Health & Place 17 (2011) 691–695

Contents lists available at ScienceDirect

Health & Place

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

Short Report

An index of unhealthy lifestyle is associated with coronary heart disease mortality rates for small areas in England after adjustment for deprivation

P. Scarborough *, S. Allender, M. Rayner, M. Goldacre

Department of Public Health, University of Oxford, Old Road Campus, Headington, Oxford, OX3 7LF, UK

ARTICLE INFO

Article history:
Received 13 April 2010
Received in revised form
31 August 2010
Accepted 7 December 2010
Available online 14 December 2010

Keywords:
Coronary disease
Synthetic estimates
Deprivation
Smoking
Fruit and vegetable consumption
Obesity
Blood pressure
Cholesterol
Small area analysis

ABSTRACT

Indices of socio-economic deprivation are often used as a proxy for differences in the health behaviours of populations within small areas, but these indices are a measure of the economic environment rather than the health environment. Sets of synthetic estimates of the ward-level prevalence of low fruit and vegetable consumption, obesity, raised blood pressure, raised cholesterol and smoking were combined to develop an index of unhealthy lifestyle. Multi-level regression models showed that this index described about 50% of the large-scale geographic variation in CHD mortality rates in England, and substantially adds to the ability of an index of deprivation to explain geographic variations in CHD mortality rates.

1. Introduction

The relationships between coronary heart disease (CHD) and smoking, poor diet, excessive alcohol consumption, physical inactivity, obesity, raised cholesterol, raised blood pressure and diabetes are well-established at the individual-level (Stamler, 2005; Yusuf et al., 2004; WHO, 2003). It does not necessarily follow that these established risk factors are powerful predictors of geographic variation of CHD rates: for example, the British Regional Heart Study found that CHD incidence rates in men were *negatively* associated with the prevalence of raised cholesterol, after other individual-level risk factors for CHD had been taken into account (Morris et al., 2001). The fact that the association between cholesterol levels and CHD has been shown to be different at the individual-level and the area-level illustrates the danger of interpreting results regarding this association using data collected at only one level—there is the potential for either 'ecological fallacy' or 'individualistic fallacy' (Robinson, 1950; Subramanian et al., 2009).

This paper explores ecological, compositional aspects of geographic variation in CHD mortality rates in England, specifically the variation that is due to differences in the behaviour of the populations in different areas. The paper uses model-based estimates (referred to here as *synthetic estimates*) of the prevalence of individual-level risk factors for CHD for all wards in England. The synthetic estimation is a technique that has been developed to allow for small-area estimation of phenomena (in the absence of direct small-area measurements of the phenomena), including health indicators, principally to refine resource allocation to the areas (Heady et al., 2003). The technique involves using data collected for a national survey to generate a logistic regression of the health indicator of interest (e.g. smoking) with both individual-level (e.g. age; sex) and area-level covariates (e.g. percentage of privately rented accommodation; geographic location). The parameters predicted by these logistic regression models are used to generate small area prevalence estimates using national census data. The technique is described in detail elsewhere (Heady et al., 2003; Bajekal et al., 2004; EURAREA Consortium, 2004; Twigg et al., 2000; Twigg and Moon, 2002). This paper introduces an index of unhealthy lifestyle, developed using synthetic estimates of the prevalence of several cardiovascular risk factors, and assesses whether this index can explain geographic variations in CHD mortality rates over and above those predicted by an index of socio-economic deprivation. The models developed for this paper use ecological data. It follows that the results can only provide information about the geographic variation in CHD rates in England, and not about the relationship between risk factors and CHD in individuals.

2. Methods

The units of analysis used for this paper are Standard Table Wards—a statistical set of boundaries based on the electoral ward

* Corresponding author. Tel.: +44 1865 289248; fax: +44 1865 617790.
E-mail address: peter.scarborough@dphpc.ox.ac.uk (P. Scarborough).

1353-8292/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.healthplace.2010.12.007

boundaries as of 1st January 2003. Henceforth these areas are referred to simply as 'wards'. There are 7,929 wards in England, which can be grouped into 355 local authorities (LAs).

