

RESEARCH ARTICLE

Open Access

Does the effect of gender modify the relationship between deprivation and mortality?

Natalia Salcedo¹, Marc Saez^{1,2*}, Basili Bragulat¹ and Carme Saurina^{1,2}

Abstract

Background: In this study we propose improvements to the method of elaborating deprivation indexes. First, in the selection of the variables, we incorporated a wider range of both objective and subjective measures. Second, in the statistical methodology, we used a distance indicator instead of the standard aggregating method principal component analysis. Third, we propose another methodological improvement, which consists in the use of a more robust statistical method to assess the relationship between deprivation and health responses in ecological regressions.

Methods: We conducted an ecological small-area analysis based on the residents of the Metropolitan region of Barcelona in the period 1994–2007. Standardized mortality rates, stratified by sex, were studied for four mortality causes: tumor of the bronchial, lung and trachea, diabetes mellitus type II, breast cancer, and prostate cancer. Socioeconomic conditions were summarized using a deprivation index. Sixteen socio-demographic variables available in the Spanish Census of Population and Housing were included. The deprivation index was constructed by aggregating the above-mentioned variables using the distance indicator, DP₂. For the estimation of the ecological regression we used hierarchical Bayesian models with some improvements.

Results: At greater deprivation, there is an increased risk of dying from diabetes for both sexes and of dying from lung cancer for men. On the other hand, at greater deprivation, there is a decreased risk of dying from breast cancer and lung cancer for women. We did not find a clear relationship in the case of prostate cancer (presenting an increased risk but only in the second quintile of deprivation).

Conclusions: We believe our results were obtained using a more robust methodology. First off, we have built a better index that allows us to directly collect the variability of contextual variables without having to use arbitrary weights. Secondly, we have solved two major problems that are present in spatial ecological regressions, i.e. those that use spatial data and, consequently, perform a spatial adjustment in order to obtain consistent estimators.

Keywords: Deprivation indexes, Distance indicator DP₂, Ecological regression, Standardized mortality ratio, Hierarchical Bayesian models, Robust

Background

Deprivation has been defined by Townsend as ‘a state of observable and demonstrable disadvantage relative to the local community or the wider society or nation to which an individual, family or group belong to’ [1]. It is a concept that has two distinguishable domains, material and social circumstances. The former relates to diet, health, clothing,

housing, household facilities, environment and work. The latter is more difficult to measure as it relates to different forms of relationships, such as community integration, leisure, and formal participation in social institutions [1].

A deprivation index (DI) is recognized as a composite measure, where no single variable can be said to measure it but rather a number of variables contribute in some way [2]. Area-level indicators are commonly used to evaluate the geographical distribution of socioeconomic inequalities in health [3–6]. Area-level indicators have been used as proxies for individual-level data when individual measures are not available (i.e. data from

* Correspondence: marc.saez@udg.edu

¹Research Group on Statistics, Econometrics and Health (GRECS), University of Girona, Girona, Spain

²Research Group on Statistics, Econometrics and Health (GRECS), CIBER of Epidemiology and Public Health (CIBERESP), University of Girona, Campus de Montilivi, 17071, Girona, Spain

census is only provided at aggregated levels for reasons of confidentiality) [7].

Approaches to ecological small-area analysis propose using the smallest practicable spatial scale for any study on inequalities in health [8,9]. This diminishes the ecological bias, as the analysis is closer to the individual level [8,10,11]. Most studies that have focused on area-level social factors use geographical boundaries, developed for censuses, as proxies for actual communities or neighborhoods [6]. Administrative boundaries allow for routinely collected data to be available at the smallest scale. In Spain, census tracts, the smallest administrative unit, tend to have a mean population of 1,000 subjects [12].

There are three main methods for constructing Deprivation Indices (DIs) [13,14]. First, standardized z-scores or log transformations of a priori selected variables. This was a very popular method up until the late 1980's [15,16]. The second method is the use of Principal Component Analysis (PCA) which has been the principal method used over the last twenty years. PCA transforms a number of possibly correlated variables into uncorrelated variables and permits the extraction of a smaller number of uncorrelated variables called principal components that collect a large percentage of the total variance contained in the original data which are then used to measure socioeconomic status [17-20]. The third method (the least common) uses feedback from health experts to assign weights to the selected variables [21-24].

The increased use of DIs has allowed for the identification of deprived areas and their association to morbidity and mortality, living in deprived areas is an effect modifier. In fact, some studies confirm that people living in deprived areas have higher rates of ill health and mortality [4,5,8,25,26]. However, material deprivation is manifested in higher rates of mortality differently by gender [27,28]. This might be due in part to the fact that men and women perceive their environment differently [29-31], they may be exposed to different local environments [32], and their vulnerability to aspects in the local environment may vary [33].

In this study we propose improvements to the method of elaborating DIs. First, in the selection of the variables, we incorporated a wider range of both objective and subjective measures which makes for a more complete DI. Several studies that analyzed deprivation-associated excess mortality found that men had higher rates of excess mortality than women [27,28]. A large proportion of excess deaths were found in diet-related causes of death in females and smoking and alcohol consumption in males [27]. However, inferences about gender differences in relation to deprivation and mortality are sensitive to the choice of inequality measures used [34-36]. Making it important to incorporate indicators that are gender sensitive (i.e. residential/ environmental

indicators as they are strongly related to women's health and individual economic activity indicators as they are associated to men's health) [32]. The strength of association between socio-economic and health measures tend to be greater for men than for women [35]. Thus, contextual indicators are crucial if we want to have a comparative measure between men and women's health.

Studies have shown that neighborhood context influences self-rated health, especially in women, beyond the effects of individual factors [33,37]. Neighborhoods with worse socioeconomic conditions have a negative effect on health [38]. Additionally residential environment may be more important for women's health while individual economic activity is more important for men's health [32]. Other studies have found that occupational factors are more important for men's health whereas the home environment is more important for women's health [39].

