2001). Furthermore, contemporary GIS techniques enable detailed mapping of estimated risk metrics and of observed health indicators, thus facilitating the study of relationships between them at different spatial scales.

This work aims at assessing whether spatial patterns of chronic exposure to ambient pollutants at substandard levels can be linked to the spatial distribution of observed health indicators (cancer incidence rates). However, the complex topographic–meteorological conditions in HBA, the relatively small and heterogeneous population and its patchy spread in the study area, and the lack of any individual-level data on possible confounders apart from a complete residential history information called for the development of a tailored method for assessing the relationship between chronic exposure to airborne pollutants and cancer incidence rates in HBA. We believe that in many regions worldwide such data limitations are common and hence the chain of procedures presented may find wide application.

## 2. Materials and methods

### 2.1. Morbidity data and health indicators

The morbidity data available for this study was an individual-level database that contained the geographic coordinates of the residences of cancer patients that were diagnosed between 1995 and 1999. Overall, this dataset included 1452 individual cases that were nearly equally distributed among the three cancer types studied (Table 1).

The Israel National Cancer Registry (INCR) is a population based central tumor registry. Reporting to the registry is mandatory since 1982 and covers all public and private medical facilities in the country (medical institutions and pathology laboratories). Cancer mortality is evaluated against data obtained from district health offices and from the central population registry of the Ministry of Interior. The most recent survey (Fishler et al., 2003) revealed that the completeness of the records was above 94%. Apart from medical data, the only demographic attributes collected are the place of birth, immigration date, detailed residence history, religion, and a few other personal data. Nonetheless, smoking habits, occupational history and other common risk factors and/or confounders (e.g. body mass index) are not recorded. The relatively small population in HBA (ca. 500,000) implies small number of cancer cases and poses difficulties in constructing reliable health indices and in obtaining significant statistical inference. Furthermore, the population at risk in HBA is

**Table 1**
Statistics of cancer incidence rates (1995–1999) and of demographic data in HBA. The complete database includes data for all the 143 wards. Wards with no cases and expected incidence rate <1 were excluded from the reduced datasets.

| | Gender | Observed | Mean | Median | Sample variance | Minimum | Maximum | Expected (based on Israel average rates) | No. of wards excluded from the reduced datasets | % of the total population in the excluded wards |
|---|---|---|---|---|---|---|---|---|---|---|
| Population | | 473,300 | 3309.8 | 3400 | 2217.8 | 200 | 7600 | | | |
| SES | | | 12.03 | 12 | 21.35 | 0 | 20 | | | |
| Lung cancer | M | 310 | 2.17 | 2 | 3.94 | 0 | 8 | 403.5 | 12 | 2.3 |
| | F | 153 | 1.07 | 1 | 1.46 | 0 | 6 | 204.9 | 31 | 11.4 |
| NHL | M | 257 | 1.80 | 2 | 3.06 | 0 | 10 | 197.7 | 21 | 5.9 |
| | F | 228 | 1.59 | 1 | 2.76 | 0 | 7 | 256.4 | 24 | 7.6 |
| Bladder cancer | M | 371 | 2.59 | 2 | 5.65 | 0 | 13 | 404.2 | 16 | 3.6 |
| | F | 97 | 0.68 | 0 | 0.92 | 0 | 5 | 104.2 | 59 | 32.1 |

very heterogeneous, comprising of many ethnic groups with diverse life styles and cultural traditions, and living in communities that are distributed very heterogeneously in the study area. The large number of immigrants from the former Soviet Union that arrived to Israel in the mid 1990s further complicated the analysis.

The database obtained from the Israel National Cancer Registry contained records of individuals that met three criteria: (a) the patients were diagnosed with either lung cancer, bladder cancer, non-Hodgkin's lymphoma (NHL), or with more than one of these cancers (<1% of the cases), in the years 1995–1999, (b) the patients lived in the study area for at list 10 years prior to diagnosis, and (c) the patients live in Jewish communities. Regarding (a), the cancer types were selected based on reported associations between these cancers and exposures to petrochemical and other industrial emissions (Sharp et al., 1996; Yang et al., 2000; Pope et al., 2002; Johnson et al., 2003; Guo et al., 2004; Soll-Johanning and Bach, 2004) as well as with traffic emissions (Nyberg et al., 2000; Reynolds et al., 2004; Jerrett et al., 2005). In addition, these cancers are relatively abundant among the HBA population and their rates are among the highest in Israel (Epstein et al., 1984). Due to differences between the adjusted rates in males and females and the possible gender-distinct exposure routes, gender-specific rates for each of the cancers were calculated. Regarding criterion (b), the requirement of at least 10 years of residency in the study area prior to diagnosis relates to the latency period of cancer. Considering only cancer patients who resided in HBA for at least 10 years prior to diagnosis works towards minimizing possible confounding factors, since it omits from the study all the immigrants from the former USSR that arrived to Israel during the 1990s and could have been exposed to uncontrolled/unknown risk factors in their former places of residence. Finally, criterion (c) excludes from the database cancer cases that resided in three non-Jewish communities in the study area due to their very distinct lifestyle. The population of these communities is less than 5% of the total population in the study area and the number of cases excluded range from <1% and up to 3% depending on the cancer type and gender. All the cases in the rest of the communities were included in the study regardless of their ethnicity. Therefore, the morbidity dataset contained a total of 1417 cases (individual-level) for the three cancers.

The nature of these data (i.e. the lack of any individual-level risk factors) and the need to standardize them and to account for possible confounders prescribed transforming the individual-level morbidity data into ward-based morbidity indicators. Standardized Incidence Rates (SIR) were calculated for the Central Bureau of Statistics census wards in the study area. The use of SIR enabled standardization for age and population density (Tamir, 1987; Ginsberg and Tulchinsky, 1992; Kokki et al., 2001; Ginsberg et al., 2003). Standardization for ethnicity (by parents' country of birth) had negligible effect on the crude SIR. Yet SIR are known to have some inherent problems, especially in small areas where the expected rates are small and result in a large variance and lack of stability of the SIR (Julious et al., 2001; Meza, 2003; Ugarte et al., 2006). Overall, out of the 143 wards in the study area between 32 and 80 (depending on the cancer type and gender) had no cases whatsoever and were assigned a zero SIR. Most of these wards contained very small population (mostly less than 750 inhabitants) and no cases were expected irrespectively of the cancer risk. Data from these last wards represent "noise" that may hamper the efforts to assess if there are relationships between the health and risk indices. Therefore, we constructed new datasets from which such wards (specific for each cancer type and gender) were removed (Table 1). Wards with SIR = 0 in which the population was large enough that cases could be expected based on the rates in the standard population (whole Israel) were retained. The concise datasets (cancer and gender specific) are termed hereinafter the reduced datasets and contain distinct lists of wards. In the complete database, SIR for each cancer and gender are reported for all the wards.

### 2.2. Spatial non-randomness

Since spatially random morbidity patterns are not likely associated with the inherently non-random air pollution data, the ward-based cancer SIR were tested for global clustering, to distinguish those that show spatial non-randomness. Testing the morbidity data for global spatial randomness enabled the elimination of those cancer types with random spatial distribution that could not be expected to relate to the coherent spatial air pollution patterns. The search for general clustering in each of the datasets is closely akin to the assessment of autocorrelation (Lawson, 2001). We used Moran's $I$ correlogram (Oden, 1984) and Tango's $C$ test (Tango, 2000; Kulldorff et al., 2003), which due to their different inner structure complement each other. Monte Carlo permutations (Good, 1994) provided the null distribution against which the actual measures were compared. A multiple testing problem arises in both methods that usually requires prescription of a more stringent significance level (Garcia, 2004), i.e. $\alpha<0.05$. Nonetheless, as we use these tests only as a pre-filter for elimination of implausible cases, we chose to use $\alpha = 0.05$ and thus to unduly elevate the chances of detecting spatial autocorrelations and increase the probability of a type II error. However, the final inference regarding possible associations between the risk metrics and cancers incidence rates with spatial distributions that do show some spatial autocorrelation is not compromised.