Data on the number of CHD (ICD-9 410-414; ICD-10 I20-I25) deaths by sex, five year age group and ward of residence for the calendar years 1999–2004 inclusive were combined with population estimates from the 2001 census by sex, five year age group and ward of residence to derive sex-specific mortality rates that were age-standardised to the European Standard Population (West Midlands Public Health Observatory, 2009) for each ward in England. The mortality rates were reasonably normally distributed, and hence suited to regression analysis. Data on the ward-level prevalence of individual-level risk factors for CHD were provided by synthetic estimates. A previous analysis has assessed the validity of 16 different sets of synthetic estimates of the prevalence of individual-level risk factors for CHD for all wards in England (Scarborough et al., 2009). It assessed face, construct and convergent validity of the estimates by exploration of the supporting regression equations, and by comparison of the synthetic estimates with small area survey findings on related health behaviours, national survey findings on health behaviours and CHD mortality rates. Only the five sets of synthetic estimates that displayed reasonable validity were included as explanatory variables in the analyses reported here:

- prevalence of consuming less than five portions of fruit and vegetables per day;
- prevalence of obesity (body mass index greater than or equal to 30 kg/m^2);
- prevalence of raised blood pressure (systolic blood pressure greater than or equal to 160 mmHg, or diastolic blood pressure greater than or equal to 95 mmHg);
- prevalence of raised cholesterol (total blood cholesterol greater than or equal to 6.5 mmol/l);
- prevalence of current smoking.

The first four sets of synthetic estimates were developed for the Health Poverty Index website, by researchers from the universities of St Andrews and Oxford, funded by the Department of Health (Dibben et al., 2004). The estimates used for the website are weighted combinations of LA-level national survey estimates and a modelled element. The synthetic estimates used for this paper, and that were used in the previous exploration of validity (Scarborough et al., 2009), are based only on the modelled element and have been produced at ward-level. The smoking synthetic estimates were developed for the Health Development Agency as part of an investigation of the impact of smoking on mortality in England (Twigg et al., 2004). All of the sets of synthetic estimates are derived from the Health Surveys for England conducted between 1998 and 2001 and a table outlining the coefficients for individual-level and area-level covariates in their supporting regression equations can be found elsewhere (Scarborough et al., 2009). The synthetic estimates have been age-standardised to the European Standard Population using ten-year age bands for the analyses reported here. Prevalence rates for men and women were generated separately. The synthetic estimates are not well-suited for use as independent variables in regression analyses as they are highly correlated. Accordingly, Principal Components Analysis (PCA) was applied to z scores of the synthetic estimates to produce a set of uncorrelated explanatory variables, described as 'unhealthy lifestyle' variables.

The Carstairs index (Carstairs and Morris, 1990), constructed using data from the 2001 census (Morgan and Baker, 2006), was used as a measure of ward-level deprivation. The index is constructed by adding the z scores of the following variables: percentage of all economically active males aged 16 and over who are unemployed; percentage of households defined as overcrowded; percentage of population without access to a car; percentage of population living in households where the head of the household is defined as social class IV or V. The index has previously been shown to be highly correlated with CHD rates for wards in England (Romeri et al., 2006).

Multi-level regression models (wards nested in LAs) and spatial error regression models of the two outcome variables (male and female CHD mortality rates) were built to explore how much of the geographic variation in CHD rates can be explained by the unhealthy lifestyle variables, and by the deprivation index, in both univariate and multivariate analyses. The multi-level models allow for an investigation of both small scale and large-scale geographic variations in CHD mortality rates simultaneously, that is the differences in CHD mortality rates over a small geographic area i.e. within a local authority, and the general differences between regions of the country i.e. the North and South of England (modelled by between-wards and between-LAs variance, respectively). The spatial error regression models were built to investigate whether the results of the multi-level models may be affected by spatial autocorrelation bias, by comparing the size and sign of estimated parameters in the two sets of models. The PCA was conducted using Stata v10 (StataCorp, 2007), the spatial error regression modelling was conducted using the GeoDa software package (Anselin, 2003), and the multi-level modelling was conducted using MLwiN v2.02 (Rasbash et al., 2003).

