Synthetic estimation of healthy lifestyles indicators: Stage 3 report

Kevin Pickering, Shaun Scholes, Madhavi Bajekal

Synthetic estimation of healthy lifestyles indicators: Stage 3 report

Kevin Pickering, Shaun Scholes, Madhavi Bajekal

Prepared for the Department of Health

February 2005

Acknowledgements

We would like to thank members of the Steering Group and the Technical Group for their helpful comments on this report. Colleagues at the Survey Methods Unit at the National Centre for Social Research contributed generously of their time in developing and refining methodology through regular seminar discussions. We would particularly like to express our appreciation of the work of Department of Health staff at all stages of the project and in particular the contribution made by Rosemary Aldridge and Tracie Kilbey.

SUMMARY

- Estimates have been produced for five health behaviours at ward and PCO level: smoking, obesity, binge drinking and fruit and vegetable consumption separately for adults and children. (Chapter 3)

- Confidence intervals have been produced for the synthetic estimates at ward and PCO level. (Chapter 3)

- The confidence intervals for wards were too wide to be able discriminate between wards with any degree of statistical confidence. However, this would be possible for PCOs. (Chapter 3)

- We would not recommend ranking the wards or to assigning them to bands. Given the large confidence intervals for the estimates, the confidence intervals around the ranks would be very wide. Assigning the wards to bands would still require the uncertainty to be represented (Bird *et al.,* 2003)– even the extreme wards/PCOs in terms of point estimates would straddle several bands. (Chapter 3)

- The internal validation did not identify any substantive problems with the synthetic estimates produced. (Chapter 4)

- The validation of the synthetic estimates with the HSfE 2003 data shows that the estimates performed well at ward level. The correlations with the direct estimates for the higher levels (PCO and SHA) were larger than for wards. This represents the larger sample sizes from the HSfE 2003 and, hence, higher precision for the synthetic estimates at these higher levels. The comparisons with the other sources of survey data confirmed the sizes of these correlations. (Chapter 5)

1 INTRODUCTION

The National Centre for Social Research (NatCen) was commissioned by the Department of Health to produce estimates of healthy lifestyle behaviours using Health Survey for England (HSfE) data. This report describes the work undertaken for the third stage (*implementation*) of that project.

1.1 Aims of the project

The main aims of the project were:

- to evaluate the technical feasibility of producing robust small-area estimates for a range of health indicators;
- to validate model-based estimates against other sources of information and local knowledge;
- to develop and apply a consistent methodology to produce synthetic estimates for small areas, focusing on a maximum of five health behaviours available from the HSfE.

The key requirement of the synthetic estimation project was to provide robust estimates that were calculated on a consistent basis for all areas of the country and which allowed meaningful comparisons within and between local areas.

There were three stages to the project – *scoping and feasibility*, *testing and validation*, and *implementation*.

1.2 Implementation

Synthetic estimates and the associated confidence intervals have been generated for five health behaviours:

- current smoking for adults (aged 16 years or more);
- obesity for adults (aged 16 years or more);
- binge drinking for adults (aged 16 years or more);
- consumption of five or more portions of fruit and vegetables for adults (aged 16 years or more); and
- consumption of three or more portions of fruit and vegetables for children (aged from 5 to 15 years inclusive).

(See Chapter 3 for more information about the derivations of these variables.)

The estimates have been produced at two levels[1]: Census Area Statistics (CAS) wards and Primary Care Organisations (PCO). To assess the accuracy of the estimates

[1] Estimates were also generated for Strategic Health Authorities (SHA) in order to validate the estimates. These estimates have not been published.

produced, they have been compared with direct measures from a number of surveys including the Health Survey for England 2003 (Chapter 5).

2 GENERATING THE SYNTHETIC ESTIMATES

The *scoping and feasibility* study (Stage 1) identified three potential methods for generating area-level estimates: from a statistical model combining individual and area-level covariates (the TWIGG method, Twigg *et al.*, 2000); from a statistical model using area-level covariates only (the ONS method, Heady *et al.*, 2003); and using indirect standardisation. (See Bajekal *et al.* (2004) for more details of these methods.) These three methods were evaluated in the *testing and validation* study (Stage 2), and an assessment made as to which was most suitable for obtaining ward-level estimates (Pickering *et al.*, 2004).

It was found that the two model-based methods (TWIGG and ONS) generated estimates that were more accurate than those from indirect standardisation. As the two model-based methods performed similarly on a range of statistical criteria, the decision to use area-level covariates only (the ONS method) to generate the synthetic estimates was reached for non-statistical reasons (see Pickering *et al*, 2004): the ONS method is easier and quicker to implement and is more simple to explain to users; the method is more transferable to other output variables (beyond the scope of the project) and can be more easily updated; and the method has already been used by the ONS to generate small area estimates of income (which are available via the Neighbourhood Statistics website).

In this chapter we give a brief overview of the ONS method, and outline how the synthetic estimates and their confidence intervals were produced for each ward in England.

2.1 Estimates for wards

2.1.1 Computing the synthetic estimates

For each health behaviour, a series of multilevel logistic regression models were fitted in the statistical package Stata (StataCorp, 2003) using the available ward-level covariates (listed in Appendix A) as predictor variables. Only covariates that were significantly associated ($p<0.05$) with the healthy lifestyle measures were retained in the models, using a selection procedure described in Appendix B.

The final (optimal) model obtained by the fitting procedure in Stata, was then re-fitted in MLwiN (Rasbash *et al.*, 2004). The MLwiN and final Stata models both allowed for the clustering within postcode sectors of the HSfE samples by fitting this as an area-level random effect term. The synthetic estimates were then generated by multiplying the fixed parameter estimates from the MLwiN model to the corresponding known covariate values at the ward level. (See Appendix B of the User Guide (Scholes *et al.,* 2005) for a worked example of generating the ward estimates.) The synthetic estimates were derived from the MLwiN model rather than the Stata model so that they were consistent with their confidence intervals which could only be derived from MLwiN (see Section 2.1.2).

2.1.2 Computing the confidence intervals

Markov Chain Monte Carlo (MCMC) methods, within a Bayesian framework, were used to generate the confidence intervals (actually referred to as 'credible intervals') for the synthetic estimates. One of the key differences between Bayesian statistics and traditional (frequentist) methods is that the parameter estimates are treated as random with corresponding (*prior*) probability distributions; there is no single point estimate for each parameter as there would be for traditional statistical methods. To generate the estimates for parameters, it is necessary to run an iterative procedure (the MCMC procedure) that generates a series of values for each parameter. The sample of values for each parameter can then be used to estimate, for example, the mean and variance for a parameter.

The confidence intervals for the synthetic estimates were derived as follows. First, we generated a set of 100,000 values for each parameter using the MCMC procedure in MLwiN. For each ward, we then generate 100,000 feasible values of the synthetic estimate – one for each set of parameter estimates. We then simulated 100,000 estimates of the *true* measure for the ward by including a random term derived from the area-level variance[2]; this is done by drawing from the normal distribution with the appropriate estimate of the variance (equal to the area-level variance) and zero mean. So, the estimate of the *true* value based on the r^{th} set of parameters for ward *k* would be:

$$E(\hat{\bar{Y}}_k^r) = logit^{-1}(\alpha^r + \underline{\beta}^r \underline{\bar{X}}_k^r + v_k^r), \text{ where } v_k^r \sim N(0,\sigma^{2(r)}).$$

The confidence interval (credible interval) for the synthetic estimate for the ward can then be derived directly from this set of feasible *true* estimates - the 95% credible interval as the range between the 2,500th and 97,500th largest feasible *true* estimate (i.e. the 2.5th and 97.5th percentiles).

2.2 Estimates for PCOs and SHAs

Estimates for PCOs and SHAs were obtained as weighted averages of the estimates for the wards within them. So, the estimate for PCO *m* would be:

$$E(\hat{\bar{Y}}_m) = \frac{\sum_{k=1}^{l} n_k E(\hat{\bar{Y}}_k)}{\sum_{k=1}^{l} n_k},$$

where $k = 1, 2, \ldots, l$ represents the wards in PCO *m* and n_k the corresponding population estimates. (See Appendix B of the User Guide (Scholes *et al.*, 2005) for a worked example of generating the PCO estimates.)

The corresponding confidence intervals were generated using a similar MCMC method as for wards (see Section 2.1.2). For each PCO, the set of MCMC draws was

[2] Note that the area-level term in the model is for postcode sectors not wards. However, we can assume that the variance across postcode sectors and wards would be fairly similar (Heady *et al.*, 2003).

used to generate a set of *true* PCO estimates as the weighted average of the *true* ward estimates. We then obtained the 95% credible interval for the synthetic estimate for a PCO directly as the range between the 2,500th and 97,500th largest simulated *true* estimate for the PCO.

Note that although the CAS wards (the principal estimation area chosen for this project) nest within higher-level administrative tiers such as Local Authority Districts and Government Office Regions they do not nest perfectly into larger health areas such as PCOs. We were provided a 'best-fit' one-to-one look-up table to uniquely attribute whole wards to a PCO and so there will be a degree of lack of fit between the two geographies, primarily at the edges. Ideally, the confidence intervals should be adjusted to allow for this. However, because there are on average more than 25 wards per PCO, we do not think the error introduced will be substantial overall[3].

[3] Research at the ONS is investigating producing small area estimates at different levels and addressing this problem. However the results of this research will not be available until 2005.

3 THE STATISTICAL MODELS

This chapter describes the details of the multilevel logistic regression models from which the synthetic estimates were derived. For each health behaviour, the characteristics of the modelling dataset are described and the optimal model shown and discussed.

The variables in the tables for the models (Tables 3.1 to 3.5) have been sorted by the magnitude of the T^2-ratio (the square of the estimate divided by the square of its standard error). The greater the T^2-ratio, the larger the statistical significance of the term and hence the greater the association between the covariate and the healthy lifestyle measure (allowing for the other covariates in the model). A value of T^2 ratio greater than 4 identifies a variable that is significantly associated with the healthy lifestyle measures, i.e. whose estimate is statistically significantly different from zero at the 0.05 level[4].

During the process of model selection only those covariates that were significantly different from zero (i.e. with a T^2-ratio greater than 4) were considered for inclusion in the models (see Appendix B). Note, however, that some covariates were retained in the model even though their T^2-ratio was below 4 if they were included in a significant interaction term.

