OPEN ACCESS

University of BRISTOL

Moon, G., Twigg, L., Jones, K., Aitken, G., & Taylor, J. (2019). The utility of geodemographic indicators in small area estimates of limiting long-term illness. *Social Science and Medicine*, *227*, 47-55. https://doi.org/10.1016/j.socscimed.2018.06.029

University of Bristol - Explore Bristol Research
General rights

The utility of geodemographic indicators in small area estimates of limiting long-term illness

Graham Moon[a],*, Liz Twigg[b], Kelvyn Jones[c], Grant Aitken[d], Joanna Taylor[e]

[a] Geography and Environment, University of Southampton, Highfield, S017 1BJ, Southampton, UK
[b] Department of Geography, University of Portsmouth, UK
[c] School of Geographical Sciences, University of Bristol, UK
[d] Information Services Division, NHS Scotland, UK
[e] Freelance Researcher, Petersfield, UK

ABSTRACT

Small area health data are not always available on a consistent and robust routine basis across nations, necessitating the employment of small area estimation methods to generate local-scale data or the use of proxy measures. Geodemographic indicators are widely marketed as a potential proxy for many health indicators. This paper tests the extent to which the inclusion of geodemographic indicators in small area estimation methodology can enhance small area estimates of limiting long-term illness (LLTI). The paper contributes to international debates on small area estimation methodologies in health research and the relevance of geodemographic indicators to the identification of health care needs. We employ a multilevel methodology to estimate small area LLTI prevalence in England, Scotland and Wales. The estimates were created with a standard geographically-based model and with a cross-classified model of individuals nested separately in both spatial groupings and non-spatial geodemographic clusters. LLTI prevalence was estimated as a function of age, sex and deprivation. Estimates from the cross-classified model additionally incorporated residuals relating to the geodemographic classification. Both sets of estimates were compared against direct estimates from the 2011 Census. Geodemographic clusters remain relevant to understanding LLTI even after controlling for age, sex and deprivation. Incorporating a geodemographic indicator significantly improves concordance between the small area estimates and the Census. Small area estimates are however consistently below the equivalent Census measures, with the LLTI prevalence in urban areas characterised as 'blue collar' and 'struggling families' being markedly lower. We conclude that the inclusion of a geodemographic indicator in small area estimation can improve estimate quality and enhance understanding of health inequalities. We recommend the inclusion of geodemographic indicators in public releases of survey data to facilitate better small area estimation but caution against assumptions that geodemographic indicators can, on their own, provide a proxy measure of health status.

1. Introduction

Small area data on the prevalence of poor health are needed to plan health services and assess the quality of care. Such data are not always available on a consistent and robust routine basis across nations, necessitating the employment of small area estimation methods to generate local-scale data, or the use of proxy measures. Both approaches have flourished globally in recent years. Small area estimation has been used to identify health needs in the US (eg. Berkowitz et al., 2016), Australia (eg. Gong et al., 2012), the UK (eg. Twigg et al., 2004) New Zealand (eg. Smith et al., 2011), India (eg. Hirve et al., 2014) and many other countries. Alongside small area estimates, numerous geographical indices have also emerged, purporting to proxy health needs. These often build on research into health inequalities. Examples include the Index of (Multiple) Deprivation and its health domain in England (Morse, 2014; Noble et al., 2006) and analogous reformulations elsewhere in the United Kingdom, all building on the early examples of the Jarman, Carstairs and Townsend indices (Carstairs and Morris, 1989; Jarman, 1983; Townsend et al., 1988). Elsewhere, the NZDep index and a more recent health-focused measure have emerged in New Zealand (Atkinson et al., 2014; Exeter et al., 2017; Salmond et al., 2006), and similar national and local indices are evident in Canada (Bell et al., 2007; Pampalon et al., 2010, 2012) and many other countries.

Geodemographic indicators are another potential proxy for health

* Corresponding author.
  E-mail address: g.moon@soton.ac.uk (G. Moon).

Please cite this article as: Moon, G., Social Science & Medicine (2018), https://doi.org/10.1016/j.socscimed.2018.06.029

needs. With origins effectively in the 1970s but with earlier antecedents, and widely used in marketing, geography and planning, geodemographic indicators, like multiple deprivation measures, describe and classify people according to the type of area within which they live (Harris et al., 2005). Input data are usually sourced from the census but may often be combined with survey information on consumer and lifestyle behaviour. Intuitive labels (e.g. 'Suburban Achievers', 'Ageing Rural Dwellers', 'Migration and Churn') are a hallmark, attempting to typify the population in each location based on distinguishing, rather than majority, characteristics. Geodemographic typologies are available in many countries from national statistics agencies and the commercial companies. Well known examples include ACORN and the Output Area Classification (UK), PRIZM (US), CAMEO and MOSAIC (multiple countries), PSYTE (Canada) and geoSmart (Australia).

The key assumption underpinning geodemographic typologies is the idea that people who live in any one type of area will have similar lifestyles and retail consumption (and health). Historically, geodemographics have been widely used in the retail and marketing industry to allow businesses to differentiate and target customers (Longley and Goodchild, 2008). This focus continues to dominate but, increasingly, there has been a recognition of the potential in the health sector, often alongside other 'public sector' applications in areas as diverse as fire risk (Corcoran et al., 2013), policing (Ashby and Longley, 2005) and education (Singleton et al., 2012). These newer applications seek to foster evidence-based policy and develop interventions in a cost effective way, but also offer insights on inequality that supplement traditional approaches based solely on deprivation. The English Department of Health has highlighted this potential in guidance recommending the use of 'Geo-demographic profiling to identify association(s) between need and utilisation and outcomes for defined target population groups, including the protected population characteristics covered by the Equality Duty' (Public Health Development Unit, 2012 pp.10–11). Commercial geodemographic companies have recognised the opportunities in the health field. CACI have developed a specialist health-focused ACORN typology that 'delivers an improved understanding of local communities' needs and delivers an ability to target health and wellbeing improvement strategies' (CACI, 2013 p.6).

Although area targeting is a key rationale for commercial applications of geodemographics in health, many applications have also focused on understanding place-based health inequalities. Here geodemographics are used to gather aspects of area disadvantage and it is argued that they offer more discriminatory power than indices of multiple deprivation because of greater multidimensionality (Abbas et al., 2009); they capture both lifestyle and prosperity (Openshaw and Blake, 1995). A swathe of literature compares the performance of deprivation indices and various geodemographic products in explaining inequalities in a range of health outcomes: sexual health (Sheringham et al., 2009), colorectal cancer screening (Nnoaham et al., 2010), birth weight (Aveyard et al., 2002), dental health (Tickle et al., 2000, 2003), and smoking prevalence (Sharma et al., 2010). Very often, geodemographic approaches perform better or equally as well as other measures of area socio-economic disadvantage.