### 2.3. Risk metrics and confounders

HBA is characterized by topographic complexity, hence utilization of standard atmospheric dispersion models is not advisable. Moreover, since a complete emissions inventory is not available for the region, an exclusive dispersion model could not be developed for HBA. However, owing to the relatively dense air quality monitoring network (the mean spatial monitoring interval ranges from 3.2 to 5 km depending on the pollutant) we spatially interpolated the monitoring data and generated long-term mean concentration maps of $SO_2$ and $PM_{10}$ (Yuval et al., 2005; Yuval and Broday, 2006). The correlation between the two concentration maps is low, unlike the

high correlation among the concentration maps of $PM_{10}$, $NO_x$ and $O_3$ (Yuval and Broday, 2006). The latter reflects the fact that $SO_2$ is emitted in HBA predominantly by the heavy industry whereas $PM_{10}$ is emitted from multiple sources: industrial, traffic-related and natural (Yuval et al., 2008). The uncorrelated $SO_2$ and $PM_{10}$ maps were used to produce two independent risk maps. Yet while it is noteworthy that $PM_{10}$ and $SO_2$ were found to have independent associations with lung cancer (Beeson et al., 1998), it should be emphasized that $SO_2$ is not a carcinogen by itself, unlike some PM components such as transition metals and adsorbed organics. Nevertheless, $SO_2$ can be regarded as a proxy to other (carcinogenic) pollutants that are co-emitted from the same sources. Moreover, it should be realized that since relative spatial patterns are used as risk metrics, the risk maps are meaningful even if the ambient concentrations are lower than the standard, and regardless of the general decrease in the sulfur content of fuels over the years and the related noticeable decline in ambient $SO_2$ concentrations.

Cancer is known to have extended latency periods that in HBA (like in many other places worldwide) goes back in time beyond the available air pollution records. Since the selected cancer types have latency periods of at least 10 years (Sharp et al., 1996), risk maps that are based on concentrations measured in 1996–2002 (for $SO_2$) or 2002–2004 (for $PM_{10}$) may not be relevant if the spatial patterns have changed over the years. Yet, since the locations of the main sources in the study area as well as the average meteorology and topography did not change over the years, spatial air pollution patterns are not expected to change much if there has not been a considerable change in the relative emission intensities among different sources. To test the validity of the pollutant concentrations maps in relation to chronic exposures that may promote health effects characterized by long latency periods, we examined the consistency of the spatial patterns over time. The variations between the concentration maps for the separate years were compared to verify that the main features of the spatial $SO_2$ and $PM_{10}$ patterns persist over long times. In particular, the Pearson correlation between any two annual average pollutant specific concentration maps was calculated at the pixel level ($N = 1{,}000{,}000$). Correlations between the yearly maps and a multi-year averaged concentrations map (based on pollutant specific data availability) were also calculated.

$SO_2$ has been monitored in HBA at 17–20 stations since 1996. Hence, maps of $SO_2$ could be produced for the years 1996–2002. $PM_{10}$, on the other hand, is monitored in 8 stations only since 2002, and valid $PM_{10}$ interpolation maps could be produced therefore only for the period 2002–2004. All the maps were produced using the kriging interpolation technique (for details see Yuval et al., 2005; Yuval and Broday, 2006). Desert dust has a large contribution to the total $PM_{10}$ levels in Israel (Dayan and Levy, 2005). Since the chemical composition and consequently the related health effects of dust particles are different than those of anthropogenic PM from combustion sources, conspicuous dust storms were removed from the $PM_{10}$ database (Yuval and Broday, 2006). To produce a ward-based risk metrics (exposure to $SO_2$ and $PM_{10}$), the interpolated concentration data were averaged over each statistical ward. Thus, the risk maps and the health indicator maps had a common spatial resolution. In general, the ward-based risk metrics tend to represent better small and regularly shaped wards in areas that are characterized by fairly spatially homogenous ambient concentrations rather than larger and irregularly shaped wards with highly heterogeneous concentrations.

The ward-based long-term average concentrations, representing metrics of cumulative exposure, were overlaid over demographic data and other ward-scale risk factors. Since no data on the actual individual-level exposures or risk factors were available, we used the group level ward-base socioeconomic status (SES) ranking, obtained from the Central Bureau of Statistics. The SES ranking reflects various socio-economical measures that characterize the ward's population, including income, parents' education, apartment size, possession of appliances, car ownership, etc. Hence, the SES ranking represents, to some extent, few individual risk factors and possible confounders. A higher SES (range from 1 to 20) signifies a higher socioeconomic status, which is known to be negatively correlated with parental smoking and child exposure to environmental tobacco smoke. It should be noted that the statistical wards are designed *a priori* to include relatively homogeneous population with respect to ethnicity, lifestyle, social class, and deprivation. Alternatively, following recent works (e.g. Dabney and Wakefield, 2005; Tzala and Best, 2008) we utilized the spatial distribution of SIR of lung cancer in males as a proxy of risk from smoking when modeling the rates of other cancers.

### 2.4. Testing for associations between risk and morbidity indices

Relationships between spatial patterns of the risk metrics and those of the spatially non-random cancer morbidity indices were estimated using a Bayesian hierarchical regression modeling framework (Best et al., 2001; Pascutto et al., 2000), implemented in WinBUGS (Spiegelhalter et al., 1998). Accordingly, the cancer risk of each ward is estimated by borrowing information from the other wards. This leads to improved risk estimates over those obtained either by non-hierarchical models that treat each ward independently or by pooling all the wards together and ignoring the between-area variability (Richardson and Best, 2003). In brief, we assume that the ward-specific baseline risk and relative risk are each realizations of distinct joint probability distributions of the unknown baseline and relative risks. The baseline risk represents the inherent risk due to area-level factors such as age distribution, ethnicity, and the residual effect of unobserved risk factors beyond the known or suspected environmental risk factors. Typically for rare diseases, a Poisson distribution is assumed for the disease risk. The Bayesian hierarchical model took the following form

$$\begin{aligned} &O_i \sim \text{Poisson}(\mu_i) \\ &\log\mu_i = \log E_i + \alpha + \sum_j \beta_j X_{ij} + \theta_i \end{aligned} \quad (1)$$

were $O_i$ and $E_i$ are the observed and expected cancer incidence rates in the $i$*th* ward, respectively ($E_i$ are the group level fixed effects), $\alpha$ is the mean log-relative risk of cancer for an unexposed person in the study region relative to the standard population, $\beta_j$ is the log-relative risk associated with a unit increase in exposure to the $j$*th* covariate $X_{ij}$, and $\theta_i$ are the random effects with exp $\theta_i$ the residual (unexplained) relative risk in area $i$. The covariates may be known risk factors, confounding variables, and effect modifiers. The latter may lead to a type I error — a false positive conclusion that the cancer rates are associated with the covariates. Due to lack of individual-level data on effect modifiers (lifestyle, diet, housing conditions, smoking, etc.), we used the between-area variation of the SES ranking as an area-level risk factor that confounds the effect of environmental exposures.

Naturally, some area-level risk factors are unobserved or unknown. Unmeasured or inadequately modeled area-level risk factors that have a spatial structure could affect the calculation of the relative risk of the environmental stressors. A possible way to assess if such a spatial structure exists in the residual random effects is to incorporate a spatial random effects component that accounts for the unexplained residual spatial structure of the health effects. For example, Pope et al. (2002) incorporated a metropolitan-based spatial random effects component and were successful in removing the spatial structure of the residual variation of the health effects model after accounting for the relative risk of particulate air pollution. On the other hand, Jerrett et al. (2005) could not remove the residual spatial autocorrelation in the random effects component after accounting for individual covariates and long-term exposure to $PM_{2.5}$. To explore this issue

we implemented a conditional autoregressive (CAR) convolution prior model (Besag et al., 1991) that accounts for both the spatial dependence and the unstructured heterogeneity of the random effects term. Namely,

$$\theta_i = S_i + H_i \quad (2)$$

where $S_i$ accounts for unmeasured/unknown risk factors that are spatially correlated and $H_i$ accounts for unmeasured/unknown risk factors that are independent across the study area.

For each cancer type and gender the models were employed twice, once using data from all the 143 wards and once using the reduced datasets. All the regression models were run with the between-area variation of the SES ranking as a risk factor. The models in which the SES ranking was the sole risk factor served as the benchmark models against which models that accounted also for environmental risk factors were compared. The results reported correspond to runs of 20,000 sweeps after a 30,000 iterations burn-in period. Two chains were run for each model, so the results are based on 40,000 samples. Brooks–Gelman–Rubin diagnostics (Brooks and Gelman, 1998) as well as graphical checks of the chains were performed to assess convergence. The relative risks (RR) reported are the quartile percentiles and the median and mean values. For comparing the different models we calculated the deviance information criterion (DIC) (Spiegelhalter et al., 2002). Although this criterion cannot be viewed as an absolute measure of the superiority of any one model, it does give an indication of the relative fit of a set of candidate models (Best et al., 2001), since models with a smaller DIC are better supported by the data (models with DIC within 1–2 of the best model are also strongly supported; Spiegelhalter et al., 2002). Finally, to assess if we accounted for all the relevant risk factors we tested if the variance of the residual log random effects term is spatially random. This was done both by performing a spatial non-randomness test on $\exp(\theta_i)$ and by calculating the ratio of the variance of the spatially structured random effects to the total variance of the random effects (the sum of the variance of the structured and the unstructured terms). If the residual relative risk,

$$\text{residual } RR_i = \exp(H_i + S_i), \quad (3)$$

shows a spatial structure, the presence of additional, unaccounted for spatially correlated risk factors cannot be overlooked and further covariates must be sought.