Each item is therefore retained in the model because it has a significant correlation with the healthy lifestyle measure (allowing for the other covariates), not because we consider there to be a direct association between them. This means that the models have been developed to explain as much of the area level variance as possible, whilst including only significant terms, and hence should not be interpreted as if explanatory models.

3.1 Current smoking

The individual-level measure of smoking was generated from the HSfE measure of "current smoking status". Adult respondents (aged 16 or more) to the HSfE were defined to be current smokers if they reported that they were a "current cigarette smoker", and not a current smoker if they reported that they had "never smoked cigarettes at all", "used to smoke cigarettes occasionally" or "used to smoke cigarettes regularly".

Respondents that did not answer, refused to answer or did not know (155 from 31,027 respondents) were excluded. Therefore, the modelling dataset consisted of completed records for 30,872 adults from the combined HSfEs from 2000 to 2002, of which 7,972 (25.8%) reported that they were current smokers. The combined samples covered 3,231 of the 7,958 wards in England, with a wide range for the number of respondents per ward - 225 wards contained only one respondent, whereas the maximum number in any ward was 60.

[4] This is (approximately) equivalent to a z-score with magnitude greater than approximately 1.96 in a standard regression table.

Table 3.1 shows the covariates in the optimal model for whether adults currently smoked in the HSfE 2000-2002.

Table 3.1 Parameter estimates for smoking model

Variable	Estimate	Standard error	T^2-ratio
Main effects only:			
Proportion 16+ residing as couple (icouple)	-2.158	*0.266*	**66.02**
Proportion female, aged 25-34	5.108	*0.744*	**47.14**
Proportion non-white (iethnic)	-0.914	*0.156*	**34.56**
Proportion professional & managerial occupations (aged 16-74) (iprofman)	-1.324	*0.275*	**23.20**
Proportion of working age with limiting longstanding illness (illsiwrk)	3.860	*1.159*	**11.09**
Attendance allowance claimant rate (aarate)	-1.637	*0.537*	**9.29**
3rd most deprived IMD decile (imd8)	0.138	*0.048*	**8.43**
North West GOR (gor_nw)	0.533	*0.189*	**7.94**
IMD education skills and training score (eduscore)	0.006	*0.002*	**6.91**
5th least deprived IMD decile (imd5)	0.119	*0.056*	**4.55**
South West GOR (gor_sw)	-0.137	*0.067*	**4.20**
Interactions:			
aarate / gor_nw	-3.583	*1.102*	**10.57**
iethnic / gor_sw	3.573	*1.256*	**8.09**
imd8 / gor_sw	0.376	*0.143*	**6.94**
Intercept	0.082	*0.311*	

NOTE: A Bold T^2 - ratio indicates statistical significance ($p < 0.05$)

The ward-level characteristics associated with increased propensity for a person to be a current smoker were: a higher proportion of females aged 25-34; a higher proportion of residents of working age who had a limiting longstanding illness; an Index of Multiple Deprivation (IMD) decile ranking of 5 or 8; being located in the North West region; and a relatively higher education, skills and training deprivation score.

The ward-level characteristics associated with decreased propensity for a person to be a current smoker were: a higher proportion of household residents over the age of 16 who were living as a couple; a higher proportion of non-white residents; a higher proportion of residents who were classified as being in managerial and professional occupations; a higher attendance allowance claimant rate; and being located in the South West region.

There were three interaction terms in the model, each a ward-level characteristic with a regional indicator. This implies that there was evidence to suggest that those characteristics had different relationships with current smoking in different regions: the association between smoking and attendance allowance claimant rate was more strongly negative in the North West region; the proportion of non-white residents was associated with an increase rate of smoking in the South West region compared with a decrease for the other regions; and the association between smoking and being in the third most deprived decile was stronger for South West region.

3.2 Obesity for adults

The individual-level measure of obesity was generated from the height and weight of the adult respondents (aged 16 or more), as measured by the HSfE interviewers. The BMI was derived as: the weight in kilograms divided by the square of the height in meters. Respondents were defined to be obese if their BMI was 30 or more and to be not obese if the BMI was less than 30.

Respondents whose height and weight were not measured, or who refused to be measured (3,907 from 31,027) were excluded. Therefore, the modelling dataset consisted of completed records for 27,120 adults from the combined HSfEs from 2000 to 2002, of which 5,991 (22.1%) were defined to be obese. The combined sample covered 3,149 of the 7,958 wards in England, with a wide range for the number of respondents per ward - 244 wards contained only one respondent, whereas the maximum number in any ward was 56.

Table 3.2 Parameter estimates for obesity model

Variable	Estimate	Standard error	T^2-ratio
Main effects:			
South West GOR (gor_sw)	-0.196	*0.056*	**12.30**
LA IMD 2004 score (laidscor)	-0.006	*0.002*	**8.35**
East of England GOR (gor_ee)	-0.398	*0.145*	**7.51**
Proportion highest qualification NVQ1 or no qualifications (aged 16-74) (iolevel)	0.919	*0.355*	**6.72**
Proportion of dwellings in council tax band G (propctxg)	-0.738	*0.353*	**4.36**
Proportion of households without central heating (hloamnty)	-0.528	*0.259*	**4.15**
Proportion in semi-routine and routine occupations (NS-SEC 6&7) (aged 16-74) (isroutin)	0.725	*0.397*	3.34
Rural indicator (rural)	-0.113	*0.117*	0.94
Interaction terms:			
rural / isroutin	0.789	*0.381*	**4.29**
laidscor / gor_ee	0.018	*0.009*	**4.30**
Intercept	-1.712	*0.101*	

NOTE: A Bold T^2 - ratio indicates statistical significance ($p < 0.05$)

Table 3.2 shows the covariates in the optimal model for whether adults were measured as obese in the HSfE 2000-2002.

The ward-level characteristics associated with an increased propensity for a person to be obese was a higher proportion of residents whose highest educational qualification was NVQ level 1 or with no qualifications at all.

The ward-level characteristics associated with lower propensity for a person to be obese were: being located in either the South West or East of England regions; a higher multiple deprivation score (at the Local Authority District level); a higher proportion of dwellings being in council tax band G (£160,001 to £320,000); and a higher proportion of households being without central heating.

There were two interaction terms in the model: the association with obesity of being in a ward with a higher proportion that were classified as being in semi-routine and routine occupations (aged 16-74) was greater in rural wards than non-rural wards; and in the East of England region, living in a ward with a higher multiple deprivation score (at the Local Authority District level) was positively associated with being obese, compared with a negative association for the other regions.

3.3 Fruit and vegetable consumption (children)

The individual-level measure for consuming three or more portions of fruit and vegetables for children (aged 5 to 15 inclusive) was generated from the data collected in the HSfE about the quantities of different types of fruit and vegetables consumed on the previous day. These measures were combined[5] to give the total number of portions of fruit and vegetables consumed (see Sproston & Primatesta (2003) for more details). Note that information about fruit or vegetable consumption was not collected in the HSfE 2000, nor for children under 5 years old in the HSfE 2001 and 2002.

Of the 8,438 children (aged 5 to 15 years) in the 2001 and 2002 HSfE, 3,163 (37.5%) had consumed three or more portions of fruit and vegetables. The combined sample covered 1,989 of the 7,958 wards in England. The number of respondents per ward varied widely - 400 wards contained only one child, whereas the maximum number in any ward was 28.

Table 3.3 shows the covariates in the optimal model for whether children consumed three or more portions of fruit and vegetables in the HSfE 2001-2002. .

The ward-level characteristics associated with an increased propensity for children to eat more than three portions of fruit and vegetables were: a higher proportion of residents who were not born in the UK, Ireland or the European Union; a higher estimated life expectancy at birth for males (at the Local Authority District level); a

---

[5] The definitions of what constituted a portion changed between the 2001 and 2002 HSfEs. The revisions stated that dried fruit and pulses were capped at 1 portion; a portion of pulses is 3 tablespoons rather than 2; and a portion of small fruit is 2 handfuls rather than 1.We have used the current definitions in this report.

higher proportion permanently sick or disabled aged 16-74; and being located in either the West Midlands or Yorkshire & The Humber regions.

Table 3.3 Parameter estimates for fruit and vegetable consumption model (children)

Variable	Estimate	Standard error	T^2-ratio
Main effects:			
Proportion highest qualification NVQ1 or no qualifications (aged 16-74) (iolevel)	-2.438	*0.401*	**36.98**
Proportion not born in UK, Ireland or EU (icobnuk)	2.928	*0.622*	**22.17**
Life expectancy at birth for males (LA-level) (lemale)	0.101	*0.022*	**21.69**
Proportion permanently sick/disabled (aged 16-74) (ipermsic)	15.920	*3.549*	**20.12**
Yorkshire & The Humberside GOR (gor_yh)	0.750	*0.232*	**10.48**
West Midlands GOR (gor_wm)	0.254	*0.090*	**8.01**
IMD employment score (empscore)	-3.095	*1.533*	**4.08**
Proportion unpaid carers >50 hours per week (iupdcr50)	-16.520	*8.265*	**4.00**
London GOR (gor_lon)	0.317	*0.168*	3.55
North West GOR (gor_sw)		*0.211*	3.13
Interactions:			
ipermsic / gor_yh	-13.170	*3.615*	**13.27**
empscore / gor_sw	-5.455	*2.277*	**5.74**
icobnuk / gor_lon	-2.093	*0.889*	**5.55**
Intercept	-7.392	*1.671*	

NOTE: A Bold T^2 - ratio indicates statistical significance ($p < 0.05$)

The ward-level characteristics associated with a decreased propensity for children to eat more than three portions of fruit and vegetables were: a higher proportion of residents whose highest educational qualification was NVQ level 1 or with no qualifications at all; being more deprived based on the IMD employment score; and a higher proportion of unpaid carers working more than 50 hours per week.

There were three interaction terms in the model, each of which were interactions of ward-level characterises with regions: in Yorkshire and the Humberside region, there was no association between fruit and vegetable consumption and the proportion permanently sick or disabled compared with a positive relationship in the other regions; in the South West region, there was a greater association between employment deprivation and a lower consumption of fruit and vegetables; and in London region, the positive association of being in a ward with a relatively higher proportion of residents who were not born in the UK, Ireland or the European Union was reduced.

3.4 Fruit and vegetable consumption (adults)

The individual-level measure for consuming five or more portions of fruit and vegetables for adults (aged 16 or more) was generated from the data collected in the HSfE about the quantities of different types of fruit and vegetables consumed on the previous day. These measures were combined[6] to give the total number portions of fruit and vegetables consumed (see Sproston & Primatesta (2003) for more details). Note that information about fruit or vegetable consumption was not collected in the HSfE 2000.