This paper sets out a novel approach to integrating both geodemographic typologies and proxy measures within a small area estimation framework, assesses claims about their relative importance, and tests the extent to which the inclusion of geodemographic indicators in small area estimation methodology can enhance small area estimates. The paper contributes to international debates on small area estimation methodologies in health research and the relevance of geodemographics indicators to the identification of health care needs. We address our aims by developing small area estimates of limiting long-term illness (LLTI) for England, Scotland and Wales, settings where there are long traditions of both geodemographic typologies and multiple deprivation measures.

Small area information on morbidity indicators, such as limiting long term illness (LLTI) are central to geographical comparisons of

health needs (Mooney and Rives, 1978). LLTI is good at identifying areas of high concentrations of ageing populations, pockets of chronic illness and revealing those areas characterised by extreme deprivation (Cohen et al., 1995; Jordan et al., 2000, 2003; Taylor et al., 2014). Power et al. (2000) argue that as life expectancy increases, dealing with chronic illness and disability consumes relatively large health and social services resources, so many countries now routinely collect information on LLTI. LLTI is also used to examine international differences in health performance (Lahelma et al., 1994) and monitoring differences in health over time (Charlton and Murphy, 1997). In the UK, a question on LLTI has been included in the census since 1991 and has been used in national funding models to allocate funding for health commissioning services (Chaplin et al., 2016). It has also been used to aid policy formulation and monitoring around population health, health inequalities and access to health services (ONS, 2010). The wording of the question is provided in the Data section below.

In the following section, we outline our approach to small area estimation, the data used to construct small area estimates, and the statistical measures used to assess our results. We then present and discuss our small area estimates of LLTI, highlighting implications for geodemographics. A short conclusion summarises our findings, acknowledges the limitations of our study and considers implications for policy.

2. Material and methods

In order to assess the utility of geodemographic indicators in the creation of small area estimates of LLTI, we developed small area estimates using two distinct approaches: one with and one without incorporation of a geodemographic indicator. The two sets of small area estimates were then compared to a perceived 'gold-standard' set of comparable data from the UK census.

2.1. Small area estimation

We used the multilevel small area estimation approach developed by Twigg et al. (2000). This approach has been widely used in subsequent UK health-related research (Moon et al., 2006; Moon et al., 2007; PHE, 2014; Szatkowski et al., 2015; Twigg and Moon, 2002; Twigg et al., 2004) and analogous approaches have been used in the US (Zhang et al., 2014, 2015) and in non-health research (Manley et al., 2017). It has been positively reviewed in governmental assessments (Bajekal et al., 2004) and comparative assessments (Berkowitz et al., 2016; Hirve et al., 2014; Jia et al., 2004). Forming part of a wider family of statistical approaches to small area estimation, it can be contrasted with the very different geocomputation approaches based on microsimulation, for which Birkin and Clarke (2012) have already outlined the potential of geodemographic typologies. The key advantage of multilevel small area estimation in comparison to other approaches, both statistical and microsimulation-based, is that it incorporates a recognition of the possibility that the general processes that predict a chosen outcome may vary locally.

The multilevel small area estimation process has been described in detail in Twigg et al. (2000). In summary, it involves modelling a survey data set using hierarchical logistic regression. A target outcome is modelled as a function of individual and area-level covariates that are co-present in a 'calibrating' data set covering all the desired small areas. An estimated logit at a chosen spatial unit for each group of individual characteristics (e.g. males aged 16–19) is adjusted by area level logits and a crucial additional adjustment using residual variance at higher geographies when the survey data set has respondents in all such higher-level areas. These adjusted logits are then untransformed to give a predicted prevalence proportion of the target outcome for each grouping at the desired small area level. These predicted proportions are then multiplied by their respective population counts to give the total numbers of the population within each grouping with the characteristic of interest. The prevalence of the population with the

characteristic of interest in each small area is then calculated by summing the group totals and dividing by the small area population total.

To date, multilevel small area estimates have all been derived from multilevel models defined using a strict hierarchy. For this example, the initial study by Twigg et al. (2000) had respondents nested within wards that were then nested within District Health Authorities. Such simple hierarchies reflect geographical groupings and assume a nested association between levels. In many cases, however, there is much more complexity in data structures. Goldstein (2011) showed how a basic two-level spatial multilevel structure of students (level 1) within schools (level 2) may be extended to additionally incorporate nesting in the administrative area (level 2) where the students live. Extending multilevel methodologies to address such structures creates two non-nested level 2 effects that need to be separated out in order to assess the effects of both schools and areas on whatever student outcome being studied.

Zaccarin and Rivellini (2002) suggest that three types of grouping can underpin these cross-classified structures: natural groupings, working groupings and theoretical groupings. The first refers to 'natural' neighbourhoods or attributes such as linguistic groups and is outwith the scope of this paper; data on natural neighbourhoods were not available and data on natural attributes were component parts of our selected deprivation and geodemographic measures (see below), and were thus incorporated separately within the small area estimation process. The second relates to settings defined by administrative geographies or locations; schools and administrative areal units fall under this heading. Theoretical groupings bring together individuals or areas using attribute data via clusters defined on theoretical grounds. Geodemographic typologies come under the banner of a theoretical grouping. This paper extends multilevel small area estimation by using a cross-classified multilevel model that combines working groupings and theoretical groupings. This is a potentially significant contribution as Jones et al. (1998) have suggested that, in the very different context of voting behaviour, greater variation may occur between theoretical groupings rather than between working groupings.

We compared two models. The first used a traditional spatial non-crossed three-level hierarchy (Model 1). We tested a model with individuals nested within Middle Layer Super Output Areas (MSOAs) in England and Wales or their equivalent Intermediate Zones (IZs) in Scotland, nested within Government Office Regions. MSOAs/IZs are official data reporting geographies with a mean population of 7787 in England and Wales, and 2500–6000 in Scotland. The region level proved superfluous, lacking significant variance, and was removed. We therefore report results for a two-level model. The choice of MSOA as the second level was prompted by the instability of models where individuals were nested within the smaller Lower Layer Super Output Areas. Other recent studies have also focused on the MSOA/IZ (Moon et al., 2017; Taylor et al., 2016). The second model explored the non-hierarchical cross-classified data structure, fitting cross-classified multilevel models to examine the relative importance of MSOAs/IZs and geodemographic groupings as sources of variation in the prevalence of LLTI (Model 2). Individuals within the same MSOAs/IZs could be in different geodemographic groupings.