Second, in the statistical methodology, we used a distance indicator (Pena Distance or DP_2) instead of the leading aggregating method PCA. DP_2 overcomes several limitations of PCA, for instance, aggregating variables expressed in different units of measurement, arbitrary weights, and duplicate information [40-45]. The greatest advantage of DP_2 with respect to PCA consists on how each handles redundant information. Thus, while in PCA the first component, i.e. the DI, is composed of those variables that capture greater variability it leaves some variables out of the index. In DP_2 the index is composed of all the variables, leaving out of the index only the redundant information but including each and every one of the variables. Thus, in principle, DP_2 would collect more variability [45].

In addition, we propose another methodological improvement, which consists in the use of a more robust statistical method to assess the relationship between deprivation and health responses in ecological regressions.

In this sense, we focused on mortality from trachea, bronchial and lung cancer and diabetes, because there is certain evidence that gender could act as an effect modifier in the relationship with deprivation, and breast and prostate cancer, given their known association with deprivation. Several studies have shown a positive association between lung cancer mortality and deprivation for men and a negative association for women [18,28,46]. This is explained because in men, lower socioeconomic status implies a higher prevalence of smoking, which leads to an increased risk of lung cancer mortality in individuals with lower socioeconomic status. On the other hand, in women, those with higher socioeconomic status are those with a higher prevalence of smoking, leading to an increased risk of dying from that cause in individuals with higher socioeconomic status. Regarding diabetes, those individuals with high educational levels have a significantly lower risk of dying from this cause

(odds ratio, OR = 0.66) compared to those with low educational levels [47]. In the Whitehall study [48], we observed a similar gradient for mortality from diabetes mellitus than with the other pathologies studied; higher mortality from diabetes mellitus in individuals who were in the lowest levels of the hierarchy. This excess mortality appeared to be due to the increased frequency of cardiovascular risk in this type of study group. In addition, variations in mortality have also been observed for diabetes mellitus type II (DM2) at the residence area level. Thus, individuals living in more deprived areas tend to have worse lifestyle (higher values of body mass index –BMI–, greater prevalence of smoking) which could lead to worse glycemic control and greater number of complications associated with diabetes mellitus. As in the case of lung cancer mortality, this association could be modified by sex.

However, the results from the Spanish literature, at least in those that use a spatial adjustment [18,28,46] do not systematically coincide with each other; particularly for prostate cancer and breast cancer. Diabetes mortality has shown a positive association with deprivation for both men [28,49] and women [49]. Borrell *et al.* [18] and Puigpinós-Riera *et al.* [46] found a positive association for breast cancer in Alicante (relative risk, RR = 1.55 95% confidence interval, 95% CI:1.04-2.19) yet the majority of other cities showed a negative association, not statistically significant, except for Vigo (RR = 0.54 95% CI 0.33-0.84). Benach *et al.* [27] found a negative association for breast cancer and deprivation with Index 1, which measured deprivation through unemployment, illiteracy and low social class; however, found no association with Index 2, measured deprivation through overcrowding and a small component of unemployment and illiteracy. For prostate cancer, some studies found no clear association between mortality risk and deprivation [46]. Benach *et al.* [27] found a negative association for prostate cancer and deprivation with Index 1; however, found no association with Index 2.

Our hypothesis is that the differences found in the literature are caused mainly by the lack of robustness of the statistical methodology applied and the lack of robustness of the spatial adjustment applied. Therefore, our objective in this paper is to evaluate a more robust methodology, in the way of developing DIs, the selection of component variables and how to build them, and, above all, in the statistical method applied to analyze how the effect of gender modifies the relationship between deprivation and mortality.

Methods

Design and study population

Most countries in the world periodically carry out population censuses that gather information on socio-

demographic characteristics of the population resident in a country. In Spain, the Spanish Census of Population and Housing (CPH) conducted in 2001 reported nearly 41 million inhabitants. The province capital of Barcelona had over 1.5 million inhabitants and an average number of 1,201.33 inhabitants per census tract (standard deviation, SD, 526.66; Median 1,085.50). For this study, we conducted an ecological small-area analysis based on the residents of the Metropolitan region of Barcelona. A total of 2,978 census tracts were examined.

Variables

Standardized mortality rates (SMR), stratified by sex, were studied for four mortality causes: tumor of the bronchial, lung and trachea (ICD-10: C33-C34), diabetes mellitus (ICD-10:E10-E14), breast cancer (ICD-10:C50), and prostate cancer (ICD-10:C61). The mortality data was provided from the Catalan Mortality Registry. We only used death certificates of residents of the Metropolitan Region of Barcelona who died between 1994 and 2007.

Deprivation index

Socioeconomic conditions were summarized using a DI in a census tract level. Sixteen socio-demographic variables available in the Spanish CPH were included. Four of these: unemployment rate, percent with low educational level, manual workers and temporary workers, have been used in several DIs for small-area studies in Spain [46,50]. The other variables (university education, mono-parental homes, activity rate, immigrants, homes without heating, no toilet/bathroom in the home, bad communications, vandalism/crime, few green areas, exterior noise, contamination/bad smells, and dirty street) were chosen based on literature that focused on the association between contextual indicators and their effect on health, particularly those that analyzed gender differences [38]. We chose contextual indicators, most of them subjective, which were representative of the neighborhood characteristics (Table 1).