Of the 23,040 adults in the 2001 and 2002 HSfE, 5,460 (23.7%) had consumed five or more portions of fruit and vegetables. The combined sample covered 2,644 of the 7,958 wards in England. The number of respondents per ward varied widely - 211 wards contained only one respondent, whereas the maximum number in any ward was 60.

Table 3.4 Parameter estimates for fruit and vegetable consumption model (Adults)

Variable	Estimate	Standard error	T^2-ratio
Main effects:			
Proportion not born in UK, Ireland or EU (icobnuk)	2.210	*0.302*	**53.62**
Proportion female aged 25-34	-3.904	*0.913*	**18.29**
South East GOR (gor_se)	1.613	*0.391*	**16.98**
Proportion female aged 16-19	-14.140	*3.459*	**16.71**
Mortality from stroke (smr_10a)	0.007	*0.002*	**14.91**
Mortality from lung cancer (smr_14b)	-0.003	*0.001*	**14.83**
Proportion in semi-routine and routine occupations (NS-SEC 6&7) (aged 16-74) (isroutin)	-1.152	*0.335*	**11.82**
Yorks & The Humber GOR (gor_yh)	0.667	*0.220*	**9.19**
2nd most deprived decile (imd9)	0.135	*0.056*	**5.87**
Built-up areas (ONS classification of wards)	0.282	*0.127*	**4.94**
Proportion unpaid carers >50 hours per week (iupdcr50)	-10.680	*4.850*	**4.85**
Interactions:			
smr_10a / gor_se	-0.015	*0.004*	**14.11**
isroutin/ gor_yh	-2.117	*0.646*	**10.73**

[6] The definitions of what constituted a portion changed between the 2001 and 2002 HSfEs. The revisions stated that dried fruit and pulses were capped at 1 portion; a portion of pulses is 3 tablespoons rather than 2; and a portion of small fruit is 2 handfuls rather than 1.We have used the current definitions in this report.

Intercept	-0.389	*0.197*

NOTE: A Bold T^2 - ratio indicates statistical significance ($p < 0.05$)

Table 3.4 shows the covariates in the optimal model for whether adults had consumed five or more portions of fruit and vegetables in the HSfE 2001-2002.

The area-level variables most strongly associated with an increase in propensity for adults to consume more than five portions of fruit and vegetables per day were: a higher proportion of residents who were not born in the UK, Ireland or the European Union; a higher standardised mortality rate for strokes (at the Local Authority District level); being located in the South East region or Yorkshire & The Humber region; being in the second most deprived IMD decile category; and living in a built-up area.

The area-level variables associated with a lower propensity for adults to consume more than five portions of fruit and vegetables per day were: a higher proportion of females aged 16-19 and 25-34; a higher standardised mortality rate for lung cancer (at the Local Authority District level); a higher proportion of residents aged 16-74 who were classified as being in semi-routine and routine occupations; and a higher proportion of residents who were unpaid carers caring for more than fifty hours per week.

There were two interaction terms in the model, each of which were interactions of ward-level characterises with regions: the negative association with being in a ward with a relatively higher proportion of residents in semi-routine and routine occupations (aged 16-74) was stronger in Yorkshire & The Humber; and in the South East region, being in a ward with a higher standardised mortality rate for strokes (at the Local Authority District level) had a negative association with adults' fruit and vegetable consumption compared with a positive association for the other regions.

3.5 Binge drinking

The individual-level measure of binge drinking was generated from the data collected in the HSfE about the quantities of all the different types of alcoholic drinks (beer, wine, spirits, sherry and alcopops) consumed on a respondent's heaviest drinking night in the previous week. These measures were combined to give the number of units of alcohol consumed on the heaviest drinking day (see Sproston & Primatesta (2003) for more details). Binge drinking was then defined separately for men and women: men were defined as having indulged in binge drinking if they had consumed 8 or more units of alcohol on the heaviest drinking day in the previous seven days; for women the cut-off was 6 or more units of alcohol.

Respondents for whom any information on drinking was either not answered, refused or not known (583 from 31,027) were excluded. Therefore, the survey data file contained 30,440 adults (aged 16+) from the combined 2000 to 2002 HSfEs, of which 5,539 (18.2%) reported binge drinking. The combined sample covered 3,230 of the 7,958 wards in England. The number of respondents per ward varied widely - 230 wards contained only one respondent, whereas the maximum number in any ward was 60.

Table 3.5 shows the covariates in the optimal model for binge drinking in the HSfE 2000-2002.

Table 3.5 Parameter estimates for binge drinking

Variable	Estimate	Standard error	T^2-ratio
Main effects:			
Proportion 16+ residing as couple (icouple)	-3.589	0.469	**58.49**
Proportion non-white (iethnic)	-1.881	0.251	**56.25**
North East GOR (gor_ne)	0.435	0.077	**31.85**
Income support claimant rate (israte)	-3.028	0.584	**26.87**
North West GOR (gor_nw)	0.284	0.059	**23.21**
Yorkshire & The Humber GOR (gor_yh)	0.268	0.061	**19.16**
Proportion male, aged 75-79	-16.700	4.513	**13.69**
South West GOR (gor_sw)	-0.832	0.261	**10.18**
Proportion female, aged 50-54	-8.756	3.122	**7.87**
Proportion female, aged 85 or more	-11.330	4.172	**7.38**
Proportion of overcrowded households (hovercr)	-1.959	0.745	**6.91**
South East GOR (gor_se)	-0.247	0.098	**6.43**
Proportion male, aged 45-49	8.515	4.049	**4.42**
Attendance allowance claimant rate (aarate)	0.529	0.630	0.70
Interactions:			
aarate / gor_sw	5.291	1.780	**8.84**
hovercr / gor_se	2.600	1.195	**4.73**
Intercept	1.430	0.332	

NOTE: A Bold T^2 - ratio indicates statistical significance ($p < 0.05$)

The area-level variables associated with increased binge drinking were: living in the North East, North West or Yorkshire & The Humber regions; and a higher proportion of males aged 45-49.

The area-level variables associated with decreased binge drink were: a higher proportion of household residents (aged 16+) who were living as a couple; a higher proportion of residents whose ethnic background was non-white; an increased income support claimant rate; a higher proportion of males aged 75-79; being in South West or South West regions; higher proportion of females aged 50-54 or 85 or more; and a higher proportion of overcrowded households.

There were two interaction terms in the model: there was a positive association between attendance allowance rates and binge drinking in South West region, whereas there was no association in the other regions; and there is no (or a small positive) association between overcrowding and binge drinking in South East region, compared with a negative relationship in the other regions.

3.6 Discussion of the models

3.6.1 Adequacy of the models

One of the conclusions of the *testing and evaluation* study (Pickering *et al.*, 2004) was that the quality of the synthetic estimate was related to both the amount of area-level variation for the healthy lifestyle measure and the proportion of that variance explained by the multilevel model. The estimates of variance for the five healthy lifestyle measures are shown in Table 3.6. We recommended that there should be adequate area-level variance (which we did not quantify) and that 40% of the area-level variance should be explained by the multilevel model as an absolute minimum to generate synthetic estimates for a healthy lifestyle measure.

The amount of area-level variance was relatively high (greater than 0.22) for all the healthy lifestyle measures, except for obesity (0.089), and the area-level variance explained was greater than 40% for all five measures adopted for this implementation stage - the minimum estimate being 47.9% for the consumption of fruit and vegetables for children. We were therefore confident that the synthetic estimates produced should be fairly good for at least four of health behaviours. To test this further we have performed both internal and external forms of validation (see Chapters 4 and 5).

Table 3.6 Proportion of area-level variance explained for each healthy lifestyle measure

	Between area variance (null model)	Residual between area variance (final model)	% Area-level variance explained
Smoking (adults)	0.228	0.094	58.8%
Obesity (adults)	0.089	0.046	48.4%
Consumption of fruit and vegetables (children)	0.305	0.159	47.9%
Consumption of fruit and vegetables (adults)	0.229	0.110	52.1%
Binge drinking (adults)	0.260	0.129	50.3%

3.6.2 Confidence intervals

The major factor influencing the width of the confidence intervals for the synthetic estimates was the extent of the between-area variation that remained unexplained once the significant area-level predictors in the model had been included (see Table 3.6). For a particular estimate of the proportion, the greater the amount of between-area variance remaining unexplained by the models the wider the confidence intervals. As Table 3.6 shows, the magnitude of the unexplained (residual) between-area variation varied across the five health healthy lifestyle measures. As a result the average width of the confidence intervals, both for wards and PCOs, also varied across the healthy lifestyle measures. Table 3.7 shows that the confidence intervals are widest for fruit and vegetable consumption for children and smallest for obesity, reflecting the magnitude of the unexplained (residual) between-area variance for these measures.

Table 3.7 Average width of the confidence intervals for the 5 healthy lifestyle behaviours (wards and PCOs)

	Average width of the confidence interval	
	Wards	*PCOs*
% currently smoke	± 11%	± 3%
% obese	± 7%	± 2%
% consuming 3 or more portions of fruit or vegetables (children)	± 19%	± 5%
% consuming 5 or more portions of fruit or vegetables (adults)	± 12%	± 3%
% binge drinking	± 10%	± 3%

The size of the confidence intervals implies that there would be little scope for discriminating between wards. As an extreme example, only 12 from the 7,958 wards were significantly different from the estimate of smoking prevalence for the median ward. Therefore, it is very important that the ward estimates are used with great caution; in the vast majority of cases, it would not be possible to state that the prevalence in one ward was higher than that in another with any degree of statistical confidence.

At PCO level the confidence intervals are narrower and, hence, comparisons between PCOs would be possible. For example, 90 of the 303 PCOs were significantly different from the estimate of smoking prevalence for the median PCO.

3.6.3 Banding/ranking

Given the width of the confidence intervals, it would not be appropriate to rank or band the ward-level estimates. Applying ranks or bands gives the false impression of precision, whereas in fact, for these estimates, the uncertainty around the bands or ranks would be so wide as to make them meaningless. As an example, consider banding the ward-level smoking estimates into three groups representing high, medium and low estimates of smoking prevalence – highest 10% of wards (band 1), middle 80% of wards (band 2) and lowest 10% of wards (band 3). When one considers the confidence intervals, each ward would cover at least two bands, with 2,209 (from 7,958) covering all three bands.

