Inequalities in behavioural risk factor prevalence between five post-war generational cohorts of working age, England

Madhavi Bajekal PhD¹, Shaun Scholes PhD²
¹Department of Applied Health Research, Institute of Epidemiology and Health Care, University College London. *Corresponding author email: m.bajekal@ucl.ac.uk
²Department of Epidemiology and Public Health, University College London.

Abstract

Background and objectives: Long-term trends in health risk factor prevalence, and in inequalities, are often summarised using age-standardised point prevalence to allow for age distribution changes over time and between groups. Policies to effectively promote health require the decomposition of social change into its age-period-cohort components. As a first step, we provide a descriptive age-cohort analysis to identify patterns of generational change in key behavioural risk factors between and within post-war cohorts of working age in England.

Data: Cross-sectional Health Survey for England data for participants aged 25-60 years was pooled between 1994-2019 (n=153,172) to construct five decennial cohorts (1940s, 1950s, 1960s, 1970s, 1980s). Socioeconomic status was classified by neighbourhood deprivation quintiles using the Index of Multiple Deprivation. Five behavioural risk factors were analysed: cigarette smoking (current- and heavy-smoking); alcohol consumption (frequency and heavy drinking); obesity; meeting recommended levels of physical activity and fruit and vegetables consumption.

Methods: Sex-specific analyses were conducted. Log-binomial regression models quantified the magnitude and direction of change in (i) prevalence ratios (PRs) between cohorts adjusting for age and deprivation and (ii) deprivation-specific PRs within- and between-cohorts to examine changes in absolute and relative inequalities between generations.

Results: In more recent cohorts, decreases in prevalence, independent of age and deprivation, were observed for current smoking, frequency of alcohol consumption and heavy drinking, resulting in decreasing absolute inequalities. However, obesity levels, particularly among females, reached their highest levels in the youngest 1980s cohort.

Relative inequalities in current smoking (most- versus least-deprived quintiles) peaked in the 1950s cohort (Males: PR 2.79 (95% CI: 2.57-3.04); Females: PR 2.81; 95% CI: 2.60-3.05)), decreased in the 1960s cohort (M: PR 2.00 (95% CI: 1.70-2.34); F: PR 2.58 (95% CI: 2.40-2.77)), and remained stable thereafter. Inequalities in heavy smoking persisted over time among current smokers.

Higher obesity levels in the most- versus least-deprived quintiles were generally persistent across all five cohorts, albeit with some suggestion of widening inequalities in the younger- versus older-cohorts in females (1940s cohort: PR 1.55 (95% CI: 1.40-1.72); 1960s cohort: PR 1.87 (95% CI: 1.73-2.01)). This pattern was also observed for mean body mass index (BMI).

For heavy drinking, relative inequality remained stable. Relative inequalities in fruit and vegetable consumption were lower in more recent cohorts. Physical activity levels were similar across cohorts, with little evidence of inequalities.

Conclusion: Our analysis of generational change reveals credible signals of behavioural risk factor changes and inequalities over successive post-war cohorts of working age.
Introduction

There is a substantial body of work that demonstrates the persistence of socioeconomic inequalities in mortality and its associated behavioural risk factors in England.\(^1,^2\) Typically, these studies rely on point estimates derived from social surveys or vital statistics data. Annual estimates are typically age-adjusted and stratified by one or more socioeconomic indicators of disadvantage to analyse changes in risk factor prevalence and trends in inequalities between different sub-groups over time. However, in this study we use a different approach. We compare (i) differences in the prevalence of known health behavioural risk factors between post-war birth cohorts and examine (ii) changes in inequalities by constructing generational (or pseudo-, or synthetic-) cohorts using a long series of annual, repeated cross-sectional health surveys.\(^3\) Unlike longitudinal birth cohorts that prospectively track changes in individuals’ health, generational cohorts offer a unique perspective by capturing the broader social changes occurring over time between cohorts using repeated cross-sectional data. This method works to alert policy makers to probable cohort effects as long as the survey data are representative of the population, and its members are not experiencing significant mortality, which could lead to selective survival bias.

Research conducted using prospective longitudinal studies has demonstrated that the adoption of healthy lifestyle behaviours during young adulthood can effectively postpone the onset and lower the incidence of chronic diseases.\(^4,^5\) However, such longitudinal studies tend to be resource-intensive and typically concentrate on a single cohort. Generational cohort studies based on annual, repeated cross-sectional data offer a valuable complement to longitudinal research by zeroing-in on the behavioural drivers of health and in estimating the magnitude of any changes in prevalence between earlier and later-born cohorts.

This approach has the potential to inform proactive public health strategies and policy making, contributing to gains in health expectancy of successive generations as they age.\(^6\) Furthermore, for pension actuaries, gaining insight into the direction of change in these pivotal determinants between successive post-war cohorts of working age provides a more solid basis for calibrating annuity models that project future life expectancies, ultimately facilitating more precise financial planning for retirement and pensions.

Using a generational cohort approach, the key research questions of the present study are:

1. Has the prevalence of key behavioural risk factors changed between cohorts, independently of age and deprivation?
2. Has the magnitude of risk factor inequalities changed between cohorts?

Methods

Data and participants

The Health Survey for England (HSE) is a series of repeated, annual cross-sectional health examination surveys designed to monitor health conditions and health-related behaviours in a nationally representative sample of the non-institutionalised population. A new sample each year is randomly selected by address. The sample size and the health topic focus varies each year with certain ‘core’ modules retained every year.\(^7\)
HSE data by single-year of age are accessible through the End User Licence (EUL) via the UK Data Service (1994-2013). However, in the more recent EUL datasets (2014-2019), participants’ age is grouped, limiting its use for constructing a generational cohort series.

**Birth cohorts and age**

We aggregated annual cross-sectional HSE data from 1994 to 2019. This pooled data allowed us to categorise participants into five decennial post-war birth cohorts: the '1940s' cohort (born between 1940-49), the '1950s' cohort (born between 1950-59), the '1960s' cohort (born between 1960-69), the '1970s' cohort (born between 1970-79), and the '1980s' cohort (born between 1980-89).

We restricted analyses to participants aged 25 to 60 at the time of the survey. This age truncation resulted in varying age ranges for each cohort, spanning from 25 to 39 years in the youngest 1980s cohort to 45 to 60 years in the oldest 1940s cohort. The selected age range of 25 to 60 years ensured that the analytical sample included participants of the same age in overlapping cohorts. The 1960s cohort covered nearly the full age range of the study, from 25 to 59 years (see Table 1). To put in context, in 2019 the central age of the five cohorts was 74, 64, 54, 44, and 34 years (plus/minus 5 years), from the oldest (1940s) cohort to the youngest (1980s). The total study sample size was 153,572. The mean age of the analytical sample was 41 years.

With the exception of the extreme tails truncated by the specified age cut-offs, we derived risk factor prevalence estimates by single-year of age by aggregating both the numerators and denominators over ten consecutive years of annual survey data for each cohort (see Supplementary Appendix S1_ Section B.1). The median sample size per single year of age for the 1960s cohort was 1,513.

Our decision to use 10-year grouping to construct the generational cohorts, rather than the more conventional 5-year grouping commonly used in demography, was mainly because of limits to sample sizes once the analytic sample is partitioned by age, sex and deprivation quintiles (see below) to assess inequalities in risk factor prevalence within- and between-cohorts.

**Behavioural risk factors**

Over the last two decades, there have been changes in survey questions or updates to measurement methods to align with policy recommendations. Our primary aim has been to maintain a consistent definition of the chosen behavioural risk factors throughout the study years, thereby ensuring that we avoid introducing artificial ‘jumps’ in prevalence levels between cohorts. However, in cases where the variables were not collected consistently and could not be harmonised (e.g. changes to units of alcohol), or when data collection commenced more recently (e.g. fruit and vegetable consumption), we initiated the analysis from the first available year when these variables were consistently collected.

We selected five modifiable behavioural risk factors for our analysis, namely the proportions of individuals who:
• Were current cigarette smokers; and among them, those who were ‘heavy smokers’, defined as smoking 20 or more cigarettes per day.