## 3. Results

### 3.1. Morbidity indices

Fig. 2 shows the morbidity maps for males. Some morbidity maps reveal a mixed pattern with what may look like a few clusters. Preliminary visual inspection does not suggest that the distance from the major industrial sources and from major roads in the study area (Fig. 1) are major risk factors. Similar maps were obtained for female SIR of the three cancers (not shown). Moran's $I$ and Tango's $C$ tests were run separately on each of the morbidity indices (cancer type and gender). Table 2 reveals that the spatial patterns of the SIR of lung cancer and of bladder cancer in males are significantly non-random whereas for NHL neither the Moran's $I$ nor the Tango's $C$ tests revealed any significant spatial non-randomness. Following these results, we did not formally consider NHL in males as a candidate cancer indicator. Indeed, the Bayesian hierarchical regression model suggests an insignificant relation between NHL and the environmental risk factors. Regarding female morbidity, only for bladder cancer the Tango's test revealed significant non-randomness. For NHL in females the result is marginal, however considering Bonferroni's adjustment for multiple testing when testing for significant associations without pre-established hypotheses (Perneger, 1998), this indicator can also be disregarded. This was supported by the results of the Bayesian regression model.

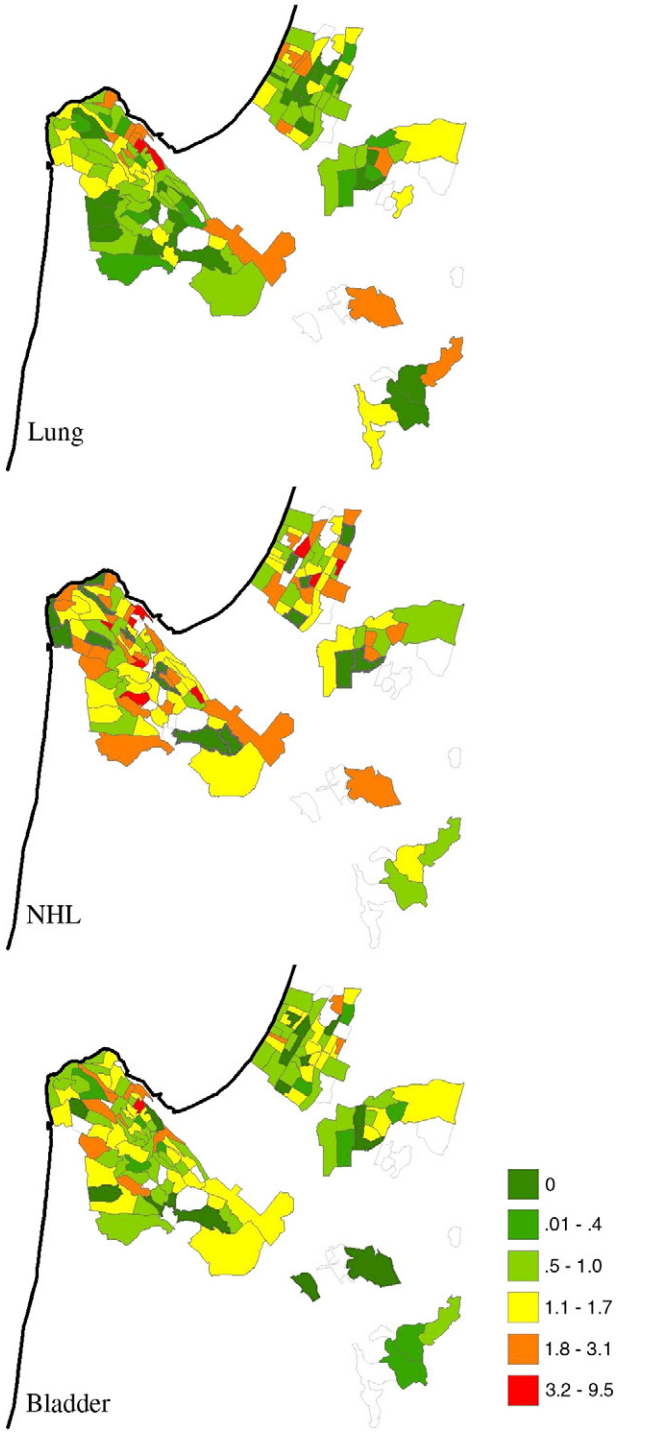


**Fig. 2.** Crude SIR maps for males: (a) lung cancer, (b) bladder cancer, and (c) NHL. All the colored wards, including those with zero reported cases, were included in the reduced datasets.

### 3.2. Risk metrics

Fig. 3 depicts the spatial patterns of mean $SO_2$ concentrations in the years 1996–2002. Mean $SO_2$ concentrations have significantly declined from 1996 to 2002, mainly due to the gradual shift towards using very

**Table 2**
*P*-values of the testing for spatial non-randomness. *P*-values below 0.05 are assumed statistically significant and imply non-random pattern.

| Risk factor | Cancer type | | Moran's *I* | Tango's *C* |
|---|---|---|---|---|
| SIR | Bladder | Male | 0.005 | 0.005 |
| | | Female | 0.2 | 0.03 |
| | Lung | Male | 0.001 | 0.001 |
| | | Female | 0.3 | 0.3 |
| | NHL | Male | 0.2 | 0.2 |
| | | Female | 0.3 | 0.06 |
| $SO_2$ | | | 0.001 | 0.001 |
| $PM_{10}$ | | | 0.001 | 0.001 |
| SES | | | 0.001 | 0.001 |

low sulfur content fuels over the years. A normalized concentration map (Fig. 4) was produced by the following transformation

$$c_{norm} = \frac{c - c_{low}}{c_{high} - c_{low}} \quad (4)$$

with $c_{low}$ and $c_{high}$ the minimum and maximum observed concentrations. The spatial pattern of the normalized concentrations is fairly constant and is characterized by elevated concentrations downwind (to the southeast) from the refineries and the petrochemical complex. Overall, there is a good correlation among the maps (Table 3) with the exception of the 1998 map that shows distinct low concentrations at downwind locations. It seems, therefore, that the relative spatial pattern of mean $SO_2$ concentrations in HBA has been relatively consistent over the years. This suggests that the average 1996–2002 $SO_2$ map (the lower right plate in Fig. 4) can be used as a chronic population exposure metric for $SO_2$ and for pollutants co-emitted with $SO_2$ in earlier years, i.e. during and before the latency period. This map is used therefore as the underlying layer of the risk map.

Similarly, the spatial patterns of the relative annual mean $PM_{10}$ levels in HBA in the years 2002–2004, and of the 3 years average, show persistence and resemble each other in a general sense (Fig. 5). Indeed, Table 4 reveals a good correlation between the $PM_{10}$ maps, suggesting that the spatial pattern of mean $PM_{10}$ concentrations is preserved over the years. This result is assumed to