For PCOs, it might seem that there would be more potential for using ranks and banding, as the confidence intervals were much narrower. Using the same example, if we banded the PCO-level estimates of smoking the same three groups – highest 10% of PCOs (band 1), middle 80% of PCOs (band 2) and lowest 10% of PCOs (band 3) – then some of the PCOs could be placed with confidence into one of the groups: the confidence intervals for six of the 30 PCOs with the lowest prevalence would cover just band 1; for the PCOs with the highest prevalence, three would cover just band 3. However, this would still mean that there would not be sufficient evidence to state with confidence (at the 0.05 level) that 51 of the 60 PCOs identified as being in one of the extreme bands were not actually in the middle band.

Given the false impression of precision that ranking and banding gives, and the inability to confidently place wards or PCOs into specific bands, we would not recommend using bands or ranking to present these data.

4 INTERNAL VALIDATION

The quality of the model-based synthetic estimates is highly dependent upon how well the models explain the observed prevalence at ward level. This is measured in part by the proportion of the variance explained by the model. However, it is also important to test whether there are any relationships between the healthy lifestyle measures and ward characteristics that have not been included in the model. To do this, we performed diagnostic checks on the models - these are described in this chapter.

4.1 Residual versus model estimates

The variance at the area level can be decomposed into the component that is explained by the model and the component that is unexplained (the residual variance). Examination of the unexplained component of the area variance allows one to investigate whether there are relationships between the healthy lifestyle measure and ward characteristics that were missed by the model. If this were the case, then we would expect to see patterns in the difference between the observed and predicted prevalence (the residuals).

To test whether there was any evidence of a linear relationship between the residuals and the predicted estimates (and hence a relationship that was not explained by the model), a regression line was fitted (see Appendix C). If the intercept and slope were not significantly different from zero, then there was no evidence of a linear pattern and hence no evidence of an unexplained relationship. (Note that this test is, in essence, a test of whether the residuals are symmetric for all values of the predicted prevalence – any asymmetry being a sign of a problem with the model.)

Note that as the HSfE sample is clustered within postcode sectors (PCSs), the multilevel model was fitted with PCSs as the second level and hence the area-level residuals were derived for PCSs rather than wards. We therefore plotted the residuals for PCSs against the model-based estimates for PCSs (generated by taking the average predicted probability estimated from the model for each HSfE respondent located within it). Although we examined the results for PCSs rather than wards, any problems with the modelled estimates at ward level should also be manifest in the estimates at PCS-level.

The plots of the residuals against the predicted area prevalence estimates for the five measures, and the accompanying linear regression estimates, are shown in Appendix C. These plots show that there was no evidence of a linear pattern in the residuals. We are confident, therefore, that we did not omit any important relationships between the healthy lifestyle measures and the ward characteristics.

4.2 Stability analysis

The criterion for the inclusion of terms in the model, namely that the significance level for that term was less than 0.05, means that some terms would have been included in the model by chance and that their apparent relationship with the healthy lifestyle measure was actually spurious. One technique to test for this is to split the dataset into two halves at random and then to re-estimate the model parameters for the covariates in the full model for both half-datasets. A non-spurious relationship between a particular covariate and the healthy lifestyle measure should be identified by an inconsistent parameter estimate in the two models.

Appendix D contains the parameter estimates for the models based on the complete sample and the two random sub-samples. Any significant differences (defined if the confidence intervals for the two parameters do not over-lap) between the parameter estimates for the two random sub-samples are identified in the last column.

The results of the stability analyses were also very encouraging. The terms in the models for the two random sub-samples were very consistent and the correlation between the resulting pairs of synthetic estimates was over 0.9 for each healthy lifestyle measure.

4.3 Comparison of the synthetic and direct estimates for HSfE 2000-02

Appendix E shows the correlations between the synthetic estiamtes at ward, PCO and SHA level against the direct estimates from the combined years of the HSfE 2000 to 2002. As the synthetic estimates were generated using these years of the HSfE, we would expect a fair level of correlation between them and the direct estimates. Table E.1 shows there were very large correlations at the SHA level, from 0.71 to 0.93. Even at ward level, for which the direct estimates become relatively imprecise because of the smaller sample sizes in each ward, the correlations were fairly large for all the healthy lifestyle measures except for obesity; including only wards which contained more than 25 HSfE respondents, the correlations ranged from 0.28 for obesity to at least 0.45 for the other healthy lifestyle measures.

We can perform an additional test of the specificity of the model, by looking at the regression line for the ward-level synthetic estimates against the direct estimates. If the model were well specified, then the regression line would be the '45 degree line'. To test this, we looked for a significant deviance from zero for the constant and from one for the slope of the line (see Appendix F).

An additional test of the specification of the model is to examine for evidence of a non-linear relationship between the synthetic and direct estimates. This is done by fitting a quadratic model for the synthetic estimates against the direct estimates and testing whether the parameter estimate for the quadratic term is significantly different to zero.

For current smoking, obesity and binge drinking the comparison of the direct estimates against the synthetic estimates revealed no significant deviation from the 45 degree line and hence no evidence of model mis-specification (see Appendix F). In the case of fruit and vegetable consumption for both children and adults there was some slight evidence of a non-linear relationship (the p-value for the quadratic term was 0.040 and 0.045 respectively).

5 EXTERNAL VALIDATION

In order to validate the synthetic estimates generated, we compared them against direct survey-based estimates from a range of external sources. For the technical evaluation stage of this project, we tested the synthetic estimates for smoking and obesity only (based on slightly different models) using external data from local surveys and the National Patient Experience Survey (see Pickering *et al*., 2004); these analyses are repeated in this chapter on the updated models for smoking and obesity.

For this stage of the project we also had access to the HSfE 2003 data, from which we could test the relationship between all five synthetic estimates and their corresponding direct HSfE 2003 estimates at three levels.

5.1 The Health Survey for England 2003

Because the synthetic estimates were developed using HSfE data from 2000 to 2002, the HSfE 2003 data could be considered to be external data, as they had not used to generate the synthetic estimates. In addition, the majority of wards covered by the HSfE in 2003 would not have been covered by the years 2000 to 2002 – hence, we would be validating the synthetic estimates in (mostly) different wards to those used to generate the estimates. This therefore gave us good tests of independent validity, as the results would not be confounded by difference in the survey design as was the case for the other sources of validation data.

Using this data we were able to test the relationship between the synthetic estimates and the direct HSfE estimates at three levels: wards, PCOs and SHAs.

5.1.1 Comparisons at SHA level

The SHA-level comparisons show positive correlations between the synthetic estimates and the direct HSfE 2003 estimates ranging from about 0.4 to 0.8 (Table 5.1). The correlations for the healthy lifestyle measures were not as strong for the HSfE 2003 data as for the direct estimates from the combined HSfE from 2000 to 2002 (see Appendix E), e.g. for smoking, 0.42 compared with 0.75. The higher correlation for the latter would have been observed because the direct estimate was from the same years of the HSfE that were used to derive the synthetic estimate and also the sample sizes for the combined samples were greater.

Table 5.1 Correlations between synthetic estimates and HSfE 2003 direct estimates: SHA level

	All SHAs
Smoking (adults)	0.42
Obesity (adults)	0.64
Consumption of fruit and vegetables (children)	0.44
Consumption of fruit and vegetables (adults)	0.75
Binge drinking (adults)	0.86
Base	*28*

5.1.2 Comparisons at PCO level

The comparisons at PCO level again showed positive correlations between the direct HSfE 2003 estimates and the synthetic Table 5.2). The correlation for smoking with the HSfE 2003 was lower that that observed for the National Patient Experience Survey 2003 (Pickering *et al.,* 2004), about 0.40 compared with 0.64. However, the sample sizes in each PCO were considerably higher for the Patients Experience Survey (about 850) compared with the HSfE 2003 (about 50 for all PCOs and 70 for PCOs with 50 or more respondents).

Table 5.2 Correlations between synthetic estimates and HSfE 2003 direct estimates: PCO level

	All PCOs	PCOs > 25[a] HSfE respondents	PCOs > 50[b] HSfE respondents
Smoking (adults)	0.37	0.39	0.41
Obesity (adults)	0.24	0.31	0.30
Consumption of fruit and vegetables (children)	0.13	0.15	0.26
Consumption of fruit and vegetables (adults)	0.41	0.42	0.50
Binge drinking (adults)	0.44	0.50	0.56
Bases:			
Smoking (adults)	*294*	*246*	*137*
Obesity (adults)	*294*	*236*	*114*
Consumption of fruit and vegetables (children)	*287*	*231*	*116*
Consumption of fruit and vegetables (adults)	*294*	*247*	*138*
Binge drinking (adults)	*294*	*246*	*136*

NOTE: For the estimate of F&V for children, the minimum number of respondents was (a) 5 and (b) 10.

5.1.3 Comparisons at ward level

Comparisons at the ward level suggest that the smoking estimates were most strongly correlated with the HSfE 2003 direct estimates compared with the other healthy lifestyle measures (Table 5.3). For all the measures, the correlations were fairly high considering that the average number of respondents in each ward was fairly low – e.g. for smoking about 9 per ward for all wards, and 24 for wards with more than 20 respondents.

Table 5.3 Correlations between synthetic estimates and HSfE 2003 direct estimates: ward level

	All wards	Wards > 10[c] HSfE respondents	Wards > 20 HSfE respondents
Smoking (adults)	0.28	0.46	0.55
Obesity (adults)	0.13	0.22	0.43
Consumption of fruit and vegetables (children)	0.14	0.47	-
Consumption of fruit and vegetables (adults)	0.19	0.37	0.41
Binge drinking (adults)	0.25	0.32	0.39
Bases:			
Smoking (adults)	*1,689*	*622*	*141*
Obesity (adults)	*1,653*	*530*	*90*
Consumption of fruit and vegetables (children)	*1,015*	*127*	*0*
Consumption of fruit and vegetables (adults)	*1,691*	*627*	*143*
Binge drinking (adults)	*1,689*	*612*	*135*

NOTES: (c) For the estimate of F&V for children, the minimum of number of respondents was 5

5.2 Local surveys

Given the relatively small numbers of respondents in each ward for the comparisons with the HSfE (Section 5.1.3), we have repeated the analyses from the technical evaluation stage that compared the estimates of smoking and obesity from those collected in two local surveys: the Camden and Islington (C&I) Health Survey 1999 and the Wigan, Bolton and Bury (WBB) Health Survey 2001. In addition, we have added comparisons for the same measures with estimates from the combined Liverpool and Sefton, and St Helens and Knowsley Lifestyle Surveys 2003[7].