Both models were initially produced using iterative generalised least squares with first order maximum quasi-likelihood estimation. A Monte Carlo Markov Chain (MCMC) approach was then used to refine the model and allow for more robust estimates and standard errors. Both MCMC models were run through 200,000 iterations, with an initial burn-in period of 50,000 iterations. The dataset was manipulated into individual and higher-level covariates using IBM SPSS Statistics version 22 and all multilevel models were computed using MLwiN version 2.24 (Browne, 2009; Rasbash et al., 2011). The two final MCMC models were used to derive two sets of small area estimates for all 7201 MSOAs and 1279 IZs in England, Wales and Scotland.

2.2. Data

The small area estimation models were developed using data from the 2011 version of the Annual Population Survey (APS), a major survey series that aims to provide data that can produce reliable estimates at the local authority level. Access to the data was controlled and was obtained via the UK Data Service Secure Data Service (Office for National Statistics Social Survey Division, 2017). We worked with the 2011 data in order to meet our objective of comparing small area estimates with census 'gold-standard' data; the last national census was held in 2011. Full details of the conduct of the APS, measurement of variables and response rates are given in the APS reports (Social Survey Division, 2012). Its 2011 sweep interviewed 331,934 individuals living in private households in Great Britain. We excluded individuals under the age of 16 (as the focus of the survey module containing the LLTI question was on individuals of a working age) and individuals aged 75+ (as they were asked a preliminary screening question as to whether they were too ill or distressed to answer the subsequent questions on health). Along with non-response, this left 213,001 individuals for analysis nested in 1520 MSOAs and 52 geodemographic groups.

Our outcome variable was LLTI. This was dichotomised to maximize concordance between its measurement in the 2011 Census and the APS (Table 1). In the Census, people were asked to define if they had any long-term illness, health problem or disability that limited their daily activities or the work they could do. If they responded 'yes, limited a lot' or 'yes, limited a little' they were coded as having a LLTI. In the APS, people were asked if they had any health problems or disabilities they expected to last more than a year. If they responded 'no' they were not coded as having a LLTI. If they responded yes, they were then asked a subsequent question about whether these problems or disabilities substantially limited their ability to carry out daily activities. If they responded 'no' they were not coded as having a LLTI and if they responded 'yes' they were coded as having a LLTI.

In the cross-classified models we used the National Statistics Output Area Classification (OAC) as our geodemographic indicator (Gale et al., 2016; Vickers and Rees, 2007). This classification uses census variables to identify groupings of census output areas (the smallest census area unit) with similar demographic, household, socio-economic and employment characteristics across England, Scotland, and Wales. The OAC is hierarchical, consisting of three tiers: super-groups, groups and sub-groups. We chose to work with the third tier of sub-groups to maximize

Table 1
Wording and coding of limiting-long term illness question for 2011 Census and the 2011 Annual Population Survey.

| Limiting Long Term Illness		Response		
2011 Census	Are your day-to-day activities limited because of a health problem or disability which has lasted, or is expected to last, at least 12 months? Include problems related to old age.	No	Yes, limited a little	Yes, limited a lot
	Coding	No LLTI	LLTI	
2011 Annual Population Survey	Do you have any health problems or disabilities that you expect will last for more than a year?	No	Yes	
	Do these health problems or disabilities, when taken singly or together, substantially limit your ability to carry out normal day-to-day activities? If you are receiving medication or treatment, please consider what the situation would be without medication or treatment.	No	Yes	
	Coding	No LLTI	LLTI if yes to second part	
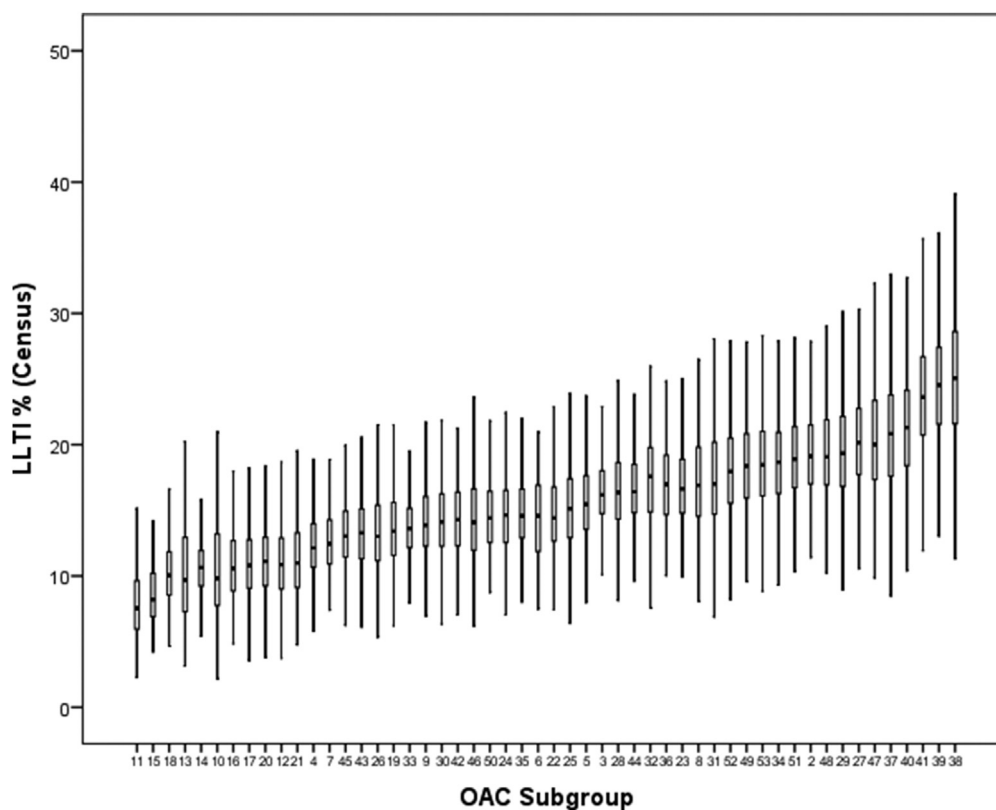
Fig. 1. LLTI and OAC sub-groups.

geodemographic variation. We worked with the 2001 OAC classification because the 2011 classification only emerged well after the 2011 APS had been conducted and we wished to use data available at the time of the 2011 census. Each individual within the APS dataset was allocated into one of the 51 OAC sub-groups; data for the 52nd sub-group (Countryside Communities A) was deleted as, with only two respondents, it destabilised the modelling. As the 51 groups all contained significant numbers of respondents (sample sizes varied from 413 to 15367) it was possible to use the OAC-level residuals from the cross-classified models to adjust the small area estimates from Model 2.

Individual-level covariates for the models were constrained by the provision of cross-tabulated UK census variables, which give individual count data. Models are usually restricted to a maximum of three individual-level covariates (Twigg et al., 2000). We worked with age and sex, drawing an analogy with age-sex standardization. Both the age and the sex variables were self-reported. Age was grouped into twelve categories, those aged 16 to 19 and five-year bandings up to the age of 74, and modelled as an orthogonal polynomial (Rasbash et al., 2017) enabling substantial model parsimony by reducing the twelve age terms to one. Both models were tested for interactions between the age and sex terms.