• Typically consumed alcohol on five or more days per week over the past 12 months (frequency); and of those who consumed alcohol last week, those who were ‘heavy drinkers’ (8+ units for males and 6+ units for females) on the heaviest drinking day in the last week (amount).

• Had a measured Body Mass Index (BMI) of 30kg/m² or more (i.e. obesity).

• Were physically active, defined as being engaged in 30 minutes or more of moderate- or vigorous-intensity activity on at least five days per week (including activity whilst at work).

• On average, consumed five or more portions of fruits and vegetables per day.

Further details of the survey protocol and the questions asked are provided in S1_Section A.

Socio-economic circumstances

The HSE includes three measures of individual socio-economic position – namely, household income, highest educational attainment, and occupation-based social class (e.g. Registrar General’s social class, and the National Statistics Socio-Economic Classification, NSSEC). However, household income and educational attainment were collected on a consistent basis only from 1997 onwards. Additionally, there was a gap in the archiving of social class information in 2010-2011, resulting in a discontinuity in the series.

Aside from being incomplete series, the measurement of individual social position was also impacted by a substantial amount of item non-response, which affected cohorts differentially. For instance, information on household income was missing for approximately 30% of the analytic sample; but this ranged from 44% in the 1940s cohort to 16% in the 1980s cohort. Likewise, the percentage of missing data for the oldest and youngest cohorts was 23% and 0% for educational attainment, respectively. The equivalent figures were 2% and 18% respectively for social class.

In contrast, the extent of missing data for a geographical measure of relative deprivation based on the area of residence of the HSE participants, known as the Index of Multiple Deprivation (IMD) at the Lower Layer Super Output Area (LSOA), exhibited the lowest overall rate of missingness, standing at 7%. Notably, this missingness pattern did not vary significantly by cohort, ranging from 8% for the 1940s cohort to 2% for the 1980s cohort (See S1_SB.2).

Throughout the study’s duration, the national IMD scores have been periodically recalibrated to reflect changes in relative deprivation, with assessments conducted at various time points (e.g., IMD 2001, IMD 2007, IMD 2010, IMD 2015). Although certain small areas may have experienced shifts in their specific rankings over time, our approach has been to use the deprivation measure that was in effect at the time of each survey. Deprivation quintiles are sizeable and encompass roughly equal fifths of the national population. Importantly, the majority of LSOAs have consistently maintained their quintile-membership across time, indicating remarkable stability. In our analysis, we classify IMD1 as the most deprived quintile (Q1) and IMD5 as the least deprived quintile (Q5).

Our pragmatic selection of IMD quintiles as the primary measure of social disadvantage in this analysis is rooted in the theory that measures of area disadvantage encompass both compositional aspects (such as the concentration of deprived individuals) and contextual factors (such as the availability of employment opportunities) related to relative disadvantage.
Statistical Analyses

Sex-specific analyses were conducted. We estimated the prevalence of each behavioural risk factor by single year of age separately in each cohort, and present prevalence using three-year moving averages. Log-binomial regression models were used to estimate prevalence ratios (PR) using the specific binary risk factor as the dependent variable with age (continuous), cohort (categorical) and IMD quintile (categorical) as independent variables.\textsuperscript{11}

Two sets of models were fitted. To examine cohort differences (research question 1), the first set of models (model 1) contained the main effects of cohort (1960s as reference), age and IMD (least deprived quintile as reference) to estimate PRs comparing the cohorts after adjustment for age and deprivation. To examine differences in the magnitude of relative inequalities between cohorts (research question 2), the second set of models (model 2) included an additional cohort by IMD interaction term (1940s cohort as reference). An overall test of the interaction term (16 terms) was used to jointly examine whether the PRs comparing the deprivation quintiles varied by cohort (after age adjustment). A more specific test of four terms was used to examine whether the relative gap in risk factor prevalence between the least and most deprived quintiles varied by cohort.

To facilitate interpretation of the patterns in inequalities, the second set of models was used to predict risk factor prevalence by deprivation quintiles within each cohort at age 40. We report on two measures of inequality – absolute and relative. Absolute inequality is the difference in prevalence between the most and least deprived quintiles (Q1-Q5); relative inequality is the ratio of the prevalence of the most deprived quintile to the least deprived (Q1/Q5). Examining changes in inequalities on both the absolute and relative scales is important. For example, if risk factor levels increase over time in the most- and least-deprived areas at about the same pace, relative inequalities between successively younger cohorts remain unchanged, but absolute inequalities could increase.

BMI was examined as a continuous measure. Linear regression models were used to estimate differences in mean BMI between cohorts after adjustment for age and deprivation (model 1) and quantify the magnitude of absolute inequalities within- and between-cohorts (model 2).

Data set preparation was performed in SPSS V.24.0 (IBM, Armonk, New York); analysis was performed in Stata V.17.1 (StataCorp LLC, College Station, Texas).

Weighting and missing data

In its early years the HSE was considered to be self-weighting, unless the sample sizes for a specific age or social group, such as children or minority ethnic groups, had been boosted in that year. Since 2003, non-response weights have been included. Notably, these non-response weights were introduced almost at the mid-point of the survey years between 1994 and 2019 analysed in this study. Consequently, nearly all data points for the 1940s cohort in our study are unweighted, while all data points for the 1980s cohort are weighted. The remaining cohorts represent a mix of weighted and unweighted data.

Given this inconsistency in cohort-specific weighting, we have opted not to incorporate survey non-response weights into this analysis (boost samples were not included and so the analytical sample is self-weighted). Previous studies that analysed generational cohorts using HSE data conducted sensitivity analyses with and without non-response weights, noting no significant differences in estimates.\textsuperscript{12,13}
We opted for a complete case analysis in our regressions for several reasons. First, both the selected outcome and predictor variables (age, sex, IMD) exhibited reasonably low levels of item non-response, in part due to data being collected at the interview stage rather than at the nurse visit. Secondly, as behavioural variables are self-reported, reliable imputation using the limited set of sociodemographic information recorded consistently was not feasible. Therefore, a complete case analysis was considered the most suitable approach for our study.

Results
Table 2 shows the overall prevalence for each risk factor by survey year for adults aged 25 to 60 in the pooled HSE study sample.

For each behavioural risk factor, we present a set of six figures, three for each sex displayed side-by-side to facilitate comparison.

An overview of these graphical representations is provided below:

**Percentage prevalence by cohort and age (top panel):** The first graph in the set shows the observed risk factor prevalence by cohort and single-year of age, smoothed using three-year moving averages. This helps to illustrate common age-patterns and, where ages overlap between successive cohorts, the between-cohort differences.

**Cohort-specific prevalence ratios with 1960 as reference cohort (middle panel):** In this, based on the first set of log-binomial models, we present the estimated PRs comparing the cohorts after adjustment for age and deprivation, with associated 95% confidence intervals (95% CIs). The 1960s cohort was chosen as reference. To facilitate comparison between cohorts, the values on the bar graph have been rebased with the 1960 reference value of 1 set to zero. Hence an estimated PR of 0.20 for a specific cohort indicates that the prevalence was estimated to be 20% higher (in relative terms) than for the 1960s cohort independently of age and deprivation.

**Predicted prevalence at age 40 by cohort and deprivation quintile (bottom panel):** The third graph in the set, based on the second set of log-binomial models (containing a cohort by IMD interaction term), displays model-predicted risk factor prevalence at age 40, by IMD quintile and cohort. Additionally, the second y-axis displays the estimated PR between the most and least deprived quintiles within each cohort (least deprived as reference). Hence, both the difference in prevalence between the most and least deprived quintiles in each cohort (absolute inequality; Q1-Q5) and the ratio between them (relative inequality; Q1/Q5) are presented.

Tables that support the data presented in each figure can be found in the Supplementary Appendix, (S1_SC and SD). A summary table of the estimated PRs (most versus least deprived quintiles) for each binary risk factor at age 40, by sex and cohort is presented in Table 3. Results for mean BMI are presented separately in S1_SD6.