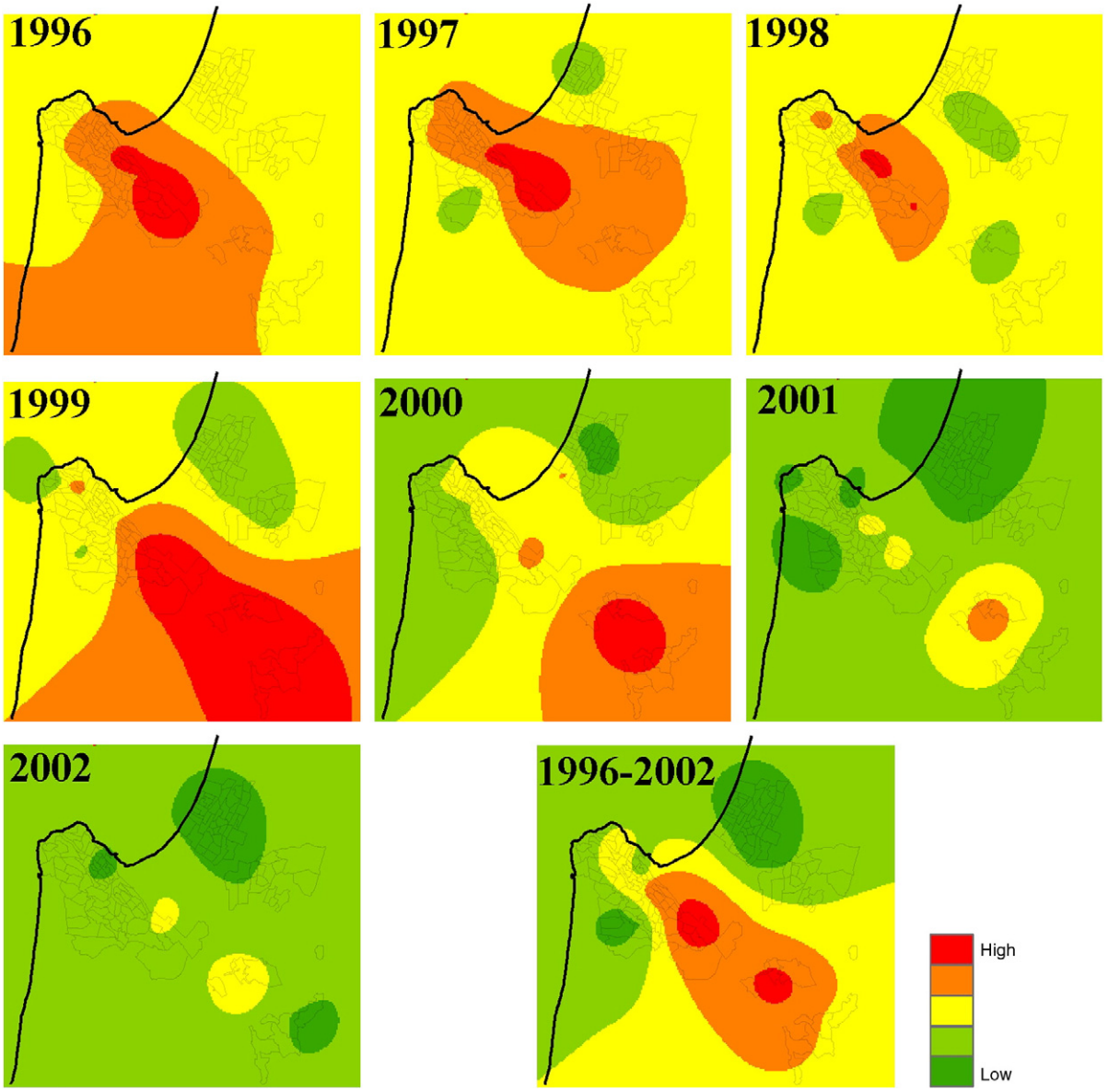


**Fig. 3.** Spatial patterns of annual mean $SO_2$ concentrations for the years 1996–2002, and of the average of the whole period (quintiles, range 1.8–14.7 μg/m$^3$). The maps were produced by an optimal spatial mapping method (Yuval et al., 2005). The decline in the concentrations results from the shutdown of the Nesher cement factory (located at the center of the study area) and from the continuous decrease of sulfur content in fuels.

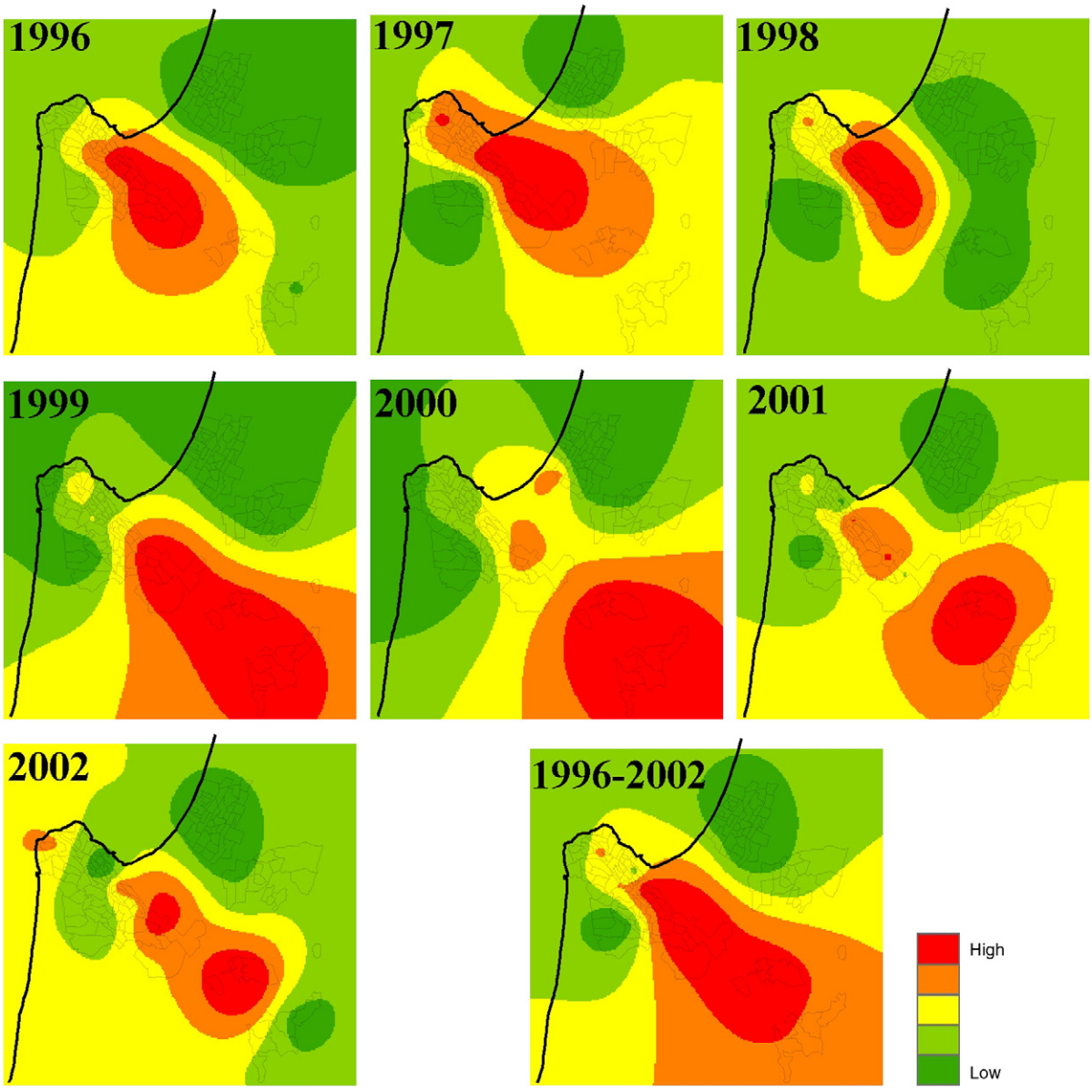


**Fig. 4.** Spatial patterns of normalized annual mean $SO_2$ levels for the years 1996–2002, and of the average of the whole period (quintiles, range 0–1). Each map was normalized relative to its own range of values.

hold also for previous years but due to lack of data from earlier years could not be verified. Hence, we use the relative mean $PM_{10}$ concentration map (the lower right plate in Fig. 5) as the underlying layer of the risk map from long-term exposure to $PM_{10}$ in HBA. It should be noted that accounting for grid-point mean concentrations is practically equivalent to using the accumulated potential chronic exposure,

$$\overline{c} = \frac{1}{T}\int_0^T c(t)dt. \quad (5)$$

**Table 3**
Pearson correlation coefficients between the yearly mean $SO_2$ maps.

| | 1996 | 1997 | 1998 | 1999 | 2000 | 2001 | 2002 | 1996–2002 |
|---|---|---|---|---|---|---|---|---|
| 1996 | 1 | | | | | | | |
| 1997 | 0.65 | 1 | | | | | | |
| 1998 | 0.73 | 0.6 | 1 | | | | | |
| 1999 | 0.66 | 0.54 | 0.29 | 1 | | | | |
| 2000 | 0.36 | 0.5 | 0.08 | 0.89 | 1 | | | |
| 2001 | 0.56 | 0.55 | 0.09 | 0.83 | 0.81 | 1 | | |
| 2002 | 0.69 | 0.49 | 0.27 | 0.54 | 0.36 | 0.76 | 1 | |
| 1996–2002 | 0.75 | 0.74 | 0.43 | 0.95 | 0.86 | 0.89 | 0.68 | 1 |

Figs. 4 and 5 were produced on an identical grid thus enabling a quantitative comparison between them. The correlation between these maps is relatively low (−0.16), suggesting that the sources of these pollutants, the processes governing their dispersion, or both are disparate. This corroborates the emissions inventory reports for HBA, which specify $PM_{10}$ emissions from multiple sources whereas $SO_2$ is mostly emitted by the heavy industry. Thus, the lower right plates in Figs. 4 and 5 can be viewed as two independent spatially distributed risk factors that represent chronic inhalation exposures to distinct sources and/or pollutants.