The DI was constructed by aggregating the above-mentioned variables using another indicator, DP_2 , instead of the standard PCA method. DP_2 is an iterative procedure that weights partial indicators depending on their correlation with the global index [41,45,51]. It is able to capture all the non-redundant variance of the indicators (i.e. avoiding multicollinearity); thus, allowing for the inclusion of a greater number of variables [41,42]. Given that DP_2 uses all the valuable information contained in the partial indicators the more complete the final DI will be, since each variable contains unique information not present in others [44]. Magnifying the goodness of DP_2 versus the extraction of an index through principal component analysis, especially when introducing a greater number of variables would not

Table 1 Mean and Percentile distribution of socioeconomic indicators over census tracts of the Barcelona Metropolitan Region, 2001

Indicators (%)	Women				Men			
	P ₂₅	P ₅₀	P ₇₅	Mean	P ₂₅	P ₅₀	P ₇₅	Mean
Manual workers	36.8	53.3	67.2	51.6	41.0	58.7	73.2	56.0
Unemployment	5.2	6.2	7.4	6.4	4.8	6.0	7.4	6.3
Temporary workers	16.7	20.1	24.4	20.9	13.8	17.2	21.2	18.0
Insufficient education	10.3	15.8	22.8	17.1	6.4	10.8	16.3	12.0
University education	26.4	32.2	37.9	32.1	25.0	31.9	39.1	32.4
Monoparental homes	10.3	20.0	32.4	23.1	0.0	16.7	33.3	22.0
Activity rate	53.4	56.8	60.6	57.1	63.1	67.1	71.3	67.3
Inmigrants	1.8	3.1	5.3	4.1	1.7	3.3	6.0	4.7
Homes without heating	64.8	79.7	90.0	75.3	64.8	79.7	90.0	75.3
No toilet/bathroom in home	0.6	1.0	1.7	1.4	0.6	1.0	1.7	1.4
Bad communications	3.3	6.1	14.3	11.5	3.3	6.1	14.3	11.5
Vandalism/Crime	19.0	26.3	38.7	31.3	19.0	26.3	38.7	31.3
Few green areas	22.0	35.0	50.6	37.0	22.0	35.0	50.6	37.0
Exterior noise	32.9	40.4	48.0	40.6	32.9	40.4	48.0	40.6
Contamination/Bad smells	17.0	23.6	31.5	25.2	17.0	23.6	31.5	25.2
Dirty streets	30.8	38.2	47.4	39.7	30.8	38.2	47.4	39.7

P₂₅: 25 Percentile; P₅₀: 50 Percentile; P₇₅: 75 Percentile; Min: Minimum; Max: Maximum.

suffice for identifying a single component for the index. More than one component should be collected when applying PCA thus resulting in an arbitrary choice of weights for aggregating the components to obtain the index. Pena [41] and Zarzosa [42] showed that DP₂ fulfills all the properties of a good composite indicator; monotony, unicity, invariance, homogeneity, transitivity, exhaustivity, existence and determination, and additivity.

The process for aggregating the variables using DP₂ is a four-step process previously described in detail [51].

Data analysis

When spatial data is available, the variability in the observed response is usually greater than the expected, resulting in over dispersion. In fact, it is important to distinguish between two sources of extra-variability [52]. The first most important source is the so-called 'spatial dependence'; it is a consequence of the correlation between the spatial unit and the neighboring spatial units, generally contiguous geographic areas. Nearby areas are more similar than those far apart. Part of this dependence is not really a structural dependence, but mainly due to variables not included in the analysis. The second source is the independent extra variability, spatially uncorrelated, called heterogeneity (not spatial); it is the result of unobserved variables, without spatial structure, which could influence the response [52,53].

To account for this extra variability, it is necessary to introduce some structure in the model. Otherwise, unless the model is linear, the estimates of the parameters

of interest and the standard errors of the estimators will be biased; therefore any inference based on them, will be wrong [54].

In this paper we have chosen to use hierarchical Bayesian models, more specifically a model based on Besag, York and Mollié [55,56] (BYM) to analyze the relationship between mortality and small-area deprivation.

As we know, the idea is to introduce two random effects in a generalized linear model with Poisson response, in order to capture the extra-variability [53].

$$O_i \sim \text{Poisson}(\mu_i \text{Pop}_i)$$

$$\text{Log}(\mu_i) = \alpha + \text{Log}(\text{Pop}_i) + \sum_{k=1}^4 \beta_k \text{Index}Q_{ki} + v_i + S_i +$$

$$\beta_5 \text{Pop}4564_i + \beta_6 \text{Pop}65m_i$$

Where O_i denotes the observed cases of the response variable in the census tract i, μ_i is the relative risk in section i, Pop_i is the population (men or women) of the census tract, v_i is the random effect which reflects the heterogeneity; S_i is the random effect that reflects the spatial dependence, α is the intercept, interpreted as the logarithm of the baseline risk; IndexQ_{ki} is the quintile of the deprivation index (standardized) in the census tract i (the first quintile was taken as a reference); the β are the parameters of the model, which can be interpreted as the logarithms of the relative risks associated with the explanatory variables; Pop4564_i is the percentage of men (or women) from 45 to 64 years in the census tract i,

and $Pob65m_i$ is the percentage of men (or women) aged 65 years or older in the census tract i .

The DI was categorized into quintiles because the influence of deprivation on the (spatial) variation of mortality could be non-linear.

The non-spatial random effect, also called heterogeneity, assumed to be normally distributed, with a mean of zero and constant variance. For the spatial random effect a conditional autoregressive model (CAR) is used [57,58]. This approach, which is the most utilized and has the lowest computational cost, approximates the spatial dependence as an average spatial effect of neighboring areas [53]. The areas considered are the census tracts of each municipality in the Metropolitan Region of Barcelona and surrounding areas are defined as adjacent census tracts, in other words they share a border.