These visualisations and tables together offer an overview of the comparative differences in prevalence and in inequalities between cohorts associated with each risk factor, separately for males and females.
Current Smoking

Age patterns and cohort differences

The age pattern of current smoking was consistent across cohorts. Prevalence was highest at age 25, and fell with increasing age for each cohort for both sexes (Fig 1, top panel). Proportions currently smoking were slightly lower in females than in males (S1_Table SC1).

As apparent from the stacking by age for overlapping cohorts, proportions of current smoking fell between successive cohorts, being lowest in the 1980s cohort and highest in the 1940s cohort (Fig 1). After adjusting for age and IMD composition of each cohort, relative to the 1960s cohort, prevalence rates for males were 24% lower in the 1980s cohort, and 33% higher in the 1940s cohort. For females, the difference was greater, at 37% lower and 47% higher, respectively.

Inequalities

Smoking prevalence was consistently higher in the most deprived compared to the least deprived quintiles across all cohorts. From its highest levels in the 1940s cohort, prevalence fell for successively younger cohorts in all deprivation quintiles and both sexes (Fig 1). Predicted prevalence at age 40 for each cohort halved for both males and females in the most deprived quintile between the oldest (1940s) and youngest (1980s) cohorts (from 60% to 31% for males; and 57% to 24% for females, see Table SD1). The percentage point fall between the 1940s and 1980s cohorts in the least deprived quintile was smaller, albeit from a lower base prevalence (males: 23% to 15%; females: 24% to 11%, respectively).

As a result, absolute inequality, as measured by the ‘gap’ or difference in smoking prevalence between the most and least deprived IMD quintiles in each successively younger cohort decreased. Relative inequality, as measured by the PR between the most and least deprived quintiles (least deprived as reference), was at its highest point in the 1950s cohort, at 2.80 (95% CI: 2.57-3.04) times higher for males and 2.82 (95% CI: 2.61-3.06) times higher for females (Fig 1, Table 2). This relative gap fell in the 1960s cohort and remained relatively stable thereafter. This suggests that the pace of fall in current smoking for successively younger cohorts was faster in the most deprived quintile compared to the least deprived quintile.

Heavy smoking

Moreover, of those who were current smokers in each cohort, the prevalence of heavy smoking (20 or more cigarettes per day) also fell dramatically across all deprivation quintiles (see Fig 2, S1_Tables SC2 and SD2). The PRs indicating the magnitude of relative inequality (above 1, indicating higher levels in the most versus least deprived quintiles) remained stable across successive cohorts, suggesting that the pace of fall in heavy smoking was of a similar magnitude across deprivation quintiles.

Alcohol consumption: frequency

Age patterns and cohort differences

The prevalence of frequent drinking (daily or up to five days a week) showed a marked decrease between successively younger cohorts for both sexes, particularly for males (Fig. 3, S1_Tables SC3).
Relative to the 1960s cohort, levels of frequent drinking were more than 50% higher for the 1940s cohort and more than 50% lower for the 1980s cohort.

<Fig 3 here>

Inequalities

The predicted prevalence of frequent drinking at age 40 was consistently higher in the least deprived quintiles when compared to the most deprived (Fig 3). Females in the most deprived quintile were about half as likely, and males about 30% less likely, as those in the least deprived quintile to be frequent drinkers. Whilst the absolute gap between the most and least deprived quintiles declined between successively younger cohorts for both sexes, the relative gap remained stable across the cohorts (Fig 3, Table 2 and S1_Table SD3).

Alcohol consumption: heavy drinking

Age patterns and cohort differences

The prevalence of heavy drinking (8+ units for males and 6+ units for females on the heaviest drinking day in the last week) among those who reported that they had consumed alcohol in the last week showed mixed results partly due to inconsistent data collection on units consumed in the first half of the survey period, which limited the age range for valid data for the older cohorts (Table 1). In comparison to the 1960s cohort, heavy drinking levels were notably lower for the 1970s and 1980s cohorts, for both sexes (see Fig 4), but were similar for the two oldest cohorts (1940s and 1950s).

<Fig 4 here>

Inequalities

In contrast to the pattern of frequent drinking, relative inequalities were evident for heavy drinking among all cohorts in both sexes, with higher levels in the most versus least deprived quintiles (Table 2; S1_Table SD4).

Whether analysed either from the perspective of consumption frequency or the amount consumed on the heaviest drinking day, it is evident that the 1980s cohort in particular showed the lowest levels of alcohol consumption across all deprivation quintiles.

Obesity

Age patterns and cohort differences

The prevalence of obesity (BMI 30+ kg/m²) showed a steady increase across cohorts, with the lowest rates observed in the 1940s cohort and the highest in the 1980s cohort. This trend was especially pronounced among females (Fig 5). As expected, obesity levels increased consistently with advancing age, and these levels did not significantly differ between the sexes. After adjustment for age and deprivation, compared to the 1960s cohort, obesity levels in relative terms were 28% higher for males in the 1980s cohort and 37% lower for those in the 1940s cohort. For females, the corresponding levels were 41% higher for the 1980s cohort and 28% lower for the 1940s cohort.

<Fig 5 here>

Inequalities

The magnitude of absolute inequality in obesity within cohorts for males showed a 3-4 percentage point (pp) difference between the most and least deprived quintiles. For females, absolute inequality
was not only higher but also increased between successively younger cohorts from 7 pp in the 1940s cohort to 20 pp in the 1980s cohort. Predicted prevalence at age 40 consistently showed higher levels of obesity in the most deprived quintile for both sexes across all cohorts, except for males in the 1980s cohort, where the estimated PR did not attain statistical significance (PR 1.13; 95% CI: 0.94-1.36). The relative gap between the most and least deprived quintiles consistently and significantly exceeded 1 for females across all cohorts, with the PR increasing between the 1940s (PR 1.55; 95% CI: 1.40-1.72) to the 1960s cohort (PR 1.87; 95% CI: 1.73-2.01), and then remaining relatively stable thereafter (Table 2; S1_Table SD5).

Mean BMI

Cohort differences

Rather than focusing solely on obesity at the upper end of the weight-to-height spectrum, an analysis of mean BMI confirmed the observed increase in higher BMI in the younger versus older cohorts (Fig 6).

<Fig 6 here>

Inequalities

In males, higher levels of BMI in the most versus least deprived quintile were significantly higher only in the oldest (1940s) cohort (difference in mean BMI: 0.31kg/m^2; 95% CI: 0.03-0.60) (S1_Table SD6). In contrast, in females, those in the most versus least deprived quintile had significantly higher mean BMI in each cohort, with the absolute gap increasing from 1.59kg/m^2 (95% CI: 1.25-1.94) in the 1940s cohort to 2.26kg/m^2 (95% CI: 1.82-2.71) in the 1980s cohort.

Physical Activity

Age and cohort differences

At all ages between 25-60 years the proportion of physically active males (engaged in 30+ minutes of at least moderate-intensity physical activity, 5+ days/week) exceeded that of females, with levels highest at age 25 and declining gradually with advancing age. Age and deprivation-adjusted levels of sufficient physical activity were notably higher for participants in the youngest versus older cohorts, for both sexes (Fig 7). The difference between the youngest and oldest cohorts was more pronounced in females than in males.

<Fig 7 here>

After adjustment for age and deprivation, prevalence rates for females were 32% higher for the 1980s cohort and 32% lower for the 1940s cohort (relative to those in the 1960s cohort). For males, the pattern was similar but the gradient between cohorts was less pronounced, with physical activity levels being 24% higher for the 1980s cohort and 18% lower for the 1940s cohort.

Inequalities

Physical activity levels increased consistently across all deprivation quintiles, in parallel with cohort progression, with little evidence of inequalities or changes in inequalities (Table 2, S1_Table SD7).
Fruit and Vegetables consumption

Age and cohort differences

Data for fruit and vegetable consumption were collected over a limited number of years in the HSE series (starting from 2001), and the proportion of participants eating five or more portions per day remained relatively constant between the cohorts (Fig 8). There was a consistent age pattern, with individuals aged 25 having slightly lower consumption levels compared to older ages. From approximately age 50 onwards, a greater proportion of females than males reported consuming five or more portions of fruits and vegetables daily.