Ward-based risk metrics were produced by transforming the full grid risk maps into ward-based risk maps using (due to lack of any

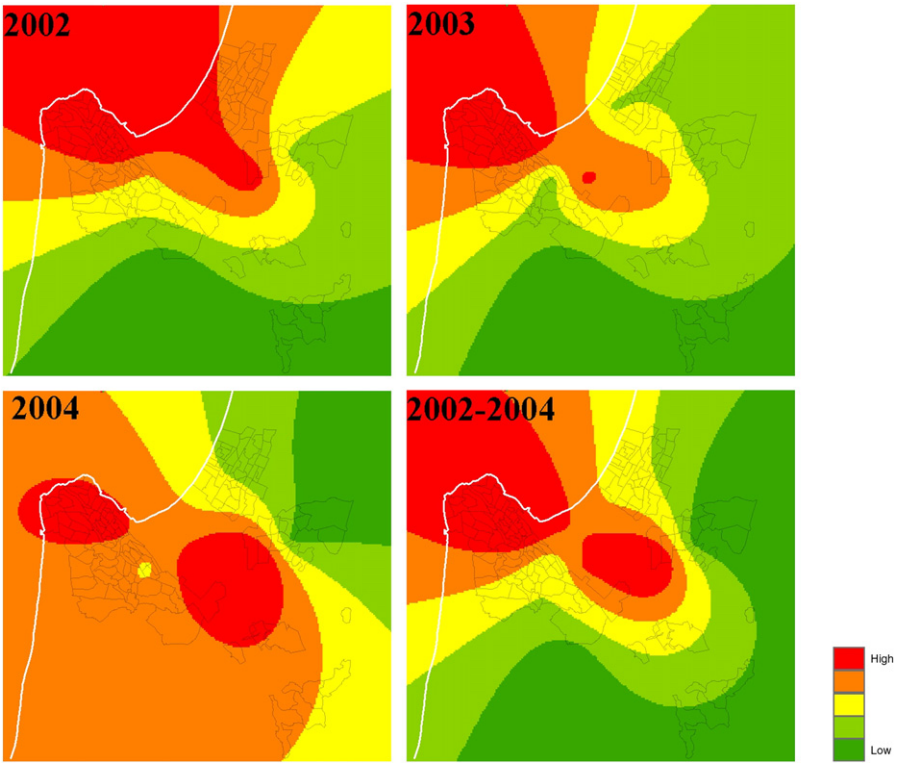


**Fig. 5.** Spatial patterns of normalized annual mean $PM_{10}$ concentrations for the years 2002–2004, and of the whole period average (quintiles, range 0–1). Observations ranged between 28.7 and 41.2 μg/m$^3$. Each map was normalized relative to its own range of values.

other information) a homogeneous (=1) averaging factor (Fig. 6). Namely, within each ward the risk from the environmental stressor was obtained by assigning the areal average exposure to the ward centroid rather than by weighting according to intra-ward population density variation. Tests for non-randomness of the spatial distribution of the ward-based risk metrics show that their spatial patterns in the study area are clearly non-random (Table 2). This result is not surprising since the pollutant sources (locations, intensities, etc.) and the physicochemical processes that govern their dispersion (meteorology, topography) are spatially non-random. Likewise, the spatial pattern of the ward-based SES rankings is also not random (Table 2), supporting the notion that it may represent a possible confounder.

### 3.3. Association between spatial patterns of cancer and risk indices

Table 5 depicts the results obtained by the Bayesian regression models when using the complete database. The disease-specific models that accounted for the between-area variation of the SES ranking as a sole risk factor served as the benchmark against which all the other modes were compared. Only for lung cancer in males, a more complex model (accounting for exposure to $PM_{10}$) performs better (i.e. has a lower DIC) than the benchmark model. In fact, all the models that accounted for both SES and $SO_2$ as risk factors had higher DICs than those of the corresponding benchmark models. Hence, none of the three cancers studied seem related to long-term environmental exposures to $SO_2$ or its co-emitted pollutants in HBA. Moreover, accounting for $PM_{10}$, $SO_2$ and SES as risk factors was not superior to the models that included only $PM_{10}$ and SES as risk factors. Therefore, Table 5 reveals that only incidence rates of lung cancer in males may be related to long-term exposure to ambient concentrations of $PM_{10}$, whereas models that accounted for other environmental exposures as risk factors were not supported by the data.

**Table 4**
Pearson correlation coefficients between the yearly mean $PM_{10}$ maps.

| | 2002 | 2003 | 2004 | 2002–2004 |
|---|---|---|---|---|
| 2002 | 1 | | | |
| 2003 | 0.91 | 1 | | |
| 2004 | 0.16 | 0.42 | 1 | |
| 2002–2004 | 0.87 | 0.95 | 0.62 | 1 |

Applying the Moran's *I* and Tango's *C* methods for assessing spatial randomness, the residual RR, Eq. (3), was found to be spatially autocorrelated, i.e. it shows spatially structured overdispersion with respect to the homogeneous Poisson model. This result was consistent for all the models. The spatial non-randomness of the residual RR suggests the presence of additional spatially varying unaccounted for risk factors. Similarly, the ratio of the spatially correlated to the total variability of the random effects (Frac.Spatial in Table 5) reveals that the marginal between-area variance of the spatial component of the log residual RR is larger than the variance of the unstructured heterogeneity. For example, this ratio for bladder cancer in both males and females is about 1 irrespectively of the model, supporting the notion that there might be spatially structured unaccounted for risk factors that may be associated with bladder cancer. Failure to account for the excess of zeros in the health data may bring on overdispersion in simple homogeneous Poisson models (Ugarte et al., 2006). It is therefore essential to assess to what extent the extra Poisson variability is due to the excess of zeros.

To evaluate the impact of the zeros in the database we ran all the models again using the reduced datasets. Results of these models are reported in Table 6. In agreement with the results presented in Table 5, the RR associated with the SES ranking for lung cancer in males is robust and <1. Namely, increased lung cancer incidence is seen in the more deprived populations. Furthermore, without exception, the RR associated with the between-area SES variation for all the models reported in Table 6 are larger than those reported in Table 5. It is, however, impossible to compare the models that use the complete vs. the reduced datasets based on their pertinent DICs. Nonetheless, for lung cancer in males the model that accounted for

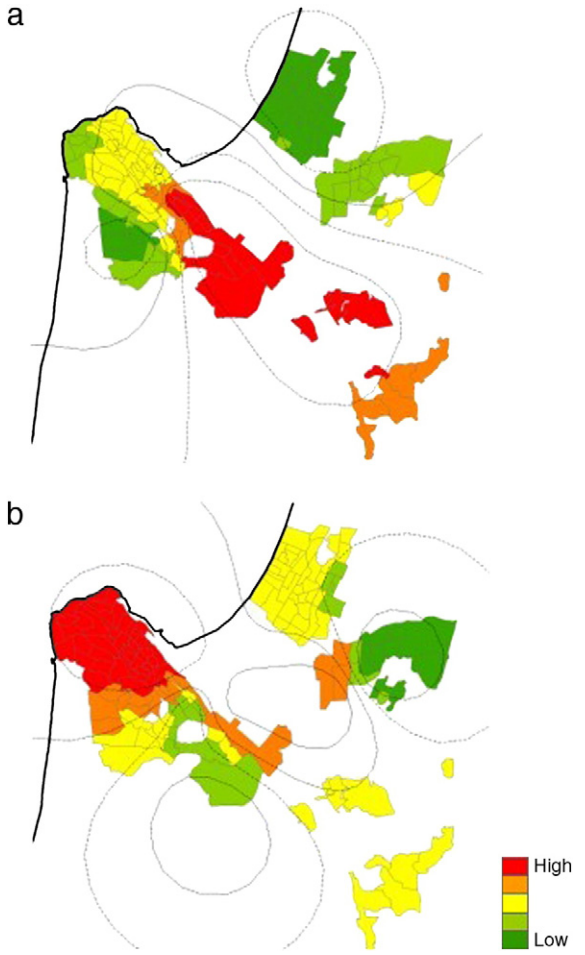


**Fig. 6.** Spatial pattern of the ward-based risk from long-term exposure to (a) $SO_2$ and (b) $PM_{10}$ ambient concentrations, and the contours of the normalized concentrations from which they were derived.

prolonged exposure to $PM_{10}$ as a risk factor obtained the lowest DIC score in both Tables 5 and 6 whereas for bladder cancer in both genders, accounting for environmental exposures as risk factors was not supported by the data. Still, even for lung cancer in males, the residual RR in the most supported ($PM_{10}$-SES) model reveals a non-random spatial structure and a larger contribution to the total variability that comes from the spatially correlated heterogeneity (Frac.Spatial changed only marginally from 0.62 to 0.59 in Tables 5 and 6, respectively). This suggests that there may be a further spatially structured risk factor in HBA. This conclusion holds for all the models reported in Table 6.