Note that although the survey ward estimates would be more precise than those obtained from the HSfE 2003, the comparisons with the synthetic estimates would be likely to be confounded by the differences in the methodology of the local surveys (sample design, question wording, mode of data collection) and differences in survey context, timing and response rates.

5.2.1 The Camden & Islington Health Survey (C&I) 1999

In 1999, a local boost survey of residents of the (erstwhile) Camden & Islington Health Authority was conducted alongside the main HSfE. The adult sample size for C&I was just under 2,000 respondents (response rate 60%) and the questionnaire coverage and survey procedures were identical to the main HSfE 1999 survey.

[7] Many thanks to Thomas Hennell from the North West Public Health Team, Department of Health for his assistance with these analyses.

However, as the survey was designed to provide robust estimates for the four PCOs in C&I, the average sample size (45 adults) per ward[8] was relatively small, although larger than the average for the HSfE 2003.

5.2.2 The Wigan, Bolton & Bury Health Survey (WBB) 2001

The Wigan and Bolton Health Survey was carried out in 2001 and consisted of a self-completion postal survey of a 5% sample of adults aged 18 and older registered with the general practices in the Wigan and Bolton Health Authority. The response rate was 71%, with 15,465 completed questionnaires and an average of 324 adults respondents per ward. The design for the Bury Health Survey was similar, although details about the number of questionnaires issued and response rate were not available.

Because the number of respondents per wards was relatively large for the WBB survey, the estimates have fairly high precision. However, because of differences between the design of be survey and the HSfE we would expect the comparisons between the WBB survey estimates and the synthetic estimates to be confounded. These would be mostly caused by the difference in the mode of the survey (postal questionnaire rather than face-to-face interview) and by height and weight being self-reported, which is known to results in an underestimate of the BMI. Despite this likely underestimation, we still used a BMI of 30 as the cut-off for obesity. In addition, the use of the GP registers for sampling would add additional biases as they would exclude people that were not registered with a GP.

5.2.3 The Liverpool & Sefton and St Helens & Knowsley Lifestyle Surveys (LSSK) 2003

The Liverpool and Sefton Lifestyle Survey and St Helens and Knowsley Lifestyle Survey were both carried out by the Health and Community Care Research Unit at the University of Liverpool. Both consisted of a postal survey of patients on the GP registers of the selected GP practices that agreed to take part. For Liverpool and Sefton, 28,000 postal questionnaires were issued and about 40% returned. For the St Helens and Knowsley Lifestyle Survey, the sample size and response rate were not available.

Because the LSSK was a postal questionnaire, we would again expect there to be biases compared with estimates derived from the HSfE survey, in particular with the estimates for obesity.

5.2.4 Current smoking

Table 5.4 shows the summary statistics for the two estimates (direct and synthetic) for the three local areas (C&I, WBB and LSSK), and the correlation between the two sets of smoking estimates.

[8] After excluding wards with a sample size less then 10

Table 5.4 Summary statistics for smoking prevalence estimates (%) at ward level: direct vs local survey estimates

Summary statistics	Direct estimates	Synthetic estimates
Camden & Islington (N=34 wards):		
Mean	32.1	31.8
Minimum	20.6	20.0
Maximum	47.1	37.9
Range	26.5	17.9
Std deviation	6.7	3.6
Correlation with direct estimates	-	0.44
Wigan, Bolton & Bury (N=60 wards):		
Mean	28.5	25.6
Minimum	18.2	18.2
Maximum	41.5	37.4
Range	23.2	19.2
Std deviation	6.0	4.2
Correlation with direct estimates	-	0.76
Liverpool etc (N=95 wards):		
Mean	26.3	31.8
Minimum	8.8	15.4
Maximum	61.5	56.8
Range	52.7	41.4
Std deviation	9.6	8.7
Correlation with direct estimates	-	0.72

The mean estimates for smoking were similar for the synthetic and direct estimates for two of the local surveys - 31.8% compared with 32.1% for C&I, and 25.6% compared with 28.5% for WBB. However, the mean of the synthetic estimates were greater than the mean of the direct estimates for LSSK – 31.8% compared with 26.3%. The correlations between the synthetic estimates and the direct estimates were fairly high – 0.44 for C&I, 0.76 for WBB and 0.72 for LSSK.

5.2.5 Obesity for adults

Table 5.5 shows summary statistics for the direct and synthetic estimates for the two areas (C&I and WBB) and the correlation between the two sets of adult obesity estimates.

The average proportion of obesity for the synthetic estimates was higher than that for the WBB and LSSK direct estimates (23.2% compared with 14.7% and 21.3% compared with 15.7%). However, this was to be expected as the measures of height and weight for WBB and LSSK were self-reported, and this tends to result in obesity being under-reported. Even given the difference in mean reporting, the correlation between the WBB direct estimates and the synthetic estimates was fairly high (0.47), although it was lower for the LSSK (0.28).

Table 5.5 Summary statistics for obesity prevalence estimates (%) at ward level: direct vs local survey estimates

Summary statistics	Direct estimates	Synthetic estimates
Camden & Islington (N=33 wards):		
Mean	13.5	16.3
Minimum	0.0	11.9
Maximum	23.7	21.6
Range	23.7	9.7
Std deviation	5.8	2.1
Correlation with direct	-	0.30
Wigan, Bolton & Bury (N=60 wards):		
Mean	14.7	23.2
Minimum	7.8	18.2
Maximum	21.5	30.0
Range	13.7	11.8
Std deviation	3.0	2.5
Correlation with direct	-	0.47
Liverpool etc (N=95 wards):		
Mean	15.7	21.3
Minimum	5.7	15.2
Maximum	33.3	27.3
Range	27.6	12.2
Std deviation	5.4	2.5
Correlation with direct estimates	-	0.28

For the C&I survey, the mean estimates were closer, reflecting the similar methodology to collect obesity. The correlation with the synthetic estimates was lower for the C&I estimates (0.30) than for WBB.

5.3 The National Patient Experience Survey (2003)

The 2003 *National Patient Survey in Primary Care Organisations* asked adults aged 16 and over about their smoking behaviour (current smoker, ex-smoker, never-smoker), using a questionnaire similar to that used on the HSfE, but asked in a self-completion postal survey rather than face-to-face as in the HSfE. The survey was conducted in all PCOs in England, with approximately 850 patients per PCO sampled from electoral registers to allow fairly precise estimates to be derived at PCO-level. The overall response rate for the survey was 49%, and varied between 26% and 65% at PCO-level. Low response rates for some PCOs raises the issue of reliability of the direct estimates derived from this survey and its value in providing a reasonable yardstick against which to compare the performance of the model-based synthetic estimates.

Other than current smoking status, the patient survey did not collect information from respondents on any of the other health lifestyle measures being evaluated in this study.

The national prevalence of smoking based on the synthetic estimates was about 4 percentage points higher than the direct estimates based on the patient survey (Table 5.5). This finding was reported in the technical report of the patient survey and can be attributed to differences in survey mode and possibly higher-levels of non-response bias in the patient survey compared to the HSfE. Despite these caveats, the correlation between the direct and the synthetic estimates was high at 0.65 (Table 5.6).

Table 5.6 Summary statistics for smoking prevalence estimates (%) at PCO level

Summary statistics	Direct	Synthetic estimates
PCOs (N=300):		
Mean	22.0	25.6
Minimum	13.5	15.6
Maximum	36.2	37.5
Range	22.7	21.9
Std deviation	4.2	4.7
Correlation with direct		0.65

5.4 The Health Development Agency smoking estimates (2004)

The Health Development Agency (HDA, 2004) recently released a set of current smoking estimates at the PCO level that were also based on synthetic estimation techniques[9]. The HDA estimates were computed using a different range of years of HSfE data (1998-2001) and were based on models that contained both individual (i.e. age, sex and social class) and area-level covariates (e.g. percentage of people non-white and the percentage of households in rented tenure).

Table 5.7 Summary statistics for synthetic smoking prevalence estimates (%) at PCO level

Summary statistics	HDA estimates	NatCen estimates
PCOs (N=303)		
Mean	28.1	25.6
Minimum	19.9	15.9
Maximum	40.0	37.5
Range	20.1	21.6
Std deviation	4.17	4.73
Correlation with HDA		0.92

[9] In relation to which estimates to use we understand that the Department of Health recommends that the HDA estimates be used for smoking-attributable mortality and the NatCen estimates (produced by this project) used for current smoking prevalence.

Given that that the years of the HSfE used to generate the estimates over-lapped, we were not able to treat the HDA estimates as external estimates and hence for external validation. However, it is reassuring how similar the two sets of estimates are. Although based on different covariates (and a different range of years of the HSfE) the two sets of smoking estimates were reassuringly similar, with a correlation of 0.92 (Table 5.7).

To compare the relative ranking of the PCO estimates we assigned each PCO into quintiles (with about 60 PCOs in each quintile group) for each measure and compared the five-by-five quintile group differences between the HDA and NatCen estimates (Table 5.8). Confirming the high positive correlation between the two sets of synthetic estimates, almost 90% of the PCOs belonged to either the same or adjacent quintile.

Table 5.8 Percentage distribution of PCOs by quintile matrix difference, HDA estimates *versus* NatCen model-based smoking prevalence estimates

	Percentage
Percentage of PCOs in:	
same quintile group	63
+/- 1 difference	35
+/- 2 difference	2
+/- 3 difference	0
+/- 4 difference	0

5.5 Conclusions

The comparisons with the HSfE 2003 direct estimates (Section 5.1) showed that the correlations at SHA were high (at least 0.4) for all five measures. The analyses at PCO level were also re-assuring, although the correlation for fruit and vegetable consumption for children was lower than those of the other measures – most probably because of the smaller sample sizes used to generate the HSfE 2003 direct estimates, rather than the synthetic estimates being less accurate. The comparison of the National Patient Experience Survey (2003) confirmed the high correlation for smoking (0.65) at PCO level for direct estimates with a relatively large sample size (about 850) in each PCO (Section 5.3).

Considering the small sample sizes for the direct estimates, the correlations between the synthetic estimates and the HSfE 2003 estimates at ward level were quite high (Section 5.1.3). The local surveys allowed additional comparisons of the smoking and obesity estimates for wards with direct estimates (Section 5.2) - the correlations for smoking were high, especially for the larger WBB (0.76) and LSSK (0.72) surveys. The results for obesity were less convincing, with correlations ranging between 0.28 to 0.47 for the three local surveys.