Fig 8 here

Inequalities

Levels of fruit and vegetable consumption at age 40 showed a consistent inequality gradient, with those in the least deprived quintile being more likely to meet the recommended number of portions, compared to their counterparts in the most deprived quintile. However, among both sexes, the absolute and relative gaps in prevalence were lower in the younger versus older cohorts. In males, for example, the PR significantly decreased from 0.52 (95% CI: 0.43-0.63) in the 1940s cohort to 0.74 (95% CI: 0.65-0.85) in the 1970s cohort (Table 2; S1_Table SD8). Likewise, among females, the PR significantly decreased from 0.56 (95% CI: 0.48-0.65) in the 1940s cohort to 0.78 (95% CI: 0.63-0.97) in the 1980s cohort. In simpler terms, the proportion of participants meeting the 5-a-day recommendations increased at a slightly faster pace in the most deprived quintile between successively younger cohorts, though starting from a lower base.

Discussion

To our knowledge, this study represents the first analysis of generational birth cohort trends which includes five of the seven major risk factors identified by the World Health Organisation to reduce premature mortality from non-communicable diseases by 25% by 2025. These five risk factors are causally linked to the future burden of morbidity, quality of life and leading causes of premature mortality. Our study maximises the utility of a long running high-quality cross-sectional and nationally-representative health data series by presenting a generational post-war cohort analysis for the population of England at working ages before survival bias substantially distorts patterns observed, particularly for health related analyses. This approach offers an alternative perspective on temporal trends in behavioural risk factor prevalence between cohorts, as well as changes in inequalities within each. Below we discuss our findings for each risk factor in turn.

Current and heavy smoking

Of the five risk factors studied, our results for current smoking were the most positive, suggesting that age- and deprivation-adjusted prevalence levels fell between the oldest and youngest cohorts, for both sexes. The absolute gap in current smoking prevalence between the most- and least-deprived quintiles narrowed, and relative inequalities fell significantly, for the 1960s cohort compared to the 1940 and 1950 cohorts, and remained at this lower level for subsequent younger cohorts.

The continuous decline in smoking prevalence at working ages in the more recent born cohorts has been attributed mainly to a decline in smoking initiation before the age of 25, and to a lesser extent to smoking cessation thereafter.
These findings align partly with a pair of previous generational studies, despite differences in survey data used and in timeframes.\textsuperscript{16,17} Basing interpretation on line graphs of observed prevalence, these studies also showed that age-for-age smoking prevalence fell for males and females in younger cohorts, but that the absolute gap between manual and non-manual groups appeared to peak for the 1950s cohort and thereafter stabilised at this higher rate, thereby implying that smoking inequalities might not diminish in more recently born cohorts. In contrast, because we have incorporated data from later rounds of surveys, and used a modelling approach which adjusted for age effects and used a finer gradation of socioeconomic circumstances (by neighbourhood deprivation quintiles), our results showed that both absolute and relative inequalities in smoking prevalence between successive cohorts have fallen markedly.

Our finding of relative inequality reduction from a cohort perspective also appears to be at odds with a recent analysis of temporal trends in socioeconomic inequalities in smoking prevalence between 2003 and 2019 for all adults aged 16 and over.\textsuperscript{18} Using a summary measure of age-standardised smoking prevalence, this study reported an increase in relative inequalities between the most- and least-advantaged groups based on measures of individual socio-economic position. However, similar to our findings, they did not observe a significant increase in either absolute- or relative-inequalities in smoking prevalence when examined by IMD quintiles. It is important to reiterate that annual age-standardised prevalence estimates summarise the average experiences of different combinations of cohorts. This highlights the importance of using different approaches to disentangle any cohort effects on temporal trends, especially when considering what future smoking trends might look like once older cohorts are replaced. In the present study, we specifically examined generational cohort effects to provide an alternative insight into how inequalities in health behaviours have evolved over the last two decades. Besides differences arising due to study design and coverage, messages from different studies can be valid and serve to provide complementary results albeit from a different perspective.

Among current smokers, the prevalence of heavy smoking (20 or more cigarettes a day) markedly decreased across all ages for younger cohorts, and at a similar pace across deprivation quintiles, especially in males. As a result, for males, absolute inequalities at age 40 reduced by about two-thirds from the level observed in the 1940s cohort compared to the 1980s cohort, but with no discernible difference in relative inequalities across cohorts. Females also experienced a sharp decline in absolute inequalities; however, relative inequalities increased over successively younger cohorts, suggesting that females living in the least deprived areas were quicker to reduce consumption compared to their counterparts living in the most deprived areas.

\textit{Alcohol consumption: frequency and amount}

With respect to alcohol consumption, longitudinal cohort studies have shown that the pattern of alcohol consumption fluctuates over the life course, with frequency of drinking occasions gradually becoming most prevalent during middle-age, stabilising, and then declining from about age 70.\textsuperscript{19} In the present study, we observed this age-pattern in the cohorts born in the 1960s or earlier. However, in the younger cohorts, this age pattern was absent, and the levels of frequent drinking remained consistently low from age 25 onwards.

The observed between-cohort differences in the frequency of alcohol consumption mirrored that of cigarette smoking, with prevalence falling dramatically for successively younger cohorts, resulting in the narrowing in absolute inequalities. However, unlike smoking, we observed that the social gradient flipped, with participants in both sexes in the least deprived areas showing higher levels of frequent drinking than their counterparts in the most deprived areas across all cohorts. The reversal
in drinking habits between more and less advantaged groups connects to the widely reported ‘alcohol harm paradox’, with levels of alcohol related harm and mortality being markedly higher in disadvantaged groups despite lower levels of alcohol consumption.20

This paradox may, in part, be explained by the patterns in the prevalence of harmful drinking, defined in the present study as drinking more than twice the daily recommended units on the heaviest drinking day in the past week. Females living in the most deprived quintiles, in particular, consistently had higher rates of harmful drinking compared to all other quintiles. For males, the heavy drinking gradient with deprivation was found to be less clear-cut. Unfortunately, changes in the definition of a standard unit of alcohol in the HSE over the study period limited deeper analysis of differences in daily consumption patterns between birth cohorts and across deprivation quintiles.

Harmful drinking prevalence levels were significantly lower in the 1970s and 1980s cohorts compared to cohorts born in 1960s or older, for both sexes. However, the signal for falls in both absolute and relative inequalities between cohorts was strong from the 1940s until the 1960s cohort, but remained static thereafter.

**Obesity**

Temporal trends in obesity among adults have shown a steady increase, with prevalence levels in England approximately doubling between 1993 and 2019 in both sexes.21 Obesity levels estimated from HSE data have been shown to be consistently higher in women than in men across all age groups, and higher for the most disadvantaged groups, as indicated by both individual (e.g. equivalised household income quintiles) and area-based measures of socio-economic position. In the present study, as expected, prevalence levels increased monotonically until the age of 60 across all cohorts. Relative to the levels in the 1960s cohort, obesity rates were significantly higher in the 1970s and 1980s cohorts, particularly for females, after adjusting for age and deprivation.

Analyses of long-run trends in inequalities in obesity prevalence are limited. In our study, against a backdrop of rising obesity prevalence, we observed modest inequalities in males in all cohorts. In contrast, the magnitude of absolute inequalities tripled between the older and younger cohorts in females from the 1940s to 1960s cohort, but then stabilised thereafter.

