Supplementary Technical Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: **Quantifying the impact of the Public Health Responsibility Deal on salt intake and cardiovascular disease and gastric cancer: interrupted time series and microsimulation study**

This supplementary technical appendix is an updated version of the one published in Kypridemos C, Guzman-Castillo M, Hyseni L, Hickey GL, Bandosz P, Buchan I, et al. Estimated reductions in cardiovascular and gastric cancer disease burden through salt policies in England: an $IMPACT_{NCD}$ microsimulation study. BMJ Open 2017;7:e013791.

Laverty AA, *et al. J Epidemiol Community Health* 2019; 0:881–887. doi: 10.1136/jech-2018-211749

TABLE OF CONTENTS

Laverty AA, *et al. J Epidemiol Community Health* 2019; 0:881–887. doi: 10.1136/jech-2018-211749

1. HIGH LEVEL DESCRIPTION OF IMPACT$_{NCD}$

IMPACT$_{NCD}$ is a discrete time, dynamic, stochastic microsimulation model.[1,2] Within IMPACT$_{NCD}$ each unit is a synthetic individual and is represented by a record containing a unique identifier and a set of associated attributes.

For this study we considered age, sex, quintile groups of index of multiple deprivation (QIMD)*, salt consumption, body mass index (BMI), systolic blood pressure (SBP), total plasma cholesterol (TC), diabetes mellitus (DM, binary variable)†, smoking status (current/ex/never smoker), pack-years, environmental tobacco exposure (ETS, binary variable), fruit and vegetable (F&V) consumption and physical activity (PA) as the set of associated attributes. A set of stochastic rules are then applied to these individuals, such as the probability of developing coronary heart disease (CHD) or dying, as the simulation advances in discrete annual steps. The output is an estimate of the burden of CHD, stroke, and gastric cancer (GCa) in the synthetic population including both total aggregate change and, more importantly, the distributional nature of the change. This allows, among others, for an investigation of the impact of different scenarios on social equity.

IMPACT$_{NCD}$ is a complex model that simulates the life course of synthetic individuals and consists of two modules: The 'population' module and the 'disease' module. Figure S1 highlights the steps of the algorithm that generates the life course of each synthetic individual. We will fully describe IMPACT$_{NCD}$ by describing the processes in each of these steps in the following chapters. The description is from an epidemiological rather than technical perspective. The source code and all parameter input files are available in https://github.com/ChristK/IMPACTncd/tree/Evaluation_of_UK_responsibility_deal under the GNU GPLv3 licence. Table S7 and Table S8 summarise the sources of the input parameters and the main assumptions and limitations, respectively.

Technical information

IMPACT$_{NCD}$ is being developed in R v3.4.3[4] and is currently deployed in a 40-core workstation with 2TB of RAM running Ubuntu Server v16.04. IMPACT$_{NCD}$ is built around the R package 'data.table'[5], which imports a new heavily optimised data structure in R. Most functions that operate on a data table have been coded in C to improve performance. Each iteration for each scenario is running independently in one of the CPU cores and the R package 'foreach'[6] is responsible for the distribution of the jobs and collection of the results. To ensure the statistical independence of the pseudo-random number

* QIMD is a measure of relative area deprivation based on the 2010 version of the Index of Multiple Deprivation[3]
† We defined as diabetics those with self-reported medically diagnosed diabetes (excluding pregnancy-only diabetes) or glycated haemoglobin (HbA1c) ≥ 6.5

generators running in parallel, the R package 'doRNG'[7] was used to produce independent random streams of numbers, generated by L'Ecuyer's combined multiple-recursive generator.[8]

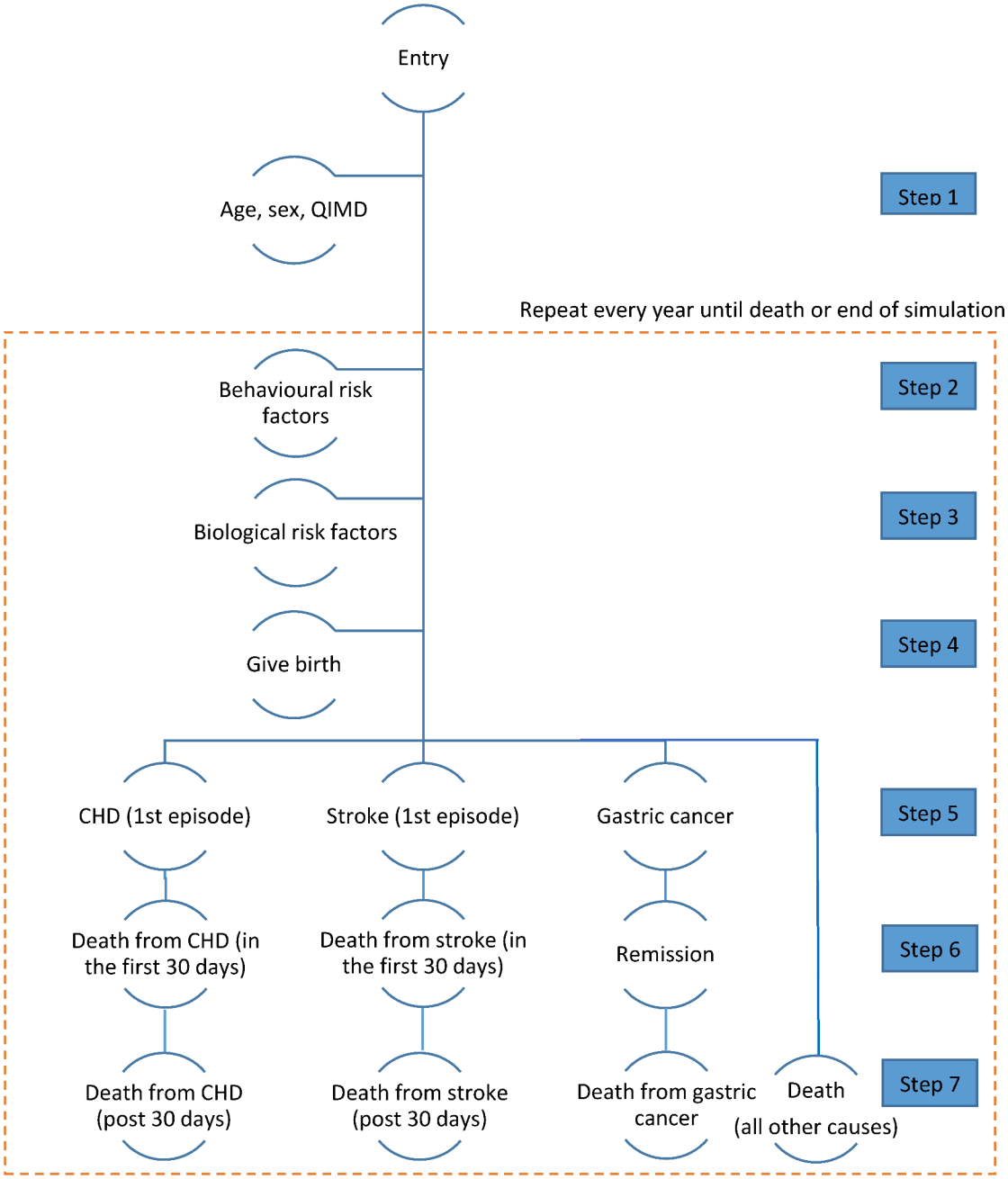

Figure S1 Simplified IMPACT$_{NCD}$ algorithm for individuals. For each step, the algorithm uses information from all appropriate previous steps. CHD denotes coronary heart disease.

2. POPULATION MODULE

The 'population' module consists of steps 1 to 4 in Figure S1. Synthetic individuals enter into the simulation in the initial year (2006 for this study). The number of synthetic individuals that enter into the simulation is user-defined and for this study was set to 500,000. The algorithm ensures that the age, sex and QIMD distribution of the sample are similar to this of the English population in mid-2006. This concludes step 1, which only happens at the beginning of each simulation. Following steps, 2-7 are calculated annually (in simulation time) for each synthetic individual until the simulation horizon is reached, or death occurs.

Estimating exposure to risk factors (steps 2-3)

In steps 2 and 3, IMPACT$_{NCD}$ estimates the exposure of the synthetic individual to the modelled risk factors. It is essential the risk profile of each synthetic individual to be similar to the risk profiles that can be observed in the real English population. For this, we first built a 'close to reality' synthetic population of England from which we sampled the synthetic individuals. Then, we used generalised linear models (GLM) for each modelled risk factor, to simulate individualised risk factor trajectories for all synthetic individuals.

Generating the 'close to reality' synthetic population for IMPACT$_{NCD}$

The 'close to reality' synthetic population ensures that the sample of synthetic individuals for the simulation is drawn from a synthetic population similar to the real one regarding age, sex, socioeconomic circumstance, and risk factors conditional distributions. In our implementation, we used the same statistical framework originally developed by Alfons *et al*.[9] and adapted it to make it compatible with epidemiological principles and frameworks.

In general, this method uses a nationally representative survey of the real population to generate a 'close to reality' synthetic population. Therefore, the method expands the, often small, sample of the survey into a significantly larger synthetic population, while preserves the statistical properties and important correlations of the original survey.

The main advantages over other approaches are: 1) it takes into account the hierarchical structure of the sample design of the original survey, and 2) it can generate trait combinations which were not present in the original survey but are likely to exist in the real population. The second is particularly important because it avoids bias from the excessive repetition of a specific combination of traits present in the original survey that results from multilevel stratification of a relatively small sample. For example, the original survey may have two 35-year-old male participants, one with a BMI of 35 and the other with a BMI of 40 and no other 35-year-old male participants with BMI between 35 and

Laverty AA, *et al. J Epidemiol Community Health* 2019; 0:881–887. doi: 10.1136/jech-2018-211749

40. Unlike other methodologies, the approach proposed by Alfons *et al*. can produce 35-year-old male synthetic individuals with a BMI between 35 and 40. This is possible because the synthetic population is produced by drawing from conditional distributions that were estimated from multinomial models fitted in the original survey data. The detailed statistical methodology and justification can be found elsewhere.[9]

Our approach consists of four stages from which the first is common with the original method by Alfons *et al*.[9] The following stages have been adapted in order to be compatible with the widely accepted 'wider determinants of health' framework.[10] The main notion of this framework is that upstream factors such as the socioeconomic conditions, influence individual behavioural risk factors (e.g. diet, smoking), which in turn, influence individual downstream risk factors such as systolic blood pressure and total cholesterol. The four stages are:

1. Setup of the household structure.
2. Generate the socioeconomic variables.
3. Generate the behavioural variables.
4. Generate the biological variables.

In each stage, information from all previous stages is used. All the variables of the synthetic population for this study were informed by the Health Survey for England 2006 (HSE06).[11] The R language for statistical computing v3.2.0 and the R package 'simPopulation' v0.4.1 were used to implement the method.[4,12]

Stage 1: household structure

The household size and the age and sex of the individuals in each household that have been recorded in HSE06 were used to inform the synthetic population, stratified by Strategic Health Authority (SHA)*.

Stage 2: socioeconomic variables

Once the basic age, sex, household and spatial information of the synthetic population was generated, other socioeconomic information was built up. QIMD for each synthetic individual was generated dependent on the household size and the age and sex of the individuals, stratified by SHA. Then, the equivalised income quintile groups[13] (EQV5) for each household was generated, dependent on five-year age groups and sex, stratified by QIMD. Finally, the employment status of the head of the household (HPNSSEC8) was generated using the National Statistics Socio-Economic Classification[14], dependent on 5-year age groups, sex and EQV5, stratified by QIMD.

* SHAs were 10 large geographic areas, part of the structure of the National Health Service in England before 2013. SHA is the only variable with spatial information in HSE06 and was used as a proxy, to roughly include some spatial information to the synthetic population.

STAGE 3: BEHAVIOURAL VARIABLES

In this stage, behavioural variables such as F&V portions per day, days achieving more than 30 min of moderate or vigorous PA per week, smoking status, exposure to ETS and salt consumption were generated, dependent on 5-year age groups, sex, HPNSSEC8 and EQV5, stratified by QIMD. Moreover, the use of statins and antihypertensive medication (two binary variables) was generated, dependent on 5-year age groups, sex and HPNSSEC8, stratified by QIMD. Other smoking-related variables like cigarettes smoked per day for smokers, years since cessation of ex-smokers and pack-years for ever-smokers were also generated in this step. Specifically for salt consumption, HSE06 contains spot-urine sodium measurements which are less reliable to 24h-urine sodium ones.[15,16] To overcome this limitation, IMPACT$_{NCD}$ adds another processing layer that is described separately (see page 9).

STAGE 4: BIOLOGICAL VARIABLES

The last stage is the generation of the biological variables. Widely accepted causal pathways that have been observed in cohort studies were used to identify associations between biological and behavioural variables. F&V consumption was used as a proxy to a healthy diet. Citations refer to specific evidence regarding the associations. BMI is associated with SBP[17–20], TC[21] and DM[22]. Thus, BMI was the first to be generated in the synthetic population dependent on 5-year age groups, sex, EQV5, F&V consumption[23] and PA[23–25], stratified by QIMD. Then, DM was generated dependent on 5-year age groups, sex, HPNSSEC8 and QIMD, stratified by BMI deciles. The TC was generated dependent on 5-year age groups, sex, deciles of BMI, use of a statin and F&V consumption, stratified by QIMD. Similarly, for the SBP the 5-year age groups, sex, deciles of BMI, smoking status[26,27] and deciles of salt consumption were used as predictors, stratified by QIMD. Socioeconomic variables were used as predictors for both behavioural and biological variables to allow for possible interaction between socioeconomic and behavioural variables.

The outcome of the method was to create a synthetic population of 55 million with similar characteristics to the non-institutionalised population of England in 2006. The synthetic population was validated against the original HSE06 sample (see p26, Synthetic population validation).

IMPACT$_{NCD}$ implementation of individualised risk factor trajectories

IMPACT$_{NCD}$ only applies the previous process for the initial year of the simulation. As the simulation evolves, all variables are recalculated to take into account age and period effects. This feature justifies the classification of IMPACT$_{NCD}$ as a dynamic microsimulation. The process depends on the nature of each variable and the available information, but generally, it uses HSE01 – HSE12[11,28–38] to capture the time trends by age, sex, and QIMD and project them into the future.

Laverty AA, *et al. J Epidemiol Community Health* 2019; 0:881–887. doi: 10.1136/jech-2018-211749

AGE, SEX AND SOCIOECONOMIC VARIABLES

As the simulation progress in annual circles the age of the synthetic individuals in the model increase by one year in each loop. The sex and socioeconomic variables remain stable though. Therefore, social mobility is not simulated in the current version of IMPACT$_{\mathbf{NCD}}$.

SALT

For this study, we assume that all consumed salt is excreted through urine and all the sodium that is excreted in urine comes from the consumed salt. HSE06 measured sodium excretion from spot urine. We used the INTERSALT equation for Northern Europe to estimate daily sodium excretion from spot urine.[20] However, while this method is acceptable to estimate the mean sodium excretion of the population, it tends to overestimate low measurements and underestimate high measurements, when compared to the golden standard of sodium estimation from 24h urine collection.[15,16] Therefore, spot urine sodium estimates are suboptimal for microsimulation because they tend to bias the distribution of salt consumption.

Additionally, sodium excretion from 24h urine collections was estimated in five nationally representative surveys between 2001 and 2014 in the UK.[39,40] The data also included the age and sex of the participants but had no information about the socioeconomic status of the participants. Below, we describe our approach that informs the synthetic population from the 24h urine collection surveys.

Stage 1: We fitted a generalised additive model for location, scale, and shape (GAMLSS) to model the location, scale, and shape of the distribution of sodium excretion conditional on time, age, sex, and the implementation of the Responsibility Deal.[41,42] A left truncated reverse Gumbel distribution was the most appropriate distribution for the GAMLSS model based on the Akaike's Information Criterion. Figure S2 depicts the distribution of salt intake in g/d observed in the surveys and the best GAMLSS model fit. The structure of the models for the parameters of the left truncated reverse Gumbel distribution was the same as in the interrupted time series analysis with the addition of a penalised beta spline for age.

Stage 2: We used individual-level data for spot urine sodium from HSE2006, and we converted the spot urine sodium to estimated 24h sodium, using the INTERSALT equation for Northern Europe.[20] Instead of using fixed coefficients for the INTERSALT equation, for each HSE participant, different coefficients were sampled from the normal distributions with mean equal to the coefficient and standard deviation (sd) equal to the standard error (S.E.) of the respective coefficient. For instance, the reported INTERSALT age coefficient for men is 0.26 (S.E. = 0.78); therefore, for each use of the INTERSALT equation in this stage we draw a new age coefficient for men from a normal distribution

Laverty AA, *et al. J Epidemiol Community Health* 2019; 0:881–887. doi: 10.1136/jech-2018-211749

with mean = 0.26 and sd = 0.78. Finally, we converted 24h sodium (in mEq/day) to salt (g/day) using the formula 1 mEq of sodium/day = 58.5 * 10^{-3} g of salt/day.

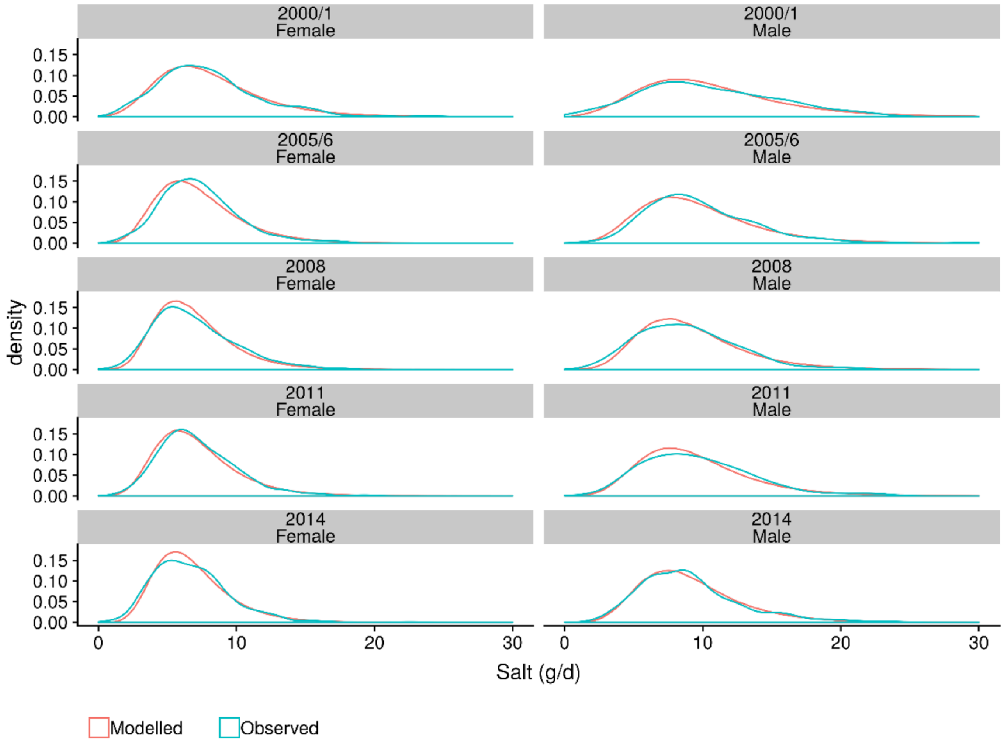

Figure S2 Modelled using GAMLSS vs. observed salt distribution by sex and survey year.

Stage 3: The rank of estimated salt for each synthetic individual is calculated by age and sex. Then, the estimated salt consumption values for each synthetic individual is calculated using the GAMLSS model conditional on the age, sex, simulation year, the rank of the synthetic individual, and the implementation or not of the Responsibility Deal. Figure S3 shows the projection of the salt distribution in the population under two scenarios; 1) with the Responsibility Deal implemented since 2010, and 2) assuming that the Responsibility Deal was never implemented, and FSA continued leading the national salt reduction strategy.

The main advantage of this approach is that uses all the available information from the 24h urine sodium surveys, while enhances it with information regarding socioeconomic gradients and correlation with other risk factors and especially SBP, from spot urine measurements. This method assumes that the socioeconomic gradient in salt did not change over time. This assumption is supported by empirical evidence[43,44] and has further confirmed in a previous study using $IMPACT_{NCD}$.[45]

Furthermore, it assumes that the transition from the FSA led salt reduction strategy to the Responsibility Deal did not alter the equity impact of the policy. Finally, for this approach, we assume that mean salt intake continues to decrease by age for ages up to 84 (extrapolating the observed trend in the 24-hour sodium surveys), based on the spot urine measurements in HSE (Figure S4).

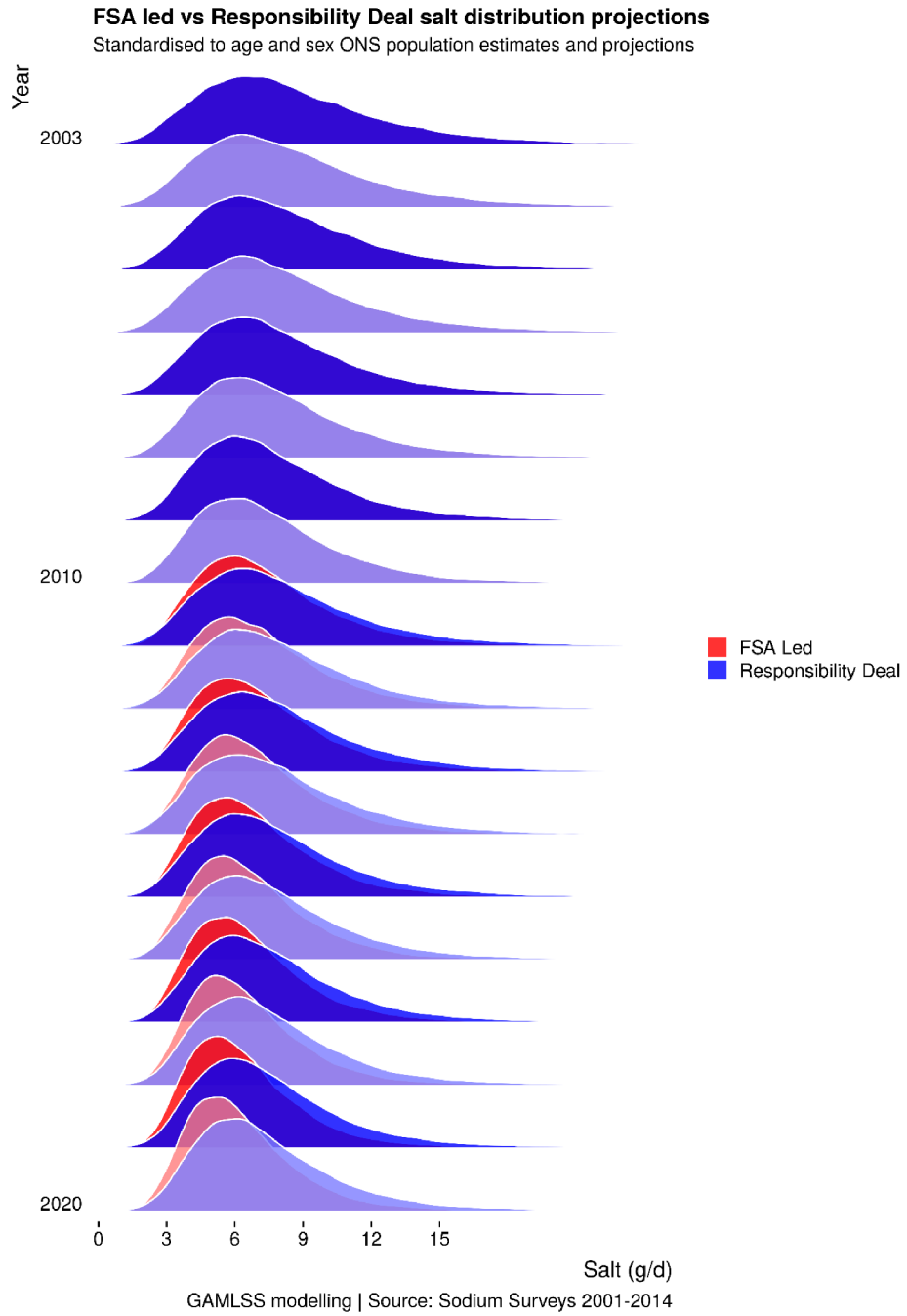

Figure S3 Projections of salt distributions with or without the implementation of the Responsibility Deal.

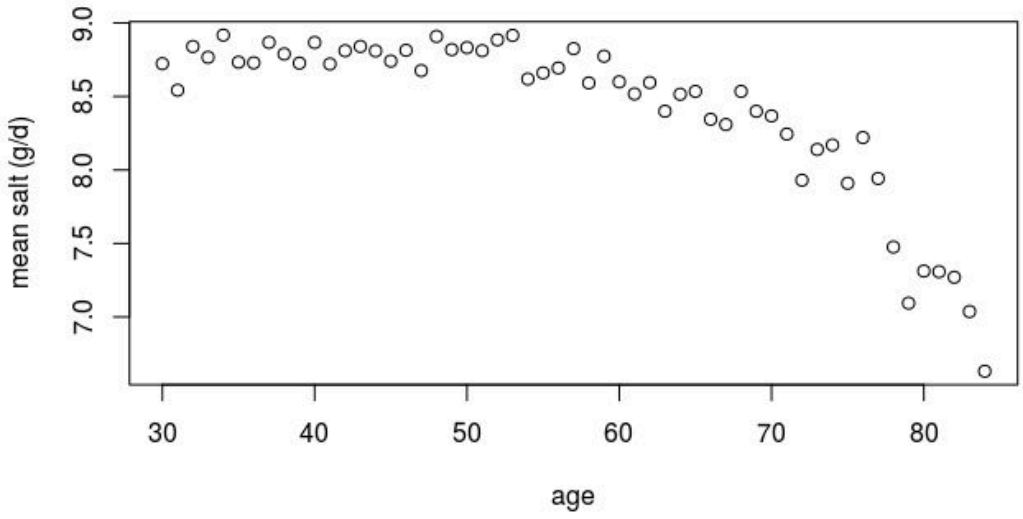

Figure S4 mean salt consumption by age in Health Survey for England. Estimated from spot urine sodium using the INTERSALT equation for Northern Europe.[20] It is evident that the decreasing trend in salt consumption continuous for ages over 64.

Fruit & veg consumption and physical activity

Both F&V consumption (portions/day) and PA (days with more than 30 min of moderate or vigorous activity/week) were modelled as ordinal factor variables. A proportional odds logistic regression model was fitted in the HSE01, HSE02, HSE04-11 individual-level data with F&V consumption as the dependent variable and year, 2nd degree polynomial of age, sex, QIMD and their 1st order interactions. Similarly, for PA a similar model was fitted in the HSE06, HSE08 and HSE12 data. These models were used for individual-level predictions about the synthetic individuals as the simulation was evolving.

Smoking

The 'close to reality' synthetic population is an accurate snapshot of active, ex-, and never smokers in 2006, as it was observed in HSE06. Then IMPACT$_{NCD}$ uses transitional probabilities for smoking initiation, smoking cessation and relapse, to generate and record smoking histories of the synthetic individuals. For smoking initiation and cessation probabilities, logistic regression models were fitted in HSE data with age, sex, and QIMD as the independent variables. A similar approach was followed for relapse probabilities with years since cessation, sex and QIMD as the independent variables.

Environmental tobacco smoking

For ETS we assumed a linear relation between smoking prevalence and ETS, stratified by QIMD. We assumed no intercept; when smoking prevalence reaches 0, ETS prevalence will be 0 too.

CONTINUOUS BIOLOGICAL VARIABLES

In IMPACT**NCD**, the value of each continuous biological risk factor (BMI, SBP, and TC) is calculated in a two-stage process for each synthetic individual and each projected year. The first stage simulates ageing effects, while the second stage simulates period effects. We follow this approach mainly for two reasons. Firstly, to simulate the physiological mechanisms of ageing. For example, the change of lipid profile in post-menopausal women, or the increase of SBP due to age-related stiffening of the arteries. Secondly, because the variance of the risk factor distributions increases with age, and we wanted to model this. Below we describe the stages:

Stage 1: Instead of tracking the actual biological risk factor values for the synthetic individuals, we track the percentile ranks* of the values by age, sex and QIMD. These percentile ranks remain fixed for each synthetic individual throughout the simulation. In each simulated year, the percentile ranks are converted back to actual risk factor values, by matching the percentile ranks of a sample of the initial synthetic population of same age group, sex, and QIMD.

For example, in 2006 a 20-year-old male synthetic individual living in a QIMD 3 area with SBP of 120 mmHg has a SBP percentile rank of 0.52. Fifty years later, the same synthetic individual has retained his percentile score for SBP. However, his SBP is now calculated to 137.6 mmHg in order to match the SBP of a 70-year old man living in a QIMD 3 area in 2006 with the same percentile rank of 0.52. Figure S5 illustrates the previous example. Despite, individuals retain their percentile for the respective risk factor throughout the simulation (vertical position in Figure S5), this stage remains stochastic because each time this stage is implemented a different sample from the synthetic population is drawn. Finally, the distance from the mean for each risk factor is calculated stratified by 5-year age group, sex, and QIMD. For instance, if a synthetic individual has SBP of 140 mmHg and the mean SBP in the respective group of same age group, sex and QIMD is 130 mmHg, the distance from the mean is 140 – 130 = 10 mmHg.

Stage 2: Similarly to the approach followed for other variables, we fitted regression models to the HSE01-12 data. For BMI, year, age, sex, QIMD and PA were the independent variables. For SBP, year, age, sex, QIMD, smoking status, BMI, and PA were the independent variables. Finally, for TC year, age,

* For the percentile rank the formula $R_{percentile} = (R-1)/(n-1)$ is used, where $R_{percentile}$ is the percentile rank and $R = (R_1, \dots, R_n)$ is the rank vector constructed from a random observation vector $(X_1, \dots, X_n)$. In IMPACT**NCD** specifically, vector X is constructed from the subset of the respective continuous risk factor values, by 5-year age group, sex and QIMD, for each year of the simulation.

