Smoking related disease risk, area deprivation and health behaviours

Barbara Eberth[1,*], Damilola Olajide[1], Peter Craig[2], and Anne Ludbrook[1]

[1] Health Economics Research Unit, Institute of Applied Health Sciences, University of Aberdeen, Foresterhill, Aberdeen AB25 2ZD, UK

[2] Chief Scientist Office, Scottish Government Health Directorates, St Andrew's House, Edinburgh EH1 3DG

[*]Correspondence to Barbara Eberth, Health Economics Research Unit, Institute of Applied Health Sciences, University of Aberdeen, Foresterhill, Aberdeen AB24 2ZD, UK, tel: +44 (0)1224 437189, fax: +44 (0)1224 437195, e-mail: b.eberth@abdn.ac.uk.

Abstract

Background: Smokers and ex-smokers are at risk of many chronic diseases. However, never smokers and never smokers exposed to environmental tobacco smoke are also at risk. Additionally, smoking behaviours and their associated disease risk are socially patterned and positively associated with health inequalities. However, other lifestyle choices also contribute to health inequalities. We aim to assess the contribution of other lifestyle behaviours pertaining to alcohol, physical inactivity and weight to smoking related disease risk across *(i)* the socioeconomic spectrum and *(ii)* smoking status. **Methods:** Smoking-related disease risk is modelled using probit analysis. The results are used to predict disease risk across the socio-economic dimension and smoking status for a set of healthy and unhealthy behaviours using the administratively linked Scottish Health Surveys and Scottish Morbidity Records. **Results:** The results confirm the deprivation gradient in disease risk regardless of smoking status group. Imposition of healthy (unhealthy) lifestyle behaviours decreases (increases) predicted risk across the deprivation distribution regardless of smoking status providing evidence of the multifaceted health behavioural determinants of disease risk across the deprivation distribution. **Conclusion:** The results are of policy interest as they suggest that to reduce inequalities in smoking related diseases, interventions reducing both smoking and other unhealthy behaviours are required.

Introduction

Smoking related diseases are an important source of preventable ill health and mortality[1] and significantly contribute to health inequalities. For example, in Scotland, the percentage of smoking attributable deaths over the period 2000- 2004 was 15% for the least deprived and 33% for the most deprived population quintile.[2] Therefore, reducing smoking prevalence amongst the deprived population could have a disproportionately large effect on population health and an important impact on health inequalities. However, risk is elevated not only by current or past tobacco smoke exposure (both active and passive) but also by other risk factors, such as obesity, physical inactivity and deprivation.

Untangling the impact of area based deprivation or individual factors, such as education or occupation, and their impact on health, is complex. They directly impact on health outcomes associated with smoking behaviour and indirectly in determining, promoting or perpetuating smoking behaviour; what the Marmot Review refers to as 'the causes of the causes'[3]. It is possible that, due to the social patterning of smoking behaviour[4] and the multifactorial influences on health outcomes, as smoking rates decline other risk factors will take over in maintaining health inequalities, albeit at a lower level of absolute risk[5]. This possibility could be avoided by explicitly targeting smoking interventions based on other health risk factors, as well as deprivation.

Recent evidence indicated that smoking behaviour makes a greater contribution to health inequalities than social position *per se*.[6] However, in the same 28 year follow up of a cohort aged 45-64 at recruitment, higher mortality rates from CHD (but not cancer) were reported for women from lower social positions who never smoked, which was associated with higher obesity prevalence.[7] Female never smokers of normal weight had low mortality rates regardless of social position.

These studies, focussing on mortality, require long-term follow up of cohorts, where both health behaviours and life circumstances may change over time.[8] There is also no accounting for exposure to environmental tobacco smoke (ETS), which would have received less attention as a risk factor at the time of data collection. The disease specific risks of ETS are well documented[9] but less is known about the general effects of exposure[10].

To address some of the gaps in the current literature, this paper uses administratively linked health records and health survey data from a general population sample to investigate the separate contributions of health behaviours, previous smoking related disease history and

deprivation to smoking related morbidity, measured by risk of hospital admission for smoking attributable diseases using probabilistic regression. The use of hospital admission rather than mortality allows for results to be obtained over a shorter follow-up period from the time at which the health behaviours were reported and allows for a broader consideration of the ill health burden. The model is used to predict the contribution of deprivation, other health behaviours and previous smoking related disease history to smoking related disease risk for the average individual in the population across smoking status groups, including never smokers reporting exposure to ETS.