However, the key question underlying the surge in obesity levels across developed countries revolves around the extent to which it can be attributed to period effects (such as increased access in more recent years to inexpensive fast food) and how much to cohort effects (through factors such as sedentary behaviour, and the obesogenic environment). Analyses of obesity patterns conducted using age-period-cohort models in both the United States (using a longitudinal cohort)22 and in England (examining generational cohort patterns)23 have concluded that cohort effects are just as significant as period effects, especially in more recent cohorts. While the estimates of age-cohort effects presented in our descriptive study incorporate the contribution of period effects, we have not disentangled the independent contributions of each. Reither et al. suggest that birth cohort membership adds an additional risk factor, distinct from period effects that impact all sections of society uniformly.18

**Physical activity**

The rise in physical activity levels in later-born cohorts observed in our study aligns with secular annual trends showing a year on year increase in physical activity at recommended levels.24

Interestingly, across deprivation quintiles, estimated activity levels (presented herein) at age 40 increased at a similar pace across successively younger cohorts, leading to no discernible differences
in inequalities. This is in contrast to age-standardised point estimates of differences in physical activity levels for all adults (i.e. aged 16 and over) using HSE 2016 data which showed a graded relationship by IMD quintile, with just over two-thirds of adults living in the least deprived quintile meeting the recommended activity levels as compared to half of those living in the most deprived quintile.25 This suggests that age trajectories of activity levels may diverge, following a pattern of decline that varies across different socioeconomic groups as individuals enter older age. Furthermore, differential survival rates between deprivation groups at older age may potentially exacerbate the social gradient observed in prevalence rates estimated among all adults, rather than focusing solely on adults of working age.

Fruit and vegetables consumption

Statistics using the more accurate food diary collection method in the National Diet and Nutrition Surveys (NDNS) support the findings from the HSE, suggesting minimal change in the proportion of the adult population consuming five or more portions of fruits and vegetables a day from 2001 onwards.26,27 But few studies have analysed point estimates of socioeconomic inequalities in fruit and vegetable consumption, and none have reported on trends in these inequalities over time.

Our analysis by generational cohorts hints at encouraging positive signals in that the absolute and relative gaps in prevalence narrowed in the younger versus older cohorts, reflecting a slightly faster increase in the most deprived areas, albeit from a lower starting base than for those living in the least deprived areas.

Similar to the challenges posed in analysing obesity trends, it is difficult to disentangle the independent contribution of period changes, such as improved availability of fruits and vegetables, from generational cohort differences in consumption influenced by raised awareness of health benefits following the introduction of the 5-a-day programme in the UK in 2003. An analysis of a single British birth cohort (1946 cohort) prospectively showed that the diet of participants became healthier in middle-age. However, as the authors noted, it is impossible to differentiate whether this is an ageing or a period effect.28

Our findings suggest that policy actions aimed at promoting awareness of health-enhancing behaviours may have positively influenced the health of successive younger cohorts. Given that the independent relative risks of mortality for cigarette smokers are roughly twice those for obesity, our findings, along with the secular declines in fatal diseases, suggest that cohort life expectancies will likely increase.29 However, this improvement is projected to be offset by higher levels of obesity-related comorbidities and functional disability in later years.30

Strengths and limitations

To the best of our knowledge this is the first UK generational cohort study to include key modifiable behavioural risk factors causally associated with a range of chronic diseases and all-cause mortality; and to describe both absolute and relative inequalities in these risk factors across post-war generational cohorts of working age. We have quantified the scale of inequality, whether these differences are statistically significant, and estimated the disparities in both absolute and relative dimensions within each cohort.

The source data used in this study are of high quality, with face to face interviews conducted by trained interviewers following standardised protocols and measurement devices. However, other than obesity (BMI calculated from measured height and weight), risk factors were self-reported; thus, measurement errors are inevitable. However, the use of repeated measures of these variables
could capture the long-term underlying social change. The continuity of the cross-sectional surveys over 26 years without interruption made it possible to analyse underlying between-cohort change over overlapping ages. Unlike longitudinal studies (of single or mixed-age cohorts), generation-cohort studies are not susceptible to attrition, data collection is less expensive and is more likely to be routinely collected.

A limitation is our focus on cohort change whilst ignoring period effects. The well-known Age-Period-Cohort (APC) identification problem arising from the perfect correlation between these three dimensions, makes it challenging to separate out the distinct contributions of each to the observed change in behavioural risk factor prevalence without employing more advanced statistical methods. Therefore, it is important to acknowledge that the cohort differences presented herein provide only a credible signal of underlying social change which may partly be attributable to period effects. For instance, changes in cigarette smoking within the population could be due, in part, to period measures such as indoor smoking bans (reducing smoking initiation in younger cohorts) or the increased availability of smoking cessation services (leading to a steeper decline with advancing age in successively younger cohorts).

Secondly, despite the large size of the pooled dataset, the scope of our analysis for some risk factors was sometimes restricted by breaks in data collection or changes in definition, adding to the small data problem once the analytic sample was stratified by age, sex, area deprivation and cohort. In addition, the constructed cohorts did not cover the same age ranges. We tried two methods to overcome this: age standardisation and a model-based approach. Using the 1960s cohort as reference (since it overlapped the majority of the entire study age range), we compared the directly age-standardised prevalence rates of the age-truncated cohorts against the corresponding rates for the 1960s cohort (results available on request). The age-adjusted prevalence ratios we obtained from the two methods were found to be very similar, providing further confidence in our opting to report results based on the model-based approach.

In our analysis of inequalities within- and between-cohorts, we only included the main effect of age, and we focused our presentation on the estimates of inequalities at age 40, the midpoint of our pooled dataset. However, it is worth noting that socioeconomic inequalities at other ages may exhibit different patterns. Inequalities could be more pronounced at younger ages due to factors such as the most advantaged groups having lower rates of smoking initiation or a more rapid pace of smoking cessation with age. The choice to present inequalities at age 40 was a practical decision for our analysis, but we acknowledge that exploring variations at different ages could provide additional insights into the dynamics of these inequalities in the working age population.

Conclusions

Our analysis of generational change reveals credible signals of behavioural risk factor changes in levels and in inequalities over successive birth cohorts of working age. In terms of absolute inequality, there is a positive health signal as smoking prevalence and alcohol consumption have reduced, while physical activity has increased across all deprivation groups. However, we observe a concerning negative health signal regarding obesity prevalence, particularly in females, which has increased for all socioeconomic groups across successively younger cohorts. These findings underscore the complex interplay of factors influencing health disparities and behaviours among different socioeconomic groups over time.
Acknowledgments

The authors thank the interviewers and nurses, and participants of the Health Surveys for England 1994-2019. NatCen Social Research, and NHS England are responsible for designing, collecting and administering the survey data and making the data available for public use via the UK Data Service.

We thank Dr Zander Crook (left UCL in 2022) for compiling the pooled survey dataset, 1994-2019.

Contributions

MB conceptualised, designed and supervised the study. SS led on the statistical modelling and advised on behavioural risk factor definitions. MB and SS were jointly responsible for conducting the analyses (including graphical presentation) and interpreting the results. MB drafted the manuscript, with input from SS on statistical methods and risk factor definitions. Both authors have read and approved the manuscript. MB is responsible for the overall content as guarantor.

Orchid:

Madhavi Bajekal https://orcid.org/0000-0003-0807-3629

Shaun Scholes https://orcid.org/0000-0002-7865-5264

Funding

MB is honorary research staff at UCL, funded by Legal & General Assurance Society Limited (L&G) as part of its wider research collaboration with UCL on longevity research. SS received no specific funding for this work. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests

MB has support from Legal and General Assurance Society Limited for the submitted work; SS has no relationship with Legal and General Assurance Society Limited in the submitted work or in the previous 3 years. The authors have no other competing interests.

Data availability

Data are available in a public, open access repository. The HSE datasets generated and analysed during the current study (age banding for participants) are available via the UK Data Service (UKDS: https://ukdataservice.ac.uk/), subject to their end user license agreement.

Ethical approval and participant consent

In the HSE, all adults in selected households are eligible for interview. Relevant committees granted research ethics approval for the survey. Participants gave verbal consent for interview.