## 4. Discussion

Due to data availability limitations, expected to be common in many places worldwide, we had to tailor common methods for studying associations between cancer incidence rates and long-term exposure to air pollutants, and fit them to the data attributes available to us. In particular, we lacked data on any individual-level risk factors, such as work place, life style, smoking habits and genetic susceptibility. Hence, we were forced to use a group-based risk metrics and confounders within an ecological approach that requires relatively low-level data but likely results in considerable exposure error. As such, our results are inherently less conclusive than those of longitudinal cohort studies, and cannot reveal etiological pathways. Actually, at the individual-level, cancer causality may have no relationship with existence or absence of an association at the group level (Rothman and Greenland, 1998) although according to Hill's criteria such associations if supported by other consistent evidence can contribute to causal associations.

Tests for non-randomness (global clustering) were used to filter out those cancers whose rates showed spatially random patterns, as these cancers are very likely not associated with the coherent patterns of the risk metrics. This step served to diminish the problem of multiple testing, commonly encountered in such situations. Although methods for assessing non-randomness and/or global clustering are well established in epidemiology (Jacquez et al., 1996), we are unaware of previous works that utilized these techniques as an intermediate step before looking for associations between air pollution and morbidity. Normally, such tools are used to test the spatial autocorrelation of the residual random effects after applying different levels of control for confounders (e.g. Jerrett et al., 2005).

The risks we tried to assess, i.e. that might have incited cancer in HBA, reflect chronic exposures from before the mid 1980s. Thus, risk metrics that are based on ambient concentrations from 2002 to 2004 ($PM_{10}$) and 1996–2004 ($SO_2$) could be irrelevant if the spatial patterns of the environmental stressors which underlie the risk have changed. Since a persistent spatial, rather than temporal, pattern of the risk factor is required, the key question is whether locations that are characterized by an increased risk remain so over long times. This question is seldom addressed but is essential for relating extremely prolonged exposures with health outcomes. In HBA, the spatial distribution of ambient concentrations of both $PM_{10}$ and $SO_2$ seem to be largely conserved in spite of the dramatic decrease in the sulfur content of fuels.

As always, the quality and resolution of the data reflect on the model results. Whereas high resolution mapping of pollutant concentrations

**Table 5**
Results of the regression models with the socioeconomic status (SES) ranking and different combinations of long-term exposure to $PM_{10}$ and $SO_2$ as risk factors. The *p*-values of the non-randomness testing of the residual relative risk (RR), the deviance information criterion (DIC), and the fraction of the spatially correlated heterogeneity from the total variance of the random effects (Frac.Spatial) are also reported. The results were obtained based on data from all the 143 wards (the complete database).

| | | Model | RR.SES (med., 95% CI) | RR.$SO_2$ (med., 95% CI) | RR.$PM_{10}$ (med., 95% CI) | DIC | Residual RR Moran's *I* | Residual RR Tango's *C* | Frac. Spatial (med., 95% CI) |
|---|---|---|---|---|---|---|---|---|---|
| Lung cancer | Male | SES | 0.21 (0.10–0.43) | | | 500.3 | 0.001 | 0.001 | 0.73 (0.28–0.99) |
| | | SES + $SO_2$ | 0.21 (0.10–0.44) | 0.78 (0.26–2.37) | | 500.9 | 0.001 | 0.001 | 0.71 (0.28–1.0) |
| | | SES + $PM_{10}$ | 0.26 (0.12–0.54) | | 2.07 (0.73–5.59) | 498.2 | 0.001 | 0.001 | 0.62 (0.19–0.99) |
| | | SES + $SO_2$ + $PM_{10}$ | 0.26 (0.12–0.56) | 0.92 (0.31–2.72) | 2.01 (0.75–6.61) | 500.2 | 0.001 | 0.001 | 0.66 (0.23–1.0) |
| Bladder cancer | Male | SES | 0.85 (0.46–1.56) | | | 483.7 | 0.001 | 0.001 | 1.0 (1.0–1.0) |
| | | SES + $SO_2$ | 0.86 (0.40–1.59) | 1.02 (0.30–2.25) | | 487.8 | 0.001 | 0.001 | 1.0 (0.95–1.0) |
| | | SES + $PM_{10}$ | 0.79 (0.40–1.53) | | 0.82 (0.37–1.70) | 484.7 | 0.001 | 0.001 | 1.0 (0.99–1.0) |
| | | SES + $SO_2$ + $PM_{10}$ | 0.80 (0.42–1.60) | 0.92 (0.40–2.21) | 0.87 (0.41–2.19) | 487.9 | 0.001 | 0.001 | 1.0 (0.99–1.0) |
| | Female | SES | 0.84 (0.26–2.80) | | | 294.1 | 0.001 | 0.001 | 1.0 (0.99–1.0) |
| | | SES + $SO_2$ | 0.88 (0.25–2.93) | 1.15 (0.22–5.27) | | 296.5 | 0.001 | 0.001 | 1.0 (0.98–1.0) |
| | | SES + $PM_{10}$ | 0.89 (0.26–3.12) | | 1.07 (0.25–5.11) | 295.7 | 0.001 | 0.001 | 1.0 (1.0–1.0) |
| | | SES + $SO_2$ + $PM_{10}$ | 0.95 (0.25–3.93) | 1.36 (0.23–9.74) | 1.38 (0.26–9.49) | 297.8 | 0.001 | 0.001 | 1.0 (0.96–1.0) |

**Table 6**
Results of the regression models with the socioeconomic status (SES) ranking and different combinations of long-term exposure to $PM_{10}$ and $SO_2$ as risk factors. The *p*-values of the non-randomness testing of the residual relative risk (RR), the deviance information criterion (DIC), and the fraction of the spatially correlated heterogeneity from the total variance of the random effects (Frac.Spatial) are also reported. The results were obtained after wards with no reported cases and with expected rates <1, due to their small population, were excluded (the reduced datasets). Hence, for each cancer type and gender a different list of wards was used (see Table 1).

| | | Model | RR.SES (med., 95% CI) | RR.$SO_2$ (med., 95% CI) | RR.$PM_{10}$ (med., 95% CI) | DIC | Residual RR Moran's *I* | Residual RR Tango's *C* | Frac. Spatial (med., 95% CI) |
|---|---|---|---|---|---|---|---|---|---|
| Lung cancer | Male | SES | 0.25 (0.14–0.48) | | | 482.3 | | | 0.72 (0.20–1.0) |
| | | SES + $SO_2$ | 0.25 (0.14–0.47) | 0.83 (0.31–2.42) | | 483.0 | | | 0.70 (0.23–1.0) |
| | | SES + $PM_{10}$ | 0.29 (0.16–0.54) | | 1.84 (0.69–4.89) | 481.6 | 0.007 | 0.021 | 0.59 (0.15–0.99) |
| | | SES + $SO_2$ + $PM_{10}$ | 0.29 (0.15–0.56) | 1.01 (0.36–2.79) | 1.85 (0.65–4.56) | 483.0 | | | 0.56 (0.16–1.0) |
| Bladder cancer | Male | SES | 0.92 (0.54–1.59) | | | 477.5 | | | 0.79 (0.20–1.0) |
| | | SES + $SO_2$ | 0.91 (0.52–1.53) | 0.99 (0.86–1.13) | | 478.4 | | | 0.77 (0.20–1.0) |
| | | SES + $PM_{10}$ | 0.90 (0.51–1.55) | | 0.90 (0.43–2.03) | 478.7 | 0.007 | 0.017 | 0.80 (0.22–1.0) |
| | | SES + $SO_2$ + $PM_{10}$ | 0.86 (0.49–1.54) | 0.97 (0.84–1.11) | 0.88 (0.39–2.17) | 479.9 | | | 0.75 (0.18–1.0) |
| | Female | SES | 1.17 (0.39–2.67) | | | 250.5 | | | 0.78 (0.27–1.0) |
| | | SES + $SO_2$ | 1.15 (0.36–2.77) | 1.10 (0.69–1.67) | | 253.9 | | | 0.80 (0.35–1.0) |
| | | SES + $PM_{10}$ | 1.09 (0.31–2.74) | | 1.54 (0.15–5.30) | 253 | 0.003 | 0.023 | 0.73 (0.26–0.99) |
| | | SES + $SO_2$ + $PM_{10}$ | 1.31 (0.37–3.28) | 1.20 (0.74–2.04) | 8.48 (0.09–43.8) | 253.7 | | | 0.82 (0.39–1.0) |

can be produced, information on historic exposures at the individual or the group levels are oftentimes missing. Individual data points for which particular risk factors are missing require standardization that could be achieved only at an aggregate level. Hence, in comparison to environmental data and the risk metrics derived from them, morbidity data oftentimes have lower resolution and are available at the ward or a coarser spatial resolution. Specific to this work, although we lacked individual risk factors and in spite of cancer being a rare disease, we attempted to work at the highest spatial resolution supported by our data. Clearly, the approach we followed can be employed at a coarser spatial scale (e.g. a district level) to increase the number of cases involved and reduce the effect of inter-ward mobility. However while this should work towards reducing the uncertainty in exposure estimation and diminishing erratic SIRs, thus helping with the inference, it oftentimes leads to loss of statistical power since large geographic areas for which composite data are used tend to exhibit considerable intra-regional variation (Elliott and Wartenberg, 2004; Nuckols et al., 2004).