3. Results

The results of the PCA are displayed in Table 1. Two of the transformed PCA variables were taken forward for further

Table 1
Transformation matrices calculated by principal components analysis for the sets of synthetic estimates for (1) male prevalence of risk factors for CHD and (2) female prevalence of risk factors for CHD, and amount of original variance explained by the transformed variables (wards, $n=7929$).

	Fruit & Veg	Obesity	Blood pressure	Cholesterol	Smoking	*Proportion of original variance*
(1) Synthetic estimates of male prevalence rates						
Unhealthy lifestyle 1	0.51	0.54	0.51	0.17	0.41	*0.63*
Unhealthy lifestyle 2	0.14	−0.08	−0.06	0.92	−0.36	*0.21*
PCA 3	0.17	−0.38	−0.45	0.22	0.76	*0.11*
PCA 4	−0.75	−0.10	0.50	0.26	0.33	*0.06*
PCA 5	−0.37	0.74	−0.54	0.13	0.10	*0.00*
(2) Synthetic estimates of female prevalence rates						
Unhealthy lifestyle 1	0.51	0.51	0.48	0.22	0.45	*0.66*
Unhealthy lifestyle 2	0.05	−0.31	−0.28	0.89	0.17	*0.19*
PCA 3	−0.26	0.27	0.46	0.39	−0.70	*0.11*
PCA 4	−0.70	−0.18	0.47	0.02	0.50	*0.04*
PCA 5	−0.42	0.73	−0.51	0.09	0.15	*0.00*

PCA3–PCA5 refer to the transformed variables that were not retained for further analysis.

analysis— these two variables explained around 85% of the total variance of the sets of synthetic estimates for both male and female prevalence estimates, and their transformation factors allowed for simple interpretation of results. Henceforth, these two PCA variables are referred to as *unhealthy lifestyle 1* and *unhealthy lifestyle 2.* These two variables broadly measure the following features: *unhealthy lifestyle 1*—increased prevalence rate of low fruit and vegetable consumption, obesity, raised blood pressure and smoking; *unhealthy lifestyle 2*—increased prevalence of raised cholesterol and reduced prevalence rate of smoking (in men) or obesity and raised blood pressure (in women). The transformation factors displayed in Table 1 show the weighting of each of the sets of synthetic estimates that make up the PCA variables. A value greater than zero for the unhealthy lifestyle variable implies that the general lifestyle of the population in the ward is more unhealthy than the England average, and vice versa.

The results of the multi-level regression models are shown in Table 2. The univariate models showed a strongly significant positive association between the unhealthy lifestyle 1 variable and the CHD outcomes (MODEL A), a negative association between the unhealthy lifestyle 2 variable and the CHD outcomes (MODEL B), and a strongly significant positive association between the deprivation index and CHD (MODEL C). The unhealthy lifestyle 2 variable explained little of the geographic variation in CHD mortality rates, and in multivariate models CHD rates were dominated by the unhealthy lifestyle 1 variable (MODEL D). For these reasons, only the unhealthy lifestyle 1 variable was retained for multivariate models including the deprivation index. Around 50% of the large scale (LA-level) geographic variation in mortality rates and between 10% and 15% of small scale (ward-level) geographic variation in mortality rates was explained by the unhealthy lifestyle 1 variable (MODEL A).

In multivariate analyses, the unhealthy lifestyle 1 variable was strongly significantly associated with CHD mortality rates at ward-level after adjustment for the deprivation level of the ward (MODEL E). Ward-level deprivation provides better explanatory power than the unhealthy lifestyle index: beta coefficients for the unhealthy lifestyle index were much more attenuated in the multivariate models compared to univariate models than the coefficients for ward-level deprivation. But inclusion of the unhealthy lifestyle index substantially increases explanatory power, particularly for large scale (LA-level) geographic variation. In comparison with univariate models of ward-level deprivation only, the multivariate models increased the amount of explained LA-level variation from 46% to 67% for male mortality rates and from 40% to 62% for female mortality rates.

Equivalent spatial error models were built with the unhealthy lifestyle index and ward-level deprivation as explanatory variables, and the results were compared with those produced by the multi-level models. In all cases the parameter estimates in the two different models were very similar, suggesting that the multi-level models have not been substantially affected by spatial autocorrelation bias (results not shown).