Supporting Information

S1. Supplementary Appendix

- Section A. Risk factor variables definitions
- Section B. Sample characteristics and distribution
- Section C. Tables underlying observed cohort prevalence rates for risk factors
- Section D. Tables of predicted prevalence and prevalence ratios by IMD at age 40 (absolute and relative inequalities)
References


### TABLES

**Table 1: Lexis diagram, sample counts age by year (persons)**

<table>
<thead>
<tr>
<th>Year</th>
<th>1940s</th>
<th>1950s</th>
<th>1960s</th>
<th>1970s</th>
<th>1980s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>60</td>
<td>59</td>
<td>58</td>
<td>57</td>
<td>56</td>
</tr>
<tr>
<td>1994</td>
<td>100</td>
<td>98</td>
<td>97</td>
<td>96</td>
<td>95</td>
</tr>
<tr>
<td>1995</td>
<td>89</td>
<td>92</td>
<td>94</td>
<td>96</td>
<td>98</td>
</tr>
<tr>
<td>1996</td>
<td>88</td>
<td>89</td>
<td>90</td>
<td>92</td>
<td>94</td>
</tr>
<tr>
<td>1997</td>
<td>87</td>
<td>88</td>
<td>89</td>
<td>91</td>
<td>93</td>
</tr>
<tr>
<td>1998</td>
<td>86</td>
<td>87</td>
<td>88</td>
<td>89</td>
<td>90</td>
</tr>
<tr>
<td>1999</td>
<td>85</td>
<td>86</td>
<td>87</td>
<td>88</td>
<td>89</td>
</tr>
<tr>
<td>2000</td>
<td>84</td>
<td>85</td>
<td>86</td>
<td>87</td>
<td>88</td>
</tr>
<tr>
<td>2001</td>
<td>83</td>
<td>84</td>
<td>85</td>
<td>86</td>
<td>87</td>
</tr>
<tr>
<td>2002</td>
<td>82</td>
<td>83</td>
<td>84</td>
<td>85</td>
<td>86</td>
</tr>
<tr>
<td>2003</td>
<td>81</td>
<td>82</td>
<td>83</td>
<td>84</td>
<td>85</td>
</tr>
<tr>
<td>2004</td>
<td>80</td>
<td>81</td>
<td>82</td>
<td>83</td>
<td>84</td>
</tr>
<tr>
<td>2005</td>
<td>79</td>
<td>80</td>
<td>81</td>
<td>82</td>
<td>83</td>
</tr>
<tr>
<td>2006</td>
<td>78</td>
<td>79</td>
<td>80</td>
<td>81</td>
<td>82</td>
</tr>
<tr>
<td>2007</td>
<td>77</td>
<td>78</td>
<td>79</td>
<td>80</td>
<td>81</td>
</tr>
<tr>
<td>2008</td>
<td>76</td>
<td>77</td>
<td>78</td>
<td>79</td>
<td>80</td>
</tr>
<tr>
<td>2009</td>
<td>75</td>
<td>76</td>
<td>77</td>
<td>78</td>
<td>79</td>
</tr>
<tr>
<td>2010</td>
<td>74</td>
<td>75</td>
<td>76</td>
<td>77</td>
<td>78</td>
</tr>
<tr>
<td>2011</td>
<td>73</td>
<td>74</td>
<td>75</td>
<td>76</td>
<td>77</td>
</tr>
<tr>
<td>2012</td>
<td>72</td>
<td>73</td>
<td>74</td>
<td>75</td>
<td>76</td>
</tr>
<tr>
<td>2013</td>
<td>71</td>
<td>72</td>
<td>73</td>
<td>74</td>
<td>75</td>
</tr>
<tr>
<td>2014</td>
<td>70</td>
<td>71</td>
<td>72</td>
<td>73</td>
<td>74</td>
</tr>
<tr>
<td>2015</td>
<td>69</td>
<td>70</td>
<td>71</td>
<td>72</td>
<td>73</td>
</tr>
<tr>
<td>2016</td>
<td>68</td>
<td>69</td>
<td>70</td>
<td>71</td>
<td>72</td>
</tr>
<tr>
<td>2017</td>
<td>67</td>
<td>68</td>
<td>69</td>
<td>70</td>
<td>71</td>
</tr>
<tr>
<td>2018</td>
<td>66</td>
<td>67</td>
<td>68</td>
<td>69</td>
<td>70</td>
</tr>
<tr>
<td>2019</td>
<td>65</td>
<td>66</td>
<td>67</td>
<td>68</td>
<td>69</td>
</tr>
</tbody>
</table>

The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC-ND 4.0 International license.
## Table 2. Risk factor prevalence: averages by year

<table>
<thead>
<tr>
<th>Year</th>
<th>Base</th>
<th>Current smokers</th>
<th>Heavy smokers</th>
<th>Usually drink on 5+ days/ per week</th>
<th>Heavy drinkers</th>
<th>Obese (BMI ≥ 30 kg/m²+)</th>
<th>Mean BMI</th>
<th>Physically active</th>
<th>5+ Fruits and Vegetables daily</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>Kg/m²</td>
<td>%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1994</td>
<td>8433</td>
<td>31.2</td>
<td>37.3</td>
<td>16.3</td>
<td>-</td>
<td>14.8</td>
<td>25.8</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1995</td>
<td>8816</td>
<td>31.4</td>
<td>37.9</td>
<td>18.8</td>
<td>-</td>
<td>16.4</td>
<td>26.1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1996</td>
<td>9272</td>
<td>32.2</td>
<td>39.1</td>
<td>20.1</td>
<td>-</td>
<td>17.2</td>
<td>26.2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1997</td>
<td>5063</td>
<td>30.8</td>
<td>37.6</td>
<td>21.1</td>
<td>-</td>
<td>18.7</td>
<td>26.5</td>
<td>31.5</td>
<td>-</td>
</tr>
<tr>
<td>1998</td>
<td>9474</td>
<td>30.4</td>
<td>36.9</td>
<td>22.3</td>
<td>-</td>
<td>19.6</td>
<td>26.6</td>
<td>35.8</td>
<td>-</td>
</tr>
<tr>
<td>1999</td>
<td>4771</td>
<td>30.0</td>
<td>34.9</td>
<td>18.1</td>
<td>-</td>
<td>20.7</td>
<td>26.7</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2000</td>
<td>5049</td>
<td>29.4</td>
<td>35.7</td>
<td>18.0</td>
<td>-</td>
<td>21.7</td>
<td>26.9</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2001</td>
<td>9709</td>
<td>28.5</td>
<td>35.0</td>
<td>19.8</td>
<td>-</td>
<td>23.2</td>
<td>27.1</td>
<td>27.3</td>
<td>-</td>
</tr>
<tr>
<td>2002</td>
<td>4508</td>
<td>29.8</td>
<td>33.6</td>
<td>18.7</td>
<td>-</td>
<td>23.8</td>
<td>27.1</td>
<td>-</td>
<td>25.0</td>
</tr>
<tr>
<td>2003</td>
<td>9183</td>
<td>28.3</td>
<td>34.8</td>
<td>18.8</td>
<td>-</td>
<td>23.7</td>
<td>27.2</td>
<td>33.8</td>
<td>25.7</td>
</tr>
<tr>
<td>2004</td>
<td>4039</td>
<td>26.6</td>
<td>32.5</td>
<td>19.2</td>
<td>-</td>
<td>24.6</td>
<td>27.2</td>
<td>33.6</td>
<td>26.3</td>
</tr>
<tr>
<td>2005</td>
<td>4644</td>
<td>28.3</td>
<td>31.3</td>
<td>19.7</td>
<td>-</td>
<td>25.1</td>
<td>27.3</td>
<td>-</td>
<td>29.4</td>
</tr>
<tr>
<td>2006</td>
<td>8563</td>
<td>25.4</td>
<td>29.2</td>
<td>17.8</td>
<td>-</td>
<td>25.3</td>
<td>27.4</td>
<td>38.0</td>
<td>31.5</td>
</tr>
<tr>
<td>2007</td>
<td>4086</td>
<td>25.2</td>
<td>25.5</td>
<td>15.6</td>
<td>33.5</td>
<td>26.1</td>
<td>27.3</td>
<td>-</td>
<td>30.9</td>
</tr>
<tr>
<td>2008</td>
<td>8945</td>
<td>24.6</td>
<td>27.5</td>
<td>16.3</td>
<td>33.3</td>
<td>26.0</td>
<td>27.4</td>
<td>39.7</td>
<td>28.7</td>
</tr>
<tr>
<td>2009</td>
<td>2663</td>
<td>25.3</td>
<td>28.3</td>
<td>14.2</td>
<td>32.9</td>
<td>25.2</td>
<td>27.4</td>
<td>-</td>
<td>26.6</td>
</tr>
<tr>
<td>2010</td>
<td>4928</td>
<td>22.6</td>
<td>27.0</td>
<td>13.7</td>
<td>32.6</td>
<td>27.9</td>
<td>27.7</td>
<td>-</td>
<td>26.3</td>
</tr>
<tr>
<td>2011</td>
<td>5048</td>
<td>23.6</td>
<td>22.2</td>
<td>12.3</td>
<td>31.7</td>
<td>25.4</td>
<td>27.5</td>
<td>-</td>
<td>27.2</td>
</tr>
<tr>
<td>2012</td>
<td>4761</td>
<td>22.5</td>
<td>22.6</td>
<td>11.9</td>
<td>30.1</td>
<td>25.7</td>
<td>27.5</td>
<td>41.5</td>
<td>-</td>
</tr>
<tr>
<td>2013</td>
<td>5127</td>
<td>23.5</td>
<td>20.6</td>
<td>11.0</td>
<td>31.1</td>
<td>26.8</td>
<td>27.6</td>
<td>-</td>
<td>27.6</td>
</tr>
<tr>
<td>2014</td>
<td>4720</td>
<td>21.2</td>
<td></td>
<td>10.1</td>
<td>28.8</td>
<td>27.5</td>
<td>27.7</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2015</td>
<td>4512</td>
<td>20.2</td>
<td></td>
<td>10.2</td>
<td>29.2</td>
<td>28.4</td>
<td>27.7</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2016</td>
<td>4414</td>
<td>20.4</td>
<td></td>
<td>10.0</td>
<td>29.5</td>
<td>28.9</td>
<td>27.8</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2017</td>
<td>4330</td>
<td>19.6</td>
<td></td>
<td>9.1</td>
<td>28.7</td>
<td>31.5</td>
<td>28.2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2018</td>
<td>4285</td>
<td>19.2</td>
<td></td>
<td>10.6</td>
<td>29.6</td>
<td>31.0</td>
<td>28.1</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2019</td>
<td>4229</td>
<td>16.8</td>
<td></td>
<td>8.4</td>
<td>31.8</td>
<td>32.0</td>
<td>28.2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>26.5</td>
<td>33.0</td>
<td>16.2</td>
<td>31.2</td>
<td>23.4</td>
<td>27.1</td>
<td>36.4</td>
<td>27.8</td>
</tr>
</tbody>
</table>