To first order approximation, aggregating dose–effect relationships from the individual to the ward level in the case of linear dose–effect relationships or for exponential dose–effect relationships with small RR (i.e. in rare diseases) can be shown to lead to a ward level RR that corresponds to the effect of exposure at the individual-level (Richardson and Best, 2003). This conclusion holds also when the exposure is nearly uniform over the ward (i.e. for small wards as found in HBA) or when the group level variance of the exposure hardly varies.

As seen in Table 5, significant RR was obtained for the between-area variation of the SES ranking only for lung cancer in males. The $RR<1$ obtained in these cases represent the well known "inverse" effect of the SES ranking on lung cancer, with low-SES ranking contributing to elevated risk of lung cancer (Su et al., 2009). Normally, this observation is attributed to an increased smoking prevalence in low-SES populations. Since smoking is known to be associated with a number of diseases, the spatial distribution of lung cancer incidence rates has been used as a risk factor that proxies smoking (Held et al., 2005). Likewise, we utilized the spatial distribution of lung cancer rates in males (LM) as a proxy of risk from smoking when assessing the RR of the environmental stressors for bladder cancer in males (DIC: 481.6, RR.SES: 0.97, RR.LM: 1.08, RR.$PM_{10}$: 0.76, RR.$SO_2$: 0.91). The improved model results (i.e. its lower DIC compared to Table 5) suggest that using the aggregated lung cancer data as a covariate for the spatial distribution of bladder cancer in males was supported by the data. Since both the SES ranking and the lung cancer rates were considered in this model, the SES ranking seems to account only partially for the impact of smoking on cancer rates. This has clear implications on the calculated RR of environmental stressors when lung cancer is studied.

The mean RR of long-term exposure to $PM_{10}$ for lung cancer in males is 1.147 (95% CI: 0.943–1.382, obtained after back-transforming the normalized figure appearing in Table 5 to physical units, since using normalized variables improved the models' convergence dramatically). Following the same procedure for the RR.$PM_{10}$ reported in Table 6, the (back transformed) mean RR associated with long-term exposure to $PM_{10}$ for lung cancer in males decreased somewhat to 1.12 (95% CI: 0.93–1.35). This result implies that enduring exposure to a 1 $\mu g/m^3$ increase of $PM_{10}$ average concentrations is expected to increase the incidence of lung cancer in males by 12%. It therefore seems that long-term exposure in HBA to $PM_{10}$ at even sub-substandard levels may contribute to a higher risk of lung cancer in males. This result is much higher than the RR reported previously for mortality from lung cancer. For example, Jerrett et al. (2005) found that the risk of mortality from lung cancer was significantly higher among people who were exposed to a 10 $\mu g/m^3$ increase in $PM_{2.5}$ (RR, 1.60, 95% CI, 1.09–2.33) and that after accounting for 44 individual covariates the RR decreased to 1.44 (95% CI, 0.98–2.11) and then hardly changed when accounting for other social factors (RR, 1.43, 95% CI, 0.96–2.13). Similarly, Pope et al. (2002) found that the risk of mortality from lung cancer was higher among those who were exposed to a 10 $\mu g/m^3$ increase in $PM_{2.5}$ (RR, 1.14, 95% CI, 1.04–1.23) and that the relative risk associated with exposure to $PM_{10}$ was higher by about 10%. The latter figure is about 10 fold lower than the RR reported here. Whereas lower RR of mortality from lung cancer than of lung cancer incidence have been reported (Reeves et al., 2007), the marginal (~3%) difference, probably reflecting the poor survival with this cancer, does not explain the large RR obtained in this study.

The relative risk associated with long-term exposure to $PM_{10}$ was found to be non-significant in all the models (Tables 5 and 6) at the $\alpha=0.05$ level. However, this significance level is often argued to be too stringent in such cases. As an alternative, Rothman and Greenland's (1998) sufficient component causes theory or other approaches (e.g. the precautionary principle) are advocated. Our findings reveal that lung cancer incidence in males is associated with extended exposure to $PM_{10}$ and that the latter is a major risk factor for the former, with attributed risk of 10.7%. However, these findings, although consistent, do not represent causation relationships, which must be assessed at the individual-level and are normally multi-factorial.

## 5. Conclusions

Based on persistent relative spatial patterns, chronic exposure to ambient $PM_{10}$ concentrations appears to be associated with lung cancer incidence rates in males in HBA. Global clustering tests were used to pre-filter out the cancers that showed random spatial distribution, thus diminishing the problem of multiple testing. The mean RR was found to

be an order of magnitude larger than previously reported in relation to mortality from lung cancer. This result was obtained even when the small-populated wards, in which no cases were expected irrespectively of the cancer risk, were excluded. The missing individual-level risk factors, the relatively low-level data available to control for confounders, and the finding that the ward-based SES ranking accounts only partially for the impact of smoking prevalence may explain the high RR.$PM_{10}$ observed. Indeed, the relatively high spatial structure in the residual RR suggests the presence of additional unaccounted for spatially correlated risk factors.

## Acknowledgement

This work was supported by the ENVIRISK consortium agreement SSPE-CT-2005, Contract no. 044232, under the Sixth Framework Program for R&D of the Research Directorate General of the European Commission, and by the Israeli Ministry of Science and Technology.