4. Discussion

The unhealthy lifestyle index developed for this paper is associated with CHD mortality rates independently of socio-economic deprivation. High prevalence rates of individual-level risk factors tend to cluster in certain areas and populations, and it is therefore not possible using this study design to attribute the geographic variation in CHD rates to specific risk factors. Therefore, in the absence of direct measures of individual-level risk factors, the index of unhealthy lifestyle is a useful tool to estimate how much of the variation in CHD rates between areas is due to differences in the lifestyle choices of the population living in those areas. The models reported here suggest that an increase in one standard deviation of this index is associated with an increase in age-standardised CHD mortality of 17.0 male deaths per 100,000, and 8.4 deaths per 100,000 in women. The difference between the best and worst ward in England is 12 standard deviations, representing a difference of 204 male deaths per 100,000 and 101 female deaths per 100,000—about the range from the 5th to 95th percentile of age-standardised CHD mortality rate.

In general, ecological analyses of chronic disease outcomes use a deprivation index as a proxy of differences in the behavioural risk factor profile of populations, since many risk factors for CHD are associated with socio-economic status. But deprivation indices are measures of the economic and social environment, rather than of the health environment. The index of unhealthy lifestyle introduced here aims to be a more direct measure of the differences in behavioural risk factor profiles of populations. An example of the difference between the two indices can be seen in inner-city wards, where the deprivation index and the unhealthy lifestyle index tend to have strongly opposing values implying that heavily urbanised, inner-city wards are heavily economically deprived, but their populations follow a reasonably healthy lifestyle (see Fig. 1). One reason for this may be the high residency of ethnic minority groups in inner-city wards (around 26% non-white in metropolitan wards at the 2001 census), who tend to have a healthier diet, lower smoking levels (particularly for females), more modest alcohol intake, lower blood pressure and lower cholesterol levels than the general population (Allender et al., 2008). Four of the five sets of

Table 2
Multi-level regression models of age-standardised CHD mortality rates per 100,000 against unhealthy lifestyle variables and against an index of socioeconomic deprivation (7929 wards nested in 355 local authorities).

	MODEL A	MODEL B	MODEL C	MODEL D	MODEL E
MEN					
Unhealthy lifestyle 1: Beta (SE)	17.9 (0.4)			17.0 (0.4)	7.2 (0.5)
Unhealthy lifestyle 2: Beta (SE)		−16.2 (0.9)		−10.2 (0.7)	
Deprivation: Beta (SE)			9.0 (0.2)		6.9 (0.2)
Ward-level variance explained (%)	16	5	24	18	25
LA-level variance explained (%)	49	0	46	57	67
WOMEN					
Unhealthy lifestyle 1: Beta (SE)	8.1 (0.2)			8.4 (0.2)	3.9 (0.3)
Unhealthy lifestyle 2: Beta (SE)		−0.7 (0.5)[a]		−3.8 (0.5)	
Deprivation: Beta (SE)			4.2 (0.1)		3.2 (0.1)
Ward-level variance explained (%)	11	0	17	11	17
LA-level variance explained (%)	45	3	40	51	62

Both unhealthy lifestyle variables and the index of deprivation are derived from z scores and are measured in standard deviations from the mean.

[a] Beta coefficients for all parameters were highly statistically significant ($p < 0.001$), with exception of unhealthy lifestyle 2 in the female model, where ($p = 0.153$).