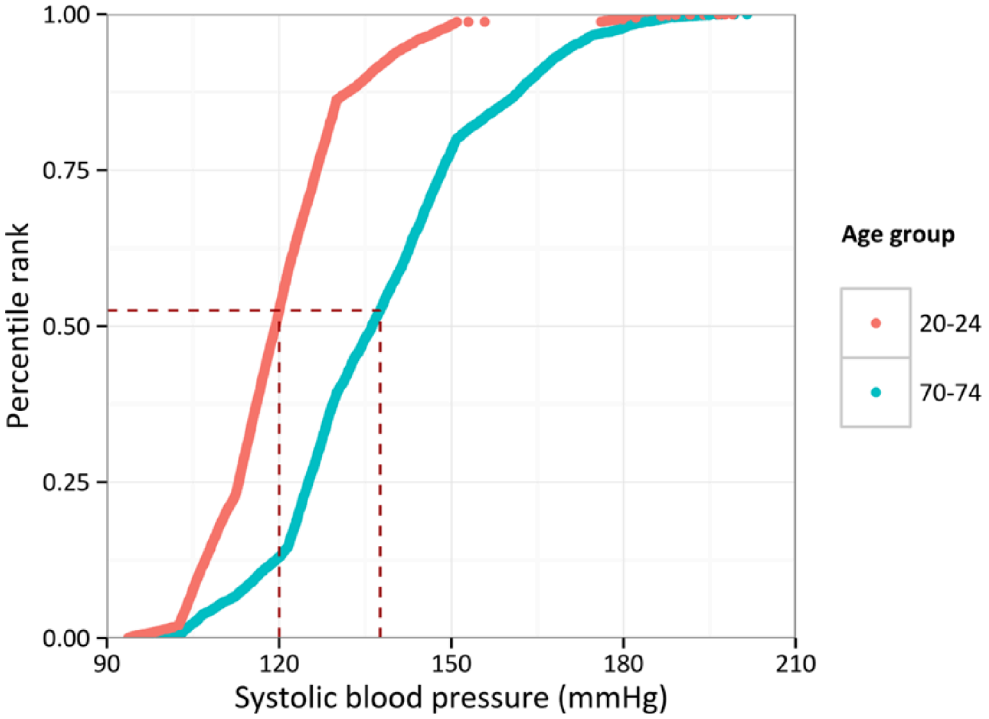

Figure S5 Plot of the percentile rank against the systolic blood pressure of male synthetic individuals living in QIMD 3 area for age groups 20-24 and 70-74.

sex, QIMD, BMI, F&V consumption and PA were the independent variables.* These models are used to predict the mean of the relevant group. These predicted means are added then, to the distances calculated in the previous stage. The result is the final value of the relevant risk factor that will be used for risk estimation. Please note that this approach may produce a healthier population especially among older ages, compared to those predicted by the regression models alone.

DIABETES MELLITUS

As with smoking, the 'close to reality' synthetic population is an accurate snapshot of diagnosed and non-diagnosed diabetics in 2006, as it was observed in HSE06. We assumed DM is an incurable chronic condition. IMPACT$_{NCD}$ uses the validated for English population Qdiabetes algorithm (ex QDscore) to calculate annual transitional probabilities of non-diabetic synthetic individuals to develop DM.[46]

Lag times

All the function that has been described above for risk factor trajectories include time and age (in years) as one of the independent variables. Therefore, lag times can be potentially considered on a

* As before, the independent variables for each risk factor were selected based on known associations from longitudinal studies. Therefore, only the magnitude of the association is informed by cross-sectional data and possibly attenuated due to reverse causality.

per risk factor basis. For instance, let us consider a 50-year-old synthetic individual in 2010 and an assumed lag time of 5 years for F&V. When IMPACT$_{NCD}$ calculates the probabilities for F&V consumption of this individual, it will use time – (lag time) = 2010 – 5 = 2005 and age – (lag time) = 50 – 5 = 45. So, when the 'disease' module of IMPACT$_{NCD}$, uses the risk exposure to F&V to estimate a disease incidence transitional probability, the lag-timed exposure will be used.

In this study, we assumed that the mean lag time between exposure and CVD is 5 years.[47–49] Similarly, the mean lag time between exposure and GCa is 8 years, except for the cumulative risk of smoking (smoking duration) which was set to follow CVD lag time. Mean lag times were roughly informed from risk reversibility trials, when available, or the median observation times of the cohort studies we used to inform the risk magnitude for each risk factor. Then for each iteration, we draw lag time values from binomial distributions with the respective means.

Birth engine (Step 4)

The Office for National Statistics (ONS) principal-assumption fertility projections for England are used to estimate the number of new synthetic individuals entering the model through birth, in every simulated year.[50] The birth engine only becomes relevant for simulations featuring a horizon of more than 30 years and its importance increases as the simulation progress further in time. The 'new-born' synthetic individuals inherit the socioeconomic position of their mother and their quantile ranks for the continuous biological risk factors from a random synthetic individual.

3. DISEASE MODULE

The disease module contains the last 3 steps of the model (Figure S1). The risk (probability) for each synthetic individual aged 30 – 84, to develop each of the modelled diseases is estimated in step 5 conditional on the exposure to relevant risk factors. The step ends by selecting synthetic individuals to develop the modelled diseases. Finally, in steps 6 and 7 the risk of dying from one of the modelled diseases or any other cause is estimated and applied. Steps 2 to 7 are then repeated for the surviving individuals until the simulation horizon is reached.

Estimating the annual individualised disease risk and incidence (Step 5)

In order to estimate the individualised annual probability of a synthetic individual to develop a specific disease conditional on his/her relevant risk exposures we follow a 3-stage approach:

1. The proportion of incidence attributable to each modelled risk factor by age group and sex is estimated, assuming a specific time lag.
2. Assuming multiplicative risks, the portion of the disease incidence attributable to all the modelled risk factors is estimated and subtracted from the total incidence.
3. For each individual in the synthetic population, the probability of developing the disease is estimated and then is used in an independent Bernoulli trial to select those who finally develop the disease.

Next, the implementation of the above method is described in more detail using CHD as an example. The same process is used for all modelled diseases.

Stage 1

The population attributable risk (PAF) is an epidemiological measure that estimates the proportion of the disease attributable to an associated risk factor.[51] It depends on the relative risk associated with the risk factor and the prevalence of the risk factor in the population. In a microsimulation context where exposure to risk factors are known to the individual level and assuming multiplicative risk factors PAF can be calculated with the formula:

$$PAF = 1 - \frac{n}{\sum_{i=1}^{n}(RR_1 * RR_2 * \ldots * RR_k)} \quad ,$$

where n is the number of synthetic individuals in the population, and $RR_{1\ldots k}$ is the relative risks of the risk factors associated with CHD. We calculated PAF based on above formula stratified by age and sex. Consistent with findings from the respective meta-analyses that were used for IMPACT$_{\mathbf{NCD}}$ (Table S7), SBP below 115 mmHg, TC below 3.8 mmol/l and BMI below 20 Kg/m^2 were considered to have a relative risk of 1. Similarly, consumption of eight or more portions of F&V and five or more days with

Laverty AA, *et al. J Epidemiol Community Health* 2019; 0:881–887. doi: 10.1136/jech-2018-211749

more than 30 minutes of moderate to vigorous activity per week were also considered to have a relative risk of 1. All the relative risks were taken from published meta-analyses and cohort studies (Table S7).

Stage 2

The incidence of CHD not attributable to the modelled risk factors can be estimated by the formula:

$$I_{Theoretical\ minimum} = I_{Observed} * (1 - PAF)$$

Where $I_{Observed}$ is the CHD incidence and PAF is from Step 1. $I_{Theoretical\ minimum}$ represents CHD incidence if all the modelled risk factors were at optimal levels. The theoretical minimum incidence is calculated by age and sex only in the initial year of the simulation and it is assumed stable thereafter.

Stage 3

Assuming that $I_{Theoretical\ minimum}$ is the baseline annual probability of a synthetic individual to develop CHD for a given age and sex due to risk factors not included in the model (i.e. genetics etc.), the individualised annual probability to develop CHD, $\mathbb{P}(\text{CHD} \mid \text{age, sex, exposures})$, given his/her risk factors were estimated by the formula:

$$\mathbb{P}(CHD \mid age, sex, exposures) = I_{Theoretical\ minimum} * RR_1 * RR_2 * RR_3 * \ldots * RR_k$$

Where $RR_{1 \ldots k}$ the relative risks that are related to the specific risk exposures of the synthetic individual, same as in stage 1. Depending on data availability this method can be further stratified by QIMD; however, data were not available for this in the current study.

The above method can be used only when the incidence of the disease in the population is known. For cancers, this information is available from the cancer registries. The true incidence of CHD (and stroke) though, is largely unknown. Several estimates exist nonetheless all have limitations. Therefore, for the estimation of CHD incidence by age and sex we opted for a modelling solution to synthesise all the available sources of information and minimise bias. Specifically, we used ONS CHD mortality (ICD10 I20-I25) for England in 2006,[52] self-reported prevalence of CHD from HSE06, the incidence of angina from primary care data[53] and incidence of acute myocardial infarction (AMI) from mortality and hospital statistics[54] to inform the World Health Organisation (WHO) DISMOD II model.[55] DISMOD II is a multi-state life table model that can estimate the incidence, prevalence, mortality, fatality and remission of a disease when information about at least three of these indicators is available. A similar approach has been followed by the Global Burden of Disease team and others.[56,57] We considered CHD an incurable chronic disease (i.e. remission rate was set to 0); therefore, the derived DISMOD II incidence refers to the first ever manifestation of angina or AMI excluding any recurrent episodes. For the DISMOD II calculations, we assumed that incidence and case-fatality had been declining by 3%

Laverty AA, *et al. J Epidemiol Community Health* 2019; 0:881–887. doi: 10.1136/jech-2018-211749

(relative), over the last 20 years. The derived CHD incidence, prevalence and fatality were used as an input for IMPACT$_{NCD}$. A similar approach was used for stroke.

For the initial year of the simulation, some synthetic individuals need to be allocated as prevalent cases for each of the modelled diseases. DISMOD II model[55] is used again to estimate the number of prevalent cases of the disease by age and sex. Then, the estimated number of prevalent cases are sampled independently from the individuals in the population with weights proportional to their relevant exposures.

Simulating disease histories (Step 6)

In the current stage of development, IMPACT$_{NCD}$ does not contain a detailed disease history module. However, Step 6 is used to simulate significant aspects of the disease. For CVD, this was used to simulate the observable spike of short-term (30 days) mortality after the first event of AMI or stroke. Data about short-term mortality were used from the 'Coronary heart disease statistics 2012 edition' report.[53]

For GCA this step is used to simulate remission cases. Once more, we used the DISMOD II model to estimate the remission rate by age and sex, using as inputs incidence, mortality, and case fatality rates by age group and sex. Specifically, the incidence and survival rates of GCa is known through the cancer registries and is reported by ONS.[58,59] From the reported first and fifth-year survival rate, assuming a Weibull survival distribution, we calculated annual case fatality and 10-year survival rate. Finally, we used the observed GCa mortality reported by ONS.[52] We assumed remission rate equals the 10-year survival rate. Furthermore, we assumed the incidence and case-fatality rate had been declining by 2% (relative) over the last 20 years, and the remission rate had been improving by 1% (relative).

Simulating mortality (Step 7)

All synthetic individuals are exposed to the risk of dying from any of their acquired modelled diseases or any other non-modelled cause. However, the algorithm behaves differently depending on the age and life course trajectory of the synthetic individual.

For ages 0 to 29, we used all-cause mortality rate by age, sex, and QIMD to inform an independent Bernoulli trial and select synthetic individuals that die every year. For years 2006 to 2013 we used the observed mortality rates as were reported from ONS.[52] For years after 2013, functional demographic models by sex and QIMD were fitted to the ONS reported annual mortality rates, from years 2002 to 2013, and then were projected to the simulation horizon using the R package 'demography'.[60] Functional demographic models are generalisations of the Lee-Carter demographic model, influenced by ideas from functional data analysis and non-parametric smoothing.[61]

The same approach as above was followed for synthetic individuals aged 85 to 100. We considered a mortality rate of 1 for all synthetic individuals reaching the age of 100. Hence, IMPACT$_{\mathbf{NCD}}$ maximum synthetic individual age is 100 years.

Finally, for synthetic individuals with ages between 30 and 84 the all-cause mortality was decomposed into modelled-diseases specific mortality and any-other cause mortality. The former applies only to the prevalent cases of each modelled disease in the synthetic population. For this, case-fatality rates by age and sex are estimated by DISMOD II for each modelled disease, as described before, and then are used in a Bernoulli trial to select prevalent cases that die from the disease in a year. For CVD, 30-day mortality of incident cases is calculated in step 6. In this step the algorithm avoids double counting of CVD deaths by allocating any additional CVD deaths from the pool of prevalent cases. For example, if we expect 100 deaths from CHD (based on annual case-fatality rates from DISMOD II) overall, and 40 of them to occur in first 30 days after CHD (from published 30-day case-fatality rates), the algorithm randomly selects 40 individuals from the incident cases of CHD and 60 from the prevalent CHD cases.

For the any-other cause mortality, a process similar to the one described for ages 0 to 29 and 85 to 100. However, this time CVD and GCa specific mortality are removed from the observed mortality and mortality projections to avoid double counting.

The case mortality and fatality rates are further parametrised and individualised based on established epidemiological evidence. The 'male British doctors' and DECODE studies have shown that smokers and diabetics had increased overall mortality even when CVD is excluded[62,63]. IMPACT$_{\mathbf{NCD}}$ adjusts for that by inflating the any-other cause mortality rate for smokers and diabetics and deflating it for non-smokers and non-diabetics, while it constrains the sum to remain the same as before the adjustments. Furthermore, we assumed that CVD and GCa case-fatality is improving by 3% and 2% annually, respectively, and that there is a constant case-fatality socioeconomic gradient of approximately 5% by QIMD level (halved for ages over 70) for CHD and GCa, and 2% for stroke. The socioeconomic gradient forces the more deprived to experience worse disease outcomes. These assumptions are based on empirical evidence.[53,64–66]

Finally, synthetic individuals who remain alive after this step progress to the next year and start again from step 1, unless the simulation horizon has been reached.

4. SCENARIOS

The process that we describe so far is the one for the 'current policy' scenario. For this study, for the current policy scenario, we assumed that the Responsibility Deal had been implemented since 2011. We also assumed, driven by the data, that the implementation has decelerated the rate of decrease of salt consumption in the population.

For our counterfactual scenario, we assumed that the Responsibility Deal has never been implemented. Therefore, we assumed that the observed declining trend in salt consumption before 2011, will also continue after 2011.

One-way sensitivity analysis

In our one-way sensitivity analysis, we assumed that the declining trends in mean salt consumption follow a logarithmic rather than linear decline. This results in slower declines for both scenarios. We calibrated the declining model to reach a mean salt consumption of 7g/d by 2020 (vs 6.5 g/d in our main analysis) and 6.6 g/day (vs 5.8 g/day) by 2025, for ages 19 to 64. Therefore, our one-way sensitivity analysis provides more conservative estimates. Figure S6 shows the modelled trends in mean salt consumption under the two scenarios, for the main and one-way sensitivity analysis. $IMPACT_{NCD}$ estimates are not directly comparable to the National Sodium Survey studies because the former is for ages 30 to 84, while the latter for ages 19 to 64. Table S1, Table S2, and Table S3 summarise the results of the one-way sensitivity analysis.

Laverty AA, *et al. J Epidemiol Community Health* 2019; 0:881–887. doi: 10.1136/jech-2018-211749

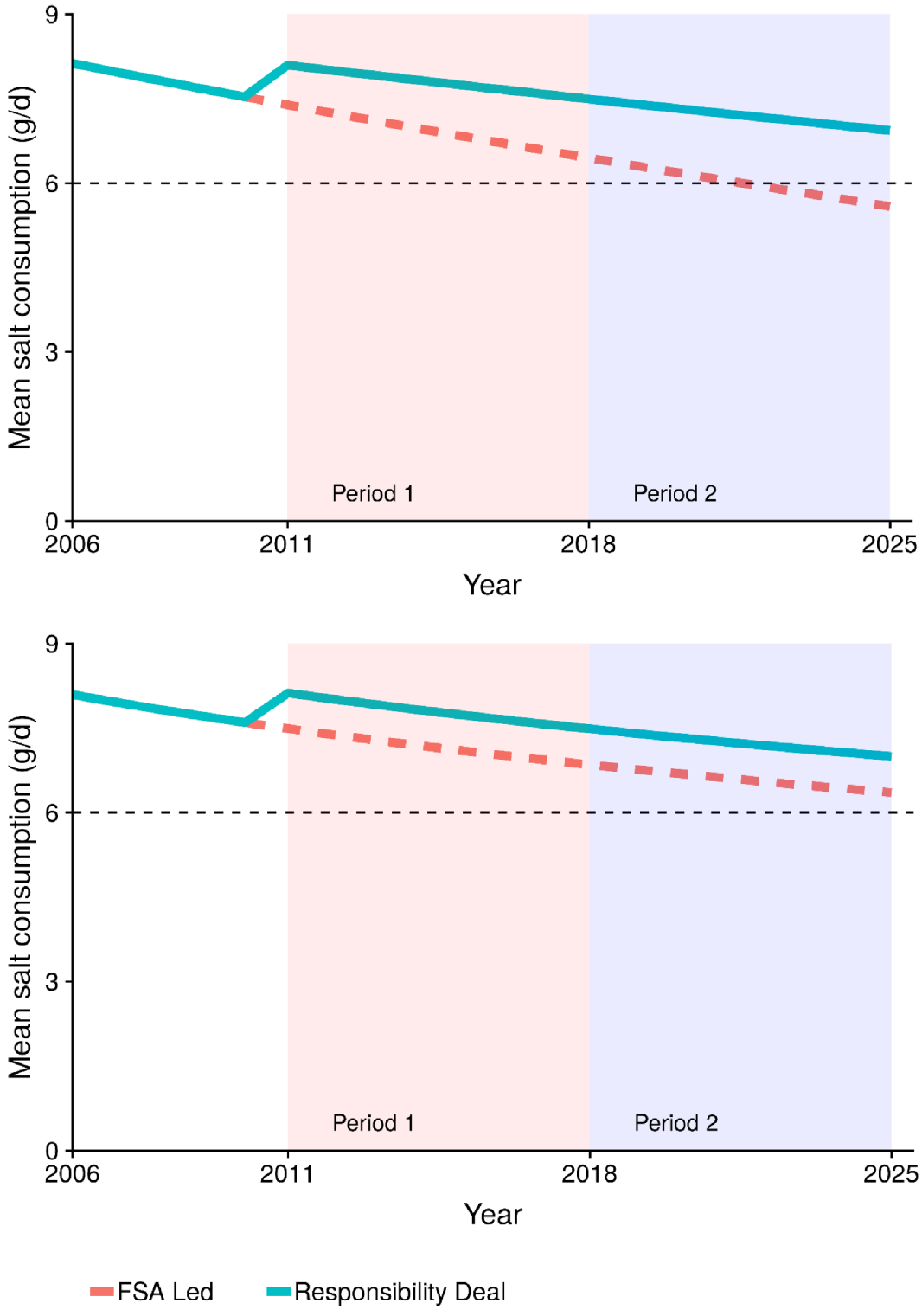

Figure S6 Mean salt consumption in the synthetic population (ages 30 to 84) under the two simulated scenarios. The top graph corresponds to the main analysis and the bottom graph to the on-way sensitivity analysis

5. UNCERTAINTY

IMPACT$_{NCD}$ implements a 2nd order Monte Carlo approach to estimate uncertainty intervals (UI) for each scenario.[67,68] Each simulation runs 1000 times. For each iteration, a different set of input parameters is used, by sampling from the respective distributions* of input parameters (Table S9), and a different sample of the synthetic population is drawn. However, the scenarios are 'paired'. For instance, the *n*th iteration of all scenarios runs with the same set of input parameters and on the same synthetic population sample for all of them. This explains why the uncertainty of in-between scenarios comparisons is significantly smaller than the uncertainty of isolated scenarios.

The framework allows stochastic uncertainty, parameter uncertainty and individual heterogeneity to be reflected in the reported UI. The following example illustrates the different types of uncertainty that were considered in IMPACT$_{NCD}$. Let us assume that the annual risk for CHD is 5%. If we apply this risk to all individuals and randomly draw from a Bernoulli distribution with p = 5% to select those who will manifest CHD, we only consider stochastic uncertainty. If we allow the annual risk for CHD to be conditional on individual characteristics (i.e. age, sex, exposure to risk factors), then individual heterogeneity is considered. Finally, when the uncertainty of the relative risks due to sampling errors is considered in the estimation of the annual risk for CHD, the parameter uncertainty is considered. From these three types of uncertainty, only the parameter uncertainty can be reduced from better studies in the future.

Due to lack of information and for computational efficiency, not all three types of uncertainty are considered in every step (Figure S1) of IMPACT$_{NCD}$. Specifically, stochastic uncertainty is included in every step, individual heterogeneity in every step except 1 and 4 and parameter uncertainty in step 5. Of course, parameter uncertainty (if any) of scenario targets are also estimated in steps 2 and 3. For example, the target of the 'Feasible' scenario is mean salt consumption of 6g/day and its uncertainty assumed to follow a PERT distribution with min = 5.8 g/day, mode = 6 g/day, and max = 7 g/day

The structure of the model is grounded in fundamental epidemiological ideas and well-established causal pathways; therefore, we considered this type of uncertainty relatively small and did not study it. However, mortality from each of the modelled diseases and any-other cause (steps 6 and 7) is calculated serially, one modelled disease at a time. To avoid bias that this approach might introduce, the order of the modelled diseases in each mortality estimation is randomised.

* We assumed log-normal distributions for relative risks and hazard ratios, normal distributions for coefficients of regression equations, and PERT distributions for other parameters. Specifically for relative risks and hazard ratios, the distributions were bounded above 1 when the mean was above 1 and vice versa.

Laverty AA, *et al. J Epidemiol Community Health* 2019; 0:881–887. doi: 10.1136/jech-2018-211749

6. HEALTH-RELATED COSTS

Healthcare costs of CVD and GCa

CHD and stroke healthcare costs were drawn from economic modelling carried out for the National Institute for Health and Care Excellence (NICE), which is generally based on the best economic evidence that is available at the time in England. We considered separate costs for year one (first year after being diagnosed), subsequent years, and fatal CVD events which reflected higher costs in the final year of life. We did not include costs for non-CVD deaths and disease. Stroke costs are from an NHS perspective and include rehabilitation but not ongoing social care costs.

No UK studies were found with specific estimates of gastric or upper GI cancer healthcare costs. Healthcare costs were estimated using the NHS Programme Budgeting Data for England[69] for 2006/07 for Upper GI cancers, combined with 5 year prevalence data for 31st December 2006. The total upper GI prevalence data excluded duodenal, gallbladder, ampulla Vater, and some biliary tract cancers. Overall this gave an average cost per year of £13,396 in 2018 prices which were applied to GCa prevalence in the model.

Costs were weighted for deprivation (Table S4) as there is good evidence that costs for the same disease show a social gradient. The weighting for deprivation was based on data from Charlton et al. who found that average disease costs vary by QIMD.[70]

The disease costs we have are averages, and it is assumed that disease costs are the same for all age and sex groups (Table S5). The cost of CHD deaths is based on costs of myocardial infarction deaths. The cost of multimorbidity is assumed to be the sum of costs of individual diseases. We inflated all costs to 2018 using UK Treasury GDP inflator tables from April 2018. We did not use the PSSRU hospital & community health services index because it only goes back to 2004 and some of the costs predate this. Table S5 shows the disease costs, the 2018 costs shown are for IMD quintile 3 (middle) quintile.

Productivity losses through CVD and GCa

There are different definitions and ways to calculate productivity losses - sometimes costs include beyond the workplace, i.e. household production, sometimes just the workplace. Sometimes employer perspective (in which case use 'friction costs' - costs of replacing someone - usually 90 days wages) or employee perspective (would include all lost wages), or both employee and employer.

Workplace productivity losses for CHD mortality and morbidity were estimated using data from Liu et al.[71] which included estimates of friction-adjusted employment productivity losses based on working years lost through early mortality, and certified incapacity days which was combined with under 65 prevalence data to get a unit cost. Productivity losses for stroke were estimated using data from Saka

Laverty AA, *et al. J Epidemiol Community Health* 2019; 0:881–887. doi: 10.1136/jech-2018-211749

et al.[72] which included income lost due to mortality and morbidity which was combined with the prevalence of stroke in people aged under 65 to get a unit cost. These estimates were inflated to 2018 prices using the ratio of average UK weekly earnings data from ONS.

No UK estimates were found for productivity losses from GCa morbidity or mortality. Workplace productivity losses from GCa mortality were derived from a study in Ireland by Pearce et al.[73] This paper also included estimates of GCa mortality related household productivity losses, but these were excluded to be more consistent with the CHD and stroke estimates. These estimates were adjusted for the ratio of average full-time earnings from Ireland (in euros) in 2011* and England (in GBP) in 2011† which were matched using EPPI cost converter using OECD purchasing power parity (PPP) values and then inflated to 2018 using average UK weekly earnings data from ONS.‡ This gave an average lost paid production of £48,124 per death from GCa.

Other costs

Other costs such as household productivity, informal care or out of pocket expenses, or future consumption and production, were not included.

Discounting

An annual discount rate of 3.5% was applied from 2018. Results from before 2018 were inversely discounted. This rate was selected based on guidance from the UK Treasury.

*https://www.cso.ie/en/releasesandpublications/ep/p-syi/psyi2017/econ/earn/

†https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/earningsandworkinghours/bulletins/annualsurveyofhoursandearnings/2017provisionaland2016revisedresults#average-earnings

‡ https://eppi.ioe.ac.uk/costconversion/

7. VALIDATION

Interrupted time series

Figure S7 depicts the residuals from our interrupted time series model over time for both men and women. There is adequate model fit and reassuringly the fit is the same both before and after the implementation of the RD.

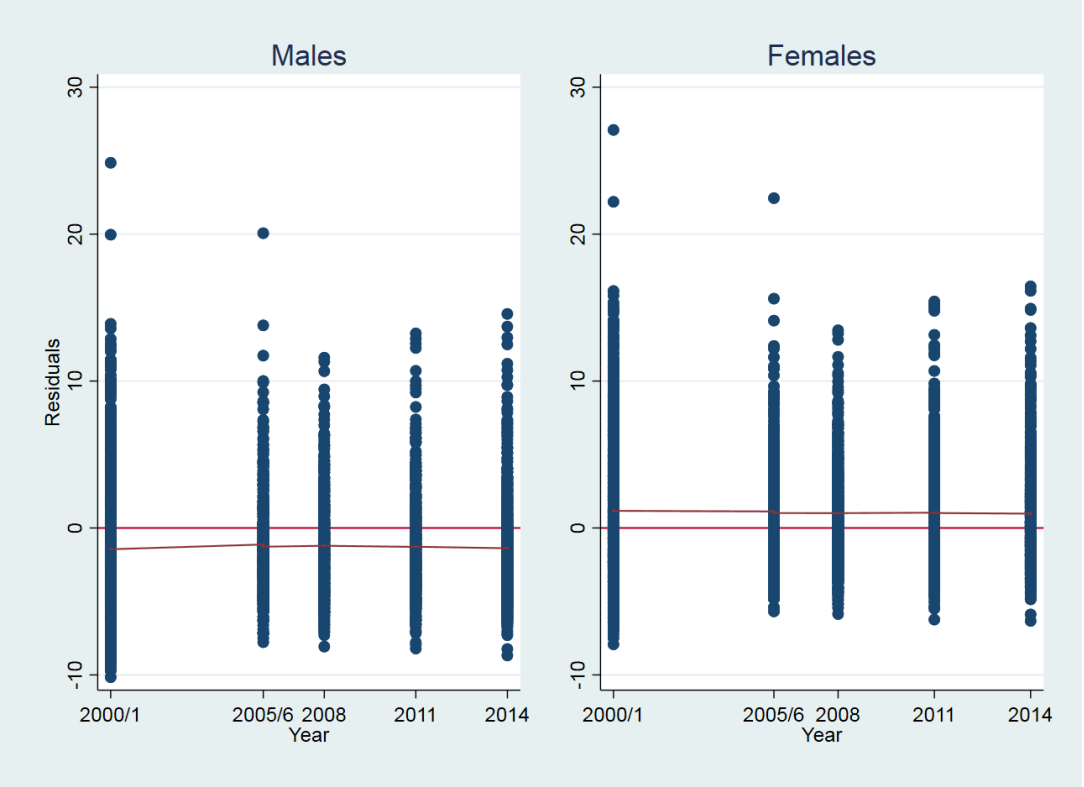

Figure S7 The residuals from our interrupted time series model over time for both men and women. The brown line depicts the local regression fit to the residuals.

IMPACT$_{NCD}$

For this study, IMPACT$_{NCD}$ is calibrated to data from 2006 or before. The only exception is the regression models that are used in steps 2 and 3 (Figure S1) for individual predictions of exposure to risk factors. These models were fitted in data from 2001 to 2012. In this chapter, we first present the internal validation of the synthetic population and the risk factor trends, as evidence that the synthetic population used in IMPACT$_{NCD}$ was similar to the English population. Then, we present the predictive validation of IMPACT$_{NCD}$ by comparing observed to predicted mortality rates for the years 2006 to

2013 by age group, sex, QIMD, and modelled disease. Specifically for GCa, we also compare observed and predicted incidence rates for the same period by age group and sex.*

Synthetic population validation

The following graphs compare a random sample of 200,000 synthetic individuals from the synthetic population to the original sample of HSE06 (n = 17,633). Mosaic plots† were used for the categorical variables, and cumulative distribution plots were used for the continuous variables. Specifically, in this document, the area of each tile of the mosaic plots is proportional to the proportion of each subgroup in the respective population. Only graphs that were relevant to the analysis for this study are presented here.