Notes: 
1. Heavy smokers: selected base of all who were current smokers.
2. Heavy drinkers: selected base of all who had consumed alcohol in the last week.
3. Overall mean for pooled sample.
Table 3: Relative Inequality in prevalence ratios (most versus least deprived quintiles) in behavioural risk factors at age 40, by sex and post-war birth cohorts (with 95% CIs).

<table>
<thead>
<tr>
<th>Cohort</th>
<th>1940s</th>
<th>1950s</th>
<th>1960s</th>
<th>1970s</th>
<th>1980s</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Current smokers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males</td>
<td>2.62</td>
<td>2.79</td>
<td>2.19</td>
<td>1.95</td>
<td>2.00</td>
</tr>
<tr>
<td>95% CI</td>
<td>2.34,2.92</td>
<td>2.57,3.04</td>
<td>2.04,2.35</td>
<td>1.78,2.14</td>
<td>1.70,2.34</td>
</tr>
<tr>
<td>Females</td>
<td>2.35</td>
<td>2.81</td>
<td>2.58</td>
<td>2.38</td>
<td>2.27</td>
</tr>
<tr>
<td>95% CI</td>
<td>2.12,2.60</td>
<td>2.60,3.05</td>
<td>2.40,2.77</td>
<td>2.16,2.63</td>
<td>1.92,2.68</td>
</tr>
<tr>
<td><strong>Heavy smoking (20+ per day)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males</td>
<td>1.23</td>
<td>1.26</td>
<td>1.37</td>
<td>1.31</td>
<td>1.34</td>
</tr>
<tr>
<td>95% CI</td>
<td>1.08,1.39</td>
<td>1.12,1.41</td>
<td>1.21,1.56</td>
<td>1.07,1.61</td>
<td>0.71,2.54</td>
</tr>
<tr>
<td>Females</td>
<td>1.34</td>
<td>1.72</td>
<td>2.01</td>
<td>2.16</td>
<td>2.55</td>
</tr>
<tr>
<td>95% CI</td>
<td>1.15,1.57</td>
<td>1.48,1.99</td>
<td>1.70,2.38</td>
<td>1.62,2.90</td>
<td>1.05,6.20</td>
</tr>
<tr>
<td><strong>Drink alcohol on 5+ days per week</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males</td>
<td>0.75</td>
<td>0.68</td>
<td>0.72</td>
<td>0.70</td>
<td>0.89</td>
</tr>
<tr>
<td>95% CI</td>
<td>0.68,0.83</td>
<td>0.63,0.75</td>
<td>0.65,0.78</td>
<td>0.61,0.81</td>
<td>0.67,1.20</td>
</tr>
<tr>
<td>Females</td>
<td>0.49</td>
<td>0.53</td>
<td>0.55</td>
<td>0.56</td>
<td>0.57</td>
</tr>
<tr>
<td>95% CI</td>
<td>0.42,0.57</td>
<td>0.47,0.60</td>
<td>0.49,0.62</td>
<td>0.46,0.67</td>
<td>0.39,0.84</td>
</tr>
<tr>
<td><strong>Heavy drinking (units drank on heaviest drinking day in the last week)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males (8+ units)</td>
<td>1.10</td>
<td>1.31</td>
<td>1.24</td>
<td>1.16</td>
<td>1.03</td>
</tr>
<tr>
<td>95% CI</td>
<td>0.69,1.75</td>
<td>1.13,1.52</td>
<td>1.11,1.39</td>
<td>1.03,1.31</td>
<td>0.90,1.19</td>
</tr>
<tr>
<td>Females (6+ units)</td>
<td>1.11</td>
<td>1.34</td>
<td>1.42</td>
<td>1.54</td>
<td>1.30</td>
</tr>
<tr>
<td>95% CI</td>
<td>0.59,2.10</td>
<td>1.10,1.62</td>
<td>1.25,1.61</td>
<td>1.35,1.75</td>
<td>1.13,1.50</td>
</tr>
<tr>
<td><strong>Obesity (BMI 30+ kg/m^2)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males</td>
<td>1.29</td>
<td>1.24</td>
<td>1.17</td>
<td>1.17</td>
<td>1.13</td>
</tr>
<tr>
<td>95% CI</td>
<td>1.15,1.45</td>
<td>1.14,1.35</td>
<td>1.08,1.27</td>
<td>1.04,1.30</td>
<td>0.94,1.36</td>
</tr>
<tr>
<td>Females</td>
<td>1.55</td>
<td>1.75</td>
<td>1.87</td>
<td>1.83</td>
<td>1.85</td>
</tr>
<tr>
<td>95% CI</td>
<td>1.40,1.72</td>
<td>1.62,1.90</td>
<td>1.73,2.01</td>
<td>1.66,2.03</td>
<td>1.58,2.16</td>
</tr>
<tr>
<td><strong>Physically active (30+ mins moderate or vigorous, 5+ days per week)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males</td>
<td>0.92</td>
<td>1.01</td>
<td>1.03</td>
<td>1.12</td>
<td>1.02</td>
</tr>
<tr>
<td>95% CI</td>
<td>0.79,1.08</td>
<td>0.90,1.14</td>
<td>0.94,1.13</td>
<td>1.00,1.25</td>
<td>0.84,1.26</td>
</tr>
<tr>
<td>Females</td>
<td>0.89</td>
<td>0.94</td>
<td>0.88</td>
<td>1.05</td>
<td>0.90</td>
</tr>
<tr>
<td>95% CI</td>
<td>0.74,1.06</td>
<td>0.84,1.06</td>
<td>0.80,0.98</td>
<td>0.92,1.19</td>
<td>0.73,1.11</td>
</tr>
<tr>
<td><strong>F&amp;V 5+ portions per day</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males</td>
<td>0.52</td>
<td>0.67</td>
<td>0.81</td>
<td>0.74</td>
<td>0.69</td>
</tr>
<tr>
<td>95% CI</td>
<td>0.43,0.63</td>
<td>0.59,0.77</td>
<td>0.71,0.91</td>
<td>0.65,0.85</td>
<td>0.52,0.92</td>
</tr>
<tr>
<td>Females</td>
<td>0.56</td>
<td>0.56</td>
<td>0.67</td>
<td>0.69</td>
<td>0.78</td>
</tr>
<tr>
<td>95% CI</td>
<td>0.48,0.65</td>
<td>0.50,0.63</td>
<td>0.60,0.74</td>
<td>0.62,0.76</td>
<td>0.63,0.97</td>
</tr>
</tbody>
</table>