Given the range of validation available and the size of these correlations, we have confidence in the validity of the synthetic estimates produced at ward and PCO level for smoking. The evidence for binge drinking and fruit and vegetable consumption was limited to comparison with the HSfE 2003, from which fruit and vegetable consumption for children did not perform as well for the other measures, although this might have been due to the reduced sample sizes for the direct measures. The results for obesity seemed mixed. The correlation with the direct estimates from the WBB was fairly large (0.47), but was smaller for the C&I and LSSK (0.30 and 0.28).

In summary, therefore, we have confidence in the synthetic estimates at PCO and ward level, although the results for obesity were less convincing than for the other measures.

REFERENCES

Bajekal M, Scholes S, Pickering K and Purdon S (2004) Synthetic estimation of healthy lifestyle indicators: Stage 1 report. (Unpublished)

Bird S, Cox D, Farewell V, Goldstein H, Holt T and Smith P (2003) *Performance Indicators: Good, Bad and Ugly*. RSS Working Paper on Performance Monitoring in the Public Service.

Heady P, Clarke P and others (2003) *Model-based small area estimation series No 2* (Small Area Estimation Project Report: Office for National Statistics).

Health Development Agency (2004) *The Smoking Epidemic in England* (Health Development Agency).

Longhurst J, Cruddas M, Goldring S and Mitchell B (2004) *Model-Based Estimates of Income for Wards, 1998/99 Technical Report.* ONS Publication.

Pickering K, Scholes S and Bajekal M (2004) Synthetic estimation of healthy lifestyle indicators: Stage 2 report. (Unpublished)

Rasbash J, Steele F, Browne W and Prosser B (2004). *A User's Guide to MLwiN, Version 2.0*. London: Institute of Education.

Scholes S, Bajekal M, and Pickering K (2005) Synthetic estimation of healthy lifestyles indicators: User guide. (Unpublished)

Sproston K and Primatesta P (eds.) (2003) *Health Survey for England 2002: volume 1: the health of children and young people,* London: The Stationary Office.

StataCorp (2003). *Stata Statistical Software: Release 8.0.* College Station, TX: Stata Corportation.

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

APPENDIX A AREA-LEVEL PREDICTOR VARIABLES AND SOURCES

A1 Ward-level covariates - Census 2001 data

Variable name	Description
Agesex*n*	All 0+, in age/sex group *n*
Hloamnty	Households, all, without central heating
Hlonepen	Households, all, pensioners living alone
Hnocar	Households, all, no car
Hovercr	Households, all, overcrowded (occupancy rating minus 1 or less)
Hownocc	Households, all, owner occupied
Hsingpar	Households, all, single parent with dependent children
Ibadhlth	All 0+, not good health
Icobnuk	All 0+, not born in UK, Ireland or EU
Icouple	Household residents, 16+, living as couple
Iethnic	All 0+, non-white (including mixed)
Illsiall	All 0+, with limiting longstanding illness
Illsiwrk	All, working age, with limiting longstanding illness
Inoqual	All, 16-74, no educational qualifications
Iolevel	All, 16-74, highest qualification NVQ1 or no qualifications
Ipermsic	All, 16-74, permanently sick/disabled
Iprofman	All, 16-74, professional and managers (NS-SEC 1 & 2)
Isroutin	All, 16-74, semi-routine and routine occupations (NS-SEC 6 & 7)
Iunemea	All, 16-74, unemployed as % of economically active
Iunemlng	All, 16-74, long term unemployed or never worked
Iupdcare	All 0+, unpaid carers
Iupdcr50	All 0+, fifty plus hours of care

A2 Ward-level covariates - Administrative data

Variable name	Source	Description
Aarate	DWP benefits data (Aug 2001)	Attendance allowance claimant rate
Dlarate	"	Disability allowance living claimant rate
Ibrate	"	Invalidity benefit claimant rate
Israte	"	Income support claimant rate
Sdarate	"	Severe disablement allowance claimant rate
Propctxa	Valuation Office Agency data (Mar 2001)	Proportion of dwellings in council tax band A
Propctxb	"	Proportion of dwellings in council tax band B
Propctxc	"	Proportion of dwellings in council tax band C
Propctxd	"	Proportion of dwellings in council tax band D
Propctxe	"	Proportion of dwellings in council tax band E
Propctxf	"	Proportion of dwellings in council tax band F
Propctxg	"	Proportion of dwellings in council tax band G
Propctxh	"	Proportion of dwellings in council tax band H
Propctxx	"	Proportion of dwellings not yet verified/allocated

A3 Other ward-level covariates

Variable name	Source	Description
Criscore	Derived from Output Area scores produced by the Office of the Deputy Prime Minister (2004)	Crime and disorder score
Eduscore	"	Education skills and training score
Empscore	"	Employment score
Envscore	"	Living environment score
Heascore	"	Health deprivation and disability score
Houscore	"	Barriers to housing and services score
Incscore	"	Income score
Wdidscor	"	IMD 2004 ward level
Rural	Countryside Agency	Rural indicator
Onssuper	National Statistics 2001 Area classification of statistical wards	1 Industrial Hinterlands 2 Traditional Manufacturing 3 Built-up Areas 4 Prospering Metropolitan 5 Student Communities 6 Multicultural Metropolitan 7 Suburbs and Small Towns 8 Coastal and Countryside 9 Accessible Countryside

A4 Higher area-level covariates

Variable name	Level of geography	Source	Description
Laidscor	LAD	Office of the Deputy Prime Minister	IMD score 2004
Lefemale	"	ONS	Life expectancy at birth, number of years, 1999-2000
Lemale	"	"	Life expectancy at birth, number of years, 1999-2000
Region	GOR	All Fields Postcode Directory (AFPD)	North East North West Yorkshire & The Humber East Midlands West Midlands East of England London South East South West
Smr_03c	LAD	Compendium of Clinical and Outcome Indicators (2003)	Mortality from all causes (ICD 10 a00-y99) indirectly standardised ratios
Smr_03d	"	"	Mortality from potentially avoidable causes indirectly standardised ratios
Smr_06a	"	"	Mortality from all circulatory diseases (icd10 i00-i99) indirectly standardised ratios
Smr_08a	"	"	Mortality from hypertensive disease (icd10 i10-i15) indirectly standardised ratios
Smr_09a	"	"	Mortality from coronary heart disease (icd10 i20-i25)
Smr_10a	"	"	Mortality from stroke (icd10 i60-i69) indirectly standardised ratios, 2001
Smr_11a	"	"	Mortality from all cancers (icd10 c00-c97) indirectly standardised ratios
Smr_14b	"	"	Mortality from lung cancer (icd10 c33-c34) indirectly standardised ratios
Smr_25a	"	"	Mortality from chronic liver disease including cirrhosis (icd10 k70, k73-k74) indirectly standardised ratios
Smr65	"	"	SMR deaths under 65 years all persons 1998-2002 data

APPENDIX B MODEL SELECTION PROCEDURE

The approach for generating the optimal multilevel model was as follows:

1. From the full set of covariates set out in Appendix A, a smaller subset of variables (**subset 1**) was obtained by running a forward stepwise selection procedure in STATA. This initial stage of modelling did not take into account the multilevel structure of the data. Previous investigations by the ONS, however, have shown that those covariates considered to be significant in a single level model are also significant in a multilevel model (Longhurst *et al.*, 2004).[10]

2. With this first subset of variables we then used an iterative selection procedure that mirrored the method used by the ONS (see Heady *et al.*, 2003). The process began with the null model fitted allowing for the area-level variance[11]. The covariate with the largest value of T^2 (where T represents the estimate divided by its standard error) was then added to the model. Any covariate with a lower T^2 than the most recently added covariate was then dropped from the model (but was eligible for inclusion again at the next iteration). This process stopped when the T^2 for the next covariate to be added to the model was less than 4. Those covariates left in the final model at this stage can be described as **subset 2**.

3. Using the forward stepwise procedure we then checked whether any of the variables not included in subset 1 could be added to the variables belonging to subset 2.[12]

4. With this set of significant main effects a number of possible pairwise interaction terms were created and tested for inclusion using forward stepwise selection.[13] Those covariates (main effects and interactions) left in the model at this stage can be described as **subset 3**.

5. The next step involved examining the extent to which the ward or higher level variables had different effects in different regions. Each region was entered into the model as a main effect. Pairwise interaction terms between the ward level variables and GOR were created and tested for inclusion using forward stepwise selection.[14] Those covariates remaining in the model at this stage can be described as **subset 4**.

---

[10] A number of covariates could not be used in the initial stepwise model because of collinearity (e.g Government Office Region A and one category each from the ward classifications (IMD and ONSSUPER). These were tested separately after the first stepwise model was obtained and added to the list of variables if statistically significant.

[11] Note that because the HSfE is clustered within postcode sector, this is the area-level term in the models rather than wards.

[12] At this stage, covariates referring to the age, sex and marital status composition of the ward were not considered for re-entry to the model. As there were 56 variables in total 3 covariates (56/0.05) would be expected to be statistically significant simply by chance alone.

[13] The covariates referring to the age, sex and marital status composition of the ward were not included in the set of interaction terms that were tested. Again, the main reason for this is that a number of the interaction terms would be expected to be significant simply by chance alone.

[14] For the reasons set out above the covariates referring to the age, sex and marital status composition of the ward were not included in the set of covariate and GOR interactions.

6. With this subset of variables we then returned to the iterative selection procedure used by the ONS that was described above. As the covariates were now reduced to a manageable number this procedure was ran in STATA whilst taking the multilevel structure of the data into account. The first step involved selecting those main effects whose T^2 in the model was greater than 4. Building on this model the interaction terms (including those with GOR) with T^2 greater than 4 were retained in the model.