Because we had secure access, we were able to link additional MSOA-level variables to the APS. Taylor et al. (2016) have shown that, for small area estimation, linking in actual area data is preferable to aggregating individual data to create areal variables. An Index of Multiple Deprivation (IMD) score was added, calibrated from the English, Scottish and Welsh versions of the IMD using the method of Payne and Abel (2012). This index ensures that the model takes into account the strong evidence associating LLTI with multiple forms of deprivation (Martin et al., 1995). It also enabled us to examine whether geodemographic indicators add to the known association of health outcomes with deprivation measures (Goodman et al., 2011; Sheringham et al., 2009). The IMD variable was grand-mean centred and tests to assess cross-level interactions with age and sex were made.

The MSOA estimates from both models were compared with 2011 Census data on LLTI. MSOA Census data for England and Wales were downloaded from the ONS official labour market statistics website (NOMIS) and information for Scotland was downloaded from the Scottish Census data warehouse website.

2.3. Statistical analyses

Diagnostic analyses were performed to assess the performance of the two multilevel small area estimation models. We report the Deviance Information Criterion (DIC) statistic (Spiegelhalter et al., 2002) and higher-level random part variance. The DIC is a generalisation of the more familiar deviance-based measures of logistic model effectiveness. It is based on the MCMC runs of the model and penalises the deviance by the number of model parameters and levels, thus capturing both the fit and the complexity of the model. Models with a lower DIC are to be preferred and comparisons can be drawn between models and with null models without predictors (Jones and Subramanian, 2016). We also consider the random variance at the MSOA and OAC levels and their implications for residual (unexplained) variance.

To test for convergent validity of data from the two models against the 2011 Census, we adopted the methodology outlined by Scarborough et al. (2009) for validating small area estimates of the risk factors for coronary heart disease. We plotted MSOA-level small area estimates of LLTI against Census data for the same target variable, examining variation around the principal diagonal where the two sets of data would match exactly. Convergent validity would be achieved if a regression line through the data had a gradient close to one and an intercept around zero, that is a regression line matching the principal diagonal. For each model, we differentiated data points for the countries of England, Scotland and Wales.

Table 2
Descriptive statistics of APS sample characteristics.

Variable		Total sample (%)	Sample with LLTI (%)
LLTI	Yes	44555 (20.9%)	–
	No	168436 (79.1%)	–
Sex	Male	102514 (48.1%)	20507 (20.0%)
	Female	110487 (51.9%)	24059 (21.8%)
Age	16–19	15287 (7.2%)	1119 (7.3%)
	20–24	15790 (7.4%)	1356 (8.6%)
	25–29	16436 (7.7%)	1483 (9.0%)
	30–34	17629 (8.3%)	1952 (11.1%)
	35–39	20514 (9.6%)	2866 (14.0%)
	40–44	22835 (10.7%)	3930 (17.2%)
	45–49	22807 (10.7%)	4582 (20.1%)
	50–54	20636 (9.7%)	5304 (25.7%)
	55–59	19754 (9.3%)	6190 (31.3%)
	60–64	19488 (9.1%)	7101 (36.4%)
	65–69	11710 (5.5%)	4363 (37.3%)
	70–74	10115 (4.7%)	4320 (42.7%)
IMD	Mean	19.78	21.92
	Median	17.14	19.66
	Standard Deviation	11.38	11.95

3. Results and discussion

To provide an initial insight on the possible association between the OAC and LLTI we created output area box-whisker plots of LLTI prevalence for each of the 52 OAC sub-profiles groups, ordered by mean prevalence (Fig. 1). This shows clearly that LLTI prevalence varies substantially within each OAC sub-profile group and that the distributions overlap for all sub-groups. The mean ordering of the sub-groups also subverts the underlying ordering of classification. For example, sub-groups 2 and 3 both form part of the wider Countryside Communities group but are not close together in the mean ordering. On the basis of this preliminary insight, it would seem that the OAC has some value but may not be a particularly successful way of classifying LLTI prevalence. Whether or not this poor prognosis applies to small area estimates at the MSOA level or is affected by controls for age, sex and deprivation, remains open to question.

Table 2 summarises our data on LLTI in relation to the proposed model covariates. In total, 44,565 individuals were recorded as having a LLTI, approximately 21% of the sample population. LLTI prevalence was higher for people who were female and older. Mean and median IMD is slightly higher for those individuals who had a LLTI compared to the rest of the sample population. The difference for the sexes is relatively small whereas older people are substantially more likely to report an LLTI. These bivariate associations confirm the established findings linking LLTI to age, sex and deprivation (Cohen et al., 1995; Jordan et al., 2000) and point to the effectiveness of the APS as a representative source for population-level analysis.

The associations suggested in the bivariate analyses persist as independent effects in the two multivariable small area estimation models

Table 3
The Two Small Area Estimation Models; logits (standard errors) and model diagnostics.

	Model 1 (Hierarchical)	Model 2 (Cross-Classified)
Intercept	−1.609 (0.010)	−1.646 (0.034)
Female	0.128 (0.011)	0.125 (0.011)
Age	2.902 (0.024)	2.954 (0.025)
IMD	1.211 (0.024)	0.649 (0.034)
Random Part		
MSOA Variance	0.076 (0.005)	0.120 (0.008)
OAC Subgroup Variance		0.056 (0.013)
Summary Statistics		
DIC	198727	197483
DIC Drop	17568	16909

(Table 3). Consistent with the literature, LLTI is higher for women, at later ages and in areas characterised by higher deprivation. The effect for age is strongest. Models 1 and 2 are broadly similar but reveal important contrasts with respect to the effect of deprivation. This reduces substantially in the cross-classified Model 2 where the OAC is included as a (random) level. This suggests that deprivation and the OAC are, to an extent, capturing common ground with regard to associations with LLTI. The independent association with deprivation does however remain strongly statistically significant on a Wald Test ($p < 0.01$). The random part of Model 2 also confirms that the cross-classified model exhibits significant unexplained residual variation between OAC sub-groups, although there is greater variation between MSOAs. On balance, we conclude that both deprivation and the OAC are necessary and independent parts of our small area estimation models.

Table 3 also reports the model summary statistics, indicating overall model quality. As regularly noted in the literature, small area estimation models are not predicated on their explanatory power as they are affected by the numerous compromises entailed in their ultimate objective, that of making small area estimates (Rao and Molina, 2015). The models reported in this paper are also illustrative and parsimonious, aiming primarily to assess the incorporation of the OAC in the small area estimation process. Nonetheless, the summary DIC statistics suggest that both models offer significant insights in variations in LLTI. The two 'drop' statistics indicate the reduction in the DIC statistic from the equivalent measure in a null model with no covariates (not reported). Drops of these magnitudes both point to sound models. The lower overall DIC for Model 2 suggests that it is, despite its greater complexity, the stronger model.