The graphs support the argument that the final synthetic population is close to reality, at least as it was captured through the HSE06, and are useful for the internal validation of the method. Alfons *et al*. used a statistical simulation approach to evaluate the process and showed that this method produces synthetic populations very similar to the original survey.[9] Of course, the method cannot overcome any limitations of the original survey, such as selection bias, or misclassification.

*For CHD and stroke, true incidence rates are rather unknown; therefore, such comparison would be meaningless.

† Mosaic plots are graphical representations of a contingency table of two or more categorical variables, using tiles with areas proportional to the frequencies in each cell of the table.[74]

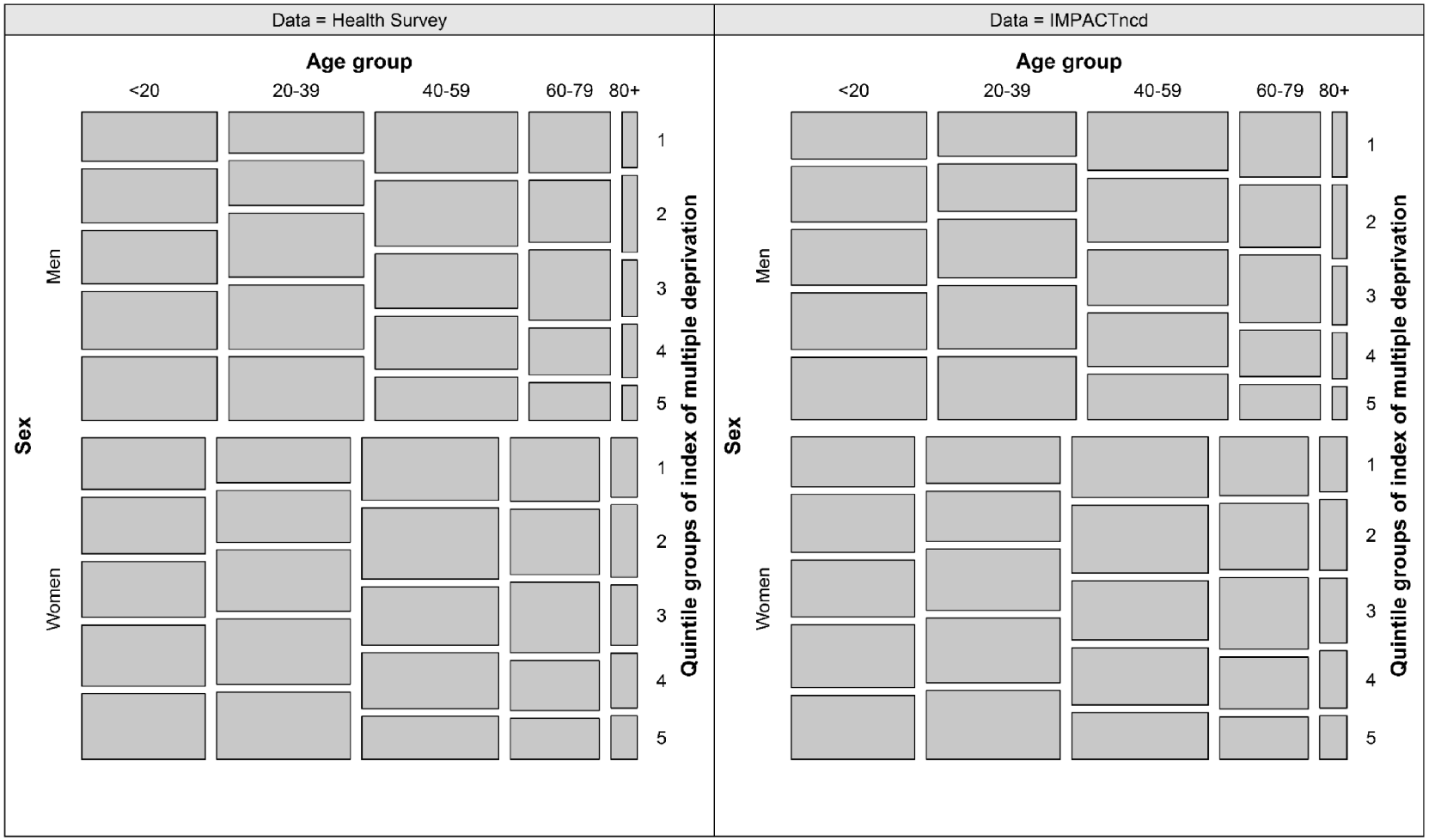

Figure S8 Comparison between the Health Survey for England 2006 (n = 17,633) and a random sample (n=200,000) from the synthetic population. Distribution of age group, sex and quintile groups of index of multiple deprivation (1=least deprived, 5=most deprived) is presented

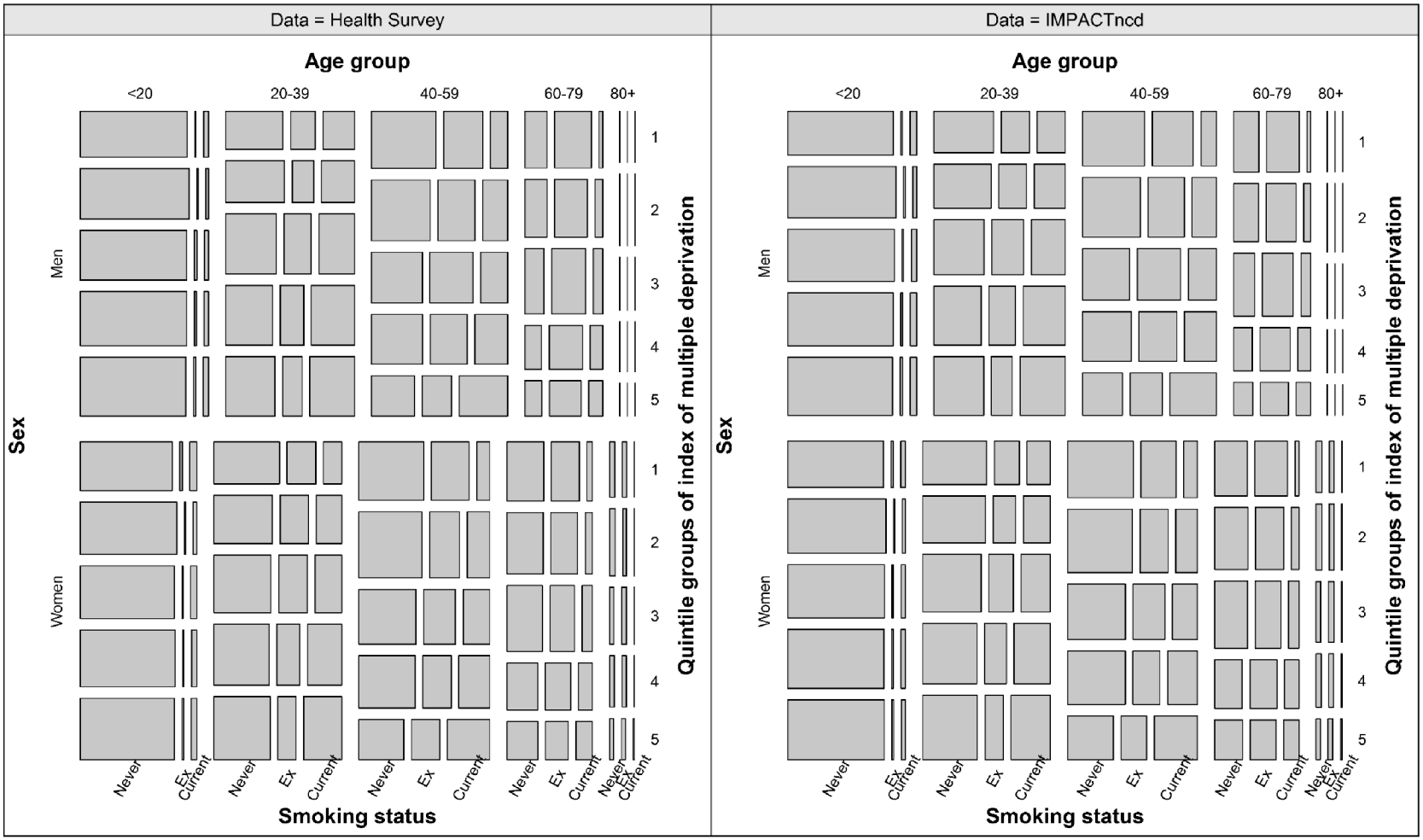

Figure S9 Comparison between the Health Survey for England 2006 (n = 17,633) and a random sample (n=200,000) from the synthetic population. Distribution of age group, sex, quintile groups of index of multiple deprivation (1=least deprived, 5=most deprived) and smoking status is presented

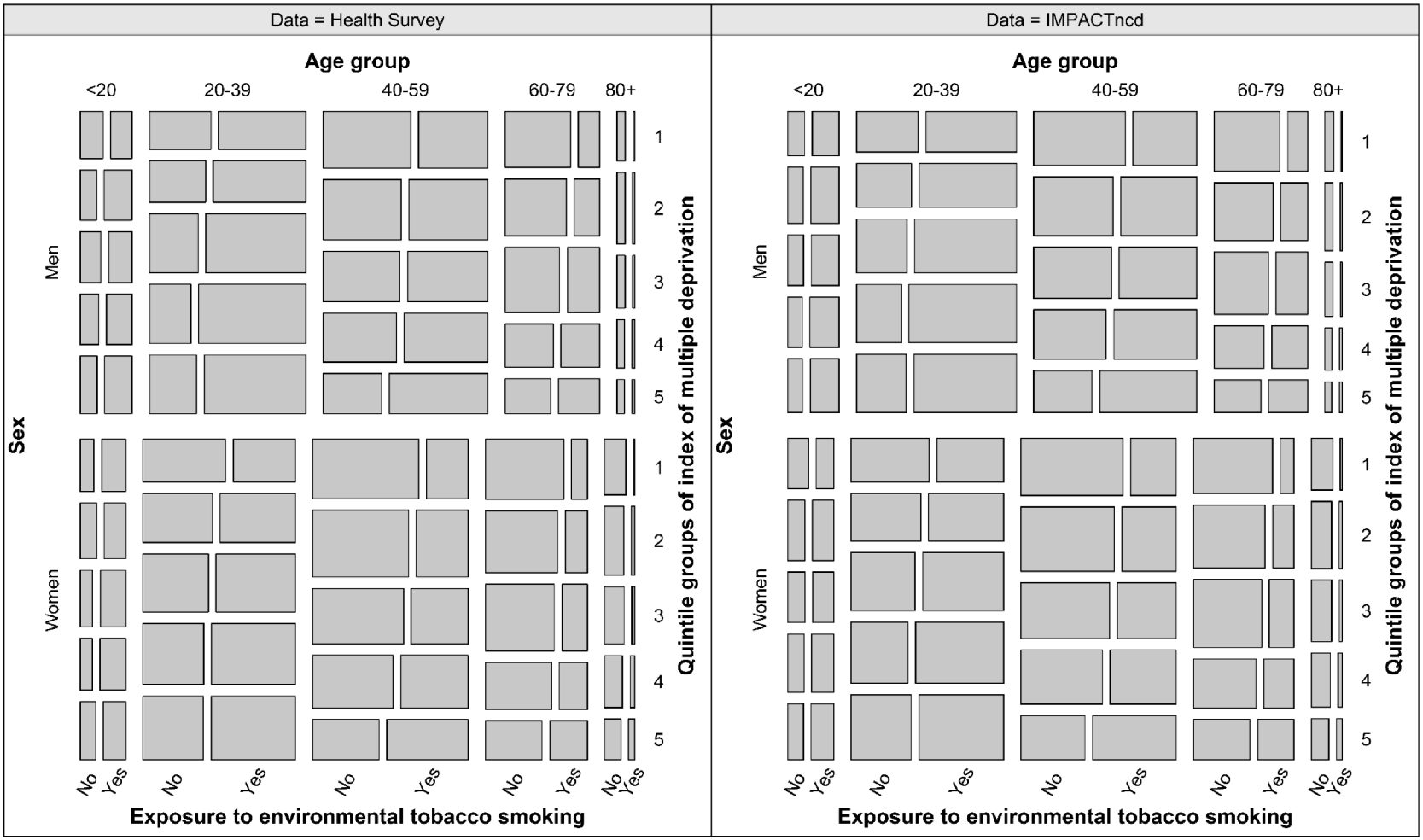

Figure S10 Comparison between the Health Survey for England 2006 (n = 17,633) and a random sample (n=200,000) from the synthetic population. Distribution of age group, sex, quintile groups of index of multiple deprivation (1=least deprived, 5=most deprived) and exposure to environmental tobacco is presented

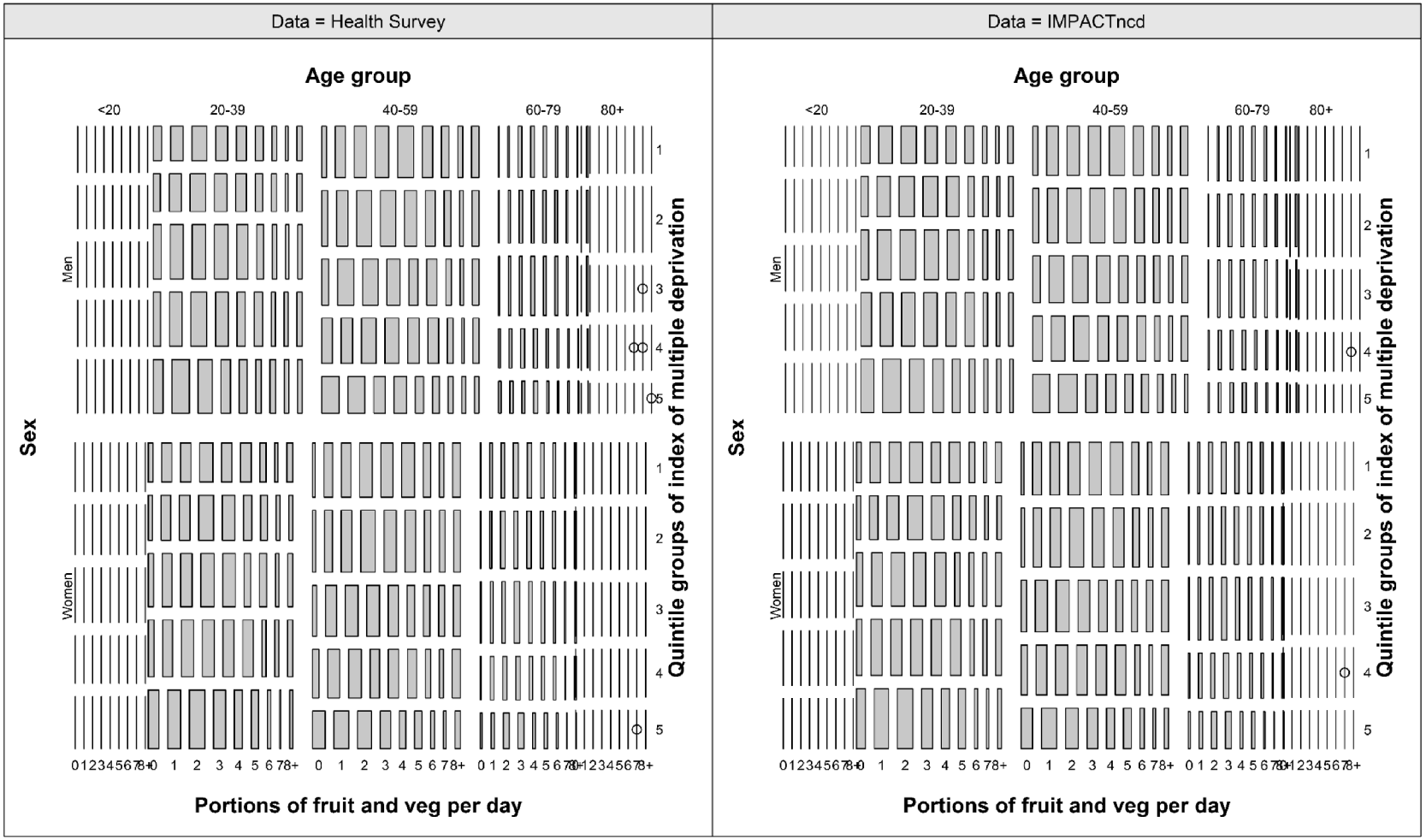

Figure S11 Comparison between the Health Survey for England 2006 (n = 17,633) and a random sample (n=200,000) from the synthetic population. Distribution of age group, sex, quintile groups of index of multiple deprivation (1=least deprived, 5=most deprived) and portions of fruit and vegetable consumed per day is presented

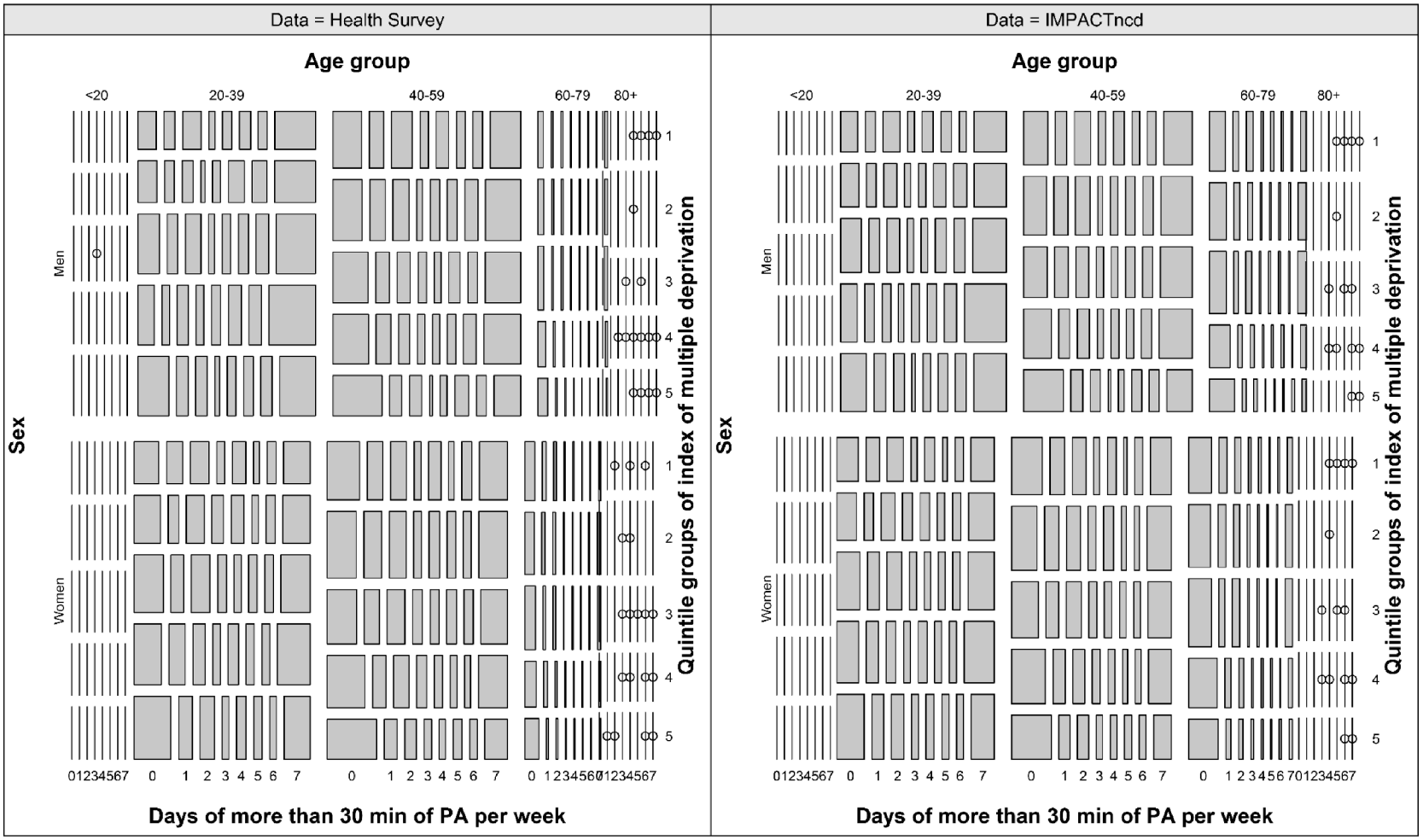

Figure S12 Comparison between the Health Survey for England 2006 (n = 17,633) and a random sample (n=200,000) from the synthetic population. Distribution of age group, sex, quintile groups of index of multiple deprivation (1=least deprived, 5=most deprived) and exposure to days of more than 30 min of physical activity (PA) per week is presented. The small circles represent sub-groups with no participants. Their number reduced in the synthetic population sample highlighting the capability of the method to create individuals with traits not present in the original survey

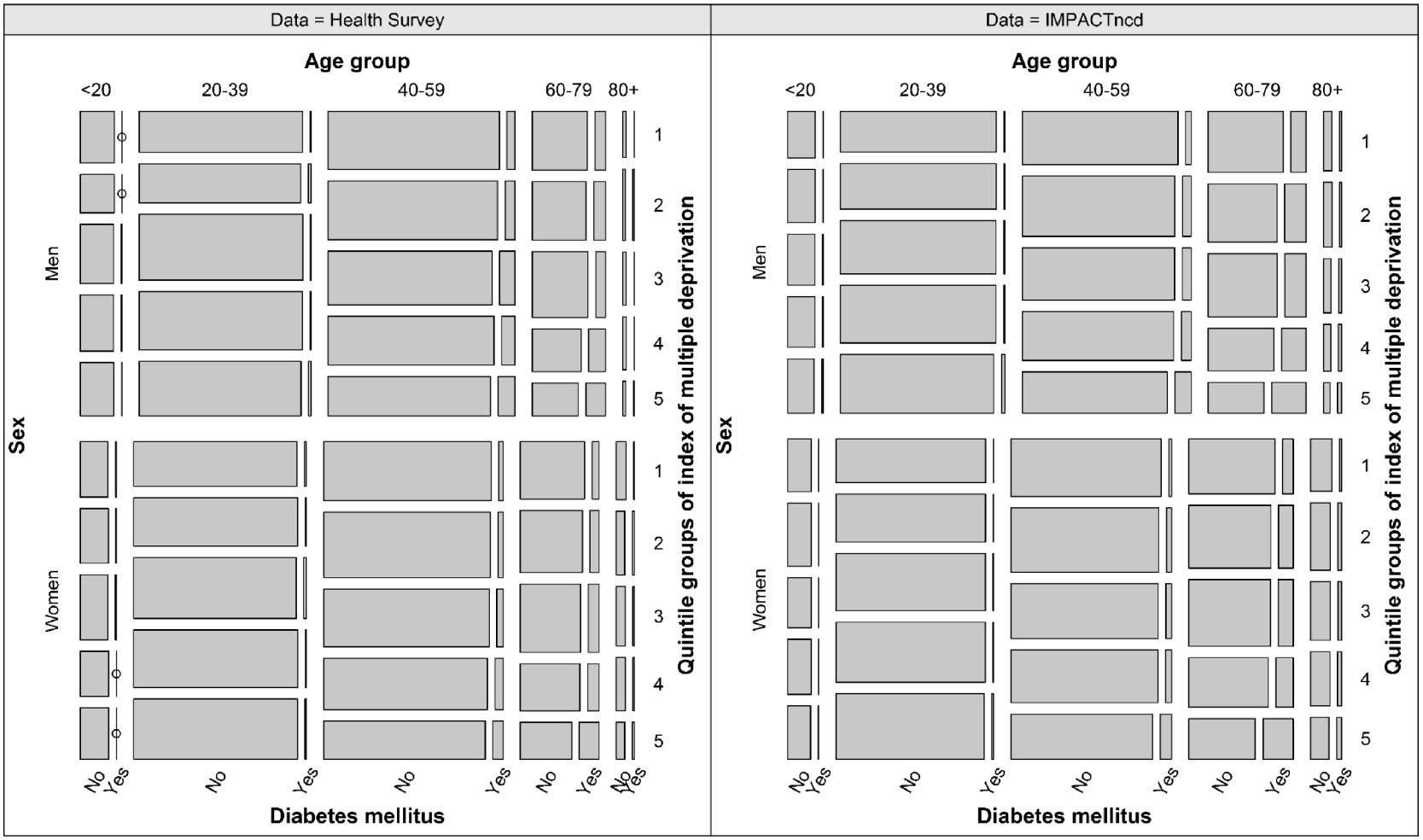

Figure S13 Comparison between the Health Survey for England 2006 (n = 17,633) and a random sample (n=200,000) from the synthetic population. Distribution of age group, sex, quintile groups of index of multiple deprivation (1=least deprived, 5=most deprived) and diabetes mellitus is presented

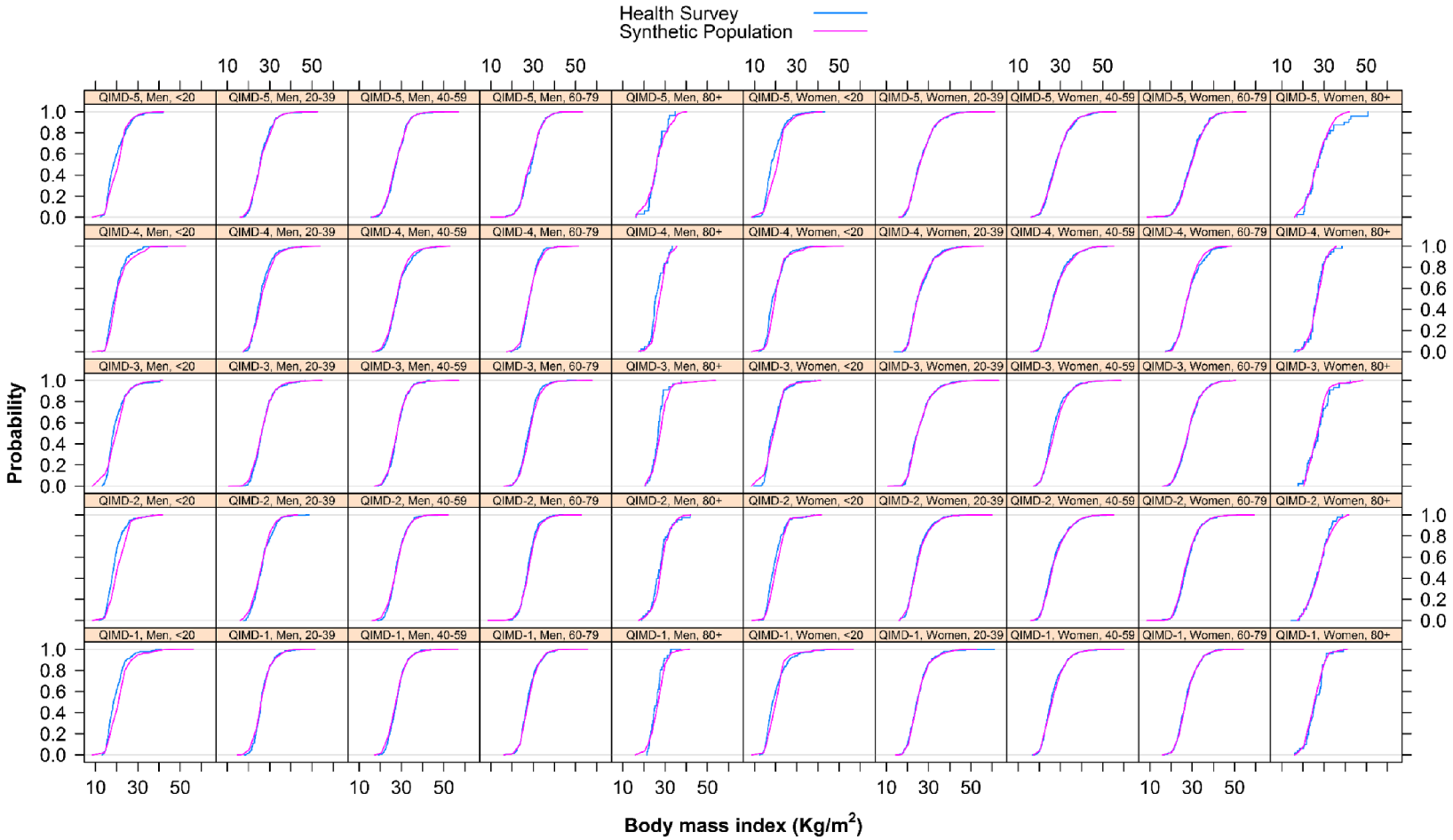

Figure S14 Comparison of body mass index cumulative distributions in Health Survey for England 2006 (n = 17,633) and a random sample (n=200,000) from the synthetic population. Each panel depicts a different subgroup of the population based on quintile groups of index of multiple deprivation (QIMD, 1=least deprived, 5=most deprived), sex and age group

