Population health impact and budget implications of reducing public health risks in England – a multi-dimensional life table analysis.

Niessen L (1,2), Jarrett J (1)*, Goryakin Y (1)

(1) University of East Anglia,
Norwich Medical School
Health Economics Group
Norwich, NR4 7TJ
*j.jarrett@uea.ac.uk (corresponding author email)
(2) John's Hopkins University
Dept International Health
615 N Wolfe Street
Baltimore, MD
USA

This paper describes an integrated multidisease model (IMM) that is a work in progress. The first version of the model was stratified by age and gender only but has now been extended to include socioeconomic status (SES) information. The first section describes the model design, its main assumptions, as well as the data used in its structure. We then follow by discussing the main model output as applied to different cohorts of the UK population of interest. We conclude by focusing on the main lessons. As this is a complex model, and space is at a premium, some calculation descriptions were not included but are available upon request.

1. The purpose of the model

The IMM can be used to model the prevalence of various risk factors and diseases in the UK population, their associated costs (from different perspectives), life years (both nominal and quality-adjusted), as well as various cost-effectiveness scenarios for health interventions. Another important application of the model is estimating the burden of several diseases, attributed to several risk factors. Incorporating socio-economic status will allow for an analysis of health inequalities as well as the impact of targeting certain socio-economic groups.

The diseases chosen for this model are several cardiovascular conditions (myocardial infarction, stroke, diabetes), lung disease (COPD), several cancers (lung, breast, colon), as well as dementia. The choice was dictated by the importance of these conditions, both in terms of their relatively high prevalence, as well as the costs they impose on people and health care systems.

2. Model design and assumptions

Model Population

On a basic level, the model simulates a health history for a single random individual with pre-defined demographic characteristics. For each "individual", a number of baseline characteristics are generated, which then determine their probability of getting certain risk factors, based on epidemiological information for a the population to which they belong. In turn, these risk factors determine their chances of getting several modeled diseases, and, ultimately, their probability of dying. The conceptual model can be found in Appendix A. The model was built in Microsoft Excel using the @Risk software add on to run the simulations.

As a default, a cohort of all people currently living in the UK is modelled. Therefore, this is a "closed cohort", in a sense that its maximum size is fixed, and in future years, no new people will be added (Rothman, Greenland et al. 2008). Consequently, in the first calendar year, it is possible to estimate the average prevalence of a certain risk factor in a population, but in all subsequent years, the averages will not exactly represent expected future prevalence. This is because in future years, new people will be born and enter the population of England. The difference between calendar–year prevalence and real population prevalence will be small in the immediately following years, but will increase with each calendar year. It is also possible to model the life course of predetermined cohorts, rather than of the whole population. For example, one can be interested in the cost of illness for males only, for people of a certain age, or for those without diabetes or high blood pressure. If these options are chosen, only people with these characteristics will enter the model.

First, demographic characteristics are generated for a random person, based on existing population distributions for England and Wales (ONS 2010). Age depends on the gender-specific data (in 10 year intervals) provided by the Office of the National Statistics for England (ONS 2010). Socioeconomic status (SES) data is based on the gender and age-specific data provided by the Health Survey for England (Craig and Mendell 2008). The HSE allocated occupation codes to the Registrar General's Social Class[1] based on current or previous employment if unemployed.

Risk Factors

Given a simulated person's age, gender and socioeconomic status, the risks of getting a certain risk factor are modelled. There are six risk factors modelled in total: hypertension, high cholesterol, high BMI, smoking, alcoholism, inactivity. These factors were chosen because of their relatively high prevalence in

[1] I-Professional, II- Managerial Technical, IIIN –Skilled Non-manual, IIIM-Skilled Manual, IV – Semiskilled Manual, V-Unskilled manual

the UK population, as well as due to their role in being an important cause of many serious illnesses. Hypertension was assumed to be present when systolic blood pressure was equal to, or greater, than 140. Hypercholesterolemia was defined as total cholesterol level of greater than 230 mg/dl.

Individual values for systolic blood pressure, body mass index, and total cholesterol were based on the random draws from the age, gender and SES specific distributions of these values in the UK population. (Craig and Mindell 2008) BMI is assumed to grow by 0.1% every year, while the annual rates of growth for total cholesterol and systolic blood pressure are age-specific, and are taken from another model developed previously (Brown, Russell et al. 2000). Such an approach necessarily means that once a person passes a certain threshold for high blood pressure or cholesterol, they will acquire a related diagnosed condition for the rest of their lives. In practice, this assumption does not appear too restrictive.