Focusing on Model 2, we next examine its implications for variations in LLTI between OAC groups (Fig. 2). As noted above, we will be able to incorporate OAC-level variation in small area estimates. Each OAC sub-group is ranked on the Figure's X axis with the residuals on the Y axis capturing how each OAC sub-group either overestimates or underestimates the probability that a respondent has an LLTI (above or below the zero line). Four bars at the extreme right of the graph stand out (highlighted in red). APS respondents in these OAC sub-groups have far greater levels of LLTI than expected given their age and sex and the deprivation of their MSOAs. These sub-groups are blue-collar urban families a and b, and struggling urban families a and b and their
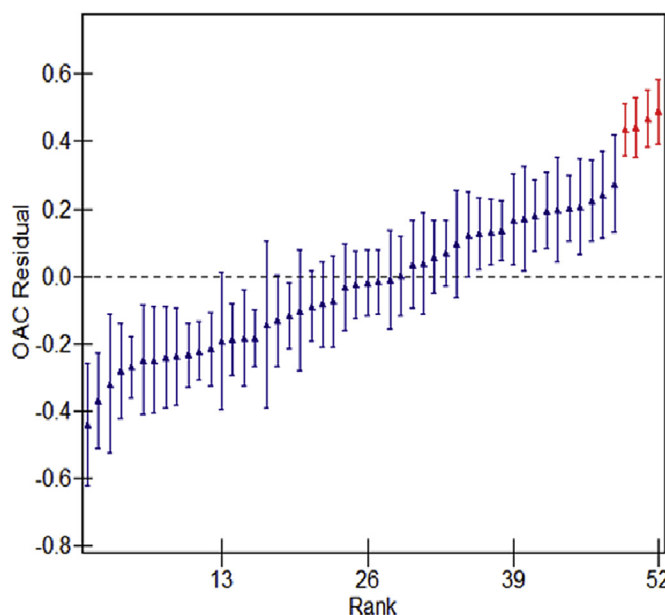
Fig. 2. OAC level residuals from the cross-classified model, 95% confidence intervals.
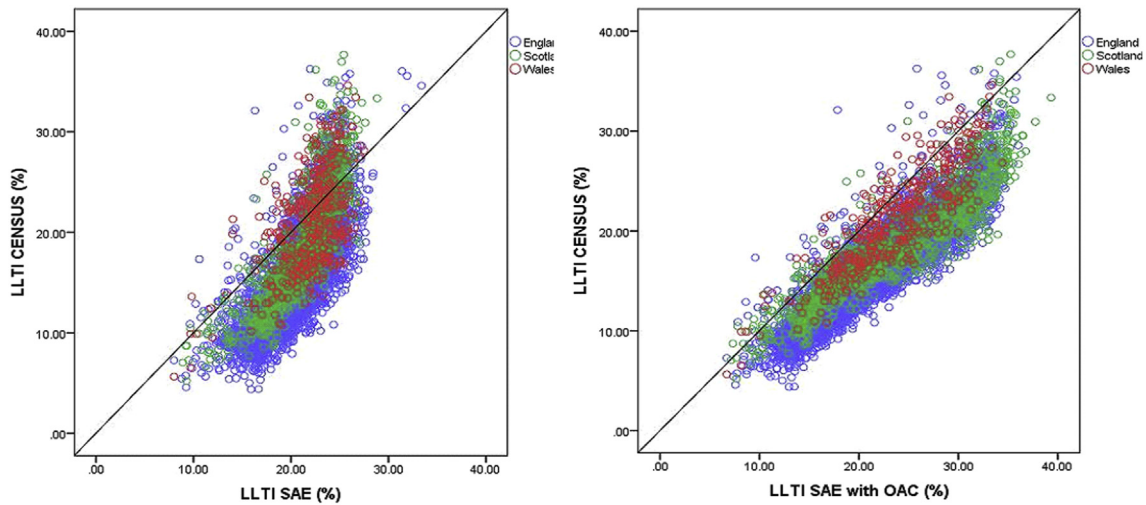
G. Moon et al.

Fig. 3. Small area estimates compared with Census data (MSOA/IZ level); Model 1 left, Model 2 right.

distinctiveness confirms that there is an urban penalty associated with LLTI (Zhang et al., 2013). Across the other OAC sub-groups there is much overlap between the bars, suggesting limited distinctiveness between sub-groups. Many also overlap zero, indicating that we cannot be sure if the OAC sub-group is associated with a lower or higher estimate of the probability of having an LLTI. The sub-groups to the left of the graph, for which LLTI is less likely than expected given age, sex and deprivation, are farming and forestry areas, educational centres, and countryside communities.

The extent to which these model-based findings play out as small area estimates depends on the distribution of the covariates (age, sex and deprivation) and the OAC across all MSOAs/IZs in England, Wales and Scotland. Fig. 3 compares the small area estimates from Models 1 and 2 with equivalent Census data indicating MSOAs/IZs in England, Scotland and Wales. The left-hand graph assesses the estimates from Model 1, the hierarchical model without OAC. A substantial departure from the principal diagonal line of equality is evident and the small area estimates exhibit a tighter range than the Census data and tend to be lower. The restricted range is characteristic of model-based small area estimates and is a consequence of their focus on fixed part (average) associations even in multilevel contexts (Twigg and Moon, 2002). The right-hand graph relates to the cross-classified Model 2 and presents a visual improvement. Incorporating consideration of the OAC-level residuals has reduced suggestions of curvilinearity in the association between the small area estimates and the Census data and brought the ranges of the two sets of estimates into greater alignment. The small area estimates remain lower, however, than those from the Census. We would not expect perfect concordance as the axes on the graphs draw on different data sources and the measurement of LLTI will be influenced by multiple aspects of survey design differing between the APS and the Census (Moon et al., 2017; Taylor et al., 2016).

These findings are confirmed in Table 4 by the assessment of concordance between the small area estimates and Census data using the Scarborough criteria (Scarborough et al., 2009). The cross-classified Model 2 exhibits a much higher coefficient of determination; including OAC residuals in the small area estimation process enables small area

estimates that capture 86% of the variation in equivalent Census indicators. The small area estimates are however clear underestimates. Neither model has a regression line that comes close to matching the desired principal diagonal; the cross-classified model is marginally closer to this goal but still, on average, produces underestimates. These tend to be greater, particularly for MSOAs with higher levels of LLTI.