Note that the specified model does not use, as an offset the number of cases expected in the census tract, but the population (men or women) of the same. This is because, unlike the standard BYM model [53], here we use the crude death rate (from the census tract) and not the standardized mortality ratio as an indicator of mortality. The reason is to avoid the problem called 'mutual standardization' [59,60]. Rosenbaun and Rubin [59] show how the use of standardized rates as the response variable in ecological regression models leads to biased results if only the answer, not predictors, are adjusted for the same confounder, usually the age distribution. When the predictor is not adjusted, it is implicitly assumed that the effect of predictor is constant for all strata of the confounding variable. This may be true for the effect of an air pollutant, in principle it is the same for all ages, but not for the deprivation index. Grisotto *et al.* [60], in line with Rosenbaun and Rubin [59] show that unbiased estimators can be obtained by adjusting the response and the predictors (the index of deprivation in this case) with the same variable (age distribution) or, even easier to use are crude rates as the response variable and entering age (as an average or structured) as an explanatory variable of the model. This is why we have introduced the age structure of the census tract (proportion of men and women aged 45 to 64 years and 65 years or more). The introduction of age also lets you control the age effect in the model. This modification allows to overcome methodological limitations of the standard implementation of the BYM model [18]. Furthermore, and also getting no significance for the coefficients of the variables that represent the population from 45 to 64 years and older population or more at 65, none of the cases studied, we can say that the risks found not differ by age group. Moreover, because the estimates associated with variables representing the population aged 45 to 64 years and older population less than 65, were not significant in any of the cases studied, we can say that risks not found differ by age group.

A second difference from the standard BYM model, not so obvious, is the use of standardized explanatory variables. The reason is that the spatial random effect, approximated by CAR, can be correlated with some (or all) explanatory variables that have a similar spatial dependence (known as concurvity). If this were the case, there would be an over-adjustment, unlike the phenomenon of multicollinearity in linear models, which could bias the estimates [61]. Simulations performed by us suggest that the problem could be solved by completely standardizing the potentially problematic explanatory variable. On the other hand, the introduction of the deprivation index (standardized) into quintiles, in addition to collecting a possible nonlinear effect, could mitigate much of the problem.

Spatial models were built as Bayesian hierarchical models with two stages [62] and estimated using the integrated nested Laplace approximation [62-64] (see Appendix: Annex).

Models were compared using the DIC (Deviance Information Criterion) [65] and the conditional predictive ordinate (CPO) for each observation (in fact $-mean(\log(cpo))$) [66,67]. CPO is a cross-validated predictive approach i.e., predictive distributions conditioned on the observed data with a single data point deleted. Asymptotically the CPO statistic has a similar dimensional penalty as DIC. In this perspective, the CPO statistic may be similar to DIC. In both cases, the lower the DIC or the CPO, the better the model.

There was not experimental research in this work. All computations were carried out using the interface INLA [68], running directly in R (version R 2.11.0) [69].

Results and discussion

According to the Spanish CPH of 2001 the average number of inhabitants per census tract in the metropolitan region of Barcelona was 1201.33 (standard deviation 526.55; median 1085.50).

Table 1 describes the distribution of the socioeconomic indicators, used to construct the DI, by gender in the Metropolitan Region of Barcelona in 2001. We observed a high percent of temporary workers (20.9% women, 18.0% men) and manual workers (51.6% women, 56.0% men) for both men and women. Women had a higher rate of insufficient education (17.1% vs. 12%) and a lower activity rate (57.1% vs. 67.3%) than men. Over a third of the individuals claimed their neighborhoods had dirty streets, exterior noise, few green areas and vandalism/crime.

Tables 2 and 3 show the associations between cause-specific mortality and deprivation by quintiles of the index, for men and women controlled for age. In the case of women (table 2), we observed a positive association for diabetes mortality. However, both lung cancer mortality (RR in Q5 = 0.78; 95%CI: 0.65-0.93) and breast

Table 2 Association between three mortality causes and the socioeconomic deprivation index in women

	Trachea, bronchial and lung cancer	Diabetes	Breast cancer
Deprivation (Quintiles)	RR (95% CI) ¹	RR (95% CI) ¹	RR (95% CI) ¹
log RR Q ₂	0.87 (0.75 - 1.01)	1.27 (1.13 - 1.42)	0.99 (0.89 - 1.11)
log RR Q ₃	0.86 (0.73 - 1.00)	1.25 (1.11 - 1.41)	1.00 (0.89 - 1.12)
log RR Q ₄	0.92 (0.78 - 1.08)	1.22 (1.07 - 1.38)	0.93 (0.82 - 1.04)
log RR Q ₅	0.78 (0.65 - 0.93)	1.24 (1.08 - 1.42)	0.81 (0.71 - 0.92)
Random Effects (σ)	Mean (SD) ²	Mean (SD) ²	Mean (SD) ²
Spatial ($\hat{\sigma}_s$)	0.35 (0.07)	0.34 (0.06)	0.31 (0.06)
Heterogeneity ($\hat{\sigma}_u$)	0.58 (0.03)	0.56 (0.02)	0.54 (0.02)
Mean (log(CPO))	1.177083	1.714764	1.736699
Deviance Information Criterion (DIC)	6875.091	10070.06	10270.72
Effective Number of Parameters	602.7834	1079.27	1054.834

¹ Mean (95% Credible Interval).

² Mean (Standard Deviation).

Credible interval does not contain the unity.

RR Relative Risk.

cancer mortality (RR in Q₅ = 0.81; 95%CI: 0.71-0.92) had an inverse relationship with socioeconomic deprivation; women with greater deprivation had less mortality risk (statistical significance was present only in the fifth quintile).

For men (Table 3), we observed a positive association for lung cancer mortality; where, at greater socioeconomic deprivation there was greater mortality risk (RR in Q₅ = 1.31; 95%CI: 1.19-1.43). For diabetes mortality we also observed a positive association with mortality risk and deprivation; however, statistical significance was found only in second and third quintiles. On the other hand, for prostate cancer mortality there is no systematic relationship between deprivation and mortality risk. Only the second quintile is significant with a RR = 1.13 (95% CI: 1.00-1.26).