Methods

The linked Scottish Health Survey (SHeS)

We use hospitalisation episodes data (Scottish Morbidity Records; SMR) that have been administratively linked to SHeS respondents (waves 1995, 1998, and 2003). The SHeS is a national representative survey of individuals living in private households in Scotland that collects respondent information on self-reported health, self-reported lifestyle, demographic and socio-economic factors. For each consenting respondent, their information is linked to hospitalisation records in Scottish NHS hospitals, cancer and death registrations, covering the period 1981 to 31 December 2008. Permission to access the linked datasets was obtained from the Privacy Advisory Committee of the Information Services Division (ISD). For detailed information on the surveys and their linkage see Gray et al.[11] and Lawder et al.[12]

Smoking related diseases

We define smoking related diseases by ICD9 and ICD10 codes (see supplement Table S1). A disease event was defined as the presence of disease-specific ICD9 and 10 codes in any of the six diagnosis fields of an SMR record or cause of death. We use the survey interview date, hospital admission date, and discharge date to determine whether a smoking-related disease event occurred pre-survey or post-survey. A pre-survey event occurs if discharge fell on or before the interview date, a post-survey event occurs if admission happened post interview date. Our primary variable of interest is the first smoking related disease event post survey to model disease risk. Smoking related disease event pre-survey controls for smoking related disease history.

Smoking status

Smoking status is categorised into never smokers, never smokers exposed to ETS, current and ex-smokers. During the interview respondents aged 16 and older were asked:

- Do you *'smoke cigarettes regularly nowadays'*. The binary variable 'Smoker' equals one if 'Yes'; zero otherwise.
- Have you *'ever smoked before but do not currently smoke'*. The binary variable 'Ex-smoker' equals one if 'Yes'; zero otherwise.
- *'Are you regularly exposed to other peoples'* tobacco/cigarette smoke at any of the six places listed: at home, at work, in other people's homes, on public transport, in pubs, or other public places. The binary variable 'Exposed' equals one if any of the listed places is selected by never smokers; zero if *'None, none of these'*. We validated this variable using respondents' cotinine values collected during a nurse visit to ensure that individuals are indeed never smokers. Those who never smoked not reporting exposure to ETS comprise the remainder of the sample.

Deprivation

Deprivation is measured by the Scottish Index of Multiple Deprivation (SIMD)[13] and categorised into ordered quintiles where SIMD51=1 indicates the least deprived and SIMD55=5 the most deprived quintile.

Lifestyle behaviours

The lifestyle indicators considered include self-reported alcohol consumption, self-reported physical activity and Body Mass Index (BMI). The BMI measure does not rely on self-reported height and weight; these measures were taken during a follow-up nurse interview.

Other covariates

Other covariates include respondent demographic, household, socioeconomic, and health characteristics. The demographic covariates include age, gender and marital status. Socio-economic status is controlled for through respondents' education level, employment status and occupational class. Self-reported health variables include general health, pre-survey hospitalisations or pre-conditions, parental medical history, and the presence of a smoking related disease event pre survey. Descriptive statistics and definitions of all variables are presented in Table 1.

Table 1 Descriptive statistics[a]

Variable	*N=20,315*	
Smoking related disease incidence post survey	23.76	
Never smoker	16.21	
Never smoker exposed to ETS	23.32	
Current smoker	36.97	
Ex-smoker	23.50	
Gender (Male = 1)	45.26	
Age in years, mean (SD)	45.91	(16.09)
Age 16 to 24	9.71	
Age 25 to 34	17.92	
Age 35 to 44	21.11	
Age 45 to 54	18.86	
Age 55 to 64	18.74	
Age 65 and above [b]	13.66	
Single	22.40	
Married/cohabiting	57.96	
Divorced/widowed/separated [b]	19.63	
Number of children (0-15) in household	0.55	(0.935)
Number of adults in household	1.95	(0.823)
No education [b]	35.85	
Low ed. - School leaving cert, 'O' grade, GCSE, etc.	21.76	
Lower mid level - SQV, 'A' level, ONC, OND/C&G, etc.	12.53	
Upper mid level – HNC, HND, etc.	13.56	
University degree or equivalent	14.69	
Employed	56.12	
Unskilled [b]	6.38	
Partly skilled	14.26	
Skilled	43.80	
Intermediate/professional occupation	32.19	
Occasionally/never drinks alcohol [b, c]	29.90	
Regular drinks under the limit [c]	45.76	
Regularly drinks over the limit [c]	23.69	
Underweight (BMI < 18.5)	1.65	
Normal weight (BMI: 20 – 24.99) [b]	38.56	
Overweight (BMI: 25 – 29.99)	37.52	
Obese (BMI ≥ 30)	22.27	
Inactive (Sports) [b, d]	49.85	
Low intensity sports [b, d]	8.89	
Moderate intensity sports	15.01	
Vigorous intensity sports	26.25	
Very good self-assessed general health[b]	33.75	
Good self-assessed general health	39.36	
Fair self-assessed general health	19.75	