Finally, we consider how the differences between the MSOA/IZ-level small area estimates and the equivalent Census data reflect specific OAC sub-groups (Fig. 4). This plot shows the 95% confidence intervals for the discrepancy between cross-classified SAEs and the census by OAC sub-group. OAC sub-groups to the left of the zero line are those for which the small area estimates of LLTI are lower than the estimates from the Census. To the right lie OAC sub-groups with small area estimates greater than the Census values. As expected, the majority of OAC sub-groups comprise MSOAs/IZs with small area estimates that are lower than the corresponding Census values and none are characterised by MSOAs/IZs that are definitively higher. A recognition that small area estimates represent expected values for the chosen indicator and an assumption that the Census provides a 'gold standard' observed value enables us to interpret this finding as confirmation that, in the case of the OAC, MSOAS/IZs categorised as Blue Collar Urban Families and Struggling Urban Families have more LLTI than expected. This interpretation can also be extended to all types of both Small Town Communities and Suburbia.

4. Conclusions

This paper has used a novel extension to multilevel small area estimation methodology incorporating a geodemographic indicator as a cross-classification alongside the more familiar spatial hierarchy of people nested within geographical settings. We draw three substantive conclusions. First, the addition of a geodemographic indicator enhances the quality of small area estimates. Second, geodemographic indicators have an independent impact on LLTI over and above that associated with multiple deprivation. Third, if we reject the reification of small area estimates as actual values of a target indicator – a common outcome when estimates are used to fill gaps in knowledge by providing information about an indicator that is otherwise not available - and recast them as expected values that can be compared with a (presumed) gold standard, areas characterised geodemographically as urban blue collar or struggling urban families not only have high levels of LLTI, they also have higher levels than expected.

Our study naturally has limitations that we must acknowledge. The geodemographics industry is global but our research focuses on England, Scotland and Wales using the OAC in the development of

Table 4
Small area estimate concordance with Census data; Scarborough criteria.

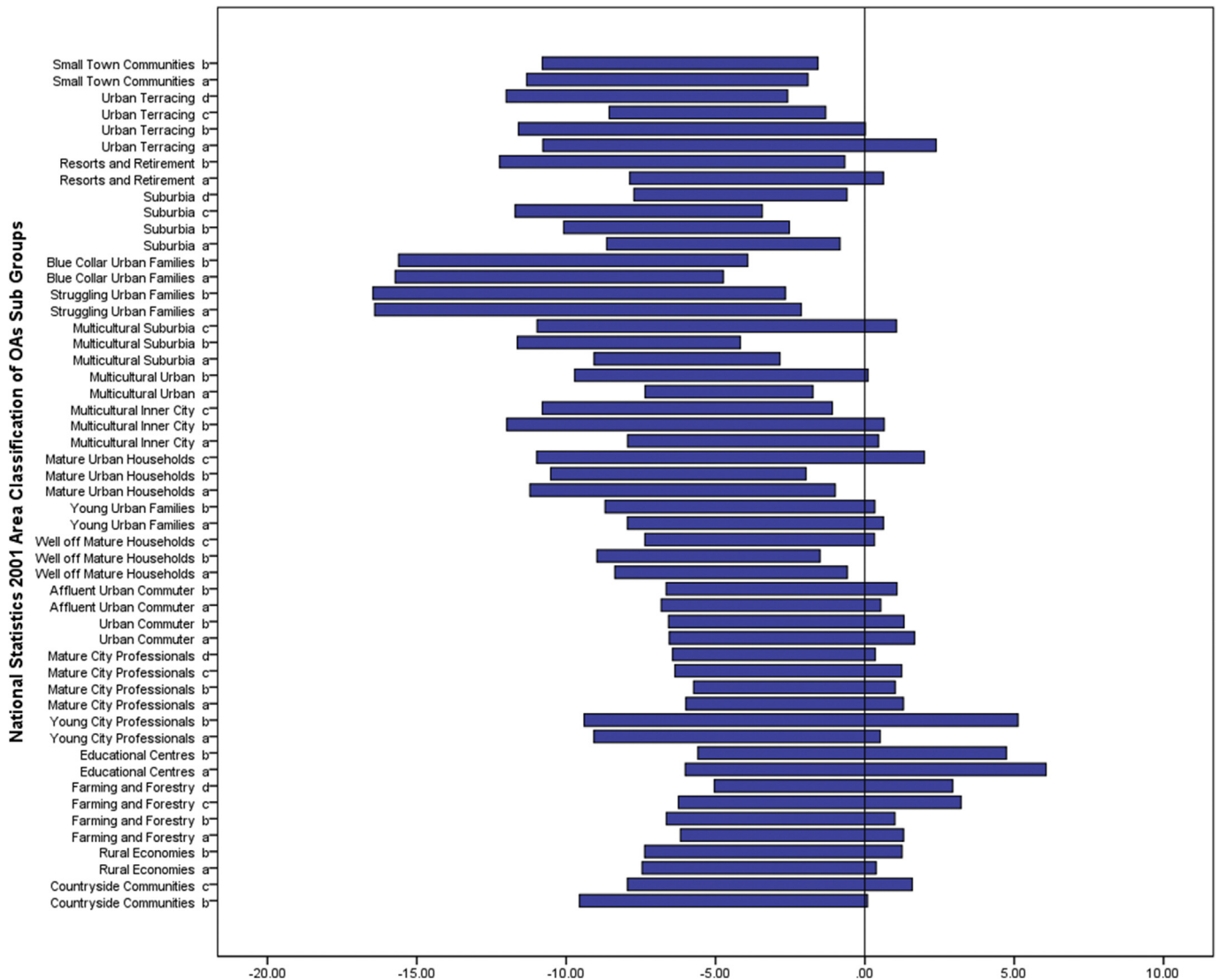|	Model 1		Model 2	
R^2	0.59		0.86	
Regression	Lower 95% CI	Upper 95% CI	Lower 95% CI	Upper 95% CI
Intercept	−12.07	−11.07	−0.62	−0.25
Slope	1.33	1.38	0.77	0.79

Fig. 4. OAC effects on small area estimates.

small area estimates of LLTI. Conclusions about the utility of geodemographic indicators may not necessarily be the same with other geodemographic indicators or with other outcomes or in other settings. We have however proceeded with a parsimonious and eminently transferable approach that would merit replication to investigate alternative inputs, outcomes and settings. More specifically, we also acknowledge that the OAC is a particular type of geodemographic classification, being dependent solely on variables from the national Census. Though Census health variables were helpfully incorporated in its construction, it does not, as yet, have a bespoke 'health' variant similar to, for example, CACI's ACORN Wellbeing product (CACI, 2013, 2017) incorporating multiple health-relevant data sources. The OAC is, however, an open source product that is freely available with a transparent methodology. We also acknowledge that more recent versions of both OAC and the deprivation measures are now available; our principles and conclusions remain transferrable however, and our focus on comparisons with the 2011 Census ' gold standard' necessitated our focus on data that were contemporaneously available with the Census.