Overall, our results are consistent with those provided by the literature. At greater deprivation, there is an

increased risk of dying from diabetes for both sexes (although in the case of men the relative risks associated with the third quintile of deprivation onwards were not statistically significant) and of dying from lung cancer for men. On the other hand, at greater deprivation, there is a decreased risk of dying from breast cancer and lung cancer for women (although in both cases only the relative risk associated with the top quintile of deprivation was statistically significant). We did not find a clear relationship in the case of prostate cancer (presenting an increased risk but only in the second quintile of deprivation).

This study has been able to improve the statistical methodology applied in building deprivation indices as well as the robustness of spatial adjustments in ecological studies. Our results were consistent with the existing literature that analyzes the association between deprivation and mortality (lung cancer, diabetes mellitus, breast cancer and prostate cancer); however, our focus

Table 3 Association between three mortality causes and the socioeconomic deprivation index in men

	Trachea, bronchial and lung cancer	Diabetes	Prostate cancer
Deprivation (Quintiles)	RR (95% CI) ¹	RR (95% CI) ¹	RR (95% CI) ¹
log RR Q ₂	1.19 (1.10 - 1.29)	1.17 (1.03 - 1.33)	1.13 (1.00 - 1.26)
log RR Q ₃	1.23 (1.13 - 1.34)	1.17 (1.03 - 1.33)	1.09 (0.96 - 1.23)
log RR Q ₄	1.22 (1.12 - 1.33)	1.08 (0.94 - 1.24)	0.96 (0.85 - 1.09)
log RR Q ₅	1.31 (1.19 - 1.43)	1.09 (0.94 - 1.26)	0.91 (0.79 - 1.04)
Random Effects (σ)	Mean (SD) ²	Mean (SD) ²	Mean (SD) ²
Spatial ($\hat{\sigma}_s$)	0.26 (0.04)	0.31 (0.06)	0.38 (0.06)
Heterogeneity ($\hat{\sigma}_u$)	0.49 (0.01)	0.57 (0.02)	0.54 (0.02)
Mean (log(CPO))	2.48789	1.507189	1.674478
Deviance Information Criterion (DIC)	14209.69	8818.788	9843.01
Effective Number of Parameters	1614.22	872.0202	1001.67

¹ Mean (95% Credible Interval).

² Mean (Standard Deviation).

Credible interval does not contain the unity.

RR Relative Risk.

was in urban small-areas. As some literature suggests, mortality differs in rural versus urban settings, particularly in women [70]; thus, it would be important to explore geographic differences in more detail in future studies. In addition, future studies should also focus on researching these same causes of death in other European cities and additional causes of death in the Barcelona Metropolitan Region.

Conclusions

We believe our results were obtained using a more robust methodology. First off, we have built a better index that allows us to directly collect the variability of contextual variables without having to use arbitrary weights, as is assumed by the aggregation of components through PCA extraction. Building an index using PCA is adequate, the results do not differ from those obtained by DP₂, when you add only a few variables because the first component allows you to collect enough variability. When using a large number of variables, like shown in this article, it is clear that a single component is not enough to collect the variability contained in the variables and the index should thus be constructed by aggregating two or more components. The arbitrariness in the choice of aggregation weights is what we avoided by using the DP₂ methodology as the algorithm used for constructing the index naturally allows you to collect all the variability contained in the variables excluding only the shared part and not the variable itself.

Secondly, we have solved two major problems that are present in spatial ecological regressions, i.e. those that use spatial data and, consequently, perform a spatial adjustment in order to obtain consistent estimators. In particular, the problem of mutual standardization and the problem of concurvity. As stated previously, the problem of concurvity is present when the spatial random effect is correlated with some (or all) of the explanatory variables that have similar spatial dependence. Using crude rates as indicators of mortality for the response variable and the introduction of age as an explanatory variable in the model allows us to ensure the collection of unbiased estimators. Furthermore, the introduction of age in the model allows us to monitor its effect on the results. On the other hand, the introduction of the index (standardized) into quintiles, in addition to collecting a possible nonlinear effect, could mitigate much of the problem of concurvity.

One limitation of this study is the limited variable selection due to lack of statistical information in censuses. Additionally, the deprivation index was computed using information from the 2001 Census of Population and Housing (the only information available) while mortality data was recorded using information from the period 1994–2007. Some characteristics

recorded in 2001 may not capture the real environment of the areas in the past, when the deaths occurred. However, the data collection period for the construction of the index corresponds to the half point of the period in which the mortality data was collected and the relative position of the census track with respect to the deprivation index (i.e. the quintile of the index where it was located) remains very stable in time. It is also important to point out that given that the data is over a decade old the present patterns of associations may have changed, primarily due to the increased amount of immigration since 2001 and the activity rate downfall due to the economic crisis.

Appendix

Annex

Spatial models were built as Bayesian hierarchical models with two stages [62,63]. The first stage was the observational model $p(y|x)$, where y denoted the vector of observations and x are the unknown parameters following a Gaussian Markov random field (GMRF) denoted as $p(x|\theta)$. The second stage was given by the hyperparameters θ and their respective prior distribution $p(\theta)$. The desired posterior marginals

$$p(x_i|y) = \int_{\theta} p(x_i|\theta, y)p(\theta|y)d\theta$$

of the GMRF were approximated using the finite sum

$$\tilde{p}(x_i|y) = \sum_k \tilde{p}(x_i|\theta_k, y)\tilde{p}(\theta_k|y)\Delta_k \quad (A1)$$

where $\tilde{p}(x_i|\theta, y)$ and $\tilde{p}(\theta|y)$ denoted approximations of $p(x_i|\theta, y)$ and $p(\theta|y)$, respectively. The finite sum (A1) was evaluated at support points θ_k using appropriate weights Δ_k .