Bad/very bad self-assessed general health	7.14
SIMD51 (least deprived) [b]	17.04
SIMD52	19.36
SIMD53	21.00
SIMD54	21.30
SIMD55 (most deprived)	21.30
Either parent died of CHD	25.70
Presence of a non-limiting longstanding illness	11.24
Pre-survey hospitalisation for IHD	6.74
Takes medicine for high blood pressure (HBP)	11.37
Whether HBP diagnosed by doctor or nurse	09.19
Presence of chronic kidney disease CKD	0.75
Pre-survey smoking related disease incidence	11.12

[a] Values are percentages for categorical variables and means (standard deviations) for continuous variables.
[b] Denotes the reference category.
[c] Regular consumption within the limit equals 1 to 14 units for women, 21 units for men. Overconsumption equals more than 14 units for women, more than 21 for men. Occasional/never consumption equals 0 units or where respondents report not to drink regularly during a week.
[d] Individuals undertaking sports at light intensity and individuals who do not participate in sports are combined as the reference category.

Statistical analysis

We model smoking related disease risk using a probit model[14, 15] and estimate this by maximum likelihood. Disease risk is modelled as a function of demographic, socio-economic, and health indicators, lifestyle behaviours and smoking status. The estimated parameters indicate how the covariates impact disease risk. A positive coefficient is interpreted as increasing the likelihood of disease risk. A negative coefficient is interpreted as decreasing the likelihood of disease risk. To summarise the impact of the covariates on disease risk we use the model outputs to predict, *post model estimation*, disease risk for each smoking status group by *i)* deprivation quintile with and without a pre-survey disease incidence, *(ii)* deprivation quintile and healthy lifestyle behaviours without a pre-survey disease incidence, and *(iii)* deprivation quintile and unhealthy behaviours with a pre-survey disease incidence. In estimation the remaining model characteristics are held at their population mean value, apart from medical history other than a pre-survey smoking related disease event which will be set to zero.

Results

The sample consists of N=20,315 individuals aged 16 and older of which N=4827 (24%) experienced a smoking related disease event post survey. At the time of interview, 16% were never smokers, 23% never smokers exposed to ETS, 37% were smokers and 24% ex-

smokers. Smoking related diseases occurred not only amongst smokers or ex-smokers but these groups have the highest prevalence. Just over 11% of the sample experienced a smoking related disease event pre-survey; ex-smokers having the highest prevalence (18%).

Probit model of smoking related disease incidence

The probit model results are presented in Table 2. The coefficients in column two show the expected gradient in disease risk with smoking status. Men are significantly more at risk than women and disease risk significantly increases with age. Education and occupation reveal the expected social patterning. Individuals with higher or university education, or higher occupational class, are at lower risk of disease. When related to an area based deprivation measure, disease risk increases across the deprivation distribution.

Underweight, overweight or obese individuals have elevated disease risk compared to normal weight individuals. This is highest for underweight individuals. Individuals consuming alcohol either under or over the recommended limit at the time of interview have significantly lower disease risk. The link between alcohol consumption and disease risk here should not be treated as causal but rather behavioural. Individuals may have changed their alcohol consumption behaviour over time due to disease occurrence. Hence, these results need to be viewed with caution. Sports participation at vigorous intensity significantly reduces disease risk relative to inactivity or sports at light intensity with moderate activity having no significant effect. Disease risk is estimated to increase significantly with worsening general health. Note that this may reflect reverse causality. In terms of respondent medical history, a pre survey smoking related disease event significantly increases disease risk post survey.

Table 2 Probit model regression coefficients and 95% CIs for smoking related disease event*