=

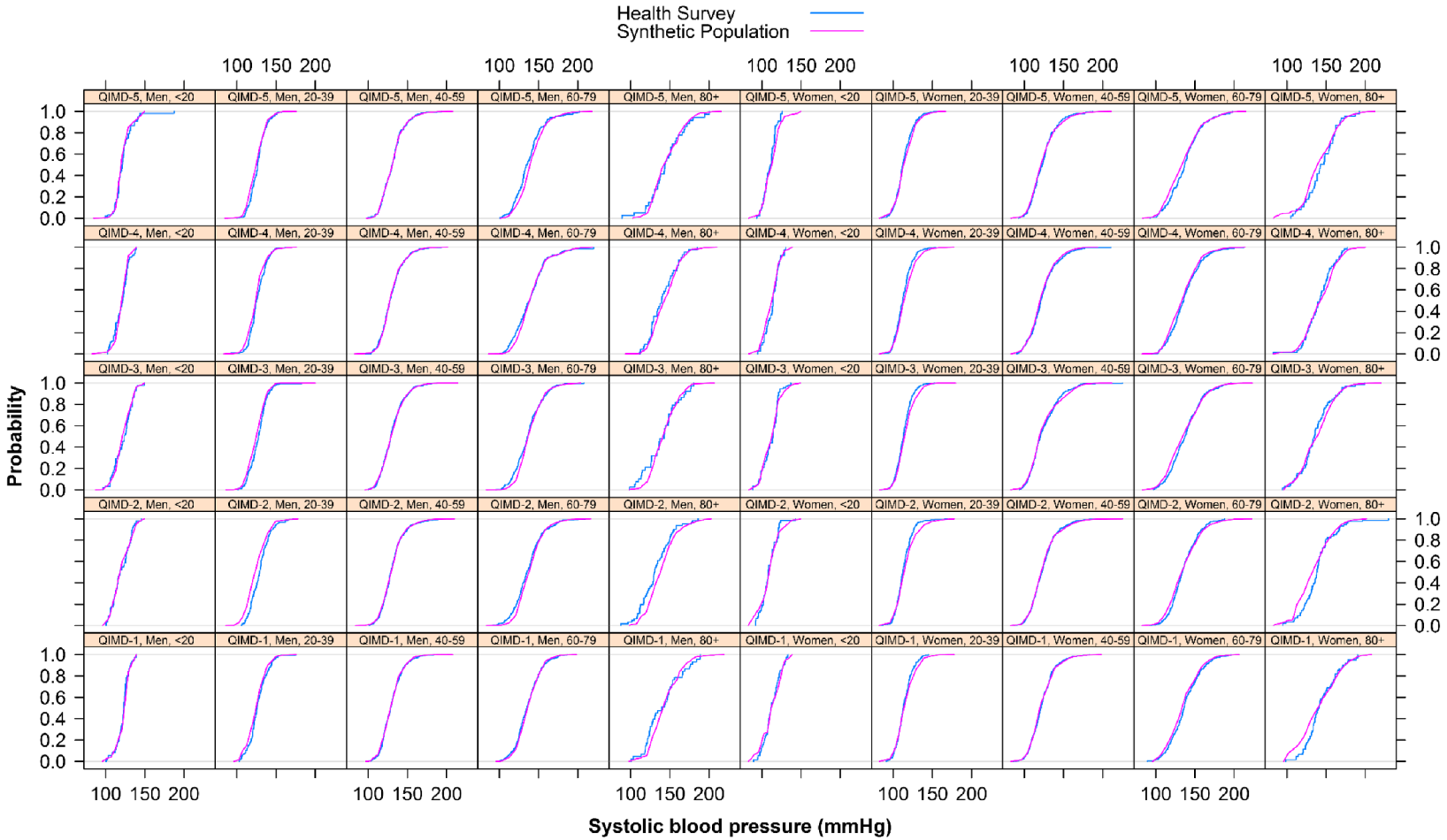

Figure S15 Comparison of systolic blood pressure cumulative distributions in Health Survey for England 2006 (n = 17,633) and a random sample (n=200,000) from the synthetic population. Each panel depicts a different subgroup of the population based on quintile groups of index of multiple deprivation (QIMD, 1=least deprived, 5=most deprived), sex and age group

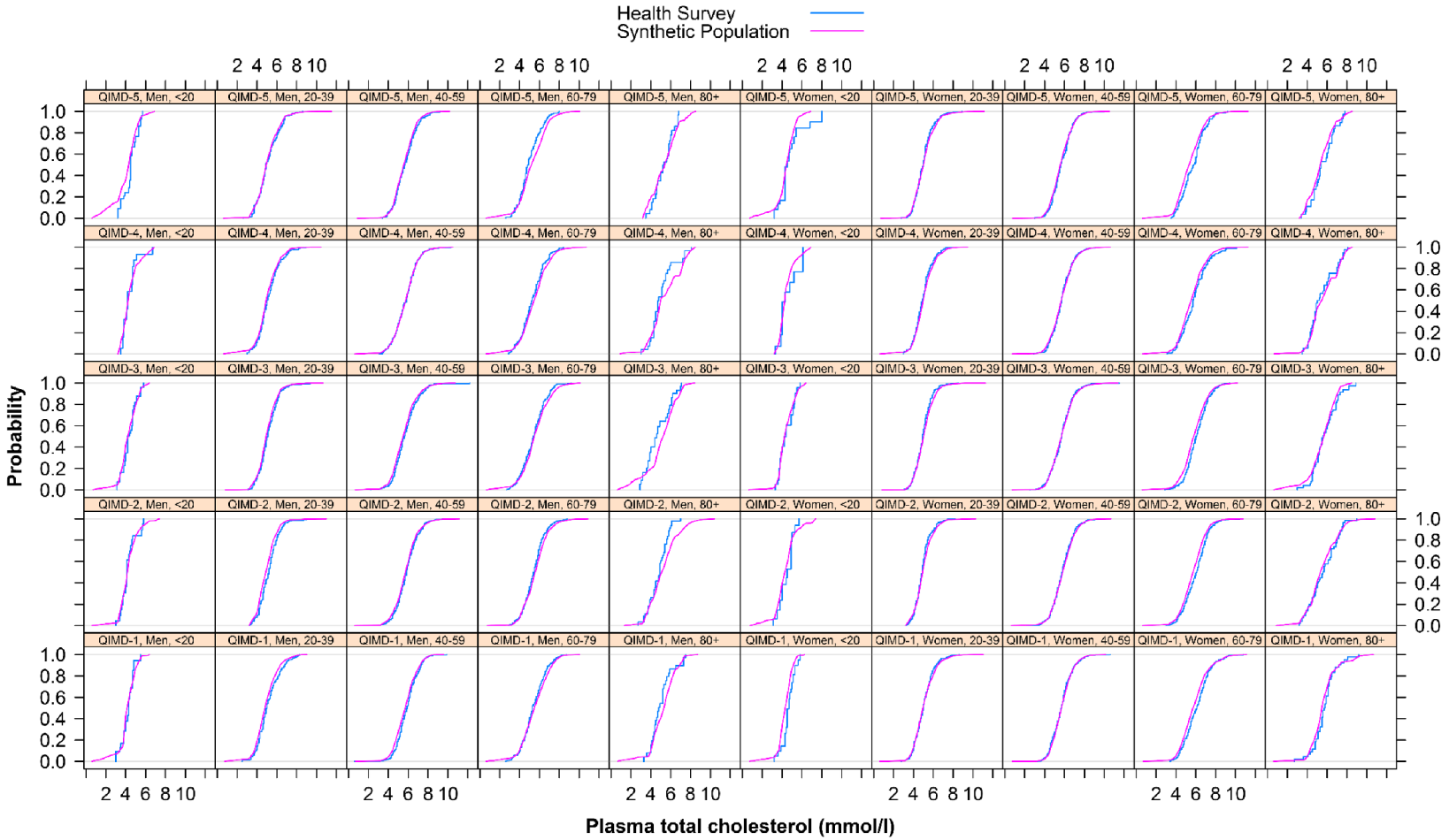

Figure S16 Comparison of plasma total cholesterol cumulative distributions in Health Survey for England 2006 (n = 17,633) and a random sample (n=200,000) from the synthetic population. Each panel depicts a different subgroup of the population based on quintile groups of index of multiple deprivation (QIMD, 1=least deprived, 5=most deprived), sex and age group

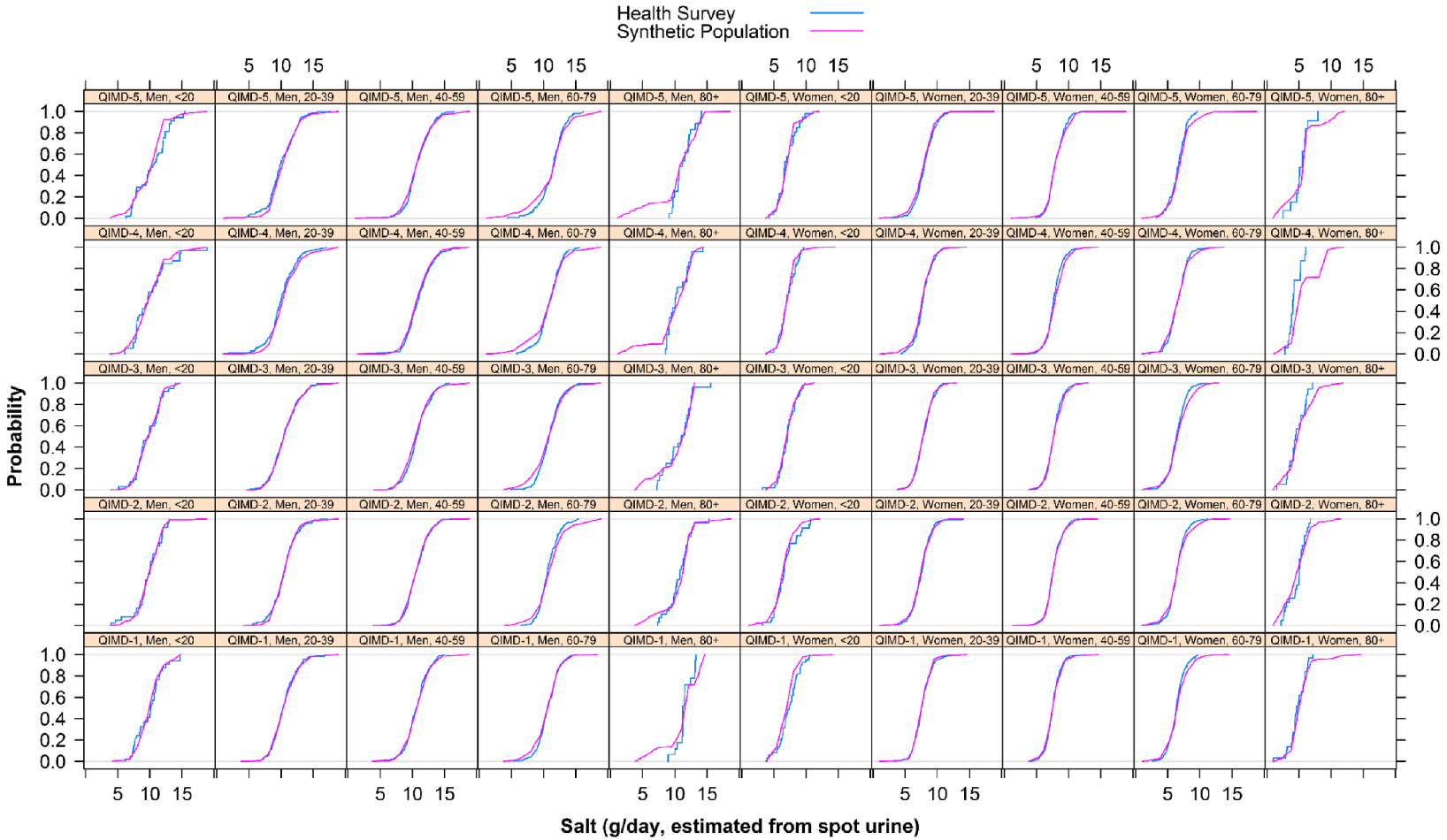

Figure S17 Comparison of salt consumption cumulative distributions in Health Survey for England 2006 (n = 17,633) and a random sample (n=200,000) from the synthetic population. Each panel depicts a different subgroup of the population based on quintile groups of index of multiple deprivation (QIMD, 1=least deprived, 5=most deprived), sex and age group. Note that IMPACT$_{NCD}$ applies another layer of processing to integrate information from 24h urine sodium measurements before risk estimation

Risk factor trends validation

Here we compare mean exposure of IMPACT$_{NCD}$ synthetic population to the observed exposure through relevant national representative surveys. We stratified by sex, age group and when data allowed by QIMD. Overall, the plots provide evidence that the regression models used in steps 2 and 3 (Figure S1) have captured trends by age, sex and QIMD well enough. Please note that the shrinking variance with time that is observed in the graphs is an artefact. For efficiency, the model only calculates lagged exposures for the years that are necessary depending on the distribution of the lag time random variables for CVD and cancer. Therefore, the number of data points for each year differs, and for the shifted binomials we chose for the lag time random variables, increases by year for the period 2001 - 2013. This is reflected in the decreasing variance on the graphs.

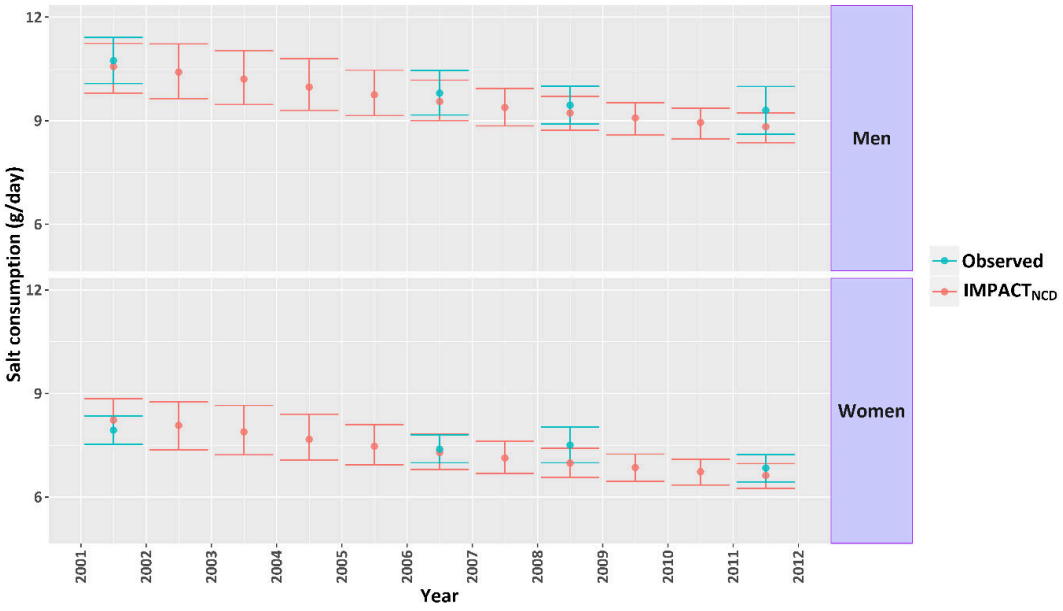

Figure S18 Mean salt consumption for ages 19 – 64 between the years 2001 and 2011. Observed in the population through surveys using 24h urine collections[75–78] vs IMPACT$_{NCD}$ synthetic population estimates. Error bars represent 95% confidence intervals of the mean.

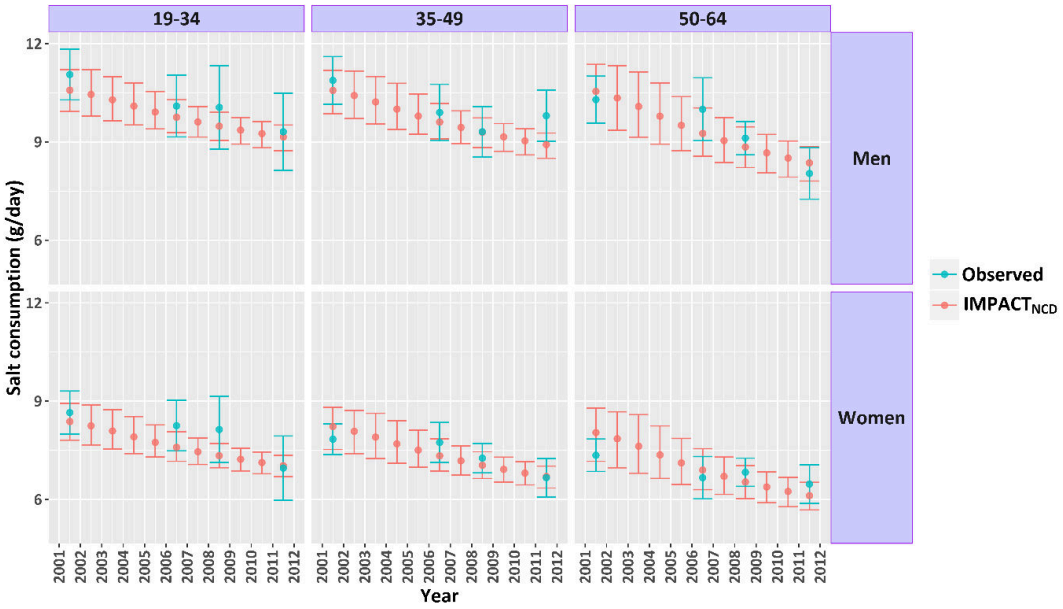

Figure S19 Mean salt consumption by age group, between the years 2001 and 2011. Observed in the population through surveys using 24h urine collections[75–78] vs IMPACT$_{NCD}$ synthetic population estimates. Error bars represent 95% confidence intervals of the mean.

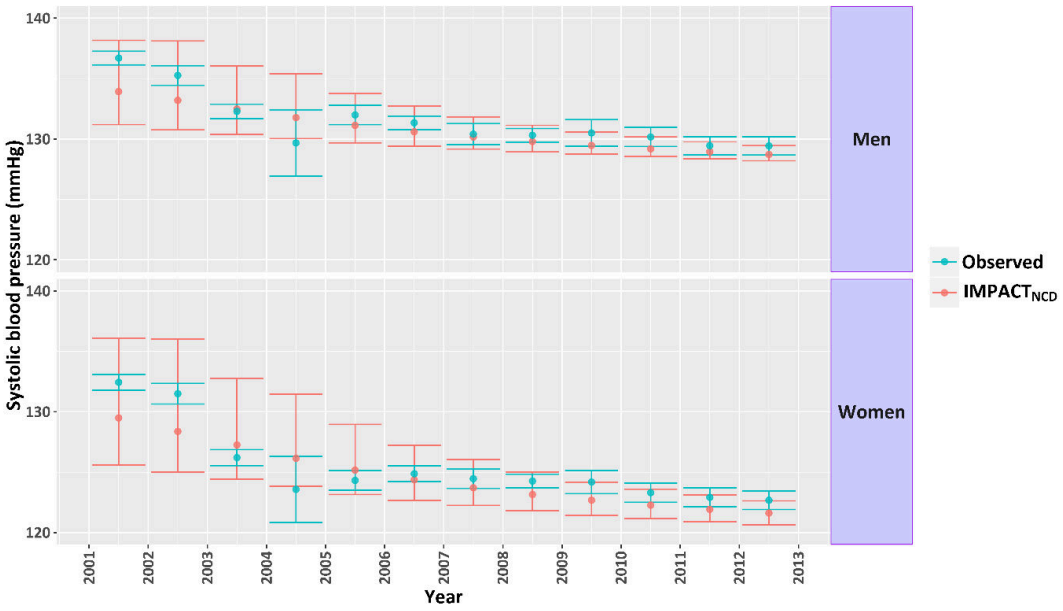

Figure S20 Mean systolic blood pressure for ages 30 – 84 between years 2001 and 2012. Observed in the population through Health Survey for England vs IMPACT$_{NCD}$ synthetic population estimates. Error bars represent 95% confidence intervals of the mean.

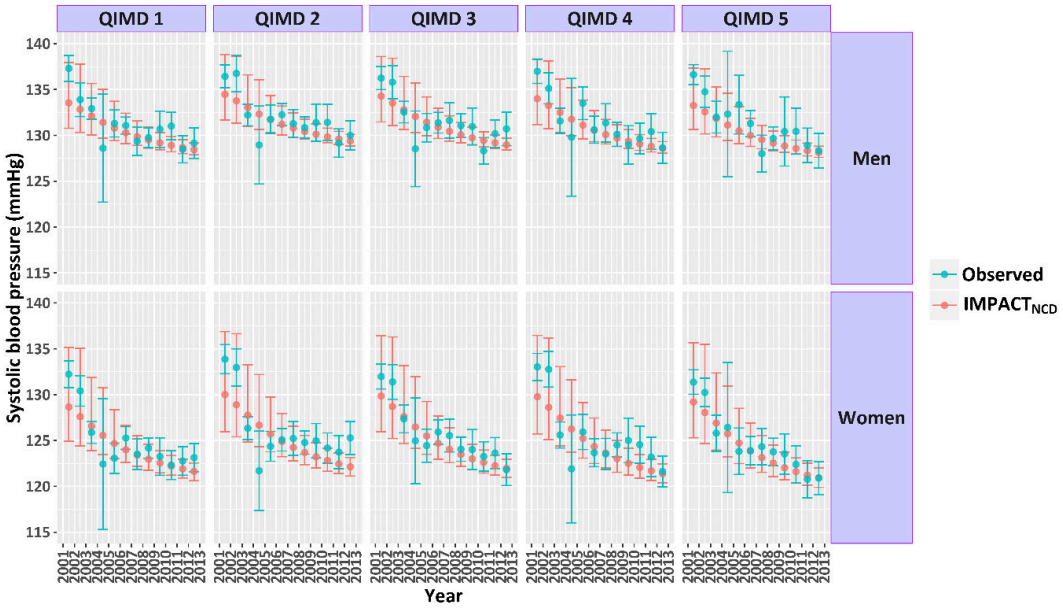

Figure S21 Mean systolic blood pressure for ages 30 – 84 by quintile group of the index of multiple deprivation (QIMD, 1 = least deprived) between years 2001 and 2012. Observed in the population through Health Survey for England vs IMPACT$_{NCD}$ synthetic population estimates. Error bars represent 95% confidence intervals of the mean.

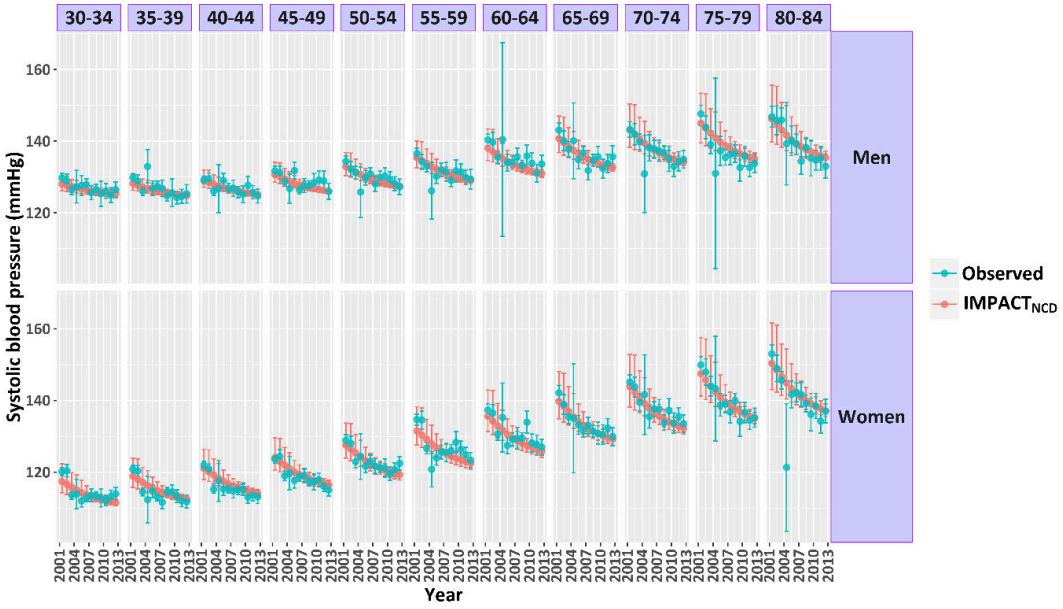

Figure S22 Mean systolic blood pressure for ages 30 – 84 by age group, between years 2001 and 2012. Observed in the population through Health Survey for England vs IMPACT$_{NCD}$ synthetic population estimates. Error bars represent 95% confidence intervals of the mean.

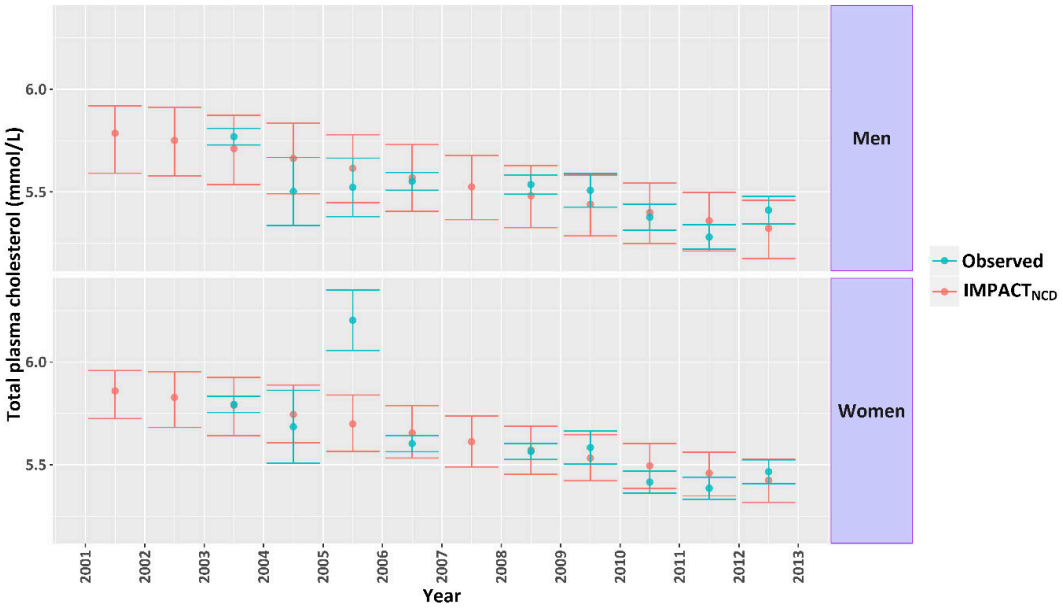

Figure S23 Mean total plasma cholesterol for ages 30 – 84 between years 2001 and 2012. Observed in the population through Health Survey for England vs IMPACT$_{NCD}$ synthetic population estimates. Error bars represent 95% confidence intervals of the mean.

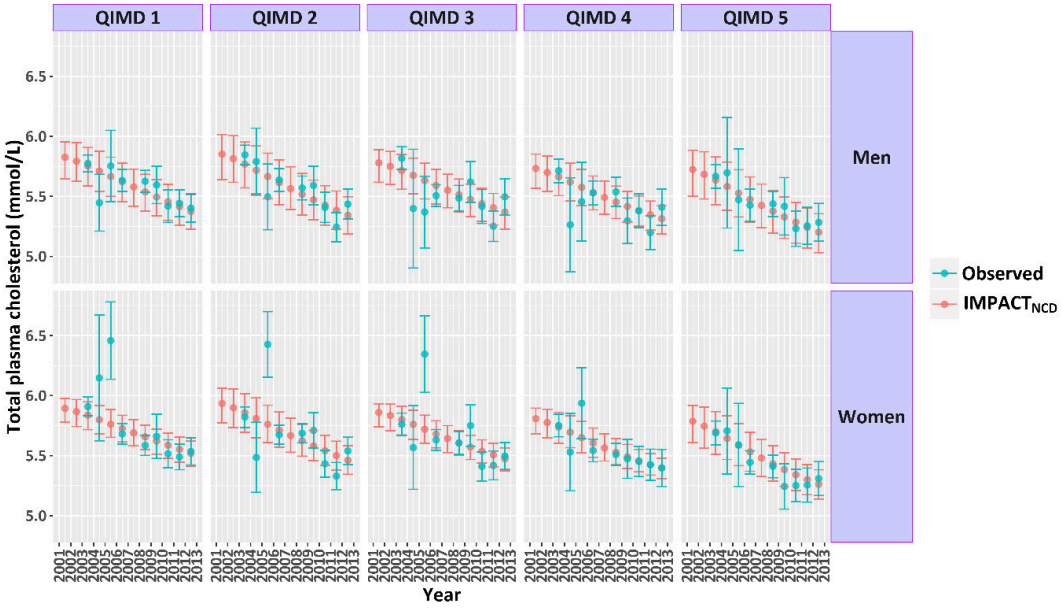

Figure S24 Mean total plasma cholesterol for ages 30 – 84 by quintile group of the index of multiple deprivation (QIMD, 1 = least deprived) between years 2001 and 2012. Observed in the population through Health Survey for England vs IMPACT$_{NCD}$ synthetic population estimates. Error bars represent 95% confidence intervals of the mean.

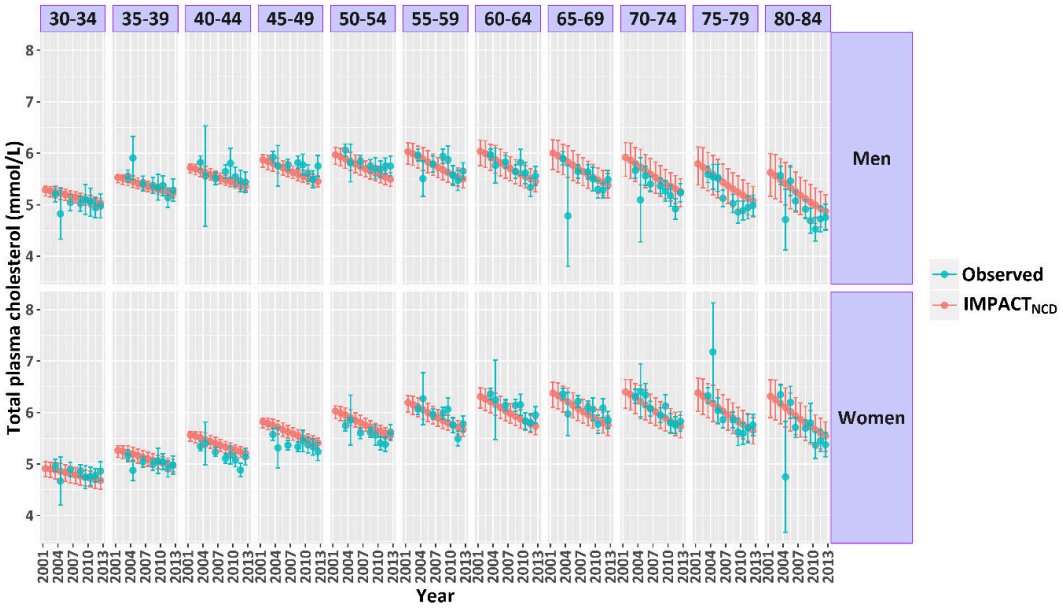

Figure S25 Mean total plasma cholesterol for ages 30 – 84 by age group, between the years 2001 and 2012. Observed in the population through Health Survey for England vs IMPACT$_{NCD}$ synthetic population estimates. Error bars represent 95% confidence intervals of the mean.