The posterior marginal $p(\theta|y)$ of the hyperparameters is approximated using a Laplace approximation [71].

$$\tilde{p}(\theta|y) \propto \frac{p(x, \theta, y)}{\tilde{p}_G(x|\theta, y)|_{x=x^*(\theta)}}$$

where the denominator $\tilde{p}_G(x|\theta, y)$ denoted the Gaussian approximation of $p(x|\theta, y)$ and $x^*(\theta)$ was the mode of the full conditional $p(x|\theta, y)$ [72].

According to Rue *et al.* [63], it is sufficient to “numerically explore” this approximate posterior density using suitable support points θ_k in (A1). In this paper, these points were defined in the h -dimensional space, using the strategy called central composite design. Here, centre points were augmented with a group of star points, which allowed for estimating the curvature of $\tilde{p}(\theta|y)$.

Here, to approximate the first component of (A1) a simplified Laplace approximation (less expensive from a

computational point of view with only a slight loss of accuracy) was used [62-64].

Abbreviations

DI: Deprivation index; PCA: Principal component analysis; DP₂: Pena distance; OR: Odds ratio; DM2: Diabetes mellitus type 2; BMI: Body mass index; RR: Relative risk; 95% CI: 95% confidence interval; CPH: Census population and housing; SD: Standard deviation; SMR: Standardized mortality rates; ICD-10: International classification of disease v. 10; BYM: Besag, York and Mollié model; CAR: Conditional autoregressive model; DIC: Deviance information criterion; CPO: Conditional predictive ordinate.

Competing interests

There are no conflicts of interest for any of the authors. All authors disclose any financial and personal relationships with other people or organizations that could inappropriately influence and/or bias their work.

Authors' contributions

All authors contributed equally to the development of the manuscript. MS had the idea of writing this manuscript, based on the Masters thesis of NS. NS and CS were responsible for choosing the variables and collect data. CS and MS specified the models. MS and BB were responsible for the statistical analyzes. Interpretation of the results were done mainly by CS, NS and MS. Drafting of the manuscript were done by CS, MS and NS. All authors revised and approved the final version of the manuscript.

Acknowledgements

We would like to the Catalan Mortality Register (*Registre de Mortalitat de Catalunya, Servei d'Informació i Estudis, Direcció General de Recursos Sanitaris*), Health Department, Catalan Government (*Generalitat de Catalunya*), for providing mortality data. This work was partially supported by the European Union, DG-SANCO, Second Programme of Community action in the field of Health (2008-2013), project A/101156 and by the FIS (Health Research Fund), Spanish Ministry of Science and Innovation, project FIS-08/0142.