	N=20,315		
	Coefficient	95% CI	
Never smoker exposed to ETS	0.071*	-0.006	0.147
Current smoker	0.359***	0.290	0.429
Ex-smoker	0.151***	0.080	0.222
Gender	0.211***	0.164	0.257
Age 16 to 24	-1.312***	-1.461	-1.163
Age 25 to 34	-1.190***	-1.301	-1.078
Age 35 to 44	-0.856***	-0.953	-0.759
Age 45 to 54	-0.578***	-0.661	-0.496
Age 55 to 64	-0.235***	-0.306	-0.164
Single	-0.096**	-0.173	-0.018
Married/cohabiting	-0.111***	-0.175	-0.047
Number of children	-0.024	-0.056	0.008
Number of adults in household	-0.068***	-0.103	-0.033
Low ed. - School leaving cert, ‘O’ grade, GCSE, etc.	0.050	-0.011	0.112
Lower mid level - SQV, ‘A’ level, ONC, OND/C&G, etc.	-0.044	-0.124	0.035
Upper mid level – HNC, HND, etc.	-0.107***	-0.187	-0.027
University degree or equivalent	-0.163***	-0.248	-0.079
Employed	-0.090***	-0.145	-0.035
Partly skilled	-0.077*	-0.165	0.010
Skilled	-0.160***	-0.237	-0.082
Intermediate/professional occupation	-0.124***	-0.211	-0.037
Regular drinks under the limit	-0.233***	-0.284	-0.181
Regularly drinks over the limit	-0.225***	-0.288	-0.162
Underweight	0.194**	0.031	0.356
Overweight	0.069***	0.017	0.122
Obese	0.108***	0.047	0.168
Moderate intensity sports	-0.028	-0.092	0.037
Vigorous intensity physical activity sports	-0.100***	-0.162	-0.038
Good self-assessed general health	0.175***	0.120	0.230
Fair self-assessed general health	0.349***	0.284	0.415
Bad/very bad self-assessed general health	0.622***	0.530	0.714

SIMD52	0.079**	0.002	0.155
SIMD53	0.114***	0.039	0.189
SIMD54	0.109***	0.032	0.185
SIMD55 (most deprived)	0.137***	0.058	0.217
Either parent died of CHD	0.112***	0.063	0.160
Presence of a non-limiting longstanding illness	-0.033	-0.102	0.036
Pre-survey hospitalisation for IHD	0.352***	0.263	0.440
Takes medicine for high blood pressure (HBP)	0.125***	0.038	0.212
Whether HBP diagnosed by doctor or nurse	0.182***	0.089	0.275
Presence of chronic kidney disease CKD	0.083	-0.147	0.313
Pre-survey smoking related disease incidence	0.533***	0.462	0.603
Constant	-0.357***	-0.493	-0.222
N	20315		
Pseudo R^2	0.259		
Log likelihood	-8259.181		
χ^2	4357.648		

* Note - * indicates significance at 10%; ** at 5%; and *** at 1% level.

Predicted disease risk post-survey

Figure 1a shows predicted risk based on average population characteristics in the presence and absence of a pre-survey smoking related disease event (SMRpre) for each smoking status group across the deprivation distribution. Predicted disease risk increases for all smoking status groups with increasing deprivation. However, disease risk and changes in disease risk across the distribution are small for never smokers and never smokers exposed to ETS relative to those for smokers and ex-smokers. Least deprived smokers have higher predicted risk (0.165; 95% CI: 0.149 to 0.183) relative to the most deprived ex-smokers (0.149; 95% CI: 0.133 to 0.164) or never smokers (0.116; 95% CI: 0.102 to 0.131). This also holds in the presence of a pre-survey smoking related disease event which always increases disease risk. This increase is highest for smokers, increasing disease risk for the least deprived smokers by 16.48 (95% CI: 0.139 to 0.191), for the most deprived smokers by 17.94 percentage points (95% CI: 0.153 to 0.206).