7. The final model was re-fitted in MLwiN to obtain the parameters estimates to derived the synthetic estimates.

APPENDIX C RESIDUAL PLOTS

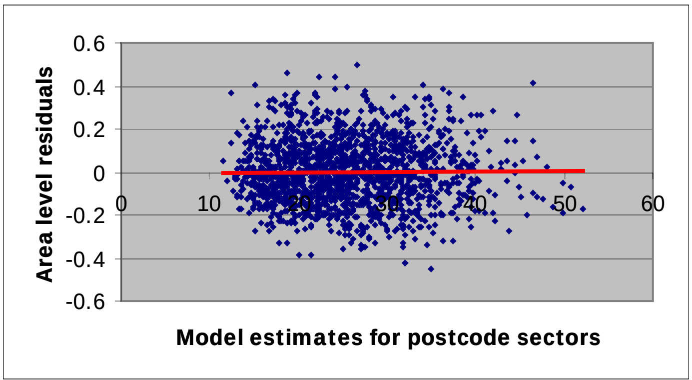
1. Current smoking

For the optimal model for smoking, the equation of the regression of the residuals (y) against the model estimates (x) (with the standard errors are displayed in parentheses) was:

$$y = -0.0002(0.013) + 0.0009(0.050)x$$

The confidence intervals for both the intercept and slope contained the value zero implying an absence of a linear pattern in the area-level residuals (Figure C1).

Figure C.1 Area-level residuals against model estimates: smoking



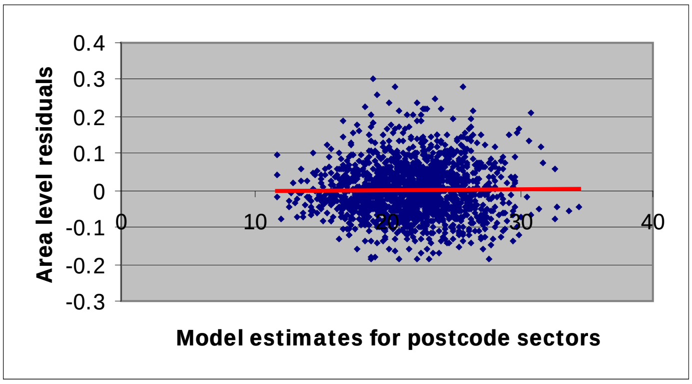
2. Obesity

For the optimal model for obesity, the equation of the regression of the residuals (y) against the model estimates (x) was:

$$y = -0.006(0.011) + 0.0289(0.049)x$$

The confidence intervals for both the intercept and slope contain the value zero implying an absence of a linear pattern in the area-level residuals (Figure C.2).

Figure C.2 Area-level residuals against model estimates: obesity



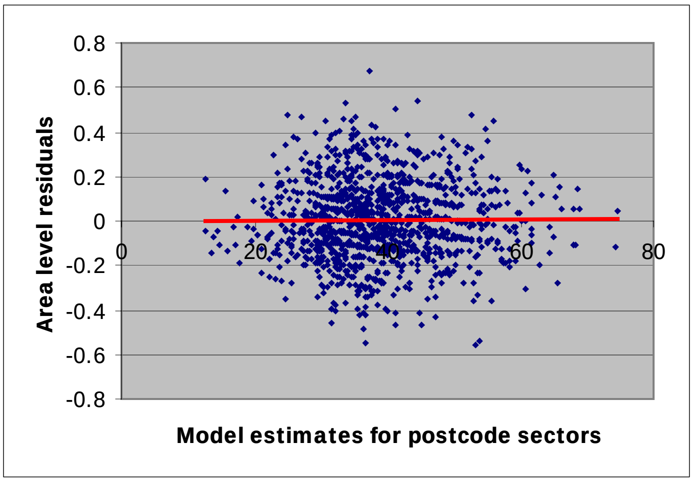
3. Fruit and vegetable consumption (children)

For the optimal model for whether children consumed three or more portions of fruit and vegetables, the equation of the regression of the residuals (y) against the model estimates (x) was:

$$y = -0.005(0.021) + 0.013(0.052)x$$

The confidence intervals for both the intercept and slope contained the value zero implying an absence of a linear pattern in the area-level residuals (Figure C.3).

Figure C.3 Area-level residuals against model estimates: children fruit & vegetable consumption



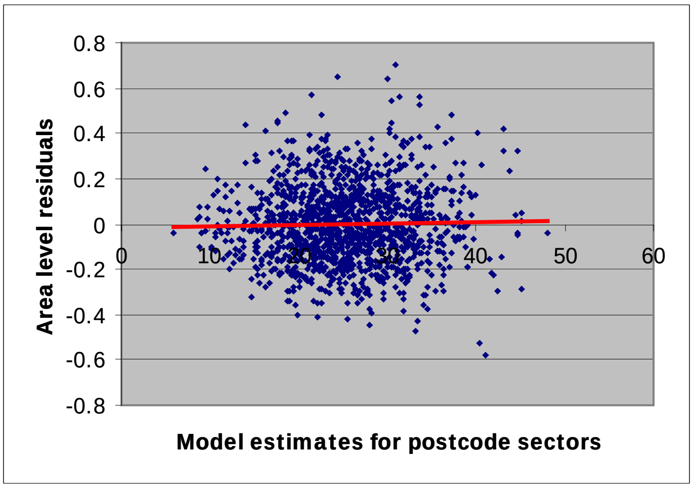
4. Fruit and vegetable consumption (adults)

For the optimal model for whether adults consumed more than five portions of fruit and vegetables, the equation of the regression of the residuals (y) against the model estimates (x) was:

$$y = -0.012(0.018) + 0.049(0.070)x$$

The confidence intervals for both the intercept and slope contained the value zero implying an absence of a linear pattern in the area-level residuals (Figure C.4).

Figure C.4 Area-level residuals against model estimates: adult fruit & vegetable consumption



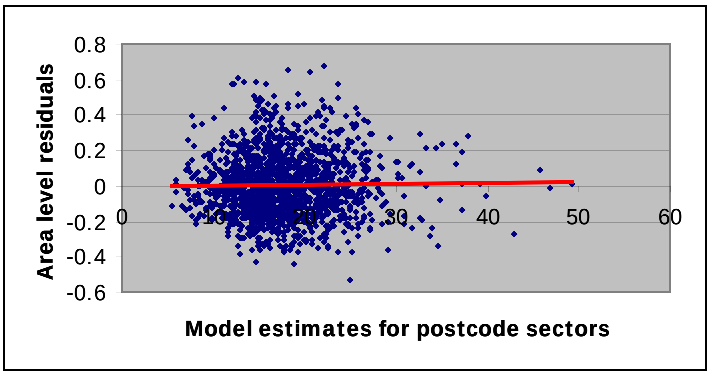
5. Binge drinking

For the optimal model for binge drinking, the equation of the regression of the residuals (y) against the model estimates (x) was:

$$y = -0.010(0.015) + 0.059(0.082)x$$

The confidence intervals for both the intercept and slope contained the value zero implying an absence of a linear pattern in the area-level residuals (Figure C.5).

Figure C.5 Area-level residuals against model estimates: binge drinking



APPENDIX D STABILITY ANALYSES

Table D.1 Stability analysis for current smoking

Parameters in the model	Full model		Dataset A		Dataset B		Does the estimate of B fall within the CI of A?
	Estimate	*Standard error*	*Estimate*	*Standard error*	*Estimate*	*Standard error*	
icouple	**-2.160**	*0.267*	**-2.048**	*0.361*	**-2.249**	*0.361*	Yes
iprofman	**-1.325**	*0.276*	**-0.970**	*0.371*	**-1.677**	*0.377*	Yes
iethnic	**-0.914**	*0.156*	**-0.929**	*0.214*	**-0.916**	*0.207*	Yes
Proportion female, aged 25-34	**5.107**	*0.747*	**5.158**	*1.012*	**5.138**	*1.019*	Yes
imd8	**0.139**	*0.048*	**0.186**	*0.065*	0.087	*0.066*	Yes
aarate	**-1.640**	*0.540*	-1.023	*0.731*	**-2.184**	*0.735*	Yes
illsiwrk	**3.862**	*1.166*	**3.342**	*1.565*	**4.304**	*1.569*	Yes
eduscore	**0.006**	*0.002*	**0.006**	*0.003*	0.005	*0.003*	Yes
imd5	**0.119**	*0.056*	**0.162**	*0.075*	0.076	*0.078*	Yes
North West GOR (gor_nw)	**0.533**	*0.190*	0.464	*0.255*	**0.560**	*0.257*	Yes
South West GOR (gor_sw)	**-0.137**	*0.067*	-0.040	*0.088*	**-0.230**	*0.091*	No
aarate/ gor_nw	**-3.586**	*1.108*	**-3.424**	*1.485*	**-3.452**	*1.509*	Yes
iethnic/ gor_sw	**3.572**	*1.266*	**3.196**	*1.616*	**4.044**	*1.693*	Yes
imd8/ gor_sw	**0.377**	*0.144*	0.328	*0.193*	**0.412**	*0.199*	Yes
Intercept	0.083	*0.313*	-0.175	*0.420*	0.302	*0.423*	

NOTE: Bold indicates statistical significance ($p < 0.05$)

Table D.2 Stability analysis for obesity

Parameters in the model	Full model		Dataset A		Dataset B		Does the estimate of B fall within the CI of A?
	Estimate	*Standard error*	*Estimate*	*Standard error*	*Estimate*	*Standard error*	
iolevel	**0.914**	*0.333*	0.900	*0.499*	**0.955**	*0.469*	Yes
rural	-0.112	*0.110*	-0.053	*0.165*	-0.170	*0.155*	Yes
South West GOR (gor_sw)	**-0.196**	*0.052*	**-0.228**	*0.079*	**-0.161**	*0.073*	Yes
propctxg	**-0.760**	*0.337*	-0.774	*0.493*	-0.788	*0.480*	Yes
laidscor	**-0.006**	*0.002*	**-0.006**	*0.003*	**-0.006**	*0.003*	Yes
East of England gor (gor_ee)	**-0.394**	*0.134*	**-0.461**	*0.203*	-0.351	*0.190*	Yes
isroutin	0.715	*0.372*	0.891	*0.556*	0.511	*0.527*	Yes
hloamnty	**-0.532**	*0.242*	**-0.853**	*0.365*	-0.237	*0.340*	Yes
rural/ isroutin	**0.790**	*0.359*	0.470	*0.540*	**1.071**	*0.504*	Yes
laidscor/ gor_ee	**0.018**	*0.008*	0.023	*0.012*	0.013	*0.011*	Yes
Intercept	-1.693	*0.095*	-1.719	*0.144*	-1.679	*0.134*	

NOTE: Bold indicates statistical significance ($p < 0.05$)

Table D.3 Stability analysis for fruit and vegetable consumption (children)