Received: 28 April 2012 Accepted: 13 July 2012

Published: 30 July 2012

References

1. Townsend P: **Deprivation**. *J Soc Policy* 1987, **16**:125-146.
2. Folwell K: **Single measure of deprivation**. *J Epidemiol Community Health* 1995, **49**(suppl 2):S51-S56.
3. Tunstall HVZ, Shaw M, Dorling D: **Places and health**. *J Epidemiol Community Health* 2004, **58**:6-10.
4. Townsend P, Phillimore P, Beattie A: *Health and deprivation: inequality and the North*. London: Routledge; 1988.
5. Carstairs V, Morris R: *Deprivation and health in Scotland*. Aberdeen: Aberdeen University Press; 1991.
6. Pickett KE, Pearl MJ: **Multilevel analysis of neighbourhood socioeconomic context and health outcomes: a critical review**. *J Epidemiol Community Health* 2001, **55**(2):111-122.
7. Galobardes B, Lynch J, Davey Smith G: **Measuring socioeconomic position in health research**. *Br Med Bull* 2007, **81** and **82**:21-37.
8. Ocaña-Riola L, Saurina C, Fernández-Ajuria A, Lertxundi A, Sánchez-Cantalejo C, Saez M, et al: **Area deprivation and mortality in the provincial capital cities of Andalusia and Catalonia (Spain)**. *J Epidemiol Community Health* 2008, **62**:147-152.
9. Noble M, Wright G, Dibben C, Smith GAN, McLennan D, Anttila C, et al: *The English Indices of Deprivation 2004 (revised)*. London, Office of the Deputy Prime Minister: Neighbourhood Renewal Unit; 2004.
10. Elliott P, Wartenberg D: **Spatial Epidemiology: current approaches and future challenges**. *Environ Health Persp* 2004, **112**:998-1006.
11. Lawson AB: *Statistical methods in spatial epidemiology*. 2nd edition. New Jersey: John Wiley and Sons; 2006.
12. Barcelò MA, Saez M, Saurina C: **Spatial variability in mortality inequalities, socioeconomic deprivation, and air pollution in small areas of the Barcelona Metropolitan Region, Spain**. *Sci Total Environ* 2009, **407**:5501-5523.
13. Carstairs V: **Socio-economic factors at area level and their relationship with Health**. In *Spatial epidemiology. Methods and applications*. Edited by Elliott P, Wakefield JC, Best NG, Briggs DJ. New York: Oxford University Press; 2000:51-67.
14. Bell N, Schuurman N, Hayes MV: **Using GIS-based methods of multicriteria analysis to construct socio-economic deprivation indices**. *Int J Health Geogr* 2007, **6**:17.
15. Townsend P: *Life and Labour in London*. London: CPAG; 1987.
16. Carstairs V: **Deprivation: explaining differences in mortality between Scotland and England and Wales**. *Br Med J* 1989, **299**:886-889.
17. Frohlich KL, Mustard C: **A regional comparison of socioeconomic and health indices in a Canadian province**. *Soc Sci Med* 1996, **42**(9):1273-1281.
18. Borrell C, Mari-Dell'olmo M, Serral G, Martínez-Beneito M, Gotsens M, MEDEA members: **Inequalities in mortality in small areas of eleven Spanish cities (the multicenter MEDEA project)**. *Health Place* 2010, **16**(4):703-711.
19. Pampalon R, Raymond G: **A Deprivation Index for Health and Welfare Planning in Quebec**. *Chronic Dis Can* 2000, **21**(3):104-113.
20. Domínguez-Berjón F, Borrell C, Cano-Serral G, Esnaola S, Nolasco A, Pasarín MI, Ramis R, Saurina C, Escolar-Pujolar A: **Constructing a deprivation index based on census data in large Spanish cities (the MEDEA Project) [In Spanish]**. *Gac Sanit* 2008, **22**(3):179-187.
21. Jarman B: **Identification of underprivileged areas**. *Br Med J* 1983, **286**:1705-1708.
22. Talbot RJ: **Underprivileged areas and health care planning: implications of use of Jarman indicators of urban development**. *Br Med J* 1991, **302**:283-286.
23. Carr-Hill R, Sheldon R: **Designing a deprivation payment for general practitioners: The UPA (8) Wonderland**. *Br Med J* 1991, **302**:393-396.
24. Davey-Smith G: **Second thoughts on the Jarman Index - Claims being made for the underprivileged area score are unproved**. *Br Med J* 1991, **302**:359-360.
25. Regidor E, Calle ME, Navarro P, Domínguez V: **Trends in the association between average income, poverty and income inequality and life expectancy in Spain**. *Soc Sci Med* 2003, **56**:961-971.
26. Mackenbach JP: *Health inequalities: Europe in profile*. Rotterdam: University Medical Center Rotterdam; 2006.
27. Benach J, Yasui Y, Borrell C, Pasarín MI, Martínez JM, Daponte A: **The public health burden of material deprivation: excess mortality in leading causes of death in Spain**. *Prev Med* 2003, **36**:300-308.
28. Benach J, Yasui Y, Borrell C, Saez M, Pasarín MI: **Material deprivation and leading causes of death by gender: evidence from a nation-wide small-area study**. *J Epidemiol Commun Health* 2001, **55**:239-245.
29. Molinari C, Ahern M, Hendryx M: **The relationship of community quality to the health of women and men**. *Soc Sci Med* 1998, **47**(8):1113-1120.
30. Mohai P: **Gender differences in the perception of most important environmental problems**. *Race Gender Class* 1997, **5**:153-169.
31. Stafford M, Bartley M, Mitchell R, Marmot M: **Characteristics of individuals and characteristics of areas: investigating their influence on health in the Whitehall II study**. *Health Place* 2001, **7**(2):117-129.
32. Stafford M, Cummins S, Maclntyre S, Ellaway A, Marmot M: **Gender differences in the associations between health and neighbourhood environment**. *Soc Sci Med* 2005, **60**:1681-1692.
33. Ellaway A, Macintyre S: **Women in their place. Gender and perceptions of neighbourhoods and health in the West of Scotland**. In *Geographies of women's health*. Edited by Dyck I, Davis Lewis N, McLafferty S. London: Routledge; 2001:265-281.
34. Mustard CA, Etches J: **Gender differences in socioeconomic inequality in mortality**. *J Epidemiol Comm Health* 2003, **57**:974-980.
35. Macintyre S, McKay L, Der G, Hiscock R: **Socio-economic position and health: what you observe depends on how you measure it**. *J Public Health Med* 2003, **25**(4):288-294.
36. Kawachi I, Kennedy BP, Gupta V, Prothrow-Stith D: **Women's status and the health of women and men: a view from the States**. *Soc Sci Med* 1999, **48**:21-32.
37. Van Praag L, Bracke P, Christians W, Lavecque K, Pattyn E: **Mental health in a gendered context. Gendered community effects on depression and problem drinking**. *Health Place* 2009, **15**(4):990-998.
38. Santos SM, Chor D, Werneck GL, Freire-Coutinho ES: **Association between contextual factors and self-rated health: a systematic review of multilevel studies [In Portuguese]**. *Cad Saude Publica* 2007, **23**(11):2533-2554.
39. Griffin J, Further R, Stansfeld S, Marmot M: **The importance of low control at work and home on depression and anxiety: do these effects vary by gender and social class?** *Soc Sci Med* 2002, **54**(5):738-798.