Figure 1 Predicted disease risk: Deprivation

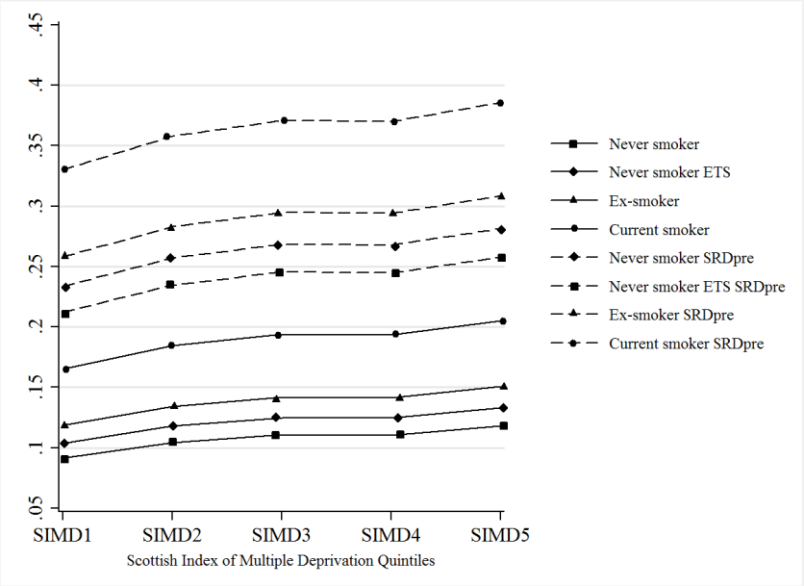


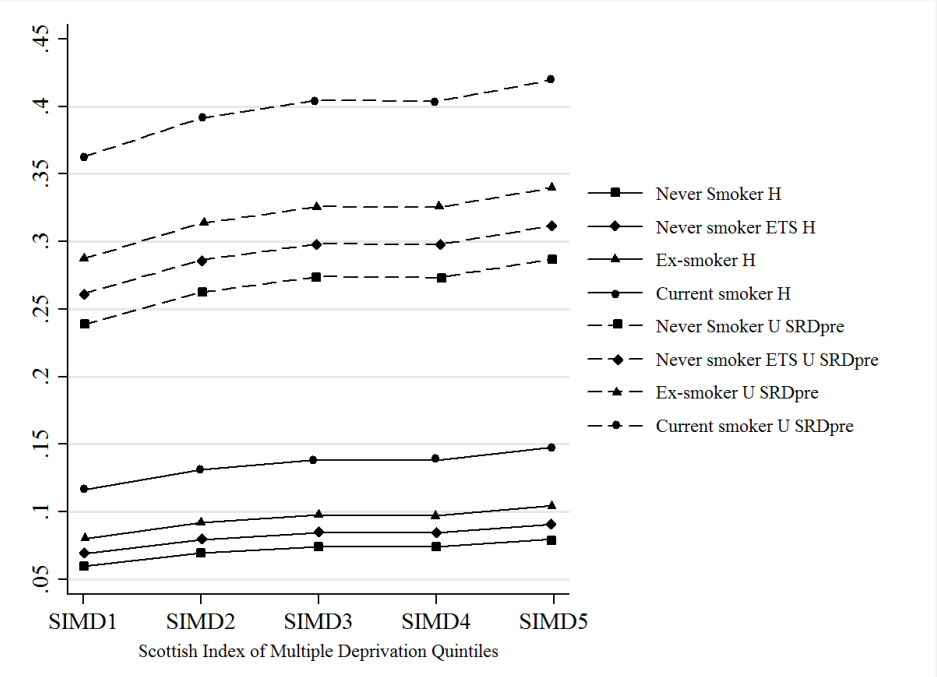
Figure 1b shows predicted risk for two extreme cases: a healthy lifestyle, H, without a pre-survey smoking related disease event and an unhealthy lifestyle, U, with a pre-survey event. We define a healthy lifestyle as a BMI of normal weight, physical activity participation at moderate or vigorous intensity and alcohol consumption within the recommended limit. An unhealthy lifestyle is defined as being overweight or obese, no sports participation or participation at light intensity, and alcohol consumption over the recommended limit. Comparing the lower half of Figures 1a and 1b, healthy behaviours generally reduce disease risks across the deprivation distribution for all smoking status groups. Predicted disease risk for the most deprived with healthy behaviours in Figure 1b is lower than for the least deprived without healthy behaviours in Figure 1a for all smoking status groups.

Predicted disease risk reductions vary by smoking status group with smokers predicted to experience the smallest relative changes in disease risk conditional on a healthy lifestyle across all deprivation groups ranging from a drop of 30% (-0.050; 95% CI: -0.067 to -0.032) for the least to a drop of 28% (-0.057; 95% CI: -0.077 to -0.037) for the most deprived quintile while healthy never smokers exposed to ETS are predicted to experience the highest decline in disease risk across all deprivation quintiles ranging from 34% (-0.035; 95% CI:-0.047 to -0.023) for the least to 32% (-0.042; 95% CI: -0.056 to -0.027) for the most deprived quintile. A healthy lifestyle is predicted to reduce smoking related disease risk across the deprivation distribution for all smoking status groups.

Considering the upper half of Figures 1a and 1b, unhealthy behaviours increase disease risk above that derived from deprivation and a previous disease event for all smoking status groups with smokers being most at risk. This suggests that unhealthy behaviours considerably impact disease risk beyond that derived from a previous disease event. Given the multi-factorial causes of smoking related diseases, in particular heart disease, the general picture emerging is that predicted risks are uniformly higher with unhealthy lifestyles and a pre-survey smoking related disease incidence and greatest for smokers. The change in risk for the least deprived smokers is 3.29 percentage points (95% CI: 0.002 to 0.064) and similar for most deprived smokers.

For completeness, supplementary Figures S1a and S1b show predicted disease risk for healthy and unhealthy behaviours with and without a smoking related disease event pre survey.

Figure 2 Predicted disease risk: Healthy lifestyle without SRDpre and unhealthy lifestyles with SRDpre



Discussion

Main finding of this study

Focusing policy efforts purely on reducing smoking prevalence may go some way to reduce smoking related disease risk and the socio-economic inequalities in disease risk but will not eliminate these completely. Other contributing risk factors need to be taken into account to reflect the multifaceted influences on and inequalities in disease risk. Our results provide

some supporting evidence indicating that, as expected, not only smoking behaviour but also pre-existing disease and other health behaviours are major factors in predicting the risk of a smoking related disease incidence. Whilst area deprivation is confirmed an important determinant of disease risk across smoking status groups, smoking status makes a bigger difference to disease than does the deprivation quintile.