The probability of having other risk factors –smoking, alcoholism, inactivity-is also based on HSE data (Craig and Mindell 2008). The category of smoker applies to current smokers only; the high alcohol consumption is based on the definition "Usual frequency of drinking alcohol in past year: almost every day OR five or six days a week"; and inactive people in the model are those classified to category "summary activity level is less than 30 minutes of moderate or vigorous activity on 1 to 4 days a week".

These probabilities were used to assign risk factors to the randomly generated individuals based on their demographic and SES characteristics. For example, a male aged 55 in SES IV/V will be assigned a probability of 0.38 of being a smoker, whereas a male aged 55 in SES I/II would be assigned a probability of .10 of being a smoker. Then, this probability will be compared with a random number between zero and one generated from uniform distribution. If the random number is smaller, the person will be a smoker in this particular year. Similar approaches are used to generate probabilities for having alcoholism and being inactive. A major current assumption of the model is that each year the probabilities of having these factors are assumed independent of each other. Therefore, it is possible for a person to be a smoker or heavy drinker in one year, but not in another. No dependence between prior smoking/alcohol/inactivity status, and the current one is assumed in the model.

In addition, inactivity is assumed to be correlated with hypertension. To impose this correlation, the adjusted odds ratios for the association between being inactive and hypertension were estimated by logistic regression (using individual-level data from the Russia Longitudinal Monitoring Dataset (CPC 2010)). Such a relationship may be of a biologic nature and thus apply to more than just the population of Russia.

The parameters of the following equation were estimated:

Ln(p/1-p)= X*b **(1)**

Where **p (probability of being inactive) = exp(X*b)/(1+exp(x*b))**,

Where X is a vector of variables including hypertension, as well as controls such as age, gender, urban residence, education, wealth, household size. Separate models were run for men and women.

Note that in this case, hypertension is not assumed to be the cause of being inactive. Rather, we are interested in estimating the correlation between these variables, controlling for potential confounders. Thus, it was found that a man with high blood pressure will also be about 36% more likely to be less active. The corresponding ratio for a woman is 1.22. Because of the structure of our model, it is much easier to put inactivity as a function of hypertension, rather than the other way around.

Note that since getting some of these risk factors for the first time (e.g., smoking and alcoholism) does not imply that the person will have them for the rest of his life, we can presume that the necessary probability should come from the prevalence data for every subsequent year. On the other hand, for certain conditions that, once acquired, become lifelong (e.g., diabetes), the entering person's probability will come from the prevalence data, and then in all following years, it will come from the incidence data. We will discuss this issue in more detail later.

Disease Prevalence/Incidence

Given a generated risk factor profile, as well as age and gender (but at this point, not SES), disease prevalence and incidence are modelled. In total, 8 diseases with high mortality and morbidity in the population are introduced here: DM, myocardial infarction, stroke, COPD, lung cancer, breast cancer, colon cancer, dementia.

There are two approaches to estimating disease incidence and prevalence in the model. The first one is used for modelling cardiovascular conditions, and is based on continuous Framingham risk equation (Anderson, Odell et al. 1991). The other uses the adjusted background incidence and prevalence to take into account the effect of certain risk factors.

Framingham equation for modelling a risk of cardiovascular disease or mortality is

$Prob(CHD - T) = 1-exp^{(-exp(\ln(T)-\mu/\sigma),}$ **(2)**

Where $\sigma = (\alpha + (\beta * \mu)$ and $\mu = X*b$

Where T is a number of years for which the cumulative probability of cardiovascular events (strokes, myocardial infarctions, and cardiovascular mortality) is modelled; μ is a linear function of the risk factors X, which include gender, log age and its interactions with gender, log SBP, cigarette consumption, log total cholesterol/HDL cholesterol, diabetes status and its interaction with gender, and b is a set of parameters specific for a modelled disease or cardiovascular death. The parameters for the equation were taken from another source (Anderson, Odell et al. 1991).

Diabetes incidence and prevalence are assumed to be correlated with age, gender, and whether people have hypertension, high BMI, and being inactive. In addition, diabetes is assumed to directly affect the incidence of having a stroke and myocardial infarction, independent of other risk factors. Lung disease and lung cancer are assumed to depend on smoking status. Breast, colon cancer, as well as dementia do not depend on any risk factors in the model, except for age and gender. The formulas are available from the author.

The probabilities of getting dementia, breast and colon cancer is only assumed to depend on the baseline incidence and prevalence of getting these diseases in the population, and are not adjusted by the existence of any risk factors.

