Social Science & Medicine 67 (2008) 1995–2006

Contents lists available at ScienceDirect

Social Science & Medicine

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

Studying place effects on health by synthesising individual and area-level outcomes☆

Christopher H. Jackson [a,*], Sylvia Richardson [b], Nicky G. Best [b]

[a] MRC Biostatistics Unit, Institute of Public Health, Forvie Site, Robinson Way, Cambridge CB2 0SR, United Kingdom
[b] Imperial College, London, United Kingdom

ARTICLE INFO

Article history:
Available online 22 October 2008

Keywords:
Multilevel modelling
Data synthesis
Ecological bias
Contextual effects
Health inequalities
UK

ABSTRACT

It is well established that there exist substantial area-level socio-demographic variations in population health. However, area-level associations between deprivation and health cannot necessarily be interpreted as place effects on individual health. We demonstrate how recently developed statistical models for combining individual and aggregate data can help to separate the effects of place of residence and personal circumstances. We apply these to two health outcomes: risk of hospitalisation for cardiovascular disease (CVD) and risk of self-reported limiting long-term illness (LLTI). A combination of small-area data from UK hospital episode statistics and the UK census and individual data from the Health Survey for England are analysed, using a new multilevel modelling method termed hierarchical related regression (HRR). The standard multilevel model for place and health explains outcomes from individual data in terms of individual and area-level characteristics. HRR models increase precision by also explaining population aggregate outcomes, in terms of the same predictors. Aggregate outcomes are modelled by averaging the individual-level exposure–outcome relationship over the area, which can alleviate the ecological bias associated with interpreting the relationship between aggregate quantities as an individual-level relationship.

We find that there are associations between area-level deprivation indicators and both area-level rates of hospital admission for CVD and area-level rates of LLTI. Multilevel models fitted to the individual data alone had insufficient power to determine whether these associations were due to compositional or contextual effects. Using HRR models which incorporate area-level outcomes in addition to individual outcomes, we found that for CVD, the area-level differences were mostly explained by individual-level effects, in particular the increased risk for individuals from non-white ethnic backgrounds. In contrast, there remained a significant association between LLTI and area-level deprivation even after adjusting for the significant increased risk associated with individual-level ethnicity and income.

Our study illustrates that extending multilevel models to incorporate both individual and area-level outcomes increases power to distinguish between contextual and compositional effects.

☆ This work was supported by ESRC award numbers R000239598 and RES-576-25-5003. The second and third authors acknowledge partial support from AFSSE RD2004004 and INSERM-ATC A03150LS grants. We are grateful to the Small Area Health Statistics Unit at Imperial College London for access to geo-coded Health Survey for England and hospital admissions data. UK census data are Crown copyright.

* Corresponding author. Tel.: +44 1223330381.
E-mail addresses: chris.jackson@mrc-bsu.cam.ac.uk (C.H. Jackson), sylvia.richardson@imperial.ac.uk (S. Richardson), n.best@imperial.ac.uk (N.G. Best).

0277-9536/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.socscimed.2008.09.041

Background

An area-level association between socio-demographic indicators and health does not necessarily imply an effect of place of residence on individual health. It can equally be caused by individual-level differences in the characteristics of residents. Some researchers (e.g. Sloggett & Joshi, 1994, 1998) have suggested that geographical variations in mortality exist solely because of variation in the numbers of disadvantaged individuals, and that individuals with similar personal characteristics have similar risks wherever they live. Others (e.g. Ben-Shlomo, White, & Marmot, 1996; Davey-Smith, Hart, Watt, Hole, & Hawthorne, 1998) have suggested that there are "contextual" effects which are common to all individuals in a particular area. As discussed by Pickett and Pearl (2001) in a systematic review of studies of neighbourhood effects, and by Macintyre, Ellaway, and Cummins (2002), the mechanism by which place can exert an effect on health is complex and difficult to isolate. Such effects may reflect characteristics of the area itself, such as quality of health services, or aggregated characteristics of the individuals in the areas, such as average incomes or income variance (Ben-Shlomo et al., 1996). Environmental exposures, such as pollution, may also contribute to area health differences.