Notes: Ratios over 1 indicate prevalence in most deprived areas was greater than in the least deprived areas; ratios under 1 (eg Alcohol 5+ days pw) indicate prevalence lower in the most deprived vs affluent areas. See S1_Tables SD1-SD8 for predicted prevalence ratios by IMD quintiles and model p values for interactions.
FIGURES

Fig 1. Current smoking

Notes: Males left panel, Females right panel.

**Top panel:** 3-year moving averages of observed prevalence (%) by age and cohort.

**Middle panel:** Prevalence ratio (value at top of bar), relative to the 1960s cohort (rebased to equal 0), with 95% confidence intervals (CIs)

**Bottom panel:** Predicted prevalence (%) at age 40, by birth cohort and deprivation quintiles (primary y-axis); prevalence ratio (PR) between most and least deprived quintiles with 95% CIs (secondary y-axis). Unlike top and middle panels, cohorts are plotted from oldest (1940s) to youngest (1980s). PR greater than 1 indicates prevalence is higher in most versus least deprived quintiles; less than indicates prevalence is lower in most versus least deprived quintiles. PR is statistically significant at the 95% level when the CIs do not span 1.
Fig 2. Heavy smoking (current smokers who smoked 20 or more cigarettes per day)

Notes: Males left panel, Females right panel.

**Top panel:** 3-year moving averages of observed prevalence (%) by age and cohort.

**Middle panel:** Prevalence ratio (value at top of bar), relative to the 1960s cohort (rebased to equal 0), with 95% confidence intervals (CIs)

**Bottom panel:** Predicted prevalence (%) at age 40, by birth cohort and deprivation quintiles (primary y-axis); prevalence ratio (PR) between most and least deprived quintiles with 95% CIs (secondary y-axis). Unlike top and middle panel, cohorts are plotted from oldest (1940s) to youngest (1980s). PR greater than 1 indicates prevalence is higher in most versus least deprived quintiles; less than 1 indicates prevalence is lower in most versus least deprived quintiles. PR is statistically significant at the 95% level when the CIs do not span 1.
Fig 3. Drinking alcohol on five or more days per week

Notes: Males left panel, Females right panel.

**Top panel:** 3-year moving averages of observed prevalence (%) by age and cohort.

**Middle panel:** Prevalence ratio (value at top of bar), relative to the 1960s cohort (rebased to equal 0), with 95% confidence intervals (CIs)

**Bottom panel:** Predicted prevalence (%) at age 40, by birth cohort and deprivation quintiles (primary y-axis); prevalence ratio (PR) between most and least deprived quintiles with 95% CIs (secondary y-axis). Unlike top and middle panel, cohorts are plotted from oldest (1940s) to youngest (1980s). PR greater than 1 indicates prevalence is higher in most versus least deprived quintiles; less than 1 indicates prevalence is lower in most versus least deprived quintiles. PR is statistically significant at the 95% level when the CIs do not span 1.
Fig 4. Heavy drinking: Drinkers consuming 8 units or more (males) or 6 units or more (females) on heaviest drinking day in the last week

Notes: Males left panel, Females right panel.

Top panel: 3-year moving averages of observed prevalence (%) by age and cohort.

Middle panel: Prevalence ratio (value at top of bar), relative to the 1960s cohort (rebased to equal 0), with 95% confidence intervals (CIs)

Bottom panel: Predicted prevalence (%) at age 40, by birth cohort and deprivation quintiles (primary y-axis); prevalence ratio (PR) between most and least deprived quintiles with 95% CIs (secondary y-axis). Unlike top and middle panel, cohorts are plotted from oldest (1940s) to youngest (1980s). PR greater than 1 indicates prevalence is higher in most versus least deprived areas; less than 1 indicates prevalence is lower in most versus least deprived areas. PR is statistically significant at the 95% level when the CIs do not span 1.
Fig 5. Obesity: BMI 30kg/m² or above

Notes: Males left panel, Females right panel.

**Top panel:** 3-year moving averages of observed prevalence (%) by age and cohort.

**Middle panel:** Prevalence ratio (value at top of bar), relative to the 1960s cohort (rebased to equal 0), with 95% confidence intervals (CIs)

**Bottom panel:** Predicted prevalence (%) at age 40, by birth cohort and deprivation quintiles (primary y-axis); prevalence ratio (PR) between most and least deprived quintiles with 95% CIs (secondary y-axis). Unlike top and middle panel, cohorts are plotted from oldest (1940s) to youngest (1980s). PR greater than 1 indicates prevalence is higher in most versus least deprived areas; less than 1 indicates prevalence is lower in most versus least deprived areas. PR is statistically significant at the 95% level when the CIs do not span 1.
Fig 6. Mean BMI

Notes: Males left panel, Females right panel.

**Top panel:** 3-year moving averages of mean BMI by age and cohort.

**Middle panel:** Absolute difference in mean BMI (value at top of bar) relative to the 1960s cohort, with 95% confidence intervals (CIs)

**Bottom panel:** Predicted mean BMI at age 40, by birth cohort and deprivation quintiles (primary y-axis); absolute difference in mean BMI between most and least deprived quintiles with 95% CIs (secondary y-axis). Unlike top and middle panel, cohorts are plotted from oldest (1940s) to youngest (1980s).
Fig 7. Physically active (spending 30 minutes or more in moderate- or vigorous-intensity activity on at least five days per week)

Notes: Males left panel, Females right panel.

Top panel: 3-year moving averages of observed prevalence (%) by age and cohort.

Middle panel: Prevalence ratio (value at top of bar), relative to cohort 1960s (rebased to equal 0), with 95% confidence intervals (CIs)

Bottom panel: Predicted prevalence (%) at age 40, by birth cohort and deprivation quintiles (primary y-axis); prevalence ratio (PR) between most and least deprived quintiles with 95% CIs (secondary y-axis). Unlike top and middle panel, cohorts are plotted from oldest (1940s) to youngest (1980s). PR greater than 1 indicates prevalence is higher in most versus least deprived areas; less than 1 indicates prevalence is lower in most versus least deprived areas. PR is statistically significant at the 95% level when the CIs do not span 1.
Fig 8. Consumed five or more portions of fruit and vegetables per day

Notes: Males on left, Females on right.

**Top panel**: 3-year moving averages of prevalence (%) by age and cohort.

**Middle panel**: Prevalence ratio (value at top of bar), relative to cohort 1960s (rebased to equal 0), with 95% confidence intervals (CIs)

**Bottom panel**: Predicted prevalence (%) at age 40, by birth cohort and deprivation quintiles (primary y-axis); prevalence ratio (PR) between most and least deprived quintiles with 95% CIs (secondary y-axis). Unlike top and middle panel, cohorts plotted from oldest (1940s) to youngest (1980s). PR greater than 1 indicates prevalence is higher in most versus least deprived areas; less than 1 indicates prevalence is lower in most versus least deprived areas. PR is statistically significant at the 95% level when the CIs do not span 1.