The following formula was used to calculate the background incidence/prevalence:

Background prevalence= Average disease prevalence*((1-prevalence RF) + prevalence RF*$RR_{disease, RF}$) **(6)**

Where average prevalence is age, and specific prevalence for an average person; prevalence RF is a risk factor prevalence in the population, and $RR_{disease, RF}$ is a risk (or odds) ratio reflecting the association between a risk factor and a disease. The formula shows, for example, that the greater the risk factor prevalence, the greater the difference between background and average disease prevalence. In the extreme, if risk factor prevalence equals zero, background prevalence will equal average prevalence. The same logic applies when incidence rather than prevalence is modelled.

All the relationships assumed in the model are shown on the Conceptual Model (Appendix A) with the arrows. An important point to highlight is that a risk for getting these conditions (except stoke and MI) by a person who just enters a cohort is assumed equal to the prevalence of a disease in the population. For all following periods, the risk is equal to the disease incidence, adjusted by the appropriate risk ratios specific for the risk factors generated for this person in the model. The reason for this is that in the first year, using incidence rather than prevalence will greatly underestimate the actual risk facing this person. On the other

hand, in all the following years, the person has already "faced" the accumulated risk specific for his or her socioeconomic status, age and gender, and using yearly incidence rates will be warranted.

Mortality Calculations

Finally, based on the disease profile of each simulated "person", their probability of dying is modelled during each yearly interval. The model has three main mortality probability types contributing to the overall risk of death:

1) Background mortality: based on the population-level data from the Life Tables (2006-2008) published by the ONS. Background mortality in our model also depends on whether people have diabetes or not, so that diabetes will have an effect on mortality not only indirectly through cardiovascular diseases, but also a direct effect. Mortality probabilities are only given for people that have no diseases other than diabetes.

2) Cardiovascular mortality: depends on the same characteristics that determine the risk of having a stroke and myocardial infarction, and is modelled in a similar way, using Framingham equations. The main inputs are age, gender, DM status, cholesterol levels (total and HDL), SBP, smoking status.

3) Disease-specific mortality (besides diabetes): It applies to people who have "acquired" certain diseases that can modify their background mortality risks. These illnesses include breast cancer, lung cancer, and colon cancer. The data is based on age and gender specific mortality tables for these diseases from the Cancer Research UK website.

It is assumed that the mortality probability is highest in the 1st year, and then it is equally spread over the next 4 years. For this reason, the data on both 1 year, and 5 year mortality, is used. After 5 years, if a person survives a disease, it is assumed that it no longer modifies his background mortality probability.

If a disease-specific mortality rate is positive, then a background probability value is subtracted from it to estimate a pure contribution of this illness to greater mortality risk. Finally, background, cardiovascular and disease-specific mortality probabilities are added to derive a total mortality probability, which will be compared each year for each person with a random number generated from a uniform distribution and bounded between 0 and 1. If the random number is smaller than the mortality probability, the "death" will occur, and the person will stop contributing life years and other information to her/his history.

Cost of Illness and Cost-effectiveness

In order to model cost of illness and cost-effectiveness estimates, one needs to enter the cost, utility as well as discount rate assumptions. Table 1 indicates the cost and disability weight assumptions currently in the model. We took the data from the following sources:

- Diabetes expenditure. Yearly per patient cost (283 GBP) was assumed to be equal to prescription costs in primary care only (NHS 2009). This, by definition, may lead to the underestimation of the total costs, as such a measure excludes treatments provided in the secondary or tertiary care (e.g., for serious complications).
- Myocardial infarction expenditure. Unfortunately, no UK- specific costs were found by authors. We decided to assume that the NHS costs per patient were equal to the adjusted average acute myocardial infarction costs incurred in the hospitals in Germany, approximately equal to 3,399 GBP per patient (Kauf, Velazquez et al. 2006). It was also assumed that this was a one-time expenditure, with no additional costs incurred in subsequent years. Again, this assumption may be quite conservative.
- Stroke expenditure. The cost was assumed to be equal to the NHS-reported 2.8 billion GBP yearly direct stroke-related costs, divided by the 300,000 people in England living with moderate to severe disabilities as a result of strokes (House of Commons 2006). This may lead to some overestimation of costs, as not all of 2.8 billion GBP expenditure is incurred by the people with moderate to severe stroke-related disabilities. On the other hand, these costs are only direct, and thus not including a range of possible indirect expenditures. This cost is assumed to be incurred every year after the initial diagnosis is made in the model, until death.
- Chronic Obstructive Pulmonary Disease (COPD) cost is assumed to be equal to 819 GBP per year, of per patient direct costs. (The National Collaborating Centre for Chronic Conditions 2004). In the model, such costs are assumed to accumulate from the initial diagnosis day until the patient dies (ie, the disease is not assumed "curable").
- Lung cancer costs. Yearly costs are assumed to be equal to 327 mln total yearly NHS expenditures (Allender, Balakrishnan et al. 2009) divided by 39,470 diagnosed with lung cancer in the UK in 2007 (Cancer Research UK 2010), or about 8,280 GBP per year. Such costs are assumed to be incurred for no more than 5 years. If the person is still alive after this period, they are assumed cured. This assumption was made for all types of cancer in the model.
- Breast cancer costs. They were assumed to be equal to 7,247 GBP per annum per patient, and include primary, secondary care costs, as well as residential care expenditures (Dolan, Torgerson et al. 1999).