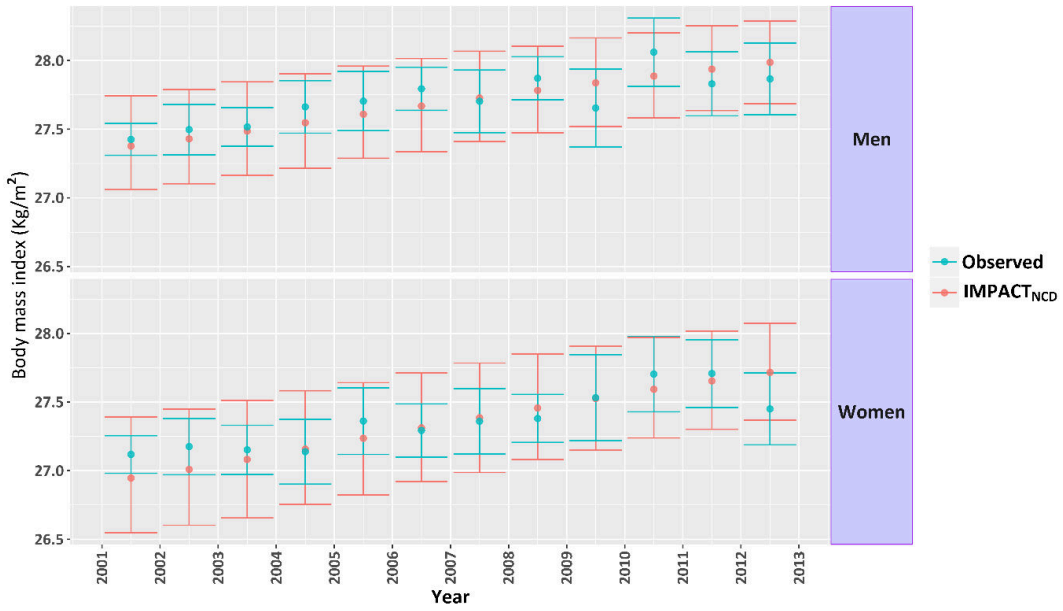

Figure S26 Mean body mass index for ages 30 – 84 between years 2001 and 2012. Observed in the population through Health Survey for England vs. IMPACT$_{NCD}$ synthetic population estimates. Error bars represent 95% confidence intervals of the mean.

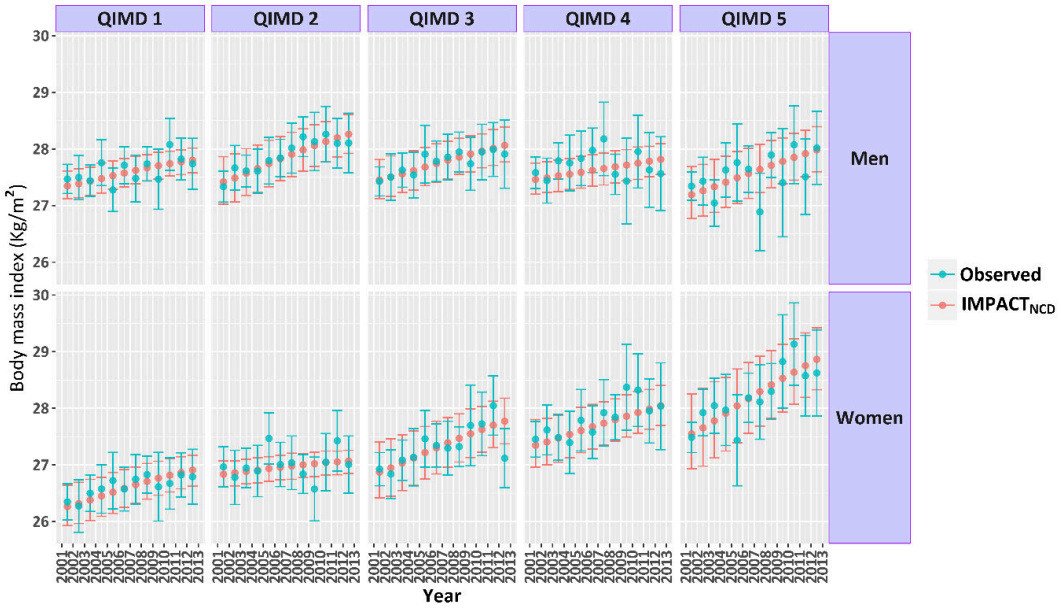

Figure S27 Mean body mass index for ages 30 – 84 by quintile group of the index of multiple deprivation (QIMD, 1 = least deprived) between years 2001 and 2012. Observed in the population through Health Survey for England vs IMPACT$_{NCD}$ synthetic population estimates. Error bars represent 95% confidence intervals of the mean.

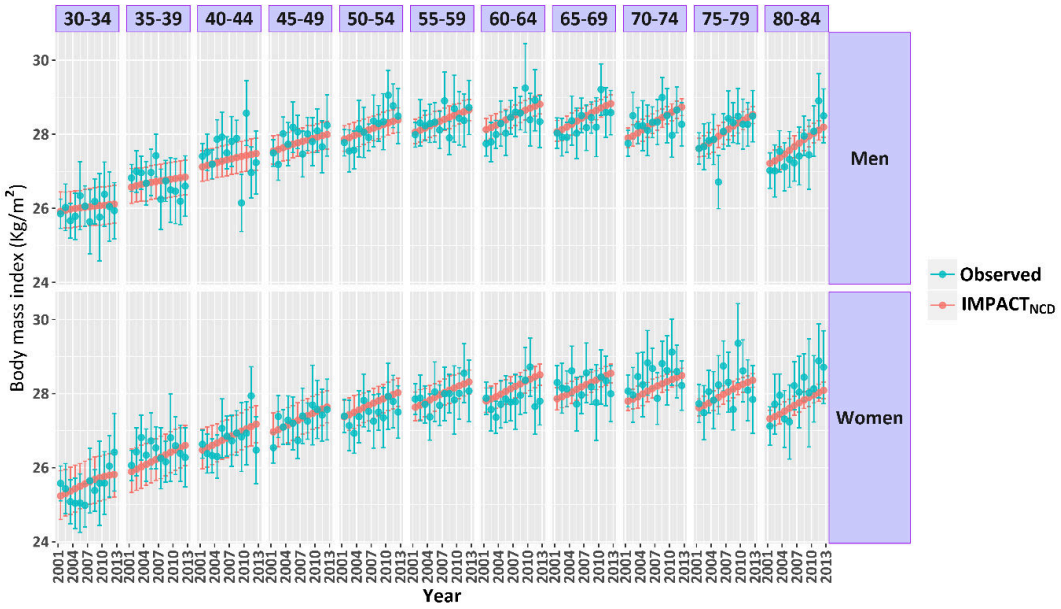

Figure S28 Mean body mass index for ages 30 – 84 by age group between the years 2001 and 2012. Observed in the population through Health Survey for England vs IMPACT$_{NCD}$ synthetic population estimates. Error bars represent 95% confidence intervals of the mean.

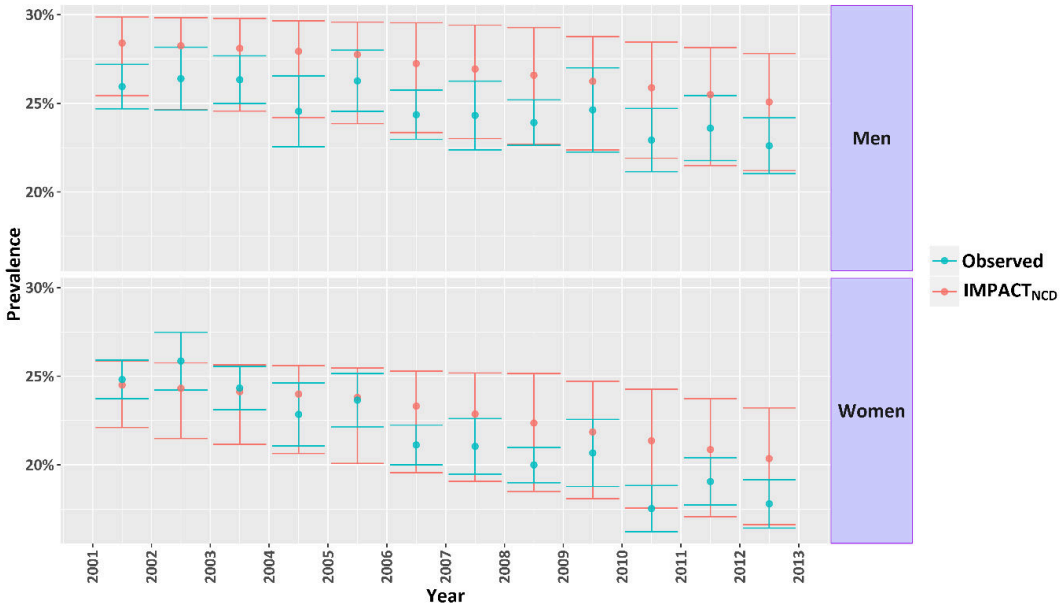

Figure S29 Smoking prevalence for ages 30 – 84 between years 2001 and 2012. Observed in the population through Health Survey for England vs IMPACT$_{NCD}$ synthetic population estimates. Error bars represent 95% confidence intervals of the mean.

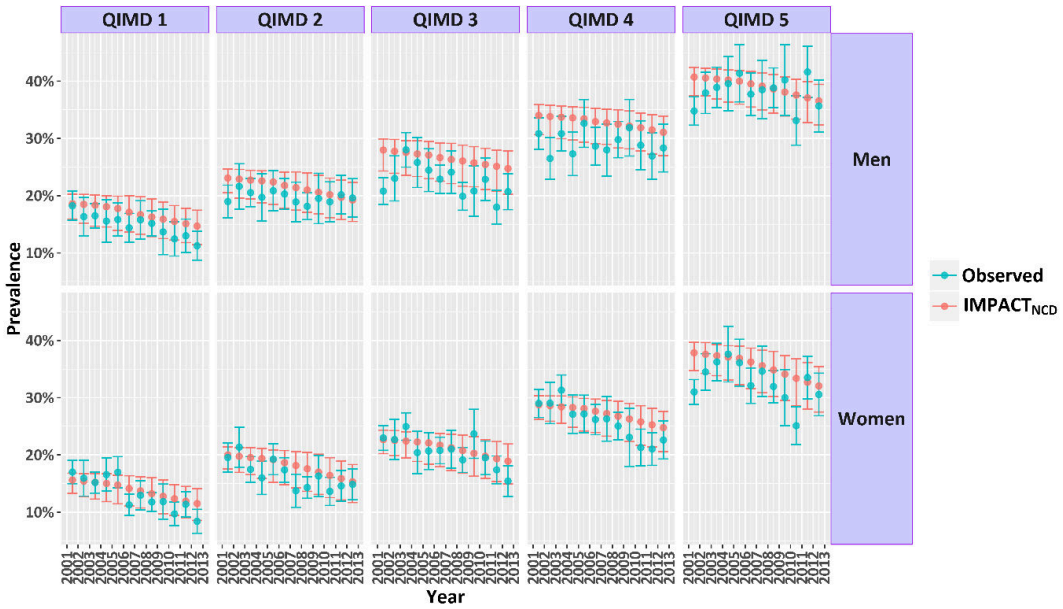

Figure S30 Smoking prevalence for ages 30 – 84 by quintile group of the index of multiple deprivation (QIMD, 1 = least deprived) between years 2001 and 2012. Observed in the population through Health Survey for England vs IMPACT$_{NCD}$ synthetic population estimates. Error bars represent 95% confidence intervals of the mean.

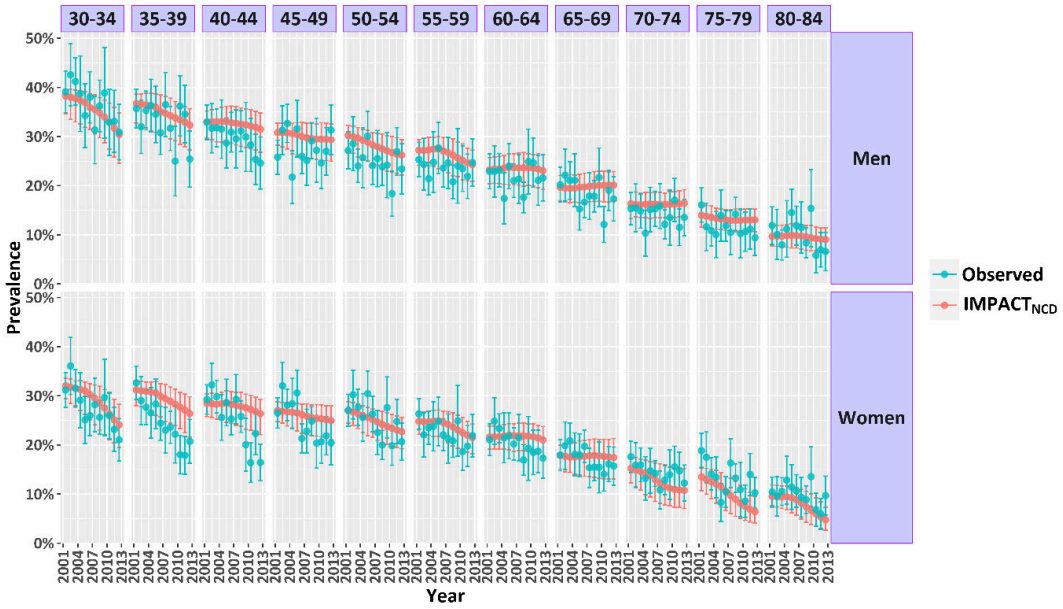

Figure S31 Smoking prevalence for ages 30 – 84 by age group between the years 2001 and 2012. Observed in the population through Health Survey for England vs IMPACT$_{NCD}$ synthetic population estimates. Error bars represent 95% confidence intervals of the mean.

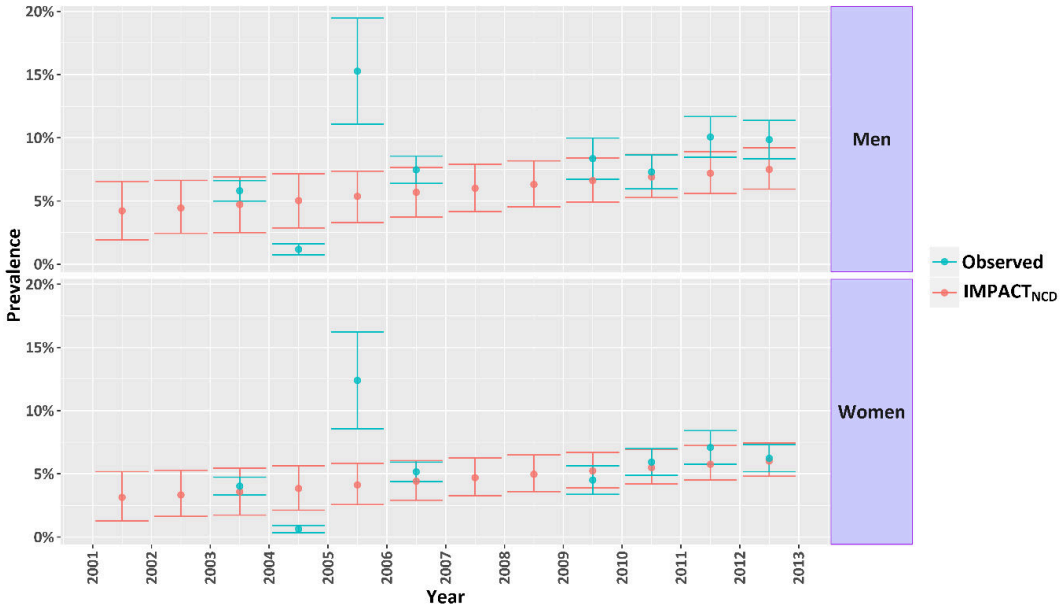

Figure S32 Diabetes mellitus prevalence for ages 30 – 84 between years 2001 and 2012. Observed in the population through Health Survey for England vs IMPACT$_{NCD}$ synthetic population estimates. Error bars represent 95% confidence intervals of the mean.

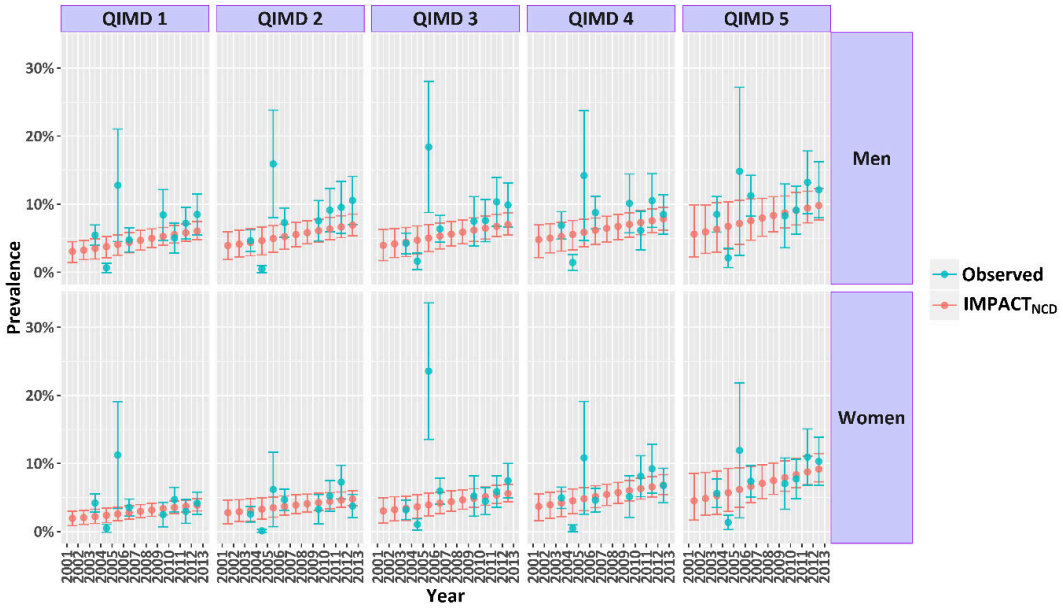

Figure S33 Diabetes mellitus prevalence for ages 30 – 84 by quintile group of the index of multiple deprivation (QIMD, 1 = least deprived) between years 2001 and 2012. Observed in the population through Health Survey for England vs IMPACT$_{NCD}$ synthetic population estimates. Error bars represent 95% confidence intervals of the mean.

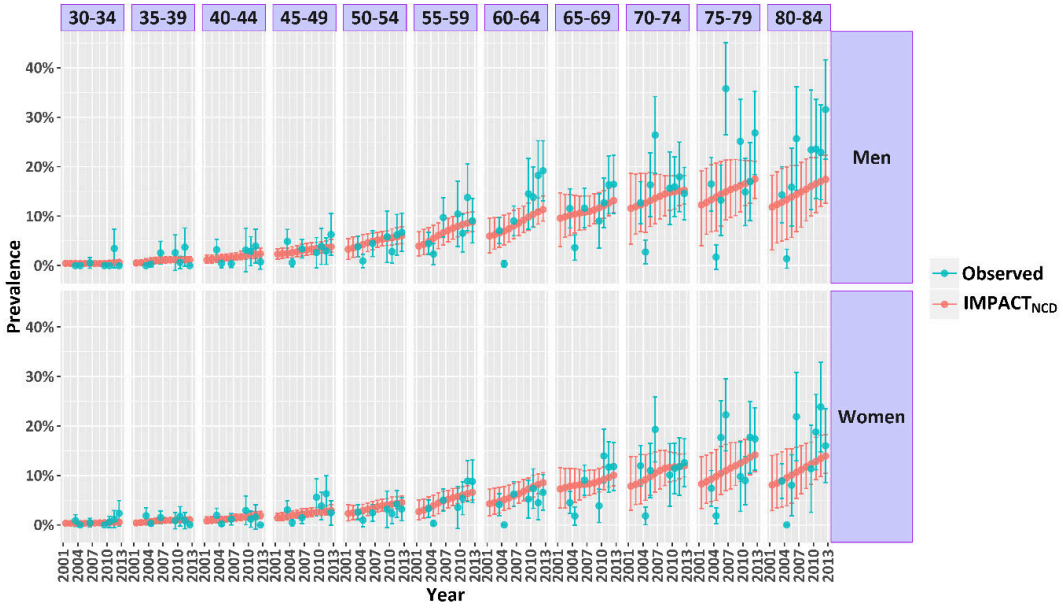

Figure S34 Diabetes mellitus prevalence for ages 30 – 84 by age group between the years 2001 and 2012. Observed in the population through Health Survey for England vs IMPACT$_{NCD}$ synthetic population estimates. Error bars represent 95% confidence intervals of the mean.

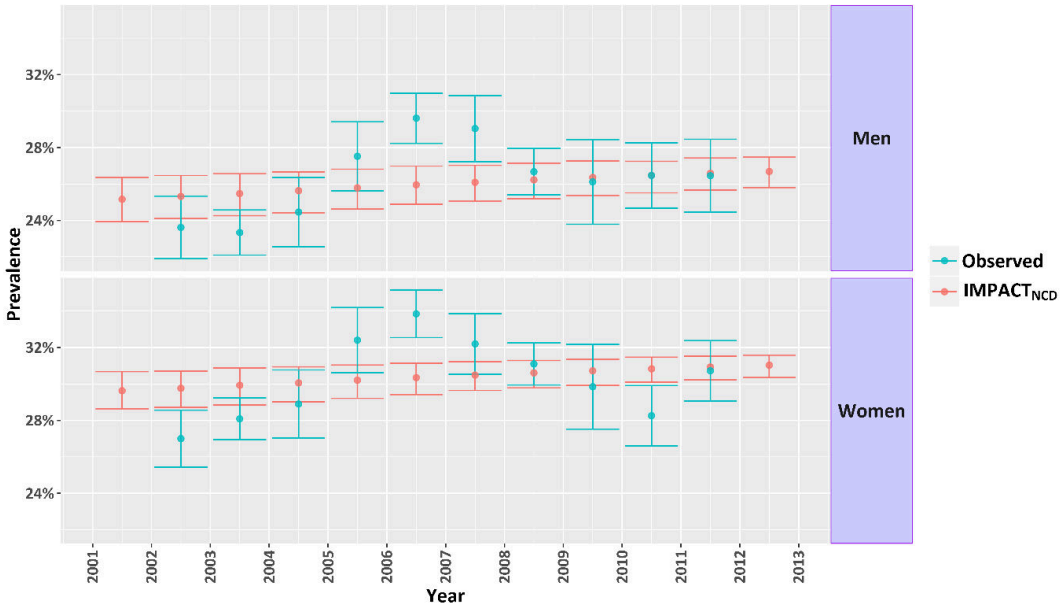

Figure S35 Five or more portions of fruit & veg per day prevalence for ages 30 – 84 between years 2001 and 2012. Observed in the population through Health Survey for England vs. IMPACT$_{NCD}$ synthetic population estimates. Error bars represent 95% confidence intervals of the mean.

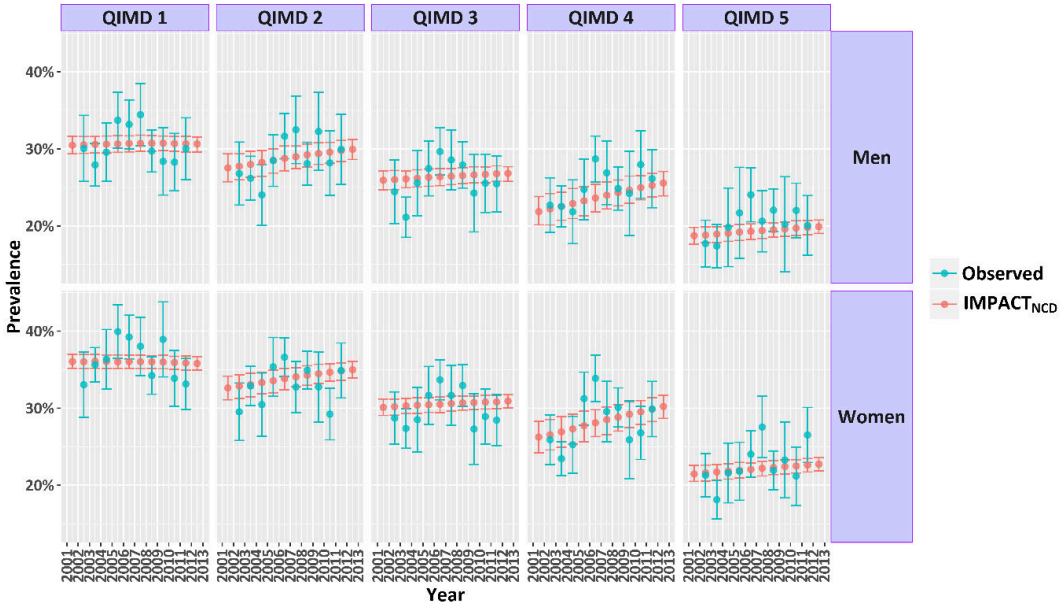

Figure S36 Five or more portions of fruit & veg per day prevalence for ages 30 – 84 by quintile group of the index of multiple deprivation (QIMD, 1 = least deprived) between years 2001 and 2012. Observed in the population through Health Survey for England vs IMPACT$_{NCD}$ synthetic population estimates. Error bars represent 95% confidence intervals of the mean.

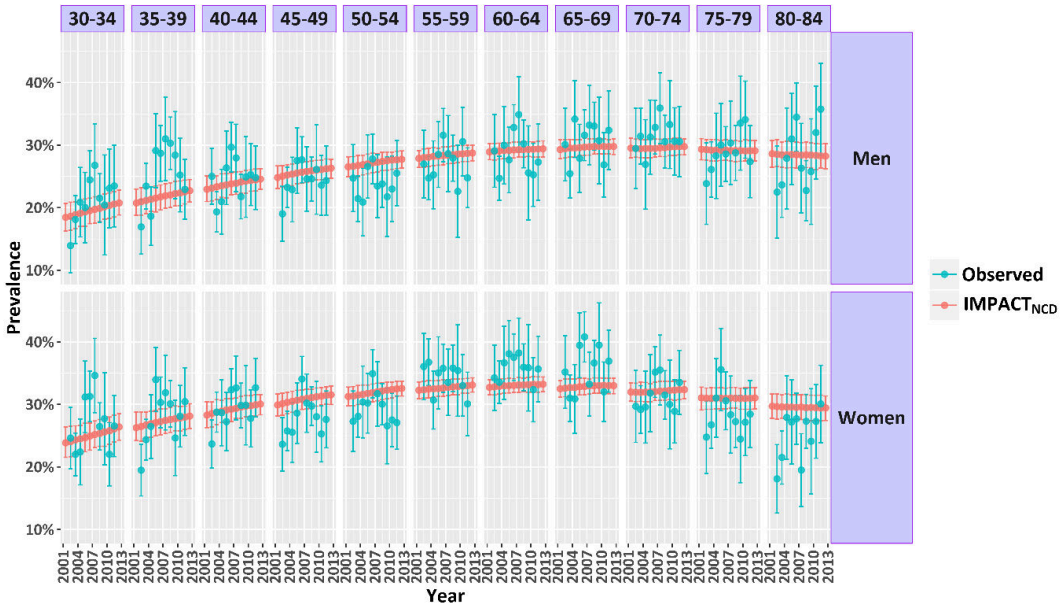

Figure S37 Five or more portions of fruit & veg per day prevalence for ages 30 – 84 by age group between the years 2001 and 2012. Observed in the population through Health Survey for England vs IMPACT$_{NCD}$ synthetic population estimates. Error bars represent 95% confidence intervals of the mean.

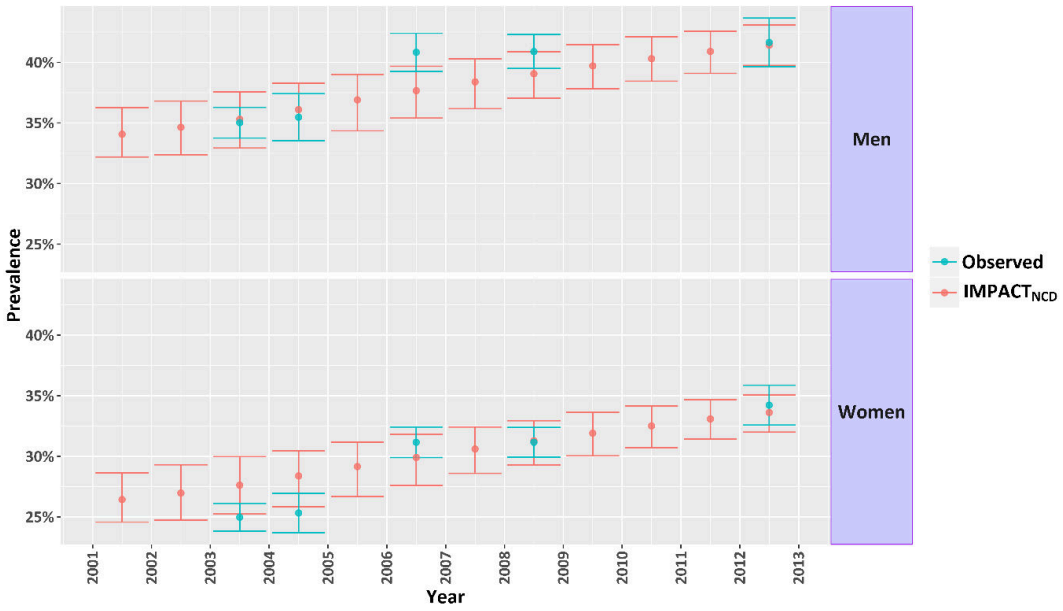

Figure S38 Five or more active days per week prevalence for ages 30 – 84 between years 2001 and 2012. Observed in the population through Health Survey for England vs IMPACT$_{NCD}$ synthetic population estimates. Error bars represent 95% confidence intervals of the mean.