40. Windenberger F, Rican S, Jouglu E, Rey G: **Spatiotemporal association between deprivation and mortality: trends in France during the nineties.** *Eur J Public Health* 2011, **22**(3):347–353.
41. Pena JB: *Problems in measuring welfare and related concepts (an application to the Spanish case)* [in Spanish]. Madrid: INE; 1977.
42. Zarzosa P: **Poverty estimate in the Spanish autonomous regions, using the P2-Distance indicator** [in Spanish]. *Estudios de Economía Aplicada* 2009, **27**(2):397–416.
43. Mishra SK: *Construction of maximin and non-elitist composite indices – alternatives to elitist indices obtained by the principal component analysis.* MPPRA Paper, 3338. Germany: University Library of Munich; 2007.
44. Montero JM, Chasco C, Larraz B: **Building an environmental quality index for a big city: a spatial interpolation approach combined with a distance indicator.** *J Geogr Syst* 2010, **12**:435–59.
45. Salcedo N, Saurina C, Saez M: *Deprivation Indices. Gender differences and geographical variations of mortality inequalities at the small area level.* Girona: GRECS, University of Girona; 2010. mimeo.
46. Puigpinós-Riera R, Mari-Dell'Olmo M, Gotsens M, Borrell C, Serral G, Ascaso C, Calvo M, Daponte A, Domínguez-Berjón FM, Esnaola S, Gandarillas A, López-Abente G, Martos CM, Martínez-Beneito MA, Montes-Martinez A, Montoya I, Nolasco A, Pasarín MI, Rodríguez-Sanz M, Saez M, Sánchez-Villegas P: **Cancer mortality inequalities in urban areas: a Bayesian small area analysis in Spanish cities.** *Int J Health Geogr* 2011, **10**:6.
47. Dray R, Gary T, Brancati F: **Educational disparities in mortality among adults with Diabetes in the US.** *Diabetes Care* 2010, **33**(6):1200–1205.
48. Chaturvedi N, Jarrett J, Shipley MJ, Fuller JH: **Socioeconomic gradient in morbidity and mortality in people with diabetes: cohort study findings from the Whitehall Study and the WHO Multinational Study of Vascular Disease in Diabetes.** *BMJ* 1998, **316**(7125):100–105.
49. Espelt A, Borrell C, Roskam AJ, Rodríguez-Sanz M, Stirbu I, Dalmau-Bueno A, Regidor E, Bopp M, Martikainen P, Leinsalu M, Artnik B, Rychtarikova J, Kalediene R, Dzurova D, Mackenbach J, Kunst AE: **Socioeconomic inequalities in diabetes mellitus across Europe at the beginning of the 21st century.** *Diabetologia* 2008, **51**(11):1971–79.
50. Feja C, Alcalá JT, Rabanaque MJ, Saez M, Marcos-Gragera R, Martos CM: **Geographical inequalities in mortality and incidence in larynx cancer in men: socioeconomic and environment factors** [In Spanish]. *Rev Esp Salud Publica* 2010, **84**(6):745–56.
51. Ivanovic B: **How to prepare a list of indicators of development.** *Revue de Statistique Appliquée* 1974, **22**(2):37–50.
52. Lawson AB, Browne WJ, Vidal-Rodeiro CL: *Disease mapping with WinBUGS and MLwiN.* Chichester: John Wiley & Sons; 2003.
53. Barceló A, Saez M, Cano-Serral G, Martínez-Beneito MA, Martínez JM, Borrell C, Ocaña-Riola R, Montoya I, Calvo M, López-Abente G, Rodríguez-Sanz M, Toro S, Alcalá JR, Saurina C, Sánchez-Villegas P, Figueiras A: **Methods to smooth mortality indicators: application to the analysis of inequalities in mortality in Spanish cities (the MEDEA Project)** [In Spanish]. *Gac Sanit* 2008, **22**(6):596–608.
54. Greene WH: *Econometric Analysis.* 5th edition. Upper Saddle River, New Jersey: Pearson Education; 2003:191–201.
55. Besag J, York J, Mollié A: **Bayesian image restoration, with two applications in spatial statistics.** *Ann I Stat Math.* 1991, **43**:1–59.
56. Mollié A: **Bayesian mapping of disease.** In *Markov Chain Monte Carlo in Practice.* Edited by Gilks WR, Richardson S, Spiegelhalter DJ. New York: Chapman & Hall; 1996:359–379.
57. Besag J: **Spatial interaction and the statistical analysis of lattice systems (with discussion).** *J R Stat Soc Ser B* 1974, **36**:192–236.
58. Clayton DG, Bernardinelli L, Montomoli C: **Spatial correlation in ecological analysis.** *Int J Epidemiol* 1993, **22**:1193–1202.
59. Rosenbaum P, Rubin D: **Difficulties with regression analyses of age-adjusted rates.** *Biometrics* 1984, **40**:437–443.
60. Grisotto L, Catelan D, Accetta G, Biggeri A: **Material deprivation as marker of health needs.** *Statistica* 2010, **70**(3):343–352.
61. Catelan D, Biggeri A, Lagazio C: **On the clustering term in ecological analysis: how do different prior specifications affect results?** *Stat Methods Appl* 2009, **18**:49–61.
62. Schrödle B, Held L: **A primer on disease mapping and ecological regression using INLA.** *Comput Stat* 2011, **26**(2):241–258 [http://www.r-inla.org/papers].
63. Rue H, Martino S, Chopin N: **Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations (with discussion).** *J R Stat Soc Ser B* 2009, **71**:319–392 [http://www.r-inla.org/papers].
64. Martino S, Rue H: **Case studies in Bayesian computation using INLA.** In *Complex Data Modeling and Computationally Intensive Statistical Methods.* Edited by Mantovan P, Secchi P. Milano: Springer; 2010:99–112 [http://www.r-inla.org/papers].
65. Gilks WR, Richardson S, Spiegelhalter DJ (Eds): *Markov Chain Monte Carlo in Practice.* New York: Chapman & Hall; 1996.
66. Gelfand AE, Dey DK, Chang H: **Model determination using predictive distributions with implementation via sampling-based methods (with discussion).** In *Bayesian Statistics 4.* Edited by Bernardo JM, Berger JO, Dawid AP, Smith AFM. Oxford: Oxford University Press; 1992.
67. Pettit LI: **The conditional predictive ordinate for the normal distribution.** *J R Stat Soc Ser B* 1990, **52**(1):175–184.
68. *R-INLA project.* 2010. URL: http://www.r-inla.org/home.
69. R Development Core Team. *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing; 2010. URL http://www.R-project.org. ISBN 3-900051-07-0.
70. Dobson A, McLaughlin D, Vagenas D, Wong KY: **Why are death rates higher in rural areas? Evidence from the Australian Longitudinal Study on Women's Health.** *Aust N Z J Public Health* 2010, **34**(6):624–628.
71. Tierney L, Kadane JB: **Accurate approximations for posterior moments and marginal densities.** *J Am Stat Assoc* 1986, **81**(393):82–86.
72. Rue H, Held L: *Gaussian Markov Random Fields.* Boca Raton-London-New York-Singapore: Chapman & Hall/CRC; 2005.

doi:10.1186/1471-2458-12-574

Cite this article as: Salcedo et al.: Does the effect of gender modify the relationship between deprivation and mortality?. *BMC Public Health* 2012 12:574.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