The impact of healthy and unhealthy behaviours on the 'pure' deprivation risk has been shown to be of considerable magnitude, especially for smokers and ex-smokers. We know that smoking behaviour is socially patterned; smokers and ex-smokers are more likely to have no educational qualifications whilst never smokers not exposed to ETS are nearly twice as likely as the population average to have a degree or equivalent. Smokers are more likely to have unskilled or partly skilled employment and to live in the most deprived areas. Thus it appears that the main effect of deprivation is realized through the differential adoption of smoking behaviour. However, the most deprived areas have been shown to have the highest predicted disease risk regardless of smoking status.

Experiencing a smoking related disease event pre survey is predicted to approximately double the risk of a subsequent event and other health behaviours also play an important role in determining predicted risk. However, for the latter this may reflect reverse causality. A healthy lifestyle affects predicted risk more than area deprivation. For unhealthy smokers, the predicted risk reduction from adopting a healthy lifestyle is greater than the gain from quitting smoking. Unhealthy never smokers face similar predicted risks to 'healthy' smokers.

What is already known on this topic

Socially patterned clustering of adverse health behaviour is seen in many countries[16-19] including Scotland[20] but the evidence to support interventions targeting multiple behaviour change is mixed[21, 22]. Most studies targeted specific chronic disease risks groups, particularly cardiovascular disease and diabetes. Whilst evidence is limited, a small number of studies suggest that multiple behaviour interventions may increase the effectiveness of smoking cessation, by including other risk factors for cancer[21], or at least not reduce it[23, 24]. The interventions would need to be designed to address relevant risk factors, at an individual or group level. Adding weight management support to smoking cessation[22], for example, would address a second, highly prevalent, health problem as well as removing a potential barrier to successful quitting. Taken together with the finding that interventions are more effective in higher risk groups[25], multiple behaviour interventions are worth investigating with smokers with a previous disease history.

What this study adds

Our data show that the predicted risk of future disease events is approximately 3 times higher for smokers with other unhealthy behaviours and a previous disease event compared with smokers with other healthy behaviours and no previous disease event. Ex-smokers with a previous disease event and other unhealthy behaviours will also benefit from adopting other healthy behaviours to reduce future disease event reoccurrence.

A novel feature of our analysis is the separation of never smokers into groups reporting exposure to ETS and those who do not. The data on exposure pre-date the smokefree legislation introduction in Scotland, providing interesting insights into patterns of exposure and effects. Compared with never smokers not reporting exposure, never smokers reporting such exposure are younger, less likely to be in an intermediate or professional occupation and more likely to live in the most deprived area. Our results reveal an increase in predicted smoking related disease risk for those reporting exposure to ETS across all deprivation groups. Whilst not significantly different, this result does point to potential bias in the measurement of relative risk of never smokers.

Limitations of this study

We expect some misclassification of the self-reported risk factors that will account for part of the association between disease risk and deprivation and other associated factors. Further, due to the cross-sectional nature of the SHeS surveys, all model covariates including the lifestyle behaviours are measured at baseline. We therefore cannot account for time-varying lifestyle behaviours. This may explain the reduced disease risk in the regression analysis for alcohol consumption over the recommended limit. Future disease risk studies and its association with lifestyle behaviours should account for lifestyle behaviour changes over time where such data are available.

Conclusion

This paper presented results intended to inform priority setting for interventions to improve health and reduce health inequalities. Previous findings for inequalities in mortality also hold for morbidity; differences in smoking related behaviours across deprivation categories are an important driver of inequalities in risk of adverse outcomes, with predicted smoking related

disease risks disproportionately concentrated amongst individuals from most deprived areas and highest for the most deprived smokers. These results have the advantage of a shorter follow up period from the initial recording of smoking behaviour than cohorts in mortality studies. The predicted risks also suggest that the impact of smoking interventions in reducing risk across the deprivation distribution can be increased by targeting those with pre-existing smoking related disease events and with other unhealthy lifestyle behaviours. Our results suggest that interventions which successfully address both smoking and other unhealthy lifestyle behaviours or promote other healthy lifestyle behaviours will have the most impact.