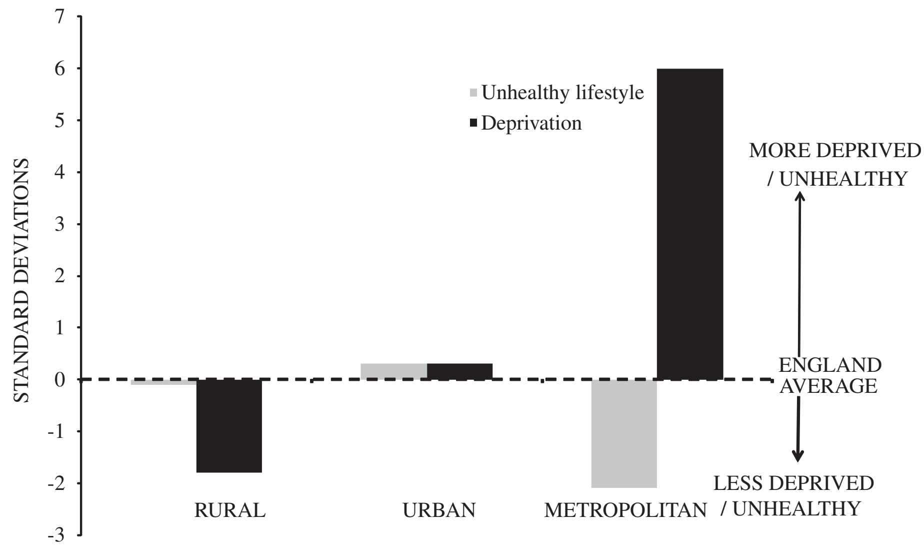

Fig. 1. Comparison of average values of Carstairs index of deprivation and index of unhealthy lifestyle in rural, urban and metropolitan wards, as defined by ONS area classification.

synthetic estimates included in the unhealthy lifestyle index include population-level measures of ethnicity as covariates in their supporting regression equations.

There are limitations of the unhealthy lifestyle index that are likely to lead to misclassification bias, and hence an underestimate of the impact of behavioural risk factors on geographic variations in CHD rates: the index is not based on direct estimates of the prevalence of risk factors; the index does not include estimates of all cardiovascular risk factors (e.g. diabetes is not included); the prevalence estimates included in the index require complex distributions of risk within an area to be collapsed to a single figure, which may obscure differences between areas (e.g. between a ward with a high percentage of elderly smokers, but few young smokers and vice versa).

A further limitation of the cross-sectional analyses reported here is that they only consider a snapshot of explanatory and outcome variables, and therefore they are poorly suited to consider conditions which have a prolonged latent period between exposure and disease, such as CHD. The effect of dismissing or inaccurately measuring risk exposure throughout the life course is that analyses will be biased towards under-estimating the cumulative impact of risk factors on health outcomes (Davey Smith and Hart, 2002). This potential bias could also be affected by migration between wards in England, which is not insignificant: 4% of people within the UK moved residence to a new location over 10 km away in 2000 (Office for National Statistics, 2006). However, an investigation of the British Regional Heart Study dataset showed that current residence was more strongly associated with CHD events than zone of birth, and migrants between towns were found to have a similar risk of CHD as individuals who have always lived in the towns under investigation (Wannamethee et al., 2002) suggesting that individuals who move tend to adopt similar risk levels to the population which they have migrated to, and lose the risk level of the population which they have migrated from. If this is the case, then the influence of migration on the findings reported here may not be substantial.

Possible areas of development of the unhealthy lifestyle index include an exploration of weighting the sets of synthetic estimates using population attributable fractions for the impact of individual-level risk factors on chronic disease developed for the Global Burden of Disease project (World Health Organisation, 2002). Also, efforts to incorporate the synthetic estimates of unhealthy behaviour developed by the National Centre for Social Research (Bajekal et al., 2004) which have been subjected to various validity assessments (Pickering et al., 2004), should be explored, which may allow for inclusion of synthetic estimates of binge drinking within the unhealthy lifestyle index.

Acknowledgement

Peter Scarborough and Steven Allender are supported by the British Heart Foundation, Grant number PPC/Jul05/5b. Mike Rayner are supported by the British Heart Foundation, Grant number PPC/JAN05/6biii. The funders had no role in study design, collection, analysis and interpretation of data, writing of the report, or decision to submit the paper for publication.