Despite these caveats, our study has implications for policy as well as for the science of small area estimation and health geography. For the latter, we have pointed to contributions that highlight the case for including both geodemographic and deprivation measures in small area estimation and the scope for this inclusion within cross-classified

models. We have also noted important conclusions about health inequality that follow. For policy, the paper demonstrates that the association between health indicators and geodemographic classifications is more complex than advocacy of geodemographics may sometimes suggest. Out research underlines the importance of routinely linking geodemographic indicators to survey data to enhance small area estimation and other analytical capabilities.

Acknowledgements

References

Abbas, J., Ojo, A., Orange, S., 2009. Geodemographics - a tool for health intelligence? Publ. Health 123, e35–e39.

Ashby, D.I., Longley, P.A., 2005. Geocomputation, geodemographics and resource allocation for local policing. Trans. GIS 9, 53–72.

Atkinson, J., Salmond, C., Crampton, P., 2014. NZDep2013 Index of Deprivation. Ministry of Health, New Zealand.

Aveyard, P., Manaseki, S., Chambers, J., 2002. The relationship between mean birth weight and poverty using the Townsend deprivation score and the Super Profile classification system. Publ. Health 116, 308–314.

Bajekal, M., Scholes, S., Pickering, K., Purdon, S., 2004. Synthetic Estimation of Healthy Lifestyles Indicators: Stage 1 Report. London.

Bell, N., Schuurman, N., Oliver, L., Hayes, M.V., 2007. Towards the construction of place-specific measures of deprivation: a case study from the Vancouver metropolitan area. Can. Geogr./Le Géographe Canadien 51, 444–461.

Berkowitz, Z., Zhang, X., Richards, T.B., Peipins, L., Henley, S.J., Holt, J., 2016. Multilevel small-area estimation of multiple cigarette smoking status categories using the 2012 behavioral risk factor surveillance system. Cancer Epidemiol. Biomark. Prev. 25, 1402–1410.

Birkin, M., Clarke, G., 2012. The enhancement of spatial microsimulation models using geodemographics. Ann. Reg. Sci. 49, 515–532.

Browne, W., 2009. MCMC Estimation in MLwiN. Centre for Multilevel Modelling. University of Bristol.

CACI, 2013. The Wellbeing ACORN User Guide: the health and wellbeing classification. London.

CACI, 2017. Wellbeing ACORN: technical guide. London.

Carstairs, V., Morris, R., 1989. Deprivation: explaining differences in mortality between Scotland and England and Wales. BMJ 229, 886–889.

Chaplin, M., Beatson, S., Lau, Y.-S., Davies, C., Smyth, C., Burrows, J., Weir, R., Tatarek-Gintowt, R., 2016. Refreshing the Formulae for CCG Allocations for Allocations to Clinical Commissioning Groups from 2016-17 Report on the Methods and Modelling. Leeds.

Charlton, J., Murphy, M., 1997. The Health of Adult Britain 1841-1994. ONS, London.

Cohen, G., Forbes, J., Garraway, M., 1995. Interpreting self reported limiting long term illness. BMJ 311, 722.

Corcoran, J., Higgs, G., Anderson, T., 2013. Examining the use of a geodemographic classification in an exploratory analysis of variations in fire incidence in South Wales, UK. Fire Saf. J. 62, 37–48.

Exeter, D.J., Zhao, J., Crengle, S., Lee, A., Browne, M., 2017. The New Zealand Indices of Multiple Deprivation (IMD): a new suite of indicators for social and health research in Aotearoa, New Zealand. PLoS One 12, e0181260.

Gale, C.G., Singleton, A.D., Bates, A.G., Longley, P.A., 2016. Creating the 2011 area classification for output areas (2011 OAC). J. Spatial Inf. Sci. 12, 1–27.

Goldstein, H., 2011. Multilevel Statistical Models. John Wiley & Sons, Chichester.

Gong, H.C., McNamara, J., Vidyattama, Y., Miranti, R., Tanton, R., Harding, A., et al., 2012. Developing spatial microsimulation estimates of small area advantage and disadvantage among older Australians. Popul. Space Place 18, 551–565.

Goodman, A., Wilkinson, P., Stafford, M., Tonne, C., 2011. Characterising socio-economic inequalities in exposure to air pollution: a comparison of socio-economic markers and scales of measurement. Health Place 17, 767–774.

Harris, R., Sleight, P., Webber, R., 2005. Geodemographics, GIS and Neighbourhood Targeting. Wiley, London.

Hirve, S., Vounatsou, P., Juvekar, S., Blomstedt, Y., Wall, S., Chatterji, S., et al., 2014. Self-rated health: small area large area comparisons amongst older adults at the state, district and sub-district level in India. Health Place 26, 31–38.

Jarman, B., 1983. Identification of underprivileged areas. Br. Med. J. 286, 1705.

Jia, H., Muennig, P., Borawski, E., 2004. Comparison of small-area analysis techniques for estimating county-level outcomes. Am. J. Prev. Med. 26, 453–460.

Jones, K., Gould, M., Watt, R., 1998. Multiple contexts as a cross-classified models: the labor vote in the British general election. Geogr. Anal. 30, 65–93.

Jones, K., Subramanian, V.S., 2016. Developing Multilevel Models for Analysing Contextuality, Heterogeneity and Change, vol.1 Centre for Multilevel Modelling. University of Bristol, Bristol.

Jordan, K., Ong, B.N., Croft, P., 2000. Researching limiting long-term illness. Soc. Sci. Med. 50, 397–405.

Jordan, K., Ong, B.N., Croft, P., 2003. Previous consultation and self reported health status as predictors of future demand for primary care. J. Epidemiol. Community Health 57, 109–113.

Lahelma, E., Manderbacka, K., Rahkonen, O., Karisto, A., 1994. Comparisons of inequalities in health: evidence from national surveys in Finland, Norway and Sweden. Soc. Sci. Med. 38, 517–524.

Longley, P.A., Goodchild, M.F., 2008. The use of geodemographics to improve public service delivery. In: Hartley, J., Donaldson, C., Skelcher, C., Wallace, M. (Eds.), Managing to Improve Public Services. CUP, Cambridge.

Manley, D., Jones, K., Johnston, R., 2017. The geography of Brexit – what geography? Modelling and predicting the outcome across 380 local authorities. Local Econ. 32, 183–203.

Martin, S., Sheldon, T.A., Smith, P., 1995. Interpreting the new illness question in the UK census for health research on small areas. J. Epidemiol. Community Health 49, 634–641.

Moon, G., Altankhuyag, G., Barnard, S., Twigg, L., 2006. Smoking Prevalence in Scotland: 2003/4 sub-national estimates. NHS Health Scotland, Glasgow.