Funding

This work was supported by the Medical Research Council National Preventive Research Initiative Phase 2 [G0701874]; see http://www.npri.org.uk. The Funding Partners relevant to this award are: British Heart Foundation; Cancer Research UK; Department of Health; Diabetes UK; Economic and Social Research Council; Medical Research Council; Research and Development Office for the Northern Ireland Health and Social Services; Chief Scientist Office; Scottish Government Health Directorates; The Stroke Association; Welsh Assembly Government and World Cancer Research Fund. The Health Economics Research Unit is funded by the Chief Scientist Office of the Scottish Government Health Directorates.

Acknowledgements

We thank Professor Luke Vale, Dr Diane Stockton and participants at the Faculty of Public Health conference, Aviemore, Scotland, November 2011 and UK Society for Behavioural Medicine conference, Stirling, Scotland, December 2011 for helpful comments. The views expressed in this article are those of the authors.

References

1 World Health Organisation. WHO report on the global tobacco epidemic. Warning about the dangers of tobacco. Geneva: World Health Organisation, 2011.

2 Taulbut M, Gordon D. Tobacco smoking in Scotland: An epidemiology briefing. NHS Health Scotland, 2008.

[3] Marmot M. Fair society, health lives. Strategic review of health inequalities in England post 2010. London: The Marmot Review, 2010.

[4] Hiscock R, Bauld L, Amos A, Fidler JA, Munafò M. Socioeconomic status and smoking: a review. Annals of the New York Academy of Sciences. 2011;1248:107-23.

[5] Mackenbach JP. What would happen to health inequalities if smoking were eliminated? BMJ. 2011;342.

[6] Gruer L, Hart CL, Gordon DS, Watt GCM. Effect of tobacco smoking on survival of men and women by social position: a 28 year cohort study. BMJ. 2009;338.

[7] Hart CL, Gruer L, Watt GCM. Cause specific mortality, social position, and obesity among women who had never smoked: 28 year cohort study. BMJ. 2011;342.

[8] Ilomäki J, Hajat A, Kauhanen J, Kurl S, Kaufman JS, Tuomainen TP et al. Relationship between alcohol consumption and myocardial infarction among ageing men using a marginal structural model. The European Journal of Public Health. 2011.

[9] IARC Handbooks of Cancer Prevention, Tobacco Control, Vol: 13. Evaluating the effectiveness of smoke-free policies. Lyon, France, 2009.

[10] Barnoya J, Navas-Acien A. Protecting the World From Secondhand Tobacco Smoke Exposure: Where Do We Stand and Where Do We Go From Here? Nicotine & Tobacco Research. 2012.

[11] Gray L, Batty GD, Craig P, Stewart C, Whyte B, Finlayson A et al. Cohort Profile: The Scottish Health Surveys Cohort: linkage of study participants to routinely collected records for mortality, hospital discharge, cancer and offspring birth characteristics in three nationwide studies. International Journal of Epidemiology. 2010;39:345-50.

[12] Lawder R, Elders A, Clark D. Using the linked Scottish health survey to predict hospitalisation & death. An analysis of the link between behavioural, biological andsocial risk factors and subsequent hospital admission and death in Scotland. Edinburgh: Information Services NHS NSS, 2007.

[13] Scottish Executive. Scottish Index of Multiple Deprivation (2006). General Report. Edinburgh: Scottish Executive, 2006.

[14] Amemiya T. Qualitative Response Models: A Survey. Journal of Economic Literature. 1981;19:1483-536.

[15] Maddala GS. Limited-dependent and qualitative variables in econometrics. Cambridge: Cambridge University Press; 1983.

[16] Chiolero A, Wietlisbach V, Ruffieux C, Paccaud F, Cornuz J. Clustering of risk behaviors with cigarette consumption: A population-based survey. Preventive Medicine. 2006;42:348-53.

[17] de Vries H, van 't Riet J, Spigt M, Metsemakers J, van den Akker M, Vermunt JK et al. Clusters of lifestyle behaviors: Results from the Dutch SMILE study. Preventive Medicine. 2008;46:203-8.

[18] Schneider S, Huy C, Schuessler M, Diehl K, Schwarz S. Optimising lifestyle interventions: identification of health behaviour patterns by cluster analysis in a German 50+ survey. Eur J Public Health. 2009;19:271-7.

[19] Conry M, Morgan K, Curry P, McGee H, Harrington J, Ward M et al. The clustering of health behaviours in Ireland and their relationship with mental health, self-rated health and quality of life. BMC Public Health. 2011;11:692-702.

[20] Lawder R, Harding O, Stockton D, Fischbacher C, Brewster D, Chalmers J et al. Is the Scottish population living dangerously? Prevalence of multiple risk factors: the Scottish Health Survey 2003. BMC Public Health. 2010;10:330-42.

[21] Goldstein MG, Whitlock EP, DePue J. Multiple behavioral risk factor interventions in primary care: Summary of research evidence. American Journal of Preventive Medicine. 2004;27:61-79.