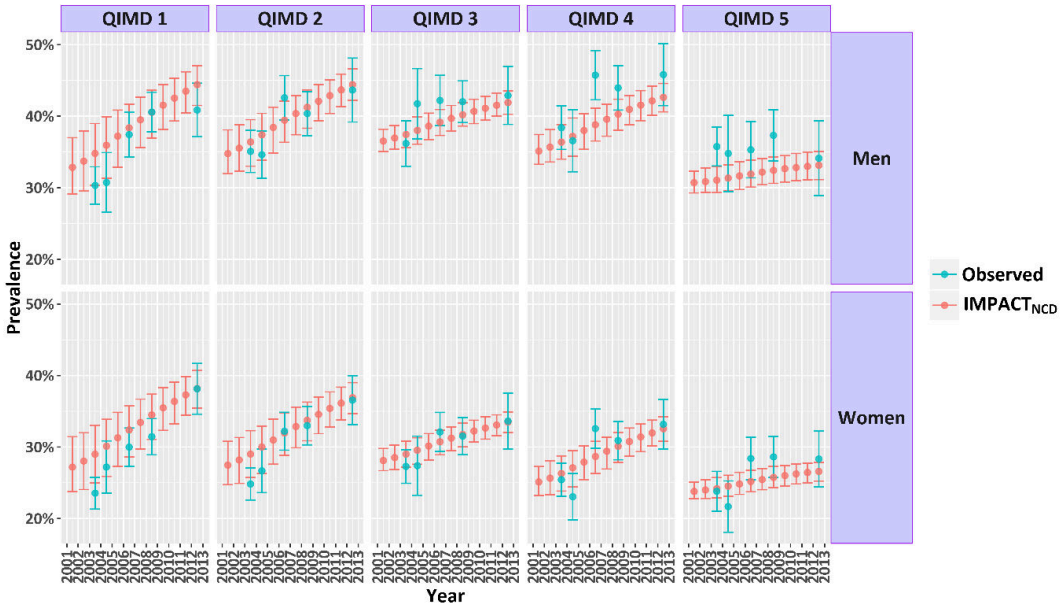

Figure S39 Five or more active days per week prevalence for ages 30 – 84 by quintile group of the index of multiple deprivation (QIMD, 1 = least deprived) between years 2001 and 2012. Observed in the population through Health Survey for England vs IMPACT$_{NCD}$ synthetic population estimates. Error bars represent 95% confidence intervals of the mean.

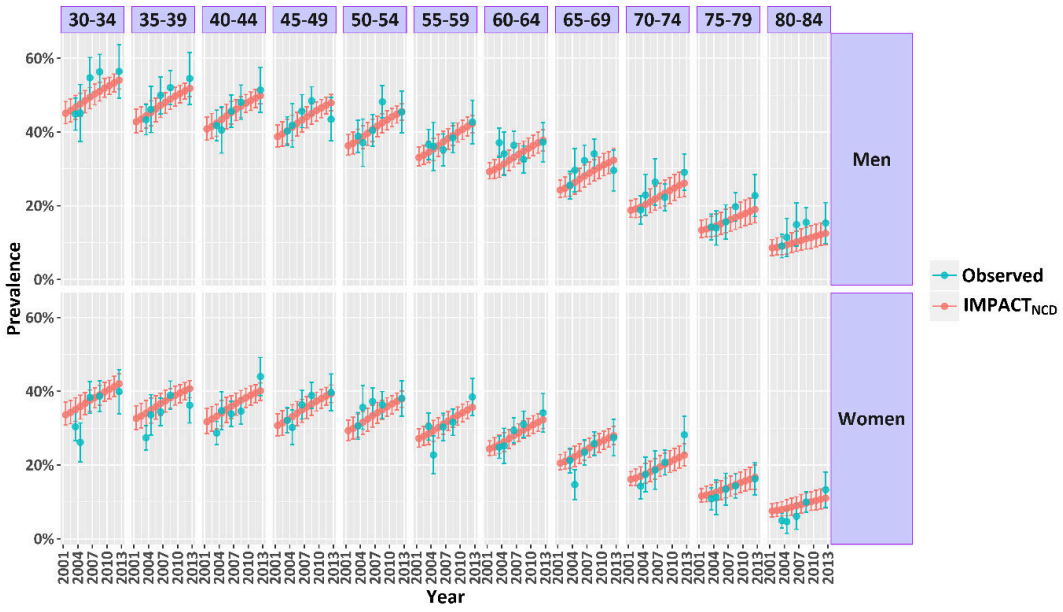

Figure S40 Five or more active days per week prevalence for ages 30 – 84 by age group between the years 2001 and 2012. Observed in the population through Health Survey for England vs IMPACT$_{NCD}$ synthetic population estimates. Error bars represent 95% confidence intervals of the mean.

Incidence predictive validation

We validated incidence only for GCa, as data of the observed incidence is known through the cancer registries. This was not possible for CVD as the true 'first ever' incidence is largely unknown.

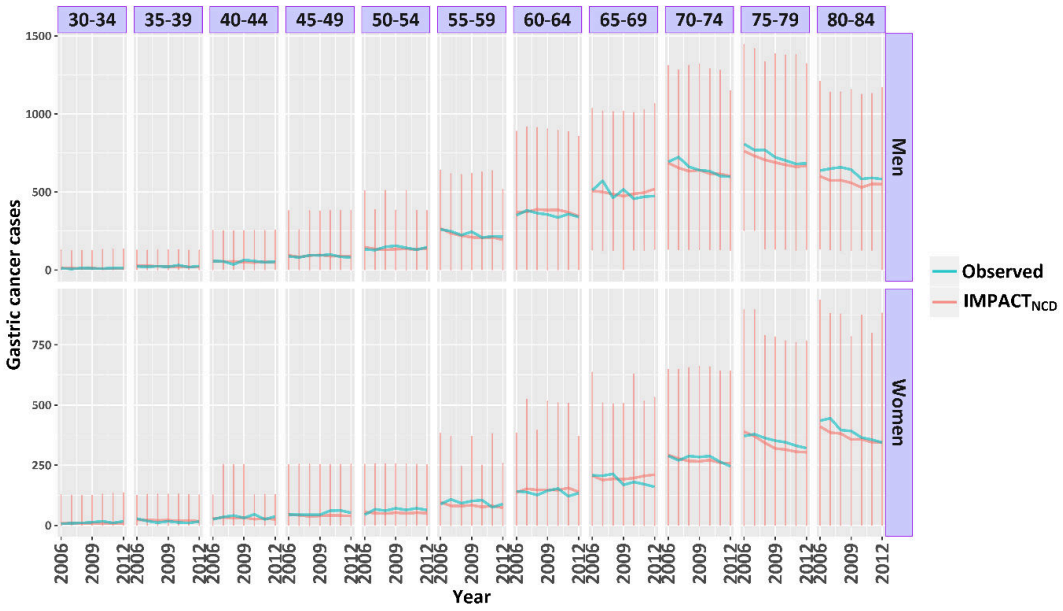

Figure S41 Gastric cancer cases in England for ages 30 – 84 by age group between the years 2006 and 2012. Observed in the population through cancer registries vs IMPACT$_{NCD}$ synthetic population estimates. Error bars represent 95% uncertainty intervals.

Mortality predictive validation

Here we validate the IMPACT$_{NCD}$ estimated mortality against the observed mortality in England between 2006 and 2013. We stratified by disease, age, sex and QIMD. Overall, the plots support the argument that IMPACT$_{NCD}$ is capable of translating changes in risk factors prevalence into changes in disease incidence and mortality, rather accurately.

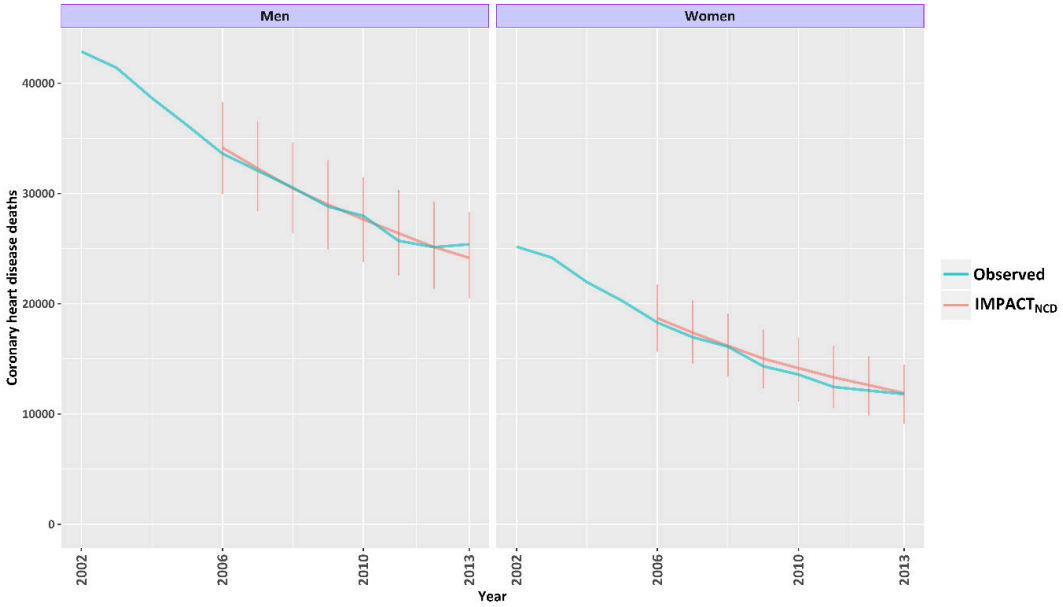

Figure S42 Number of deaths from coronary heart disease in England, by year and sex for ages 30 to 84. Office for National Statistics reported deaths (observed) vs IMPACT$_{NCD}$ estimated

Laverty AA, *et al. J Epidemiol Community Health* 2019; 0:881–887. doi: 10.1136/jech-2018-211749

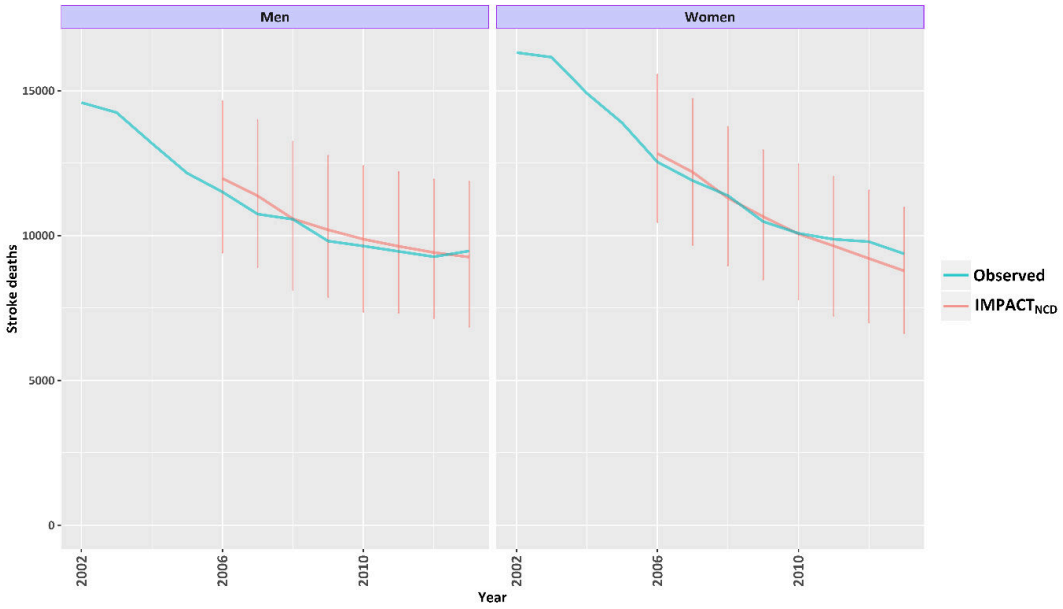

Figure S43 Number of deaths from stroke in England, by year and sex for ages 30 to 84. Office for National Statistics (ONS) reported deaths (observed) vs IMPACT$_{NCD}$ estimated. Observed deaths after 2010 were adjusted to account for changes in the ICD-10 version used by ONS since 201. Error bars represent interquartile ranges.

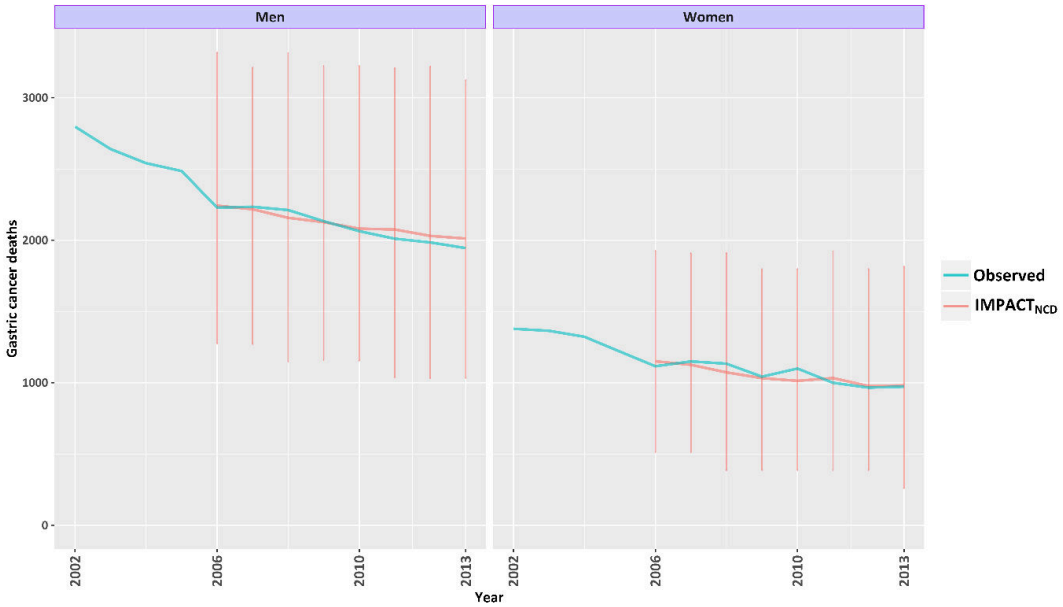

Figure S44 Number of deaths from gastric cancer in England, by year and sex for ages 30 to 84. Office for National Statistics reported deaths (observed) vs IMPACT$_{NCD}$ estimated.

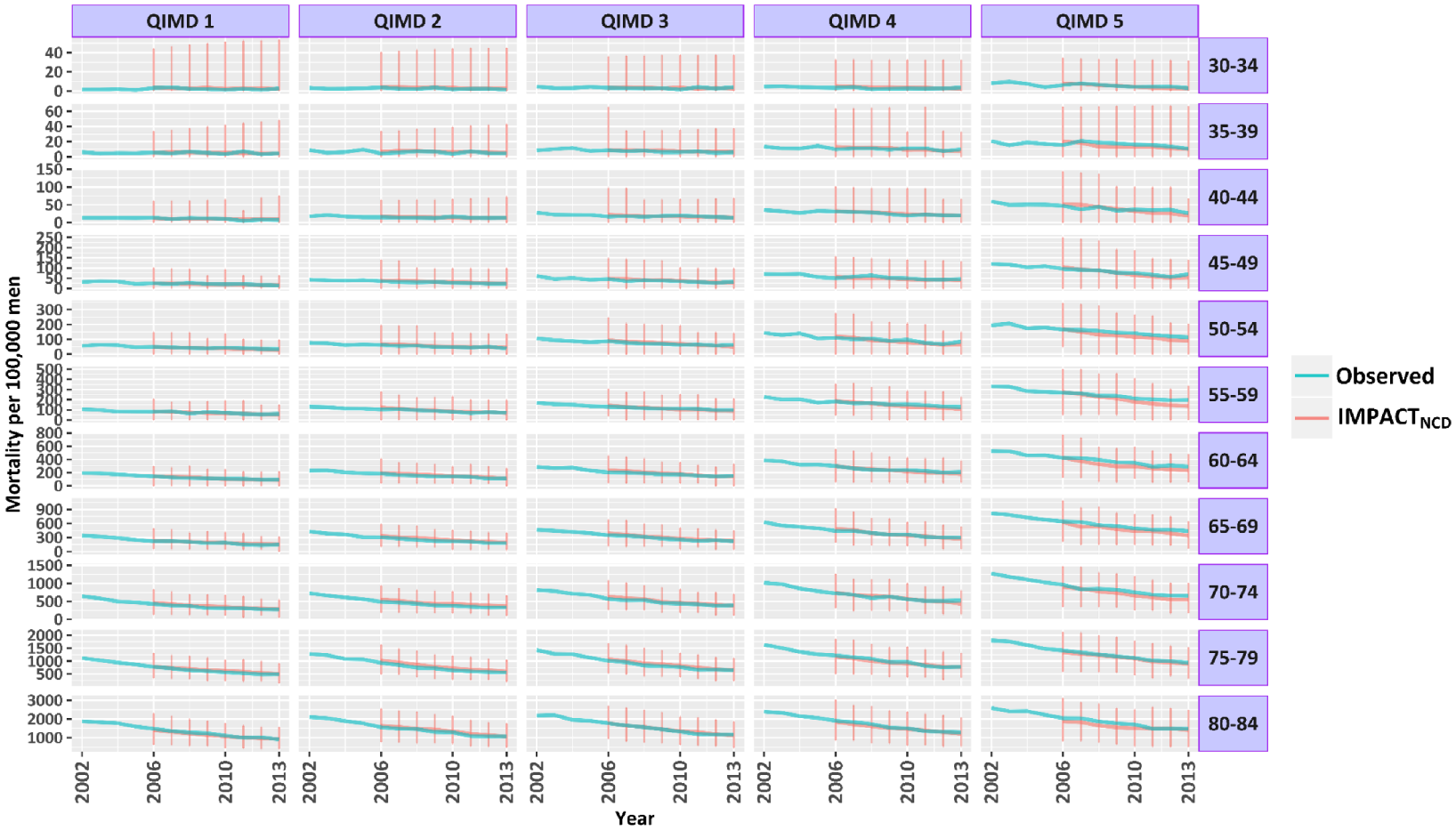

Figure S45 Coronary heart disease mortality (ICD10: I20 – I25) for men by age group and quintile group of index of multiple deprivation (QIMD, 1 = least deprived) between years 2002 and 2013. Observed in the population through mortality registries vs. IMPACT$_{NCD}$ synthetic population estimates. Whiskers represent 95% uncertainty intervals.

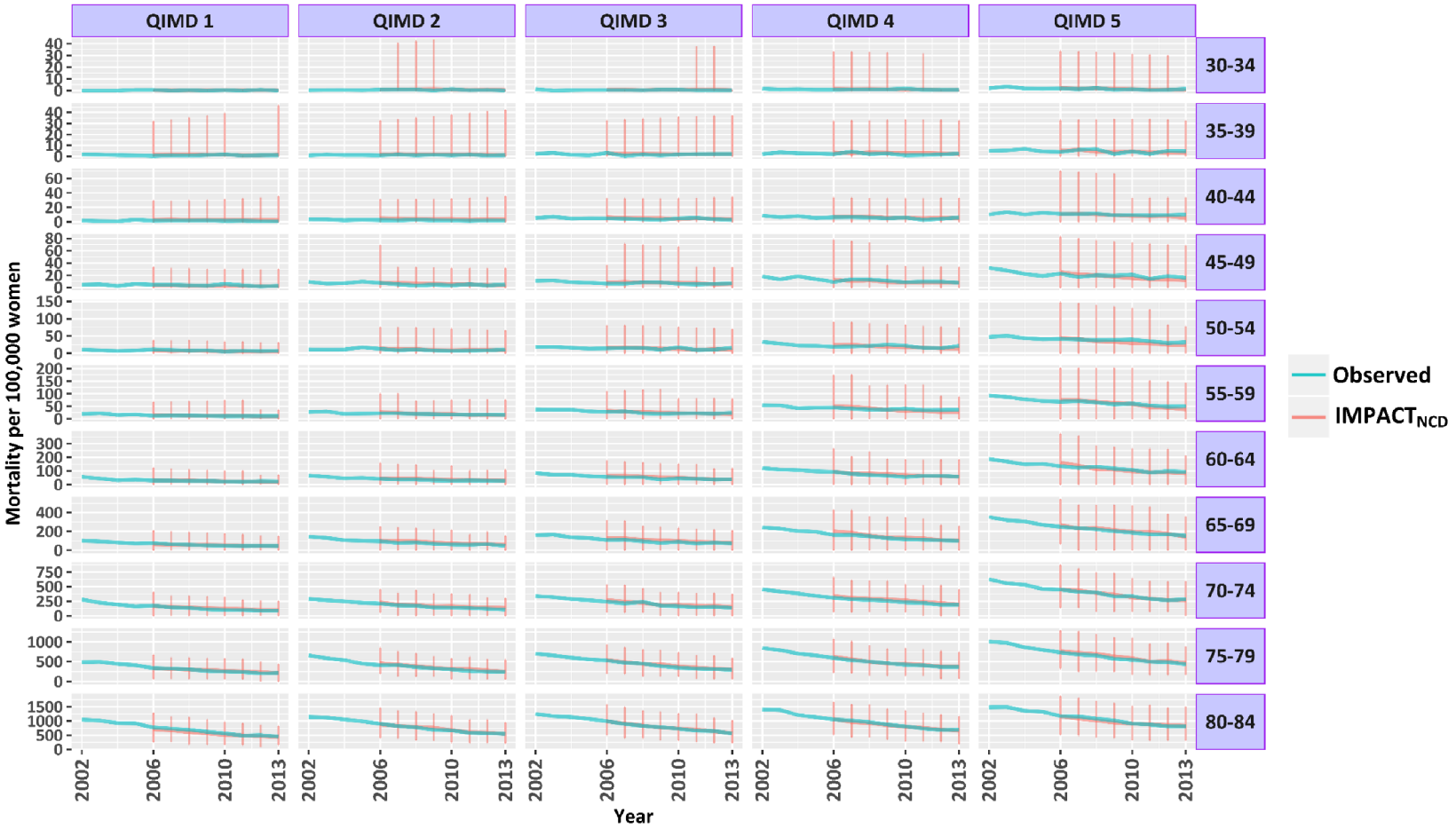

Figure S46 Coronary heart disease mortality (ICD10: I20 – I25) for women by age group and quintile group of index of multiple deprivation (QIMD, 1 = least deprived) between years 2002 and 2013. Observed in the population through mortality registries vs. IMPACT$_{NCD}$ synthetic population estimates. Whiskers represent 95% uncertainty intervals.

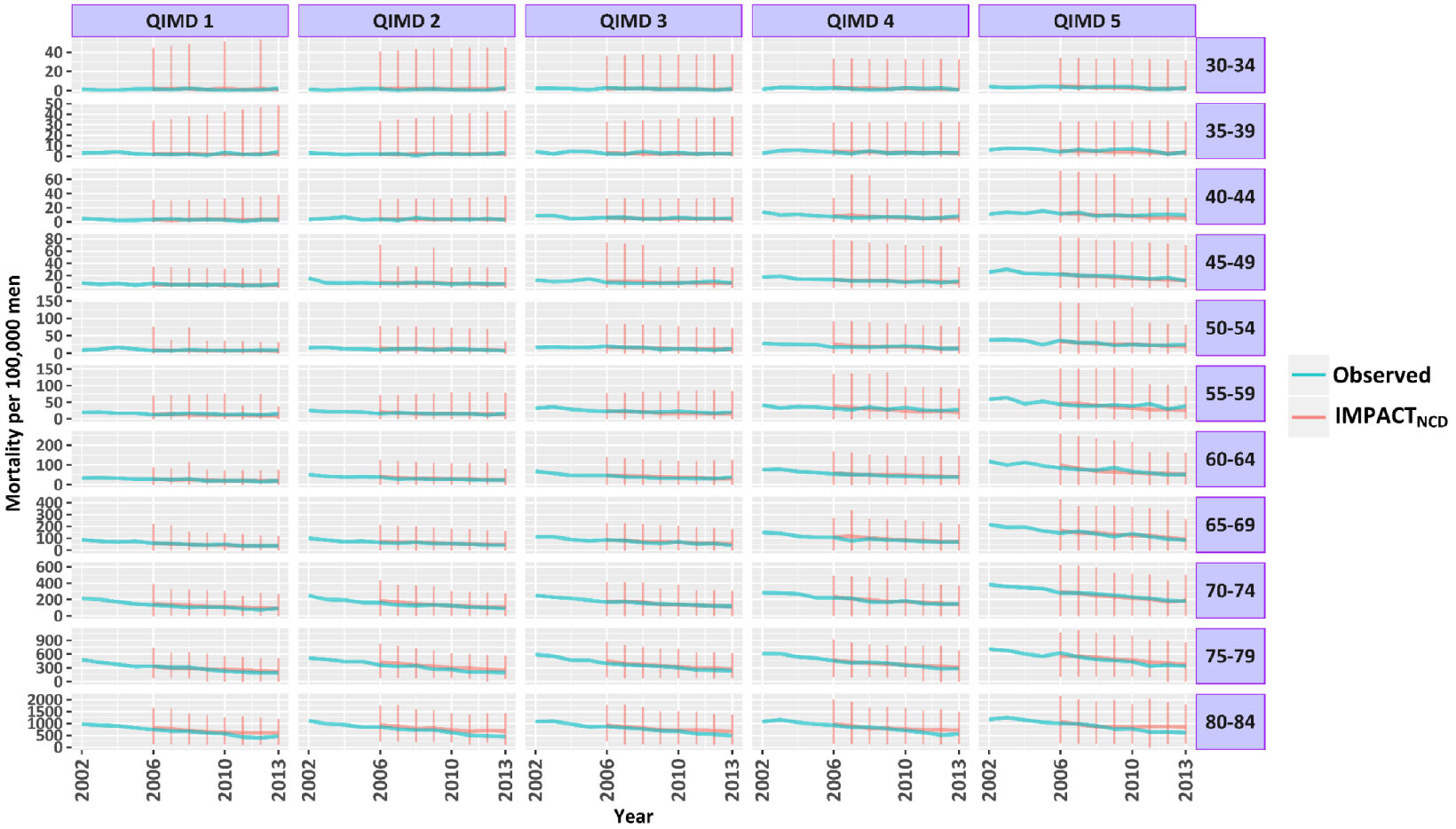

Figure S47 Stroke mortality (ICD10: I60 – I69) for men by age group and quintile group of index of multiple deprivation (QIMD, 1 = least deprived) between years 2002 and 2013. Observed in the population through mortality registries vs. IMPACT$_{NCD}$ synthetic population estimates. Whiskers represent 95% uncertainty intervals.

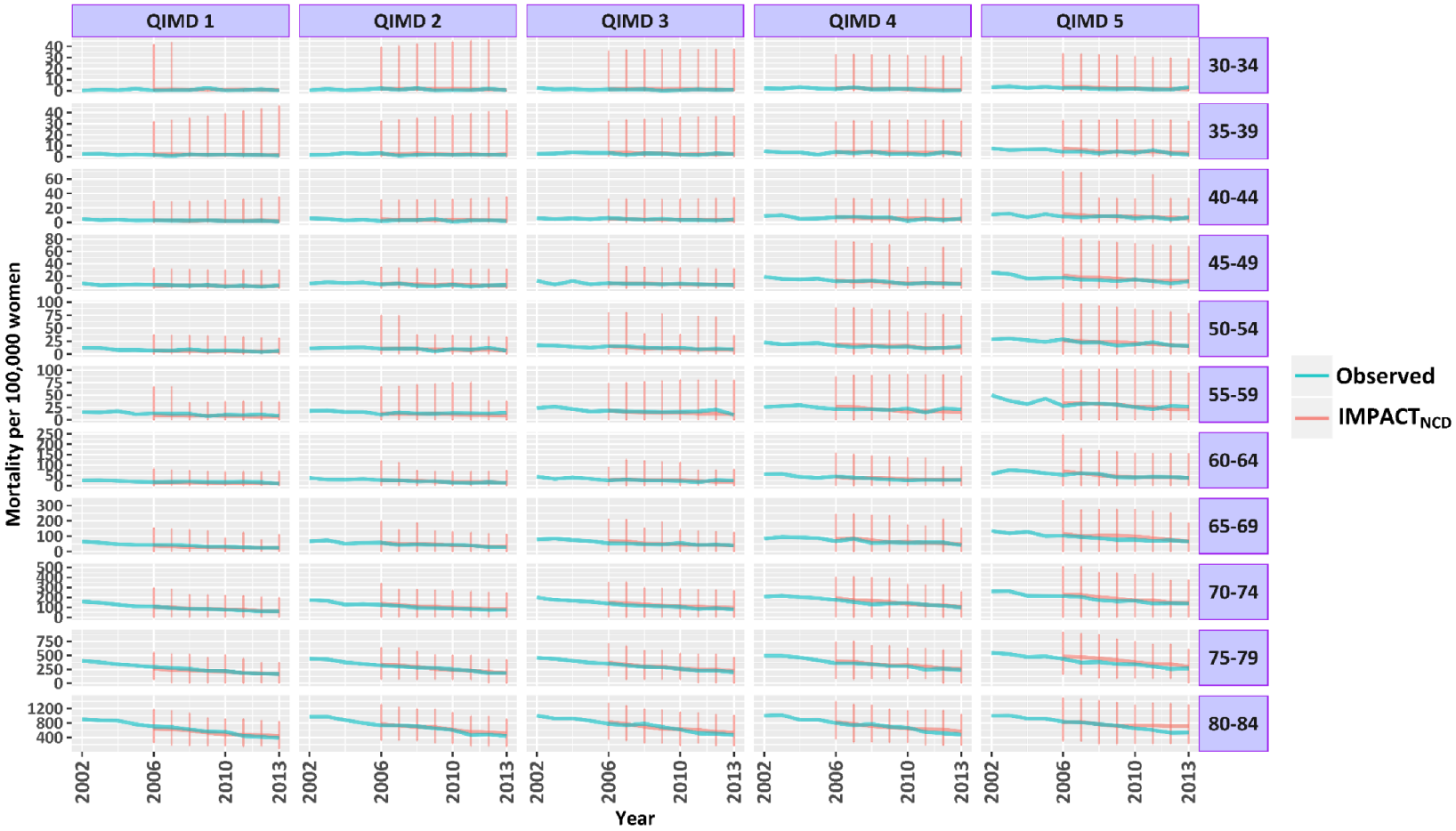

Figure S48 Stroke mortality (ICD10: I60 – I69) for women by age group and quintile group of index of multiple deprivation (QIMD, 1 = least deprived) between years 2002 and 2013. Observed in the population through mortality registries vs. IMPACT$_{NCD}$ synthetic population estimates. Whiskers represent 95% uncertainty intervals.