Moon, G., Quarendon, G., Barnard, S., Twigg, L., Blyth, B., 2007. Fat nation: deciphering the distinctive geographies of obesity in England. Soc. Sci. Med. 65, 20–31.

Moon, G., Aitken, G., Taylor, J., Twigg, L., 2017. Integrating national surveys to estimate small area variations in poor health and limiting long-term illness in Great Britain. BMJ Open 7, e016936.

Mooney, A., Rives, N., 1978. Measures of community health status for health planning. Health Serv. Res. 13, 129–145.

Morse, S., 2014. Index of multiple deprivation (IMD)(UK). In: Encyclopedia of Quality of Life and Well-being Research. Springer, pp. 3165–3168.

Nnoaham, K.E., Frater, A., Roderick, P., Moon, G., Halloran, S., 2010. Do geodemographic typologies explain variations in uptake in colorectal cancer screening? An assessment using routine screening data in the south of England. J. Publ. Health 32, 572–581.

Noble, M., Wright, G., Smith, G., Dibben, C., 2006. Measuring multiple deprivation at the small-area level. Environ. Plann. 38, 169–185.

Office for National Statistics Social Survey Division, 2017. Annual Population Survey 2004-2016: Secure Access [data Collection], sixth ed. UK Data Service SN: 671.

ONS, 2010. Final recommended questions for the 2011 census in England &Wales: health. Titchfield.

Openshaw, S., Blake, M., 1995. Geodemographic segmentation systems for screening health data. J. Epidemiol. Community Health 49, S34–S38.

Pampalon, R., Hamel, D., Gamache, P., 2010. Health inequalities in urban and rural Canada: comparing inequalities in survival according to an individual and area-based deprivation index. Health Place 16, 416–420.

Pampalon, R., Hamel, D., Gamache, P., Philibert, M.D., Raymond, G., Simpson, A., 2012. An area-based material and social deprivation index for public health in Québec and Canada. Can. J. Public Health/Rev. Can. Sante Publique S17–S22.

Payne, R., Abel, G., 2012. UK indices of Multiple Deprivation - a Way to Make Comparisons across Constituent Countries Easier. Office for National Statistics, London.

PHE, 2014. Public Health England National Cardiovascular Intelligence Network: Chronic Kidney Disease (CKD) Prevalence Model.

Power, C., Li, L., Manor, O., 2000. A prospective study of limiting longstanding illness in early adulthood. Int. J. Epidemiol. 29, 131–139.

Public Health Development Unit, 2012. Guidance to Support the Provision of Healthcare Public Health Advice to Clinical Commissioning Groups. London.

Rao, J., Molina, I., 2015. Small Area Estimation. John Wiley & Sons, London.

Rasbash, J., Steele, F., Browne, W.J., Goldstein, H., 2011. A User's Guide to MLwiN. Centre for Multilevel Modelling. University of Bristol v2.24.

Rasbash, J., Charlton, C., Jones, K., Pillinger, R., 2017. Manual Supplement for MLwiN. Centre for Multilevel Modelling, Bristol Version 3.01.

Salmond, C., Crampton, P., King, P., Waldegrave, C., 2006. NZiDep: a New Zealand index of socioeconomic deprivation for individuals. Soc. Sci. Med. 62, 1474–1485.

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, 596–605.

Sharma, A., Lewis, S., Szatkowski, L., 2010. Insights into social disparities in smoking prevalence using Mosaic, a novel measure of socioeconomic status: an analysis using a large primary care dataset. BMC Publ. Health 10.

Sheringham, J., Sowden, S., Stafford, M., Simms, I., Raine, R., 2009. Monitoring inequalities in the National Chlamydia Screening Programme in England: added value of ACORN, a commercial geodemographic classification tool. Sex. Health 6, 57–62.

Singleton, A.D., Wilson, A., O'Brien, O., 2012. Geodemographics and spatial interaction: an integrated model for higher education. J. Geogr. Syst. 14, 223–241.

Smith, D.M., Pearce, J.R., Harland, K., 2011. Can a deterministic spatial microsimulation model provide reliable small-area estimates of health behaviours? An example of smoking prevalence in New Zealand. Health Place 17, 618–624.

Social Survey Division, 2012. Annual population survey: April 2011-March 2012. In: Office for National Statistics, (Colchester, Essex, UK Data Archive).

Spiegelhalter, D., Best, N., Carlin, B., van der Linde, A., 2002. Bayesian measures of model complexity and fit. Stat. Meth. 64, 583–639.

Szatkowski, L., Fahy, S.J., Coleman, T., Taylor, J., Twigg, L., Moon, G., et al., 2015. Small area synthetic estimates of smoking prevalence during pregnancy in England. Popul. Health Metrics 13.

Taylor, J., Twigg, L., Moon, G., 2014. The convergent validity of three surveys as alternative sources of health information to the 2011 UK census. Soc. Sci. Med. 116, 187–192.

Taylor, J., Moon, G., Twigg, L., 2016. Using geocoded survey data to improve the accuracy of multilevel small area synthetic estimates. Soc. Sci. Res. 56, 108–116.

Tickle, M., Blinkhorn, A.S., Brown, P.J.B., Matthews, R., 2000. A geodemographic analysis of the Denplan patient population in the North West Region. Br. Dent. J. 189, 494–499.

Tickle, M., Milsom, K.M., Jenner, T.M., Blinkhorn, A.S., 2003. The geodemographic distribution of caries experience in neighboring fluoridated and nonfluoridated populations. J. Publ. Health Dent. 63, 92–98.

Townsend, P., Phillimore, P., Beattie, A., 1988. Health and Deprivation: Inequality and the North. Croom Helm, London.

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

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

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

Vickers, D., Rees, P., 2007. Creating the UK national statistics 2001 output area classification. J. Roy. Stat. Soc. 379–403.

Zaccarin, S., Rivellini, G., 2002. Multilevel analysis in social research: an application of a cross-classified model. Stat. Meth. Appl. 11, 95–108.

Zhang, X., Cook, P.A., Lisboa, P.J., Jarman, I.H., Bellis, M.A., 2013. The effects of deprivation and relative deprivation on self-reported morbidity in England: an area-

level ecological study. Int. J. Health Geogr. 12.

Zhang, X., Holt, J.B., Lu, H., Wheaton, A.G., Ford, E.S., Greenlund, K.J., et al., 2014. Multilevel regression and poststratification for small-area estimation of population health outcomes: a case study of chronic obstructive pulmonary disease prevalence using the behavioral risk factor surveillance system. Am. J. Epidemiol. 179,

1025–1033.

Zhang, X., Holt, J.B., Yun, S., Lu, H., Greenlund, K.J., Croft, J.B., 2015. Validation of multilevel regression and poststratification methodology for small area estimation of health indicators from the behavioral risk factor surveillance system. Am. J. Epidemiol. 182, 127–137.