Parameters in the model	Full model		Dataset A		Dataset B		Does the estimate of B fall within the CI of A?
	Estimate	*Standard error*	*Estimate*	*Standard error*	*Estimate*	*Standard error*	
iolevel	**-2.445**	0.409	**-2.358**	0.535	**-2.477**	0.548	Yes
icobnuk	**2.944**	0.637	**2.377**	0.804	**3.318**	0.863	Yes
lemale	**0.102**	0.022	**0.078**	0.029	**0.115**	0.029	Yes
ipermsic	**15.994**	3.621	**15.828**	4.818	**14.924**	4.819	Yes
Yorks & The Humber GOR (gor_yh)	**0.755**	0.236	**1.001**	0.303	0.447	0.318	Yes
West Midlands GOR (gor_wm)	**0.256**	0.092	0.128	0.120	**0.361**	0.119	Yes
London GOR (gor_lon)	0.322	0.173	0.160	0.226	0.421	0.221	Yes
iupdcr50	**-16.509**	8.416	-20.115	11.244	-13.290	11.334	Yes
empscore	**-3.096**	1.562	-3.272	2.064	-2.603	2.092	Yes
ipermsic/gor_yh	**-13.252**	3.670	**-15.447**	4.752	**-10.398**	4.929	Yes
icobnuk/gor_lon	**-2.120**	0.912	-1.109	1.170	**-2.767**	1.198	Yes
South West GOR (gor_sw)	0.374	0.215	0.365	0.285	0.367	0.284	Yes
empscore/gor_sw	**-5.466**	2.316	-4.343	3.101	**-6.482**	3.087	Yes
Intercept	-7.491	1.716	-5.600	2.233	-8.559	2.280	

NOTE: Bold indicates statistical significance ($p < 0.05$)

Table D.4 Stability analysis for fruit and vegetable consumption (adults)

Parameters in the model	Full model		Dataset A		Dataset B		Does the estimate of B fall within the CI of A?
	Estimate	*Standard error*	*Estimate*	*Standard error*	*Estimate*	*Standard error*	
isroutin	**-1.152**	*0.336*	**-1.232**	*0.446*	**-1.106**	0.455	Yes
icobnuk	**2.211**	*0.303*	**2.507**	*0.395*	**1.883**	0.412	Yes
Proportion female, aged 25-34	**-3.905**	*0.916*	**-5.304**	*1.223*	**-2.470**	1.243	No
Proportion female, aged 16-19	**-14.152**	*3.472*	**-14.391**	*4.591*	**-13.737**	4.759	Yes
iupdcr50	**-10.683**	*4.864*	**-15.099**	*6.530*	-6.304	6.569	Yes
smr_14b	**-0.003**	*0.001*	**-0.004**	*0.001*	**-0.003**	0.001	Yes
smr_10b	**0.007**	*0.002*	**0.008**	*0.002*	**0.005**	0.002	Yes
South East GOR (gor_se)	**1.612**	*0.394*	**1.557**	*0.511*	**1.719**	0.525	Yes
imd9	**0.135**	*0.056*	**0.232**	*0.075*	0.029	0.076	No
Yorks & Humber GOR (gor_yh)	**0.667**	*0.221*	**1.040**	*0.291*	0.323	0.295	No
Built-up areas	**0.282**	*0.127*	0.210	*0.173*	0.327	0.174	Yes
smr_10a/gor_se	**-0.015**	*0.004*	**-0.015**	*0.005*	**-0.016**	0.006	Yes
isroutin/gor_yh	**-2.116**	*0.647*	**-3.263**	*0.876*	-1.133	0.857	No
Intercept	-0.389	0.198	-0.273	*0.258*	-0.513	0.268	

NOTE: Bold indicates statistical significance ($p < 0.05$)

Table D.5 Stability analysis for binge drinking

Parameters in the model	Full model		Dataset A		Dataset B		Does the estimate of B fall within the CI of A?
	Estimate	*Standard error*	*Estimate*	*Standard error*	*Estimate*	*Standard error*	
iethnic	**-1.881**	*0.250*	**-1.796**	*0.327*	**-1.991**	*0.341*	Yes
Proportion male, aged 75-79	**-16.726**	*4.500*	**-13.697**	*6.036*	**-19.462**	*6.110*	Yes
icouple	**-3.587**	*0.469*	**-3.131**	*0.617*	**-4.025**	*0.627*	Yes
North East GOR (gor_ne)	**0.435**	*0.077*	**0.385**	*0.100*	**0.478**	*0.099*	Yes
North West GOR (gor_nw)	**0.284**	*0.059*	**0.304**	*0.076*	**0.256**	*0.077*	Yes
Yorks & The Humber (gor_yh)	**0.268**	*0.061*	**0.229**	*0.079*	**0.299**	*0.080*	Yes
israte	**-3.026**	*0.585*	**-2.816**	*0.770*	**-3.337**	*0.779*	Yes
Proportion female, aged 50-54	**-8.761**	*3.115*	**-12.641**	*4.168*	-5.256	*4.213*	Yes
hovercr	**-1.957**	*0.744*	-1.608	*0.974*	**-2.235**	*1.002*	Yes
Proportion female, aged 85+	**-11.317**	*4.164*	**-18.607**	*5.682*	-5.182	*5.572*	No
Proportion male, aged 45-49	**8.515**	*4.045*	9.044	*5.451*	7.070	*5.476*	Yes
aarate	0.529	*0.629*	0.195	*0.839*	0.928	*0.842*	Yes
South East GOR (gor_se)	**-0.247**	*0.097*	**-0.262**	*0.130*	-0.229	*0.129*	Yes
hovercr/ gor_se	**2.602**	*1.195*	2.073	*1.578*	3.022	*1.611*	Yes
South West GOR (gor_sw)	**-0.831**	*0.260*	**-0.683**	*0.338*	**-1.051**	*0.363*	Yes
aarate/ gor_sw	**5.286**	*1.776*	**4.541**	*2.318*	**6.513**	*2.461*	Yes
Intercept	1.429	*0.332*	1.353	*0.435*	1.565	*0.440*	

NOTE: Bold indicates statistical significance ($p < 0.05$)

APPENDIX E COMPARISON OF DIRECT HSFE (2000-02) ESTIMATES AND SYNTHETIC ESTIMATES

Table E.1 Comparison of HSfE 2000-2002 direct estiamtes and synthetic estimates (SHA level)

	Current Smoking		Obesity		F&V (children)		F& V (adults)		Binge Drinking	
SHA Average:	%	*(s.d.)*	%	*(s.d.)*	%	*(s.d.)*	%	*(s.d.)*	%	*(s.d.)*
Synthetic Est	25.8	*(2.6)*	21.7	*(2.1)*	37.7	*(6.2)*	23.4	*(3.9)*	17.4	*(3.6)*
HSfE (2003)	25.9	*(2.5)*	22.0	*(2.6)*	38.1	*(6.7)*	23.8	*(4.6)*	18.1	*(4.0)*
Correlation:	r	*SHAs*	r	*SHAs*	r	*SHAs*	r	*SHAs*	r	*SHAs*
All	0.75	*28*	0.71	*28*	0.83	*28*	0.84	*28*	0.93	*28*

Table E.2 Comparison of HSfE 2000-2002 direct estiamtes and synthetic estimates (PCO level)

	Current Smoking		Obesity		F&V (children)		F& V (adults)		Binge Drinking	
PCO Average:	%	*(s.d.)*	%	*(s.d.)*	%	*(s.d.)*	%	*(s.d.)*	%	*(s.d.)*
Synthetic Est	25.6	*(4.7)*	21.9	*(2.6)*	36.7	*(7.1)*	23.2	*(4.9)*	17.4	*(4.0)*
HSfE (2003)	26.0	*(7.8)*	21.7	*(6.0)*	38.1	*(16.4)*	23.9	*(9.5)*	18.0	*(7.0)*
Correlations:	r	*PCOs*	r	*PCOs*	r	*PCOs*	r	*PCOs*	r	*PCOs*
All	.49	*302*	.39	*302*	.36	*300*	.56	*302*	.62	*302*
>25 in PCO	.57	*260*	.51	*250*	.59	*150*	.69	*223*	.66	*259*
>50 in PCO	.67	*137*	.56	*102*	.72	*28*	.67	*71*	.74	*134*

Table E.3 Comparison of HSfE 2000-2002 direct estimates and synthetic estimates (ward level)

	Current Smoking		Obesity		F&V (children)		F& V (adults)		Binge Drinking	
PCO Average:	%	*(s.d.)*	%	*(s.d.)*	%	*(s.d.)*	%	*(s.d.)*	%	*(s.d.)*
Synthetic Est	24.9	*(97.7)*	21.8	*(3.8)*	37.6	*(9.0)*	23.9	*(6.5)*	17.1	*(5.1)*
HSfE (2003)	25.0	*(23.4)*	21.8	*(21.4)*	38.7	*(34.2)*	25.1	*(25.0)*	16.8	*(19.4)*
Correlations:	r	*wards*	r	*wards*	r	*wards*	r	*wards*	r	*wards*
All	0.31	*3,231*	0.19	*3,149*	0.26	*1,989*	0.25	*2,645*	0.24	*3,230*
>25 in PCO	0.54	*1,274*	0.28	*1,070*	0.45	*176*	0.48	*909*	0.45	*1,253*
>50 in PCO	0.56	*359*	0.41	*256*	-		0.49	*223*	0.56	*345*

APPENDIX F REGRESSION OF DIRECT AGAINST SYNTHETIC ESTIMATES

Table F.1 Parameter estimates from linear and quadratic regression of the synthetic estimates against the direct estimates (ward level)

	Constant	Linear parameter	Quadratic parameter
Smoking:			
Linear model	0.014 (0.013)	0.945 (0.051)	-
Quadratic model	0.060 (0.038)	0.579 (0.295)	0.675 (0.536)
Obesity:			
Linear model	-0.017 (0.022)	1.069 (0.100)	-
Quadratic model	-0.062 (0.091)	1.484 (0.815)	-0.927 (1.808)
Fruit & vegetables (children):			
Linear model	0.010 (0.032)	1.003 (0.083)	-
Quadratic model	-0.114 (0.098)	1.665 (0.503)	-0.834 (0.625)
Fruit & vegetables (adults):			
Linear model	0.020 (0.018)	0.966 (0.073)	-
Quadratic model	-0.072 (0.049)	1.769 (0.403)	-1.632 (0.806)
Binge drinking:			
Linear model	0.015 (0.012)	0.895 (0.065)	-
Quadratic model	0.009 (0.025)	0.956 (0.259)	-0.154 (0.636)