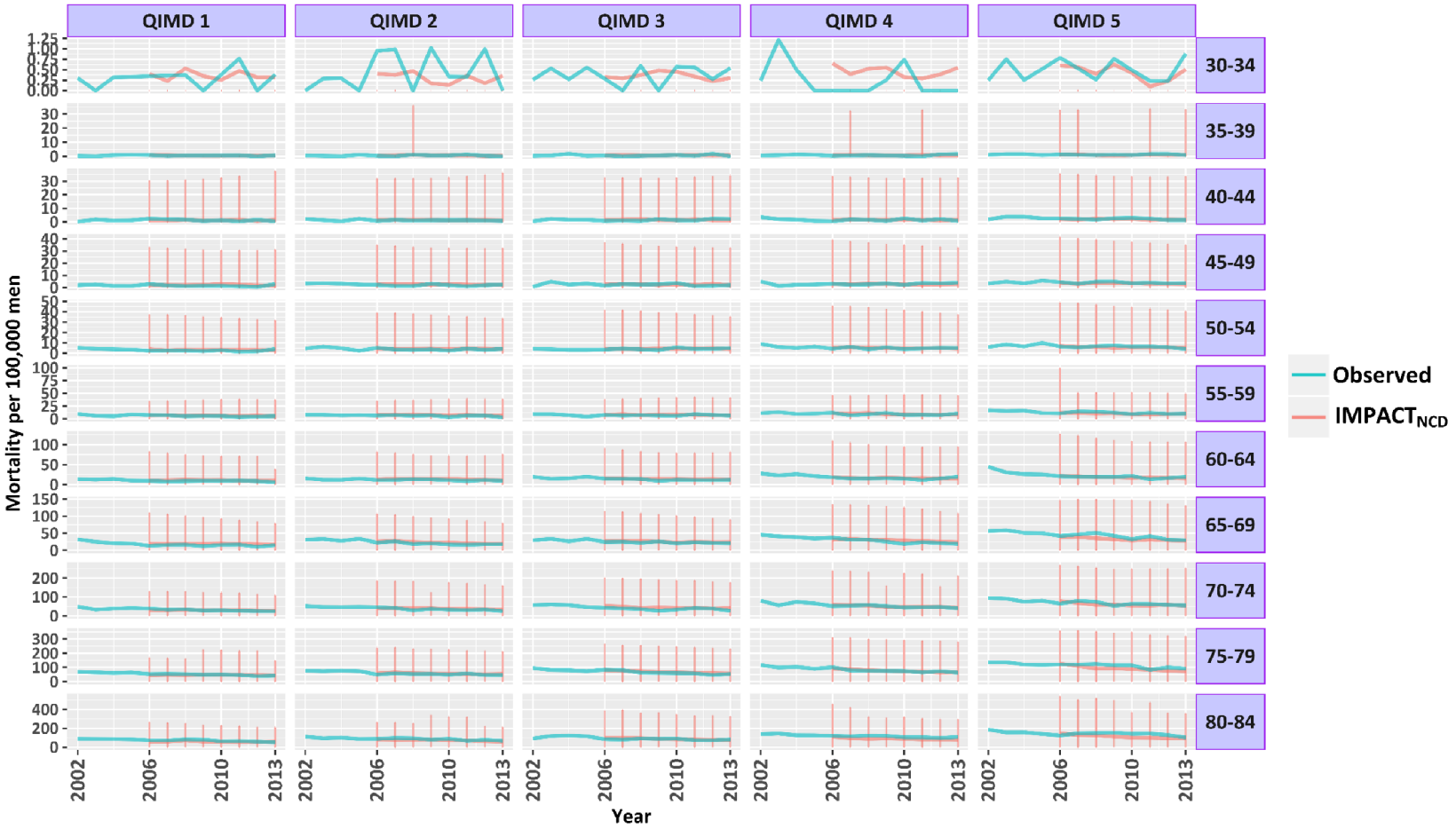

Figure S49 Gastric cancer mortality (ICD10: C16) for men by age group and quintile group of index of multiple deprivation (QIMD, 1 = least deprived) between years 2002 and 2013. Observed in the population through mortality registries vs. IMPACT$_{NCD}$ synthetic population estimates. Whiskers represent 95% uncertainty intervals. Uncertainty intervals could not be estimated for younger age groups due to small number of events.

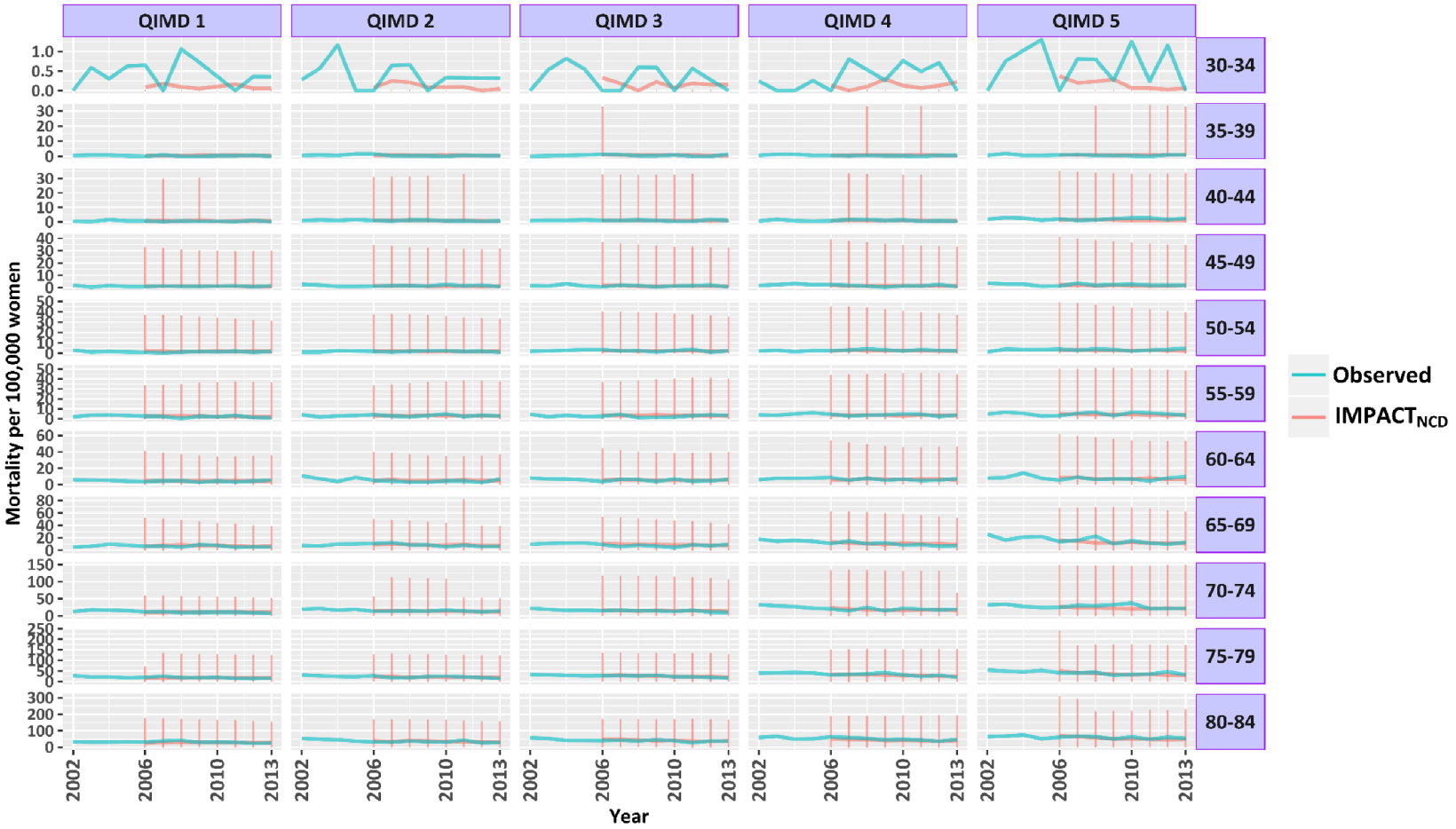

Figure S50 Gastric cancer mortality (ICD10: C16) for women by age group and quintile group of index of multiple deprivation (QIMD, 1 = least deprived) between years 2002 and 2013. Observed in the population through mortality registries vs. IMPACT$_{NCD}$ synthetic population estimates. Whiskers represent 95% uncertainty intervals. Uncertainty intervals could not be estimated for younger age groups due to small number of events.

TABLES

Table S1 IMPACT$_{NCD}$ estimates from the one-way sensitivity analysis, assuming logarithmic salt decline. Additional cases in the Responsibility Deal scenario. Ps denotes the probability of superiority

Disease	Period of exposure	The absolute number of additional cases (IQR)	The absolute number of additional deaths (IQR)	Ps
CVD	2011-2018	6,100 (2,700 to 9,500)	660 (-820 to 2,100)	9.8%
	2019-2025	14,000 (8,600 to 19,000)	3,100 (510 to 5,800)	3.0%
	2011-2025	20,000 (13000 to 26000)	3,700 (820 to 6,500)	1.1%
Gastric cancer	2011-2018	1,000 (100 to 1,800)	360 (-510 to 1,200)	24.0%
	2019-2025	2,100 (920 to 3,200)	1,000 (0 to 2,100)	12.0%
	2011-2025	3,100 (1,500 to 4,700)	6,600 (3,400 to 10,000)	9.4%

Laverty AA, *et al. J Epidemiol Community Health* 2019; 0:881–887. doi: 10.1136/jech-2018-211749

Table S2 IMPACT$_{NCD}$ estimates from the one-way sensitivity analysis, assuming logarithmic salt decline. Additional cases in the Responsibility Deal scenario, by quintile group of Index of Multiple Deprivation.

2011 – 2018 exposure period				
Disease	**QIMD (5 = most deprived)**	**Absolute number of additional cases (IQR)**	**Rate per 100,000 person-years (IQR)**	**Rate per 100,000 new CVD cases (IQR)**
CVD	1	920 (-610 to 2600)	2.3 (-1.3 to 5.7)	940 (-500 to 2300)
	2	1200 (-710 to 3100)	2.1 (-1.3 to 5.9)	760 (-500 to 2200)
	3	1100 (-710 to 3200)	2.2 (-1.7 to 5.7)	800 (-640 to 2100)
	4	1200 (-920 to 3100)	2 (-1.3 to 6.2)	750 (-510 to 2300)
	5	1100 (-710 to 3300)	2.8 (-1.2 to 6.6)	990 (-410 to 2400)
Gastric cancer	1	200 (-410 to 820)	0.38 (-0.75 to 1.5)	1100 (-3800 to 5700)
	2	200 (-410 to 920)	0.38 (-0.75 to 1.7)	0.0071 (-4000 to 5300)
	3	260 (-410 to 920)	0.47 (-0.75 to 1.7)	1100 (-3700 to 6000)
	4	200 (-410 to 820)	0.38 (-0.76 to 1.5)	0.017 (-5000 to 5200)
	5	200 (-410 to 820)	0.39 (-0.79 to 1.6)	0.002 (-4700 to 6000)
2019 – 2025 exposure period				
CVD	1	2400 (-610 to 5700)	5 (-1.3 to 12)	2000 (-560 to 4900)
	2	3000 (-230 to 5600)	6.3 (-0.48 to 12)	2400 (-190 to 4600)
	3	2800 (-510 to 6100)	5.8 (-1.1 to 13)	2100 (-380 to 4900)
	4	3000 (-410 to 6100)	6.3 (-0.86 to 13)	2400 (-310 to 5000)
	5	2700 (-330 to 6000)	5.8 (-0.71 to 13)	2100 (-240 to 4700)
Gastric cancer	1	310 (-410 to 1000)	0.65 (-0.86 to 2.2)	3500 (-5400 to 14000)
	2	410 (-310 to 1200)	0.88 (-0.65 to 2.6)	5000 (-3400 to 14000)
	3	510 (-310 to 1200)	1.1 (-0.65 to 2.6)	5000 (-4300 to 15000)
	4	410 (-310 to 1200)	0.86 (-0.65 to 2.6)	6000 (-4600 to 17000)
	5	410 (-310 to 1100)	0.86 (-0.64 to 2.4)	5100 (-4900 to 16000)

Laverty AA, *et al. J Epidemiol Community Health* 2019; 0:881–887. doi: 10.1136/jech-2018-211749

Table S3 IMPACT$_{NCD}$ estimates from the one-way sensitivity analysis, assuming logarithmic salt decline. Incremental healthcare and workplace productivity loss costs in the Responsibility Deal scenario compared with the counterfactual FSA Trend. Costs in 2018 GBP.

2011 – 2018 exposure period (Total costs in million £100 (£28 to £170, Ps = 16.5%))		
Disease	**Healthcare costs in million (IQR)**	**Workplace productivity costs in million (IQR)**
CVD	£52 (£19 to £87)	£25 (-£1.9 to £50)
Gastric cancer	£19 (-£16 to £54)	£5.5 (-£19 to £31)
2019 – 2025 exposure period (Total costs in million £560 (£320 to £790, Ps = 4.6%))		
CVD	£290 (£150 to £410)	£170 (£59 to £290)
Gastric cancer	£84 (£7.2 to £160)	£14 (-£21 to £50)

Table S4 Long-term condition costs from Charlton et al.[70]

QIMD	Cost per year - one morbidity (£)	Cost ratio relative to least deprived
1 Least deprived	744	1
2	785	1.05
3	797	1.07
4	830	1.11
5 Most deprived	917	1.23

Laverty AA, *et al. J Epidemiol Community Health* 2019; 0:881–887. doi: 10.1136/jech-2018-211749

Table S5 Disease costs used in the model

Disease	Event Type	Original costs	Cost Year	Source	2018 Costs	Table in the report
Ischemic Heart Disease	Non-fatal event – year of event	£2,274	2000	NICE (2015) NG28[79]	£3,196	Table 61
Ischemic Heart Disease	Non-fatal event – subsequent years	£751	2000	NICE (2015) NG28[79]	£1,056	Table 61
Stroke	Non-fatal event – year of event	£8,274	2009	NICE (2010a) PH25[80]	£9,481	Table 4
Stroke	Non-fatal event – subsequent years	£3,660	2009	NICE (2010a) PH25[80]	£4,194	Table 4
Myocardial Infarction	Fatal event – year of event	£1,152	2000	NICE (2015) NG28[79]	£1,627	Table 61
Stroke	Fatal event – year of event	£3,383	2000	NICE (2015) NG28[79]	£4,778	Table 61
Gastric Cancer	Prevalence	£10,513	2006	NHS England Programme Budgeting Data for Upper GI	£13,396	

Table S6 Productivity losses used in the model

Disease	Original costs	Cost per	Cost Year	Source	2018 Costs	Table or page in the report
CHD	£3,240	prevalent case aged <65	2002	Liu et al 2002	£4,903	Table 2
Stroke	£3,484	prevalent case aged <65	2005	Saka et al 2008	£4,575	Table 2
Gastric cancer	£44,580	death (all ages)	2011	Pearce et al. 2016	£48,124	Table 1

Table S7 IMPACT$_{NCD}$ data sources

Parameter	Outcome	Details	Comments	Source
Fertility rates	Births	Principal-assumption fertility projections for England	Stratified by age	National Population Projections, 2012-based Statistical Bulletin [Internet]. Office for National Statistics; 2013 [cited 2014 Nov 11]. Available from: http://www.ons.gov.uk/ons/rel/npp/national-population-projections/2012-based-projections/index.html
Mortality rates	Deaths from non-modelled causes	Mortality and mid-year population estimates for England	Stratified by age, sex, QIMD and cause of death. Years 2002-2013.	Data requested and obtained by the Office for National Statistics. Available from: http://www.ons.gov.uk/ons/about-ons/business-transparency/freedom-of-information/what-can-i-request/published-ad-hoc-data/health/december-2014/number-of-registered-deaths-by-sex--cause--year--the-adjusted-index.xls
Exposure to risk factors (except 24h urine sodium)	Exposure of individuals	Health survey for England	Anonymised, individual-level datasets. Years 2001-2012.	Health survey for England 2001-2012. Data available to researchers from http://ukdataservice.ac.uk/
24h urine sodium	Exposure of individuals	National Diet and Nutrition Survey	Anonymised, individual-level datasets. Years 2001, 2006, 2008, 2011, 2014	NatCen Social Research MRC Elsie Widdowson Laboratory. National Diet and Nutrition Survey: Assessment of Dietary Sodium in Adults, 2006/09 and 2011/15 [Internet]. UK Data Service; 2018 [cited 2018 Aug 10]. Available from: http://discover.ukdataservice.ac.uk/doi?sn=8233#7 Office For National Statistics. Social And Vital Statistics Division And Food Standards Agency. National Diet and Nutrition Survey : Adults Aged 19 to 64 Years, 2000-2001. Colchester, Essex: UK Data Archive; 2005.

Parameter	Outcome	Details	Comments	Source
Relative risk for salt consumption	Gastric cancer incidence (ICD10: C16)	Meta-analysis of 2 cohort studies	Both studies adjusted for age, sex, and smoking. One also adjusted for non green/yellow vegetable intake and the other for education, stomach disorders and history of stomach cancer in the family.	World Cancer Research Fund, American Institute for Cancer Research. Food, nutrition, physical activity, and the prevention of cancer: a global perspective. Washington, DC: WCRF/AICR; 2007. (Figure 4.6.1)
Effect of salt consumption on systolic blood pressure	Systolic blood pressure change	Meta-analysis/meta-regression of 103 trials	Only trials with duration > 7 days were analysed.	Mozaffarian D, Fahimi S, Singh GM, Micha R, Khatibzadeh S, Engell RE, *et al*. Global Sodium Consumption and Death from Cardiovascular Causes. New England Journal of Medicine 2014;371:624–34. (Text S1 in the appendix)
Setting reference level of salt consumption	Ideal salt consumption below which no risk was considered	Evidence from ecologic studies, randomized trials and meta-analyses of prospective cohort studies	Intake levels associated with lowest risk ranged from 1.5 to 6 g/day. The lowest observed mean national intakes were ~3.8 g/day. Thus a PERT (1.5, 3.8, 6) distribution was used.	Mozaffarian D, Fahimi S, Singh GM, Micha R, Khatibzadeh S, Engell RE, *et al*. Global Sodium Consumption and Death from Cardiovascular Causes. New England Journal of Medicine 2014;371:624–34. (Text S4 in the appendix and Table S3)

Parameter	Outcome	Details	Comments	Source
Relative risk for active smoking	CHD and stroke (ICD10: I20 – I25 and I60 – I69)	Re-analysis of American Cancer Society's Cancer Prevention Study II. Prospective cohort study, 6 years of follow up	Stratified by age and sex. Adjusted for age, race, education, marital status, "blue collar" employment in most recent or current job, weekly consumption of vegetables and citrus fruit, vitamin (A, C, and E) use, alcohol use, aspirin use, body mass index, exercise, dietary fat consumption, hypertension and diabetes at baseline.	Ezzati M, Henley SJ, Thun MJ, Lopez AD. Role of Smoking in Global and Regional Cardiovascular Mortality. Circulation 2005;112:489–97. (Table 1 Model B)
	Gastric cancer incidence (ICD10: C16)	EPIC prospective cohort study	Stratified by country. Adjusted for sex, consumption of vegetables, fresh fruits, processed meat, alcohol, body mass index and educational level.	González CA, Pera G, Agudo A, Palli D, Krogh V, Vineis P, *et al*. Smoking and the risk of gastric cancer in the European Prospective Investigation Into Cancer and Nutrition (EPIC). Int J Cancer 2003;107:629–34. (HR of the $\log_2$ of cigarette-years = 1.040)
	Other mortality (except CHD and stroke)	Male British doctors prospective cohort study	Age-standardised	Doll R, Peto R, Boreham J, Sutherland I. Mortality in relation to smoking: 50 years' observations on male British doctors. BMJ 2004;328:1519. (Table 1)
Relative risk for ex-smoking	CHD (ICD10: I20 – I25)	Meta- analysis. Multiple-adjusted pooled estimates from 19 prospective studies	Multiply-adjusted	Huxley RR, Woodward M. Cigarette smoking as a risk factor for coronary heart disease in women compared with men: a systematic review and meta-analysis of prospective cohort studies. The Lancet 2011;378:1297–305. (Web-figure 8)

Parameter	Outcome	Details	Comments	Source
	Stroke (ICD10 I60 – I69)	The Framingham study. Prospective cohort study	Stroke risk decreased significantly by two years and was at the level of non-smokers by five years after cessation of cigarette smoking.	Wolf PA, D'Agostino RB, Kannel WB, Bonita R, Belanger AJ. Cigarette smoking as a risk factor for stroke: The Framingham study. JAMA 1988;259:1025–9.
	Gastric cancer incidence (ICD10: C16)	EPIC prospective cohort study	Stratified by country. Adjusted for sex, consumption of vegetables, fresh fruits, processed meat, alcohol, body mass index and educational level.	González CA, Pera G, Agudo A, Palli D, Krogh V, Vineis P, *et al*. Smoking and the risk of gastric cancer in the European Prospective Investigation Into Cancer and Nutrition (EPIC). Int J Cancer 2003;107:629–34. (Table IV. Continuous RR)
Relative risk for environmental tobacco smoking	CHD (ICD10: I20 – I25)	Meta-analysis of 10 cohort and case-control studies	Adjusted for important CHD risk factors.	He J, Vupputuri S, Allen K, Prerost MR, Hughes J, Whelton PK. Passive Smoking and the Risk of Coronary Heart Disease — A Meta-Analysis of Epidemiologic Studies. N Engl J Med 1999;340:920–6. (Table 3. Adjusted RR)
	Stroke (ICD10 I60 – I69)	Meta-analysis of 20 prospective, case-control and cross-sectional studies	13 studies adjusted for important CHD risk factors. The overall effect from all 20 studies was used.	Oono IP, Mackay DF, Pell JP. Meta-analysis of the association between second hand smoke exposure and stroke. J Public Health 2011;33:496–502. (Figure 1)

Parameter	Outcome	Details	Comments	Source
Relative risk for systolic blood pressure	CHD and stroke (ICD10: I20 – I25 and I60 – I69)	Meta-analysis of individual data from 61 prospective studies	Stratified by age and sex. Adjusted for regression dilution and total blood cholesterol and, where available, lipid fractions (HDL and non-HDL cholesterol), diabetes, weight, alcohol consumption, and smoking at baseline.	Age-specific relevance of usual blood pressure to vascular mortality: a meta-analysis of individual data for one million adults in 61 prospective studies. The Lancet 2002;360:1903–13. (Figures 3 and 5)
Relative risk for total cholesterol	CHD and stroke (ICD10: I20 – I25 and I60 – I69)	Meta-analysis of individual data from 61 prospective studies	Stratified by age and sex. Adjusted for regression dilution and age, sex, study, systolic blood pressure and smoking.	Prospective Studies Collaboration. Blood cholesterol and vascular mortality by age, sex, and blood pressure: a meta-analysis of individual data from 61 prospective studies with 55 000 vascular deaths. The Lancet 2007;370:1829–39. (Web-table 6 fully adjusted and Figure 3)
Relative risk for body mass index	CHD and stroke (ICD10: I20 – I25 and I60 – I69)	Meta-analysis of 58 prospective studies	Stratified by age. Adjusted for age, sex, smoking status, systolic blood pressure, history of diabetes, and total and HDL cholesterol.	The Emerging Risk Factors Collaboration. Separate and combined associations of body-mass index and abdominal adiposity with cardiovascular disease: collaborative analysis of 58 prospective studies. The Lancet 2011;377:1085–95. (Table 1 and Figure 2)
	Gastric cancer incidence (ICD10: C16)	Meta-analysis of 7 studies	Non-linear dose-response meta-analysis for risk of cardia gastric cancer. Adjusted for age, sex, and smoking.	World Cancer Research Fund International/American Institute for Cancer Research. Continuous Update Project report: diet, nutrition, physical activity and stomach cancer. AICR/WCRF 2016. wcrf.org/stomach-cancer-2016 (Table 8 p37).
Relative risk for diabetes mellitus	CHD and stroke (ICD10: I20 – I25 and I60 – I69)	Meta-analysis of 102 prospective studies	Stratified by age. Adjusted for age, smoking status, body-mass index, and systolic blood pressure.	The Emerging Risk Factors Collaboration. Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease: a collaborative meta-analysis of 102 prospective studies. The Lancet 2010;375:2215–22. (Figure 2)

Parameter	Outcome	Details	Comments	Source
	Other mortality (except CHD and stroke)	DECODE. A collaborative prospective study of 22 cohorts in Europe	Adjusted for BMI, blood pressure, smoking and serum cholesterol.	The DECODE Study Group. Is the current definition for diabetes relevant to mortality risk from all causes and cardiovascular and noncardiovascular diseases? Diabetes Care 2003;26:688–96.
Relative risk for physical activity	CHD and stroke (ICD10: I20 – I25 and I60 – I69)	Meta-analysis of 18 cohort studies for CHD and 8 cohort studies for ischaemic stroke	Stratified by age and sex. Adjusted for measurement error, age, sex, smoking, blood pressure and cholesterol.	Bull FC, Armstrong TP, Dixon T, Ham S, Neiman A, Pratt M. Comparative quantification of health risks. Chapter 10: physical inactivity. Geneva: World Health Organisation; 2004. (Tables 10.19 and 10.20)
Relative risk for fruit and vegetable consumption	CHD (ICD10: I20 – I25)	Meta-analysis of 9 cohort studies	RR per portion of F&V. Multiply-adjusted.	Dauchet L, Amouyel P, Hercberg S, Dallongeville J. Fruit and Vegetable Consumption and Risk of Coronary Heart Disease: A Meta-Analysis of Cohort Studies. J Nutr 2006;136:2588–93.
	Stroke (ICD10: I60 – I69)	Meta-analysis of 7 cohort studies	RR per portion of F&V. Multiply-adjusted.	Dauchet L, Amouyel P, Dallongeville J. Fruit and vegetable consumption and risk of stroke A meta-analysis of cohort studies. Neurology 2005;65:1193–7.
	Gastric cancer incidence (ICD10: C16)	Reanalysis of the Netherlands Cohort study	Stratified by age group. Estimates are based on the Netherlands Cohort study. Adjusted for age, sex, smoking, education, stomach disorders, and family history of stomach cancer. We considered a risk only for <2 portions/day consumption.[81]	Lock K, Pomerleau J, Causer L, McKee M. Comparative quantification of health risks. Chapter 9: Low fruit and vegetable consumption [Internet]. Geneva: World Health Organisation; 2004. Available from: http://www.who.int/publications/cra/en/ (Table 9.28)

Table S8 IMPACT$_{NCD}$ assumptions and limitations

Population module	Immigration is not considered.
	Social mobility is not considered.
	Quintile groups of index of multiple deprivation (QIMD) is a relative marker of (area) deprivation with several versions since 2003. We considered all version of QIMD identical.
	We assume all salt that is consumed is excreted from urine and all urine sodium origins from salt consumption.
	We assume that the surveys used, are truly representative of the population. For example, the adjustments for selection bias in the Health Surveys for England are perfect.
	We assumed the decline in 24hour-urine sodium by age that we observed in the 24-hour urine surveys will continue for ages older than 64 years, based on the Health Survey for England spot urine sodium.
Disease module	We assume multiplicative risk effects.
	We assume log-linear dose-response for the continuous risk factors.
	We assume that the effects of the risk factors on incidence and mortality are equal and risk factors are not modifying survival.
	We assume 5-year mean lag time for CVD and 8-year for GCa (except for the cumulative effect of smoking on GCa where lag was assumed similar to CVD one).
	We assume 100% risk reversibility.
	We assume that trends in disease incidence are attributable only to trends of the relevant modelled risk factors.
	Only well accepted associations between upstream and downstream risk factors that have been observed in longitudinal studies are considered. However, the magnitudes of the associations are extracted from a series of nationally representative cross-sectional surveys (Health Survey for England).
	For GCa, we assume that survival 10 years after diagnosis equals remission.
Policy module	We assume that the change in salt decline after 2011 is fully attributed to the Responsibility.
	We assume that the transition from the FSA led salt reduction strategy to the Responsibility Deal did not alter the equity impact of the policy.