## References

Barchana M. Geographical mapping of malignant diseases in Israel 1984–1999. Israel National Cancer Registry, Tel Aviv; 2001 [Hebrew].
Bascom R. Health effects of outdoor air pollution. Am J Respir Crit Care Med 1996;153: 3-50.
Beeson WL, Abbey DE, Knutsen SF. Long-term concentrations of ambient air pollutants and incident lung cancer in California adults: results from the AHSMOG study. Environ Health Perspect 1998;106(12):813–23.
Besag J, York J, Mollie A. Bayesian image restoration with two applications in spatial statistics. Ann Inst Stat Math 1991;43:1-59.
Best N, Cockings S, Bennett J, Wakefield J, Elliott P. Ecological regression analysis of environmental benzene exposure and childhood leukaemia: sensitivity to data inaccuracies, geographical scale and ecological bias. J R Statist Soc A 2001;164(1): 155–74.
Brooks SP, Gelman A. Alternative methods for monitoring convergence of iterative simulations. J Comput Graph Stat 1998;7:434–55.
Brunekreef B, Forsberg B. Epidemiological evidence of effects of coarse airborne particles on health. Eur Respir J 2005;26:309–18.
Dabney AR, Wakefield JC. Issues in the mapping of two diseases. Stat Meth Med Res 2005;14:83-112.
Dayan U, Levy I. The influence of seasonal meteorological conditions and atmospheric circulation types on $PM_{10}$ and visibility in Tel-Aviv, Israel. J Appl Meteorol 2005;44(5): 606–19.
Dockery DW, Pope CA, Xu X, Spengler JD, Ware JH, Fay ME. An association between air pollution and mortality in six cities. N Engl J Med 1993;329:1753–9.
Doll R, Peto R. The causes of cancer: quantitative estimates of avoidable risks of cancer in the United States today. J Natl Cancer Inst 1981;66:1191–308.
Elliott P, Wartenberg D. Spatial epidemiology: current approaches and future challenges. Environ Health Perspect 2004;112(9):998-1006.
Epstein LM, Katz L, Tamir A, Rishpon S. Incidence of lung cancer in five towns in Israel, 1960–74. Isr J Med Sci 1984;20:27–32.
Fishler Y, Chitrit A, Barchana M, Modan B. Examination of Israel national cancer data accumulation completeness for 1991. The National Center for Disease Control, Publication No. 230, Tel Hashomer, Israel; 2003. [Hebrew].
Garcia L. Escaping the Bonferroni iron claw in ecological studies. Oikos 2004;105(3): 657–63.
Georgopoulos PG, Lioy PJ. Conceptual and theoretical aspects of human exposure and dose assessment. J Expo Anal Environ Epidemiol 1994;4(3):253–85.
Ginsberg GM, Tulchinsky TH. Regional differences in cancer incidence and mortality in Israel: possible leads to occupational causes. Isr J Med Sci 1992;28:534–42.
Ginsberg GM, Tulchinsky TH, Salahov E, Clayman M. Standardized mortality ratios by region of residence, Israel, 1987–1994: a tool for public health policy. Public Health Rev 2003;31(2):111–31.
Good P. Permutation tests, a practical guide to resampling methods for testing hypotheses. New York: Springer-Verlag; 1994.
Guo J, Kauppinen T, Kyyronen P, Heikkila P, Lindbohm ML, Pukkala E. Risk of esophageal, ovarian, testicular, kidney and bladder cancers and leukemia among Finnish workers exposed to diesel or gasoline engine exhaust. Int J Cancer 2004;111(2):286–92.
Held L, Natário I, Fenton SE, Rue H, Becker N. Towards joint disease mapping. Stat Meth Med Res 2005;14:61–82.
Jacquez GM, Grimson R, Waller LA, Wartenberg D. The analysis of disease clusters. 2. Introduction to techniques. Infect Control Hosp Epidemiol 1996;17(6):385–97.
Jerrett M, Burnett RT, Ma R, Pope CA, Krewski D, Newbold KB, et al. Spatial analysis of air pollution and mortality in Los Angeles. Epidemiology 2005;16(6):727–36.
Johnson KC, Pan S, Fry R, Mao Y. Residential proximity to industrial plants and non-Hodgkin lymphoma. Epidemiology 2003;14:687–93.
Julious SA, Nicholl J, George S. Why do we continue to use standardized mortality ratios for small area comparisons? J Public Health Med 2001;23(1):40–6.
Kokki E, Ranta J, Penttinen A, Pukkala E, Pekkanen J. Small area estimation of incidence of cancer around a known source of exposure with fine resolution data. Occup Environ Med 2001;58:315–20.
Kulldorff M, Tango T, Park PJ. Power comparisons for disease clustering tests. Comput Stat Data Anal 2003;42:665–84.
Lawson AB. Statistical methods in spatial epidemiology. New York: Wiley; 2001.
Maslia ML, Aral MM, Williams RC, Susten AS, Heitgerd JL. Exposure assessment of populations using environmental modeling, demographic analysis, and GIS. Water Resour Bull 1994;30(6):1025–41.
Meza JL. Empirical Bayes estimation smoothing of relative risks in disease mapping. J Stat Plann Infer 2003;112(1–2):43–62.
Nisbet ICT, Schneiderman MA, Karch NJ, Siegel DM. Review and evaluation of the evidence for cancer associated with air pollution. Final Report EPA-450/5-83-006R. Research Triangle Park, NC: US EPA; 1984.
Nuckols JR, Ward MH, Jarup L. Using geographic information systems for exposure assessment in environmental epidemiology studies. Environ Health Perspect 2004;112(9):1007–15.
Nyberg F, Gustavsson P, Jarup L, Bellander T, Berglind N, Jakobsson R, et al. Urban air pollution and lung cancer in Stockholm. Epidemiology 2000;11(5):487–95.
Oden NL. Assesing the significance of a spatial correlogram. Geog Anal 1984;16(1):1-16.
Pascutto C, Wakefield JC, Best NG, Richardson S, Bernardinelli L, Stainesv A, et al. Statistical issues in the analysis of disease mapping data. Stat Med 2000;19: 2493–519.
Perneger TV. What's wrong with Bonferroni adjustments. Br Med J 1998;316:1236–8.
Phillips DL, Lee EH, Herstrom AA, Hogsett WE, Tingey DD. Use of auxiliary data for spatial interpolation of ozone exposure in southeastern forests. Environmentrics 1997;8:43–61.
Pope CA, Thun MJ, Namboodiri MM, Dockery DW, Evans JS, Speizer FE, et al. Particulate air pollution as a predictor of mortality in a prospective study of U.S. adults. Am J Respir Crit Care Med 1995;151:669–74.
Pope CA, Burnett RT, Thun MJ, Calle EE, Krewski D, Thurston GD. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. JAMA 2002;287(9):1132–41.
Pope CA, Burnett RT, Thurston GD, Thun MJ, Calle EE, Krewski D, et al. Cardiovascular mortality and long-term exposure to particulate air pollution: epidemiological evidence of general pathophysiological pathways of diseases. Circulation 2004;109:71–7.
Pott F, Stöber W. Carcinogenicity of airborne combustion products observed in subcutaneous tissue and lungs of laboratory rodents. Environ Health Perspect 1983;47:293–303.
Reeves GK, Pirie K, Beral V, Green J, Spencer E, Bull D. Cancer incidence and mortality in relation to body mass index in the Million Women Study: cohort study. Br Med J 2007;335(7630):1134.
Reynolds P, Von Behren J, Gunier RB, Goldberg DE, Hertz A. Residential exposure to traffic in California and childhood cancer. Epidemiology 2004;15(1):6-12.
Richardson S, Best N. Bayesian hierarchical models in ecological studies of health–environment effects. Environmetrics 2003;14:129–47.
Rothman KJ, Greenland S. Modern epidemiology. 2nd ed. Philadelphia: Lippincott Williams & Wilkins; 1998.
Samet JM, Dominici F, Curriero FC, Coursac I, Zeger SL. Fine particulate air pollution and mortality in 20 US cities, 1987–1994. N Engl J Med 2000;343(24):1742–9.
Schwartz J. Harvesting and long term exposure effects in the relation between air pollution and mortality. Am J Epidemiol 2000;151(5):440–8.
Sharp L, Black RJ, Harkness EF, McKinney PA. Incidence of childhood leukaemia and non-Hodgkin's lymphoma in the vicinity of nuclear sites in Scotland, 1968–93. Occup Environ Med 1996;53:823–31.
Soll-Johanning H, Bach E. Occupational exposure to air pollution and cancer risk among Danish urban mail carriers. Int Arch Occup Environ Health 2004;77:351–6.
Spiegelhalter D, Thomas A, Best N. WinBUGS: Bayesian inference using Gibbs sampling. Manual v1.2. Cambridge: Medical Research Council Biostatistics Unit; 1998 (available from www.mrc-bsu.cam.ac.uk\bugs).
Spiegelhalter DJ, Best NG, Carlin BP, Van der Linde A. Bayesian measures of model complexity and fit. J R Statist Soc B 2002;64(4):583–639.
Su JG, Morello-Frosch R, Jesdale BM, Kyle AD, Shamasunder B, Jerrett M. An index for assessing demographic inequalities in cumulative environmental hazards with application to Los Angeles, California. Environ Sci Technol 2009;43(20):7626–34.
Sunyer J, Ballester F, Le Tertre A, Atkinson R, Ayres JG, Forastiere F, et al. The association of daily sulfur dioxide air pollution levels with hospital admissions for cardiovascular diseases in Europe (The Aphea-II study). Eur Heart J 2003;24: 752–60.
Tamir A. (1987). Possible causes for the differences in incidence of lung cancer in towns in Israel (a case–control study). Ph.D. Dissertation, Medical School, Technion, Haifa.
Tango T. A test for spatial disease clustering adjusted for multiple testing. Stat Med 2000;19:191–204.
Tzala E, Best N. Mortality Bayesian latent variable modelling of multivariate spatio-temporal variation in cancer. Stat Meth Med Res 2008;17:97-118.
Ugarte MD, Ibáñez B, Militino AF. Modelling risks in disease mapping. Stat Meth Med Res 2006;15:21–35.
Vineis P, Hoek G, Krzyzanowski M, Vigna-Taglianti F, Veglia F, Airoldi L, et al. Lung cancers attributable to environmental tobacco smoke and air pollution in non-smokers in different European countries: a prospective study. Environ Health 2007;6:7-14.
Yang CY, Cheng BH, Hsu TY, Tsai SS, Hung CF, Wu TN. Female lung cancer mortality and sex ratios at birth near a petroleum refinery plant. Environ Res 2000;83(1):33–40.
Yuval, Broday DM, Carmel Y. Mapping spatio-temporal variables: the impact of the time-averaging window width on the spatial accuracy. Atmos Environ 2005;39:3611–9.
Yuval, Broday DM. High resolution spatial patterns of long-term mean air pollutants concentrations in Haifa Bay area. Atmos Environ 2006;40:3653–64.
Yuval, Flicstein B, Broday DM. The impact of a forced reduction in traffic volumes on urban air pollution. Atmos Environ 2008;42:428–40.