[22] Prochaska JJ, Prochaska JO. A Review of Multiple Health Behavior Change Interventions for Primary Prevention. American Journal of Lifestyle Medicine. 2011;5:208-21.

[23] Prochaska JJ, Velicer WF, Prochaska JO, Delucchi K, Hall SM. Comparing intervention outcomes in smokers treated for single versus multiple behavioral risks. Health Psychology. 2006;25:380-8.

[24] Leslie W, Koshy P, MacKenzie M, Murray H, Boyle S, Lean M et al. Changes in body weight and food choice in those attempting smoking cessation: a cluster randomised controlled trial. BMC Public Health. 2012;12:389.

[25] Ebrahim S, Taylor F, Ward K, Beswick A, Burke M, Davey SG. Multiple risk factor interventions for primary prevention of coronary heart disease. Cochrane Database of Systematic Reviews: John Wiley & Sons, Ltd; 2011.

Supplementary Table S1 ICD 9 and ICD 10 group codes for smoking-related diseases

	ICD 9	ICD 10
Tuberculosis	010 – 012	A15, A16
Cancer of lip, oral cavity, pharynx	140 – 149	C00 – C14
Cancer of the oesophagus	150	C15
Cancer of the stomach	151	C16
Cancer of the liver	155	C22
Cancer of the pancreas	157	C25
Cancer of the larynx	161	C32
Cancer of the lung	162	C33X, C34
Cancer of the cervix	180	C53
Cancer of the bladder	188	C67
Cancer of the kidney	189	C64X, C65X, C66X, C68
Cancer, unspecified site	199	C80X
Leukaemia	204 – 208	C91, C92, C93, C94, C95
Rheumatic heart disease	390 – 398	I00X, I01, I02, I05, I06, I08, I09
Hypertensive diseases	401 – 405	I10X, I11, I12, I13, I15
Coronary artery disease	410 – 414	I20 – I25
Other heart diseases	415 – 429	I26 – I52
Cerebrovascular diseases	430 – 437, 438	G45, I60 – I68, I69
Atherosclerosis	440	I70
Aortic aneurysm	441	I71
Other circulatory diseases	442 – 448	I72 – I78, M30, M31
Pneumonia	480 – 486	J12 – J18
Influenza	487	J10, J11
COPD	490 – 492, 496	J40 – J44
Asthma	493	J45, J46
Gastric & duodenal ulcer	531 – 533	K25 – K27

Figure S1a: Predicted disease rik:Healthy and unhealthy lifestyle without SRDpre

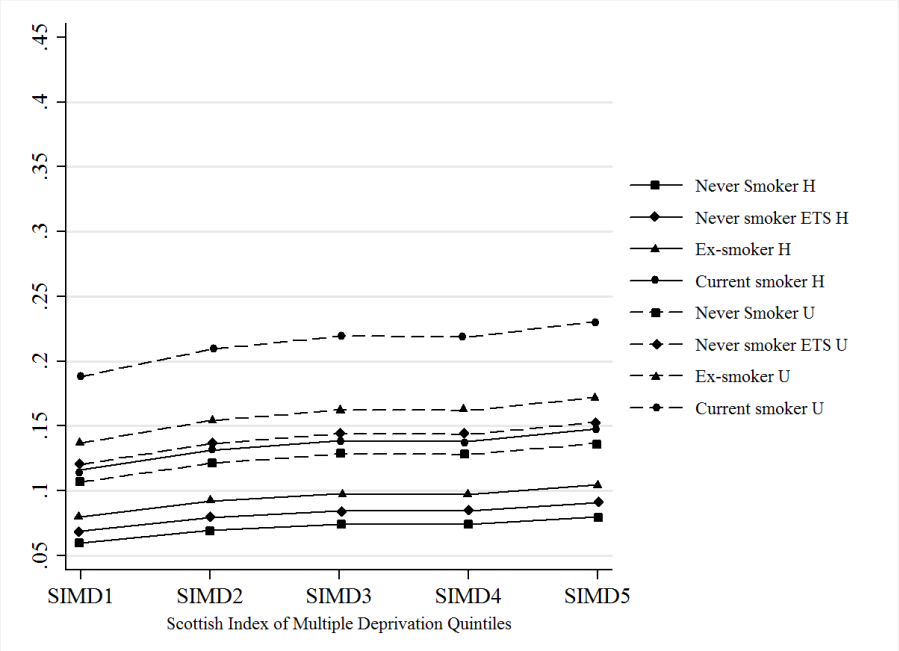


Figure S1b: Predicted disease risk: Healthy and unhealthy lifestyle with SRDpre