We compare the part played by individual socio-demographic factors and contextual effects on two important health outcomes. Firstly, we study the relationship of rates of hospital admission for cardiovascular disease (CVD) to various socio-demographic variables. There is evidence that area-level socioeconomic disadvantage (Leyland, 2005) or income inequality (Diez-Roux, Link, & Northridge, 2000) are associated with cardiovascular disease and its risk factors, and with coronary heart disease in particular (Diez-Roux et al., 2001). Secondly, we examine self-reported limiting long-term illness (LLTI). This was the only health outcome recorded at the 1991 UK census. This outcome is associated with several chronic illnesses such as arthritis, asthma, chronic bronchitis, heart disease and diabetes (Cohen, Forbes, & Garraway, 1995). Again, several studies have suggested that place of residence influences the risk of LLTI (Gould & Jones, 1996; Shouls, Congdon, & Curtis, 1996) and self-rated health (Subramanian, Kawachi, & Kennedy, 2001).

In order to distinguish between individual and area-level effects, suitable data and methodology are needed. Administrative data collected on small areas are useful for ascertaining variations in health between small areas, and evaluating the need for service provision. Examples of such data in the UK include birth, death and cancer registers (Office for National Statistics, UK), small-area statistics from the UK census, and Hospital Episode Statistics (Department of Health, UK). However, such aggregate data are often unsuitable for investigating individual determinants of health, due to ecological bias and lack of confounder information (Greenland & Morgenstern, 1989; Richardson, Stucker, & Hémon, 1987). In particular, they cannot usually be used to distinguish between individual and area-level effects (Greenland, 2001). To evaluate risk factors at the individual level, cohort and survey studies are more appropriate. However, these can often be too sparse to accurately determine geographical variations. To compare the influence of individual-level and area-level determinants of an individual's health, it is desirable to combine the information from individual and area-level data. Statistical power can be increased by "borrowing strength" between the two data sources, whilst ecological bias can be alleviated by appropriate specification of the underlying statistical model for the aggregate data.

Multilevel models separate the variation in an outcome into individual and group-level components, accounting for the correlation in the outcome within groups (Duncan, Jones, & Moon,1998; Merlo et al.,2005). But the commonly used multilevel model for place and health analyses only health outcomes from individual-level survey and cohort studies, explaining them in terms of individual and area-level socio-demographic predictors. In this paper we apply a new class of multilevel model in which population aggregate health outcomes are modelled alongside individual outcomes, in terms of the same predictors. A joint likelihood is specified for the two types of outcome, and this allows us to use outcome information, as well as just predictor information, from small-area administrative data, to increase statistical power. We have previously demonstrated theoretically when and how these models can reduce ecological bias and increase power (Jackson, Best, & Richardson, 2006), and used them in a study of socioeconomic risk factors for CVD, discussing the practical issues involved in this kind of data synthesis (Jackson, Best, & Richardson, 2008).

In this paper, we show how compositional and contextual effects on two major health outcomes can be better distinguished using these models. We compare with the results of a) ecological regressions of area-level rates on area-level variables alone, and b) standard multilevel models for the individual-level outcomes alone, neither of which are able to satisfactorily explain area-level associations as either an individual or place effect.

Data

Table 1 lists the sources of data used in the analyses of CVD and LLTI, whether they are individual or area-level (aggregate), their population coverage and whether the dataset includes the health outcome, socio-demographic predictors or both.

Area-level data sources

Hospital admissions for cardiovascular disease

Small-area counts of adults (individuals 16 years or older) admitted to hospital overnight or longer during the year 1998, with a primary diagnosis of CVD (ICD10 codes beginning with "I"), were examined. These were obtained from the Hospital Episode Statistics database for the 759 electoral wards (defined according to 1991 ward boundaries) in Greater London, which excludes some wards with small resident population.

Limiting long-term illness

The number of adults between 45 and 59 years of age reporting LLTI in the same study region, by ward, were obtained from the 1991 UK census.