Table S9 Distributions that were used as inputs for the simulations. Numbers are rounded

Variable	Sex	Ages	Distribution
Relative risks of relevant risk factors for CHD			
Active smoking[68 table 1 model B]	Men	30 - 44	Log-Normal (mean = ln(5.51), sd = ln(12.3 / 5.51) / 1.96)
		45 - 59	Log-Normal (mean = ln(3.04), sd = ln(3.48 / 3.04) / 1.96)
		60 - 69	Log-Normal (mean = ln(1.88), sd = ln(2.08 / 1.88) / 1.96)
		70 - 79	Log-Normal (mean = ln(1.44), sd = ln(1.63 / 1.44) / 1.96)
	Women	30 - 44	Log-Normal (mean = ln(2.26), sd = ln(6.14 / 2.26) / 1.96)
		45 - 59	Log-Normal (mean = ln(3.78), sd = ln(4.62 / 3.78) / 1.96)
		60 - 69	Log-Normal (mean = ln(2.53), sd = ln(2.87 / 2.53) / 1.96)
		70 - 79	Log-Normal (mean = ln(1.68), sd = ln(1.93 / 1.68) / 1.96)
		80 - 84	Log-Normal (mean = ln(1.38), sd = ln(1.77 / 1.38) / 1.96)
Ex-Smoking[69 web-figure 8]	Men	30 - 84	Log-Normal (mean = ln(1.25), sd = ln(1.32 / 1.25) / 1.96)
	Women	30 - 84	Log-Normal (mean = ln(1.2), sd = ln(1.34 / 1.2) / 1.96)
ETS[70 table 3 adjusted RR]	Both	30 - 84	Log-Normal (mean = ln(1.26), sd = ln(1.38 / 1.26) / 1.96)
SBP[71 figure 5]	Men	30 - 49	Log-Normal (mean = ln(0.5), sd = ln(0.54 / 0.5) / 1.96)
		50 - 59	Log-Normal (mean = ln(0.5), sd = ln(0.52 / 0.5) / 1.96)
		60 - 69	Log-Normal (mean = ln(0.55), sd = ln(0.57 / 0.55) / 1.96)

Variable	Sex	Ages	Distribution
		70 - 74	Log-Normal (mean = ln(0.62), sd = ln(0.64 / 0.62) / 1.96)
		80 - 84	Log-Normal (mean = ln(0.69), sd = ln(0.73 / 0.69) / 1.96)
	Women	30 - 49	Log-Normal (mean = ln(0.4), sd = ln(0.49 / 0.4) / 1.96)
		50 - 59	Log-Normal (mean = ln(0.49), sd = ln(0.54 / 0.49) / 1.96)
		60 - 69	Log-Normal (mean = ln(0.5), sd = ln(0.61 / 0.5) / 1.96)
		70 - 74	Log-Normal (mean = ln(0.55), sd = ln(0.58 / 0.55) / 1.96)
		80 - 84	Log-Normal (mean = ln(0.64), sd = ln(0.68 / 0.64) / 1.96)
TC[72 web-table 6]	Both	30 - 49	Log-Normal (mean = ln(0.49), sd = ln(0.52 / 0.49) / 1.96)
		50 - 59	Log-Normal (mean = ln(0.62), sd = ln(0.65 / 0.62) / 1.96)
		60 - 69	Log-Normal (mean = ln(0.74), sd = ln(0.76 / 0.74) / 1.96)
		70 - 74	Log-Normal (mean = ln(0.84), sd = ln(0.86 / 0.84) / 1.96)
		80 - 84	Log-Normal (mean = ln(0.87), sd = ln(0.9 / 0.87) / 1.96)
BMI[73 table 1 and figure 2]	Both	30 - 59	Log-Normal (mean = ln(1.21), sd = ln(1.28 / 1.21) / 1.96)
		60 - 69	Log-Normal (mean = ln(1.06), sd = ln(1.12 / 1.06) / 1.96)
Diabetes[74 figure 2]	Both	40 - 59	Log-Normal (mean = ln(2.51), sd = ln(2.8/ 2.51) / 1.96)
		60 - 69	Log-Normal (mean = ln(2.01), sd = ln(2.26/ 2.01) / 1.96)
		70 - 84	Log-Normal (mean = ln(1.78), sd = ln(2.05/ 1.78) / 1.96)

Variable	Sex	Ages	Distribution
PA[75 table 10.19]	Both	30 - 69	No active days: Log-Normal (mean = ln(1.71), sd = ln(1.85/ 1.71) / 1.96)
			1 – 4 active days: Log-Normal (mean = ln(1.44), sd = ln(1.62/ 1.44) / 1.96)
		70 - 79	No active days: Log-Normal (mean = ln(1.5), sd = ln(1.61/ 1.5) / 1.96)
			1 – 4 active days: Log-Normal (mean = ln(1.31), sd = ln(1.48/ 1.31) / 1.96)
		80 - 84	No active days: Log-Normal (mean = ln(1.4), sd = ln(1.41/ 1.4) / 1.96)
			1 – 4 active days: Log-Normal (mean = ln(1.2), sd = ln(1.35/ 1.2) / 1.96)
F&V[76]			Log-Normal (mean = ln(0.96), sd = ln(1.0.99/ 0.96) / 1.96)
Relative risks of relevant risk factors for stroke			
Active smoking[68 table 1 model B]	Men	30 - 59	Log-Normal (mean = ln(3.12), sd = ln(4.64 / 3.12) / 1.96)
		60 - 69	Log-Normal (mean = ln(1.87), sd = ln(2.44 / 1.87) / 1.96)
		70 - 79	Log-Normal (mean = ln(1.39), sd = ln(1.77 / 1.39) / 1.96)
	Women	30 - 59	Log-Normal (mean = ln(4.61), sd = ln(6.37 / 4.61) / 1.96)
		60 - 69	Log-Normal (mean = ln(2.81), sd = ln(3.58 / 2.81) / 1.96)
		70 - 79	Log-Normal (mean = ln(1.95), sd = ln(2.45 / 1.95) / 1.96)
ETS[77 figure 1]	Both	30 - 84	Log-Normal (mean = ln(1.25), sd = ln(1.38 / 1.25) / 1.96)
SBP[71 figure 3]	Men	30 - 49	Log-Normal (mean = ln(0.33), sd = ln(0.38 / 0.33) / 1.96)
		50 - 59	Log-Normal (mean = ln(0.34), sd = ln(0.37 / 0.34) / 1.96)

Variable	Sex	Ages	Distribution
		60 - 69	Log-Normal (mean = ln(0.41), sd = ln(0.44 / 0.41) / 1.96)
		70 - 74	Log-Normal (mean = ln(0.48), sd = ln(0.51 / 0.48) / 1.96)
		80 - 84	Log-Normal (mean = ln(0.68), sd = ln(0.75 / 0.68) / 1.96)
	Women	30 - 49	Log-Normal (mean = ln(0.41), sd = ln(0.49 / 0.41) / 1.96)
		50 - 59	Log-Normal (mean = ln(0.45), sd = ln(0.5 / 0.45) / 1.96)
		60 - 69	Log-Normal (mean = ln(0.47), sd = ln(0.51 / 0.47) / 1.96)
		70 - 74	Log-Normal (mean = ln(0.53), sd = ln(0.56 / 0.53) / 1.96)
		80 - 84	Log-Normal (mean = ln(0.65), sd = ln(0.71 / 0.65) / 1.96)
TC[72 figure 3]	Both	40 - 49	Log-Normal (mean = ln(0.87), sd = ln(1 / 0.87) / 1.96)
		50 - 59	Log-Normal (mean = ln(0.91), sd = ln(0.97 / 0.91) / 1.96)
		60 - 69	Log-Normal (mean = ln(0.93), sd = ln(0.97 / 0.93) / 1.96)
BMI[73 table 1 and figure 2]	Both	30 - 59	Log-Normal (mean = ln(1.18), sd = ln(1.26 / 1.18) / 1.96)
		60 - 69	Log-Normal (mean = ln(1.08), sd = ln(1.15 / 1.08) / 1.96)
Diabetes[74 figure 2]	Both	40 - 59	Log-Normal (mean = ln(3.74), sd = ln(4.58/ 3.74) / 1.96)
		60 - 69	Log-Normal (mean = ln(2.06), sd = ln(2.58/ 2.06) / 1.96)
		70 - 84	Log-Normal (mean = ln(1.8), sd = ln(2.27/ 1.8) / 1.96)
PA[75 table 10.20]	Both	30 - 69	No active days: Log-Normal (mean = ln(1.53), sd = ln(1.79/ 1.53 / 1.96)

Variable	Sex	Ages	Distribution
		70 - 79	No active days: Log-Normal (mean = ln(1.38), sd = ln(1.6/ 1.38) / 1.96)
		80 - 84	No active days: Log-Normal (mean = ln(1.24), sd = ln(1.45/ 1.24) / 1.96)
F&V[78]			Log-Normal (mean = ln(0.95), sd = ln(0.97/ 0.95) / 1.96)
Relative risks of relevant risk factors for GCa			
Active smoking (duration in years)[82 table III]	Both	30 - 84	Normal (mean = 0.03, sd = 0.002)
Ex-smoking (years since cessation)[82 table IV]	Both	30 - 84	Log-Normal (mean = ln(0.96), sd = ln(1/ 0.96) / 1.96)
BMI[81 table 8]	Both	30 - 84	Normal (mean and sd is a function of BMI)
F&V[83 table 9.28]	Both	30 - 69	Log-Normal (mean = ln(0.94), sd = ln(1/ 0.94) / 1.96)
	Both	70 - 79	Log-Normal (mean = ln(0.96), sd = ln(1/ 0.96) / 1.96)
	Both	80 - 84	Log-Normal (mean = ln(0.97), sd = ln(1/ 0.97) / 1.96)
Salt[84]	Both	30 - 84	Log-Normal (mean = ln(1.08), sd = ln(1.08/ 1) / 1.96)
Other inputs			
CVD lag time	Both	30 - 84	1 + Binomial(n = 9, p = (5-1)/9)
GCa lag time	Both	30 - 84	1 + Binomial(n = 9, p = (8-1)/9)
Optimal salt consumption[85 appendix Text S4]	Both	30 - 84	PERT(min = 1.5, mode = 3.8, max = 6, shape = 4)
Stricter salt policy target	Both	30 - 84	PERT(min = 5.8, mode = 6, max = 7, shape = 4)

REFERENCES

1. Williamson P. The role of the International Journal of Microsimulation. Int J Microsimulation. 2007;1:1–2.

2. Zucchelli E, Jones AM, Rice N. The evaluation of health policies through dynamic microsimulation methods. Int J Microsimulation. 2012;5:2–20.

3. Department for Communities and Local Government. English indices of deprivation 2010 - Publications - GOV.UK [Internet]. 2011 [cited 2014 Aug 26]. Available from: https://www.gov.uk/government/statistics/english-indices-of-deprivation-2010

4. R Core Team. R: a language and environment for statistical computing [Internet]. R Foundation for Statistical Computing; 2017. Available from: http://www.R-project.org/

5. Dowle M, Srinivasan A, Lianoglou S, Srinivasan A. data.table: Extension of data.frame [Internet]. 2018. Available from: https://github.com/Rdatatable/data.table/

6. Revolution Analytics, Weston S. foreach: Foreach looping construct for R [Internet]. 2014. Available from: http://CRAN.R-project.org/package=foreach

7. Gaujoux R. doRNG: Generic Reproducible Parallel Backend for foreach Loops [Internet]. 2014. Available from: http://CRAN.R-project.org/package=doRNG

8. L'Ecuyer P. Good Parameters and Implementations for Combined Multiple Recursive Random Number Generators. Oper Res. 1999;47:159–64.

9. Alfons A, Kraft S, Templ M, Filzmoser P. Simulation of close-to-reality population data for household surveys with application to EU-SILC. Stat Methods Appl. 2011;20:383–407.

10. Dahlgren G, Whitehead M. Policies and strategies to promote social equity in health [Internet]. Copenhagen: World Health Organisation; 1992 [cited 2015 May 5]. Available from: http://core.ac.uk/download/pdf/6472456.pdf

11. National Centre for Social Research, University College London. Department of Epidemiology and Public Health. Health Survey for England, 2001 [computer file]. 3rd Edition. Colchester, Essex: UK Data Archive [distributor] [Internet]. 2010 [cited 2014 May 1]. Available from: http://dx.doi.org/10.5255/UKDA-SN-4628-1

12. Alfons A, Kraft S. simPopulation: simulation of synthetic populations for surveys based on sample data [Internet]. 2013. Available from: http://CRAN.R-project.org/package=simPopulation

13. Giles Horsfield SSD. Chapter 3: Equivalised Income [Internet]. Office for National Statistics. 2012 [cited 2015 May 4]. Available from: http://www.ons.gov.uk/ons/rel/family-spending/family-spending/family-spending-2012-edition/art-chapter-3--equivalised-income.html

14. Economic Labour and Social Analysis Office for National Statistics. SOC2010 volume 3: the National Statistics socio-economic classification (NS-SEC rebased on the SOC2010) [Internet]. Office for National Statistics. 2010 [cited 2015 Jan 6]. Available from: http://www.ons.gov.uk/ons/guide-method/classifications/current-standard-classifications/soc2010/soc2010-volume-3-ns-sec--rebased-on-soc2010--user-manual/index.html

15. Ji C, Sykes L, Paul C, Dary O, Legetic B, Campbell NRC, et al. Systematic review of studies comparing 24-hour and spot urine collections for estimating population salt intake. Rev Panam Salud Pública. 2012;32:307–15.

16. Ji C, Miller MA, Venezia A, Strazzullo P, Cappuccio FP. Comparisons of spot vs 24-h urine samples for estimating population salt intake: Validation study in two independent samples of adults in Britain and Italy. Nutr Metab Cardiovasc Dis. 2014;24:140–7.

17. Mulrow CD, Chiquette E, Angel L, Cornell J, Summerbell C, Anagnostelis B, et al. Dieting to reduce body weight for controlling hypertension in adults. Cochrane Database Syst Rev. 2000;CD000484.

18. Goldstein DJ. Beneficial health effects of modest weight loss. Int J Obes Relat Metab Disord J Int Assoc Study Obes. 1992;16:397–415.

19. MacMahon S, Cutler J, Brittain E, Higgins M. Obesity and hypertension: epidemiological and clinical issues. Eur Heart J. 1987;8 Suppl B:57–70.

20. Dyer AR, Elliott P. The INTERSALT study: relations of body mass index to blood pressure. INTERSALT Co-operative Research Group. J Hum Hypertens. 1989;3:299–308.

21. Dattilo AM, Kris-Etherton PM. Effects of weight reduction on blood lipids and lipoproteins: a meta-analysis. Am J Clin Nutr. 1992;56:320–8.

22. Kahn SE, Hull RL, Utzschneider KM. Mechanisms linking obesity to insulin resistance and type 2 diabetes. Nature. 2006;444:840–6.

23. Lahti-Koski M, Pietinen P, Heliövaara M, Vartiainen E. Associations of body mass index and obesity with physical activity, food choices, alcohol intake, and smoking in the 1982–1997 FINRISK Studies. Am J Clin Nutr. 2002;75:809–17.

24. DiPietro L. Physical activity in the prevention of obesity: current evidence and research issues. Med Sci Sports Exerc. 1999;31:S542-546.

25. Fogelholm M, Kukkonen-Harjula K. Does physical activity prevent weight gain--a systematic review. Obes Rev Off J Int Assoc Study Obes. 2000;1:95–111.

26. Dochi M, Sakata K, Oishi M, Tanaka K, Kobayashi E, Suwazono Y. Smoking as an independent risk factor for hypertension: a 14-year longitudinal study in male Japanese workers. Tohoku J Exp Med. 2009;217:37–43.

27. Bowman TS, Gaziano JM, Buring JE, Sesso HD. A Prospective Study of Cigarette Smoking and Risk of Incident Hypertension in Women. J Am Coll Cardiol. 2007;50:2085–92.

28. National Centre for Social Research, University College London. Department of Epidemiology and Public Health. Health Survey for England, 2002 [computer file]. 2nd Edition. Colchester, Essex: UK Data Archive [distributor] [Internet]. 2010 [cited 2014 May 1]. Available from: http://dx.doi.org/10.5255/UKDA-SN-4912-1

29. National Centre for Social Research, University College London. Department of Epidemiology and Public Health. Health Survey for England, 2003 [computer file]. 2nd Edition. Colchester, Essex: UK Data Archive [distributor] [Internet]. 2010 [cited 2014 May 1]. Available from: http://dx.doi.org/10.5255/UKDA-SN-5098-1

30. National Centre for Social Research, University College London. Department of Epidemiology and Public Health. Health Survey for England, 2004 [computer file]. 2nd Edition. Colchester, Essex: UK Data Archive [distributor] [Internet]. 2010 [cited 2014 May 1]. Available from: http://dx.doi.org/10.5255/UKDA-SN-5439-1

31. National Centre for Social Research, University College London. Department of Epidemiology and Public Health. Health Survey for England, 2005 [computer file]. 3rd Edition. Colchester, Essex: UK Data Archive [distributor] [Internet]. 2011 [cited 2014 May 1]. Available from: http://dx.doi.org/10.5255/UKDA-SN-5675-1

32. National Centre for Social Research, University College London. Department of Epidemiology and Public Health. Health Survey for England, 2006 [computer file]. 4th Edition. Colchester, Essex: UK Data Archive [distributor] [Internet]. 2011 [cited 2014 May 1]. Available from: http://dx.doi.org/10.5255/UKDA-SN-5809-1

33. National Centre for Social Research, University College London. Department of Epidemiology and Public Health. Health Survey for England, 2007 [computer file]. 2nd Edition. Colchester, Essex: UK Data Archive [distributor] [Internet]. 2010 [cited 2014 May 1]. Available from: http://dx.doi.org/10.5255/UKDA-SN-6112-1

34. National Centre for Social Research, University College London. Department of Epidemiology and Public Health. Health Survey for England, 2008 [computer file]. 4th Edition. Colchester, Essex: UK Data Archive [distributor] [Internet]. 2013 [cited 2014 May 1]. Available from: http://dx.doi.org/10.5255/UKDA-SN-6397-2

35. National Centre for Social Research, University College London. Department of Epidemiology and Public Health. Health Survey for England, 2009 [computer file]. 2nd Edition. Colchester, Essex: UK Data Archive [distributor] [Internet]. 2011 [cited 2014 May 1]. Available from: http://dx.doi.org/10.5255/UKDA-SN-6732-1

36. NatCen Social Research, Royal Free, University College London. Department of Epidemiology and Public Health. Health Survey for England, 2010 [computer file]. 2nd Edition. Colchester, Essex: UK Data Archive [distributor] [Internet]. 2012 [cited 2014 May 1]. Available from: http://dx.doi.org/10.5255/UKDA-SN-6986-2

37. NatCen Social Research, University College London. Department of Epidemiology and Public Health. Health Survey for England, 2011 [computer file]. Colchester, Essex: UK Data Archive [distributor] [Internet]. 2013 [cited 2014 May 1]. Available from: http://dx.doi.org/10.5255/UKDA-SN-7260-1

38. NatCen Social Research, University College London. Department of Epidemiology and Public Health. Health Survey for England, 2012 [computer file]. Colchester, Essex: UK Data Archive [distributor] [Internet]. 2014 [cited 2014 May 1]. Available from: http://dx.doi.org/10.5255/UKDA-SN-7480-1

39. NatCen Social Research MRC Elsie Widdowson Laboratory. National Diet and Nutrition Survey: Assessment of Dietary Sodium in Adults, 2006/09 and 2011/15 [Internet]. UK Data Service; 2018 [cited 2018 Aug 10]. Available from: http://discover.ukdataservice.ac.uk/doi?sn=8233#7

40. Office For National Statistics. Social And Vital Statistics Division And Food Standards Agency. National Diet and Nutrition Survey : Adults Aged 19 to 64 Years, 2000-2001. Colchester, Essex: UK Data Archive; 2005.

41. Stasinopoulos MD, Rigby RA, Heller GZ, Voudouris V, De Bastiani F. Flexible regression and smoothing: using GAMLSS in R. Boca Raton: CRC Press/Taylor & Francis Group; 2017. 549 p.

42. Rigby RA, Stasinopoulos DM. Generalized additive models for location, scale and shape. J R Stat Soc Ser C Appl Stat. 2005;54:507–54.

43. Ji C, Cappuccio FP. Socioeconomic inequality in salt intake in Britain 10 years after a national salt reduction programme. BMJ Open. 2014;4:e005683.

44. Millett C, Laverty AA, Stylianou N, Bibbins-Domingo K, Pape UJ. Impacts of a national strategy to reduce population salt intake in England: serial cross sectional study. PLoS ONE. 2012;7:e29836.

45. Kypridemos C, Guzman-Castillo M, Hyseni L, Hickey GL, Bandosz P, Buchan I, et al. Estimated reductions in cardiovascular and gastric cancer disease burden through salt policies in England: an IMPACTNCD microsimulation study. BMJ Open. 2017;7:e013791.

46. Hippisley-Cox J, Coupland C, Robson J, Sheikh A, Brindle P. Predicting risk of type 2 diabetes in England and Wales: prospective derivation and validation of QDScore. BMJ. 2009;338:b880.

47. Capewell S, O'Flaherty M. Rapid mortality falls after risk-factor changes in populations. The Lancet. 2011;378:752–3.

48. Capewell S, O'Flaherty M. Can dietary changes rapidly decrease cardiovascular mortality rates? Eur Heart J. 2011;32:1187–9.

49. Comparative quantification of health risks [Internet]. Geneva: World Health Organisation; 2004 [cited 2014 Jan 30]. Available from: http://www.who.int/publications/cra/en/

50. National Population Projections, 2012-based Statistical Bulletin [Internet]. Office for National Statistics; 2013 Nov [cited 2014 Nov 11]. Available from: http://www.ons.gov.uk/ons/rel/npp/national-population-projections/2012-based-projections/index.html

51. Levin ML. The occurrence of lung cancer in man. Acta - Unio Int Contra Cancrum. 1953;9:531–41.

52. Office for National Statistics. Published ad hoc data: health, requests during December 2014 [Internet]. Published Ad hoc Data: Health, Requests During December 2014. 2014 [cited 2015 Oct 14]. Available from: http://www.ons.gov.uk/ons/about-ons/business-transparency/freedom-of-information/what-can-i-request/published-ad-hoc-data/health/december-2014/number-of-registered-deaths-by-sex--cause--year--the-adjusted-index.xls

53. Townsend N, Wickramasinghe K, Bhatnagar P, Smolina K, Nichols M, Leal J, et al. Coronary heart disease statistics 2012 edition [Internet]. British Heart Foundation; 2012 [cited 2016 Nov 14]. Available from: https://www.bhf.org.uk/-/media/files/publications/research/2012_chd_statistics_compendium.pdf

54. Smolina K, Wright FL, Rayner M, Goldacre MJ. Incidence and 30-day case fatality for acute myocardial infarction in England in 2010: national-linked database study. Eur J Public Health. 2012;22:848–53.

55. Barendregt JJ, van Oortmarssen GJ, Vos T, Murray CJ. A generic model for the assessment of disease epidemiology: the computational basis of DisMod II. Popul Health Metr. 2003;1:4.

56. Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. The Lancet. 2012;380:2224–60.

57. Boshuizen HC, Lhachimi SK, Baal PHM van, Hoogenveen RT, Smit HA, Mackenbach JP, et al. The DYNAMO-HIA Model: An Efficient Implementation of a Risk Factor/Chronic Disease Markov Model for Use in Health Impact Assessment (HIA). Demography. 2012;49:1259–83.

58. Office for National Statistics. Cancer registrations in England [Internet]. Cancer Registrations in England. 2015 [cited 2015 Sep 28]. Available from: http://www.ons.gov.uk/ons/rel/vsob1/cancer-registrations-in-england/index.html

59. Office for National Statistics. Cancer survival rates - cancer survival in England: patients diagnosed 2006–2010 and followed up to 2011 [Internet]. Cancer Survival Rates - Cancer Survival in England: Patients Diagnosed. 2012 [cited 2015 Sep 28]. Available from: http://www.ons.gov.uk/ons/rel/cancer-unit/cancer-survival/2006---2010--followed-up-to-2011/index.html

60. Hyndman RJ. demography: Forecasting mortality, fertility, migration and population data [Internet]. 2017. Available from: http://CRAN.R-project.org/package=demography

61. Hyndman RJ, Shahid Ullah M. Robust forecasting of mortality and fertility rates: A functional data approach. Comput Stat Data Anal. 2007;51:4942–56.

62. Doll R, Peto R, Boreham J, Sutherland I. Mortality in relation to smoking: 50 years' observations on male British doctors. BMJ. 2004;328:1519.

63. The DECODE Study Group. Is the current definition for diabetes relevant to mortality risk from all causes and cardiovascular and noncardiovascular diseases? Diabetes Care. 2003;26:688–96.

64. Stephens MR, Blackshaw GRJC, Lewis WG, Edwards P, Barry JD, Hopper NA, et al. Influence of socio-economic deprivation on outcomes for patients diagnosed with gastric cancer. Scand J Gastroenterol. 2005;40:1351–7.

65. Siemerink EJM, Hospers GAP, Mulder NH, Siesling S, van der Aa MA. Disparities in survival of stomach cancer among different socioeconomic groups in North-East Netherlands. Cancer Epidemiol. 2011;35:413–6.

66. Allemani C, Weir HK, Carreira H, Harewood R, Spika D, Wang X-S, et al. Global surveillance of cancer survival 1995–2009: analysis of individual data for 25 676 887 patients from 279 population-based registries in 67 countries (CONCORD-2). The Lancet. 2015;385:977–1010.

67. Koerkamp BG, Stijnen T, Weinstein MC, Hunink MGM. The combined analysis of uncertainty and patient heterogeneity in medical decision models. Med Decis Making. 2011;31:650–61.

68. Briggs AH, Weinstein MC, Fenwick EAL, Karnon J, Sculpher MJ, Paltiel AD. Model parameter estimation and uncertainty: a report of the ISPOR-SMDM modeling good research practices Task Force-6. Value Health. 2012;15:835–42.

69. NHS England. Programme Budgeting [Internet]. 2018 [cited 2018 Aug 14]. Available from: https://www.england.nhs.uk/resources/resources-for-ccgs/prog-budgeting/

70. Charlton J, Rudisill C, Bhattarai N, Gulliford M. Impact of deprivation on occurrence, outcomes and health care costs of people with multiple morbidity. J Health Serv Res Policy. 2013;18:215–23.

71. Liu JL, Maniadakis N, Gray A, Rayner M. The economic burden of coronary heart disease in the UK. Heart. 2002;88:597–603.

72. Saka Ö, McGuire A, Wolfe C. Cost of stroke in the United Kingdom. Age Ageing. 2009;38:27–32.

73. Pearce A, Bradley C, Hanly P, O'Neill C, Thomas AA, Molcho M, et al. Projecting productivity losses for cancer-related mortality 2011 – 2030. BMC Cancer [Internet]. 2016 [cited 2018 Aug 10];16. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5069877/

74. Friendly M. Mosaic Displays for Multi-Way Contingency Tables. J Am Stat Assoc. 1994;89:190–200.

75. Sadler K, Nicholson S, Steer T, Gill V, Bates B, Tipping S, et al. National Diet and Nutrition Survey - Assessment of dietary sodium in adults (aged 19 to 64 years) in England, 2011 [Internet]. London: Department of Health; 2012 [cited 2016 Nov 14]. Available from: http://webarchive.nationalarchives.gov.uk/20130402145952/http://transparency.dh.gov.uk/2012/06/21/sodium-levels-among-adults/

76. National Centre for Social Research. An assessment of dietary sodium levels among adults (aged 19-64) in the UK general population in 2008, based on analysis of dietary sodium in 24-hour urine samples [Internet]. 2008. Available from: http://tna.europarchive.org/20110116113217/http://www.food.gov.uk/multimedia/pdfs/08sodiumreport.pdf

77. National Centre for Social Research. An assessment of dietary sodium levels among adults (aged 19-64) in the general population in England, based on analysis of dietary sodium in 24-hour urine samples [Internet]. 2006 [cited 2016 Nov 14]. Available from: http://tna.europarchive.org/20110116113217/http://www.food.gov.uk/multimedia/pdfs/englandsodiumreport.pdf

78. Henderson L, Irving K, Gregory J, Bates CJ, Prentice A, Perks J, et al. National Diet and Nutrition Survey: adults aged 19 to 64 years. Volume 3: Vitamin and mineral intake and urinary analytes. [Internet]. London: The Stationery Office; 2003 [cited 2015 Jun 5]. Available from: http://www.ons.gov.uk/ons/rel/lifestyles/the-national-diet-and-nutrition-survey/2001-edition/the-national-diet-and-nutrition-survey---volume-3.pdf

79. National Institute for Health & Care Excellence. NG 28; Type 2 diabetes in adults: management. Economic modelling report [Internet]. London: NICE; 2015 [cited 2017 Mar 17]. Available from: https://www.nice.org.uk/guidance/ng28/evidence/appendix-f-full-health-economics-report-2185320355

80. National Institute for Health & Care Excellence. PH25; Cardiovascular disease prevention. Costing report [Internet]. London; 2010 Jun [cited 2016 Jul 26]. Available from: https://www.nice.org.uk/guidance/ph25/resources/costing-report-67331053

81. World Cancer Research Fund International/American Institute for Cancer Research. Continuous Update Project report: diet, nutrition, physical activity and stomach cancer [Internet]. AICR/WCRF; 2016 [cited 2016 Apr 22]. Available from: http://wcrf.org/sites/default/files/Stomach-Cancer-2016-Report.pdf

82. González CA, Pera G, Agudo A, Palli D, Krogh V, Vineis P, et al. Smoking and the risk of gastric cancer in the European Prospective Investigation Into Cancer and Nutrition (EPIC). Int J Cancer. 2003;107:629–34.

83. Lock K, Pomerleau J, Causer L, McKee M. Comparative quantification of health risks. Chapter 9: low fruit and vegetable consumption [Internet]. Geneva: World Health Organisation; 2004. Available from: http://www.who.int/publications/cra/en/

84. World Cancer Research Fund, American Institute for Cancer Research. Food, nutrition, physical activity, and the prevention of cancer: a global perspective. Washington, DC: WCRF/AICR; 2007.

85. Mozaffarian D, Fahimi S, Singh GM, Micha R, Khatibzadeh S, Engell RE, et al. Global sodium consumption and death from cardiovascular causes. N Engl J Med. 2014;371:624–34.