References

Allender, S., Peto, V., Scarborough, P., Kaur, A., Rayner, M., 2008. Coronary heart disease statistics 2008.. British Heart Foundation, London.
Anselin L., 2003. GeoDa version 0.9. University of Illinois, Illinois.
Bajekal, S., Scholes, S., Pickering, K., Purdon, S., 2004. Synthetic estimation of healthy lifestyle indicators: stage 1 report. Department of Health, London.
Carstairs, V., Morris, R., 1990. Deprivation and health in Scotland. Health Bulletin 48, 162–175.
Davey Smith, G., Hart, C., 2002. Life-course socioeconomic and behavioural influences on cardiovascular disease mortality: the collaborative study. American Journal of Public Health 92 (8), 1295–1298.
Dibben, C., Sims, A., Watson, J., Barnes, H., Smith, T., Sigala, M., Hill, A., Manley, D., 2004. The Health Poverty Index. South East Public Health Observatory, Oxford. 〈http://www.hpi.org.uk/index.php〉 (accessed January 2007).
EURAREA Consortium, 2004. Enhancing small area estimation techniques to meet European needs. Final project report. Office for National Statistics, London.
Heady, P., Clarke, P., Brown, G., Ellis, K., Heasman, D., Hennell, S., Longhurst, J., Mitchell, B., 2003. Model-based small area estimation series no. 2 Small area estimation project report. Office for National Statistics, London.
Morgan, O., Baker, A., 2006. Measuring deprivation in England and Wales using 2001 Carstairs scores. Health Statistics Quarterly 31, 28–33.
Morris, R., Whincup, P., Lampe, F., Walker, M., Wannamethee, S., Shaper, A., 2001. Geographic variation in incidence of coronary heart disease in Britain: the contribution of established risk factors. Heart 86, 277–283.
Office for National Statistics, 2006. NHSCR Inter-regional migration movements, within the UK during the year ending December 2005. Office for National Statistics, London.
Pickering, K., Scholes, S., Bajekal S., 2004. Synthetic estimation of healthy lifestyle indicators: stage 2 report. Department of Health, London.
Rasbash, J., Browne, J., Goldstein, H., 2003. MLwiN v2.02. Centre for Multilevel Modelling, Institute of Education, London.
Robinson, W., 1950. Ecological correlations and the behaviour of individuals. American Sociological Review 15, 351–357.
Romeri, E., Baker, A., Griffiths, C., 2006. Mortality by deprivation and cause of death in England and Wales, 1999–2003. Health Statistics Quarterly 32, 19–34.

Scarborough, P., Allender, S., Rayner, M., Goldacre, M., 2009. Validation of model-based estimates (synthetic estimates) of the prevalence of risk factors for coronary heart disease for wards in England. Health & Place 15 (2), 596–605.

Stamler, J., 2005. Established major coronary risk factors: historical overview. In: Marmot, M., Elliott, P. (Eds.), Coronary heart disease epidemiology: from aetiology to public health second ed. Oxford University Press, Oxford, pp. 18–31.

StataCorp, 2007. Stata statistical software: release 10. StataCorp LP, College station, Texas.

Subramanian, S., Jones, K., Kaddour, A., Krieger, N., 2009. Revisiting Robinson: the perils of individualistic and ecologic fallacy. International Journal of Epidemiology 38, 342–360.

Twigg, L., Moon, G., Jones, K., 2000. Predicting small area health-related behaviour: a comparison of smoking and drinking indicators. Social Science & Medicine 50, 1109–1120.

Twigg, L., Moon, G., 2002. Predicting small area health-related behaviour: a comparison of multilevel synthetic estimation and local survey data. Social Science & Medicine 54, 931–937.

Twigg, L., Moon, G., Walker, S., 2004. The smoking epidemic in England. Health Development Agency, London.

Wannamethee, S., Shaper, A., Whincup, P., Walker, M., 2002. Migration within Great Britain and cardiovascular disease: early life and adult environmental factors. International Journal of Epidemiology 31, 1054–1060.

West Midlands Public Health Observatory (WMPHO). European Standard Population. 〈http://www.wmpho.org.uk/localprofiles/metadata.aspx?id=META_EUROSTD〉 (Accessed January 2009).

World Health Organization, 2002. World Health Report 2002: Reducing risk, promoting healthy life. World Health Organization, Geneva.

World Heath Organization, 2003. Diet, nutrition and the prevention of chronic diseases. WHO Technical Report 916. World Health Organization, Geneva.

Yusuf, S., Hawken, S., Ounpuu, S., Dans, T., Avezum, A., Lanas, F., McQueen, M., Budaj, A., Pais, P., Varigos, J., Lisheng, L., on behalf of the INTERHEART study investigators, 2004. Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. The Lancet 364, 937–952.