- Colon cancer costs. The annual cost was assumed equal to the mean cost per bowel cancer treatment of 8,808 GBP (Trueman 2007).
- Dementia costs. They are assumed to be by far the largest in the model- 25,472 GBP per person annually (Knapp, Prince et al. 2007). 41% of these costs are accounted for by accommodation expenses, and 36%- by informal care inputs (Knapp, Prince et al. 2007). In the model, once a person is diagnosed with dementia, they are assumed to incur such costs till the end of their lives.

Various diseases can contribute costs (discounted and undiscounted), as well as morbidity years that can be used to adjust the utility of life for people living with these conditions. Currently the model uses life years adjusted by disability scores, summed up for each person over their lives, to give an estimate of the quality adjusted life years, that can be used in burden of illness, or cost effectiveness calculations.

Disability weights for Years of Life Disabled (YLD) calculations were taken from the Global Burden of Disease report, 2004 update (Mathers, Fat et al. 2008). The weight for myocardial infarction was taken for the acute stage only, as it was assumed to be a one-time event. For diabetes and stroke, the chronic stage utility weight was chosen, as the patient was assumed to leave with these conditions from the year of diagnosis, until the end of their lives. For COPD, utility weight for a chronic stage with mild and moderate symptoms was taken. The most disabling disease was assumed to be dementia, with the weight of 0.666 (Mathers, Fat et al. 2008).

Discount rate for both costs and QALYs was assumed to be equal to 3.5% per year. The model allows to easily changing this number to alternative values. For the purposes of calculating years of life lost (YLL) for disability adjusted life years (DALYs), the threshold assumed is 100 years. Again, this can be changed within the model.

Table 1: Annual cost and disability weights

Disease	Annual Cost (2010GBP)	Disability Weights
Diabetes	283	0.015
Myocardial Infarction	3399	0.439
Stroke	9333	0.266
Lung Disease	819	0.17
Lung Cancer	8280	0.15
Breast Cancer	7247	0.09
Colon cancer	8808	0.2
Dementia	25472	0.666

5. Preliminary findings

Descriptive statistics

The average age of entering cohort (representative of all people living in England in 2009, of all ages) was estimated to be 40.3 years. The proportion of males in this cohort was about 49.3%, approximately equal to the officially estimated 49.25% of males in the population of England.

Average age at death for the whole cohort was estimated to be 74.9 years (Table 2). Note that this number differs from the officially estimated life expectancy at birth (currently at 79.4 years (CIA Factbook, 2009)). Our model predicted it to be about 79.61, but only when death risk used for mortality function was assumed equal to background mortality risk only. However, when total mortality risk also included disease-specific mortality risks, on top of background mortality risk (e.g., for diabetes, cancers, cardiovascular disease), estimated life expectancy at birth decreased to 75 years. This is inevitable since background mortality risk already includes all relevant risks. However, since our goal is to estimate the effect of change in different risk factors on *change* in life expectancy and on other outcomes, rather than the *level* of these outcomes per se, this may not be a very serious concern.

The prevalence of hypertension (e.g., SBP>=140) in the whole cohort at entry was estimated to be around 25%. This differs from the age and gender adjusted official prevalence estimate of about 29%, but note that in the latter case, the sample was restricted to people aged 16 and older. Since the prevalence of hypertension in 1-15 years age group is very low, our estimated average value is correspondingly smaller. The estimated lifetime prevalence of hypertension is about 65%. On average, people in the cohort live 10.1 years with this condition.

When high cholesterol was defined to be greater or equal to be 230 mg/dl, the prevalence of the condition in the entering cohort was about 16%, and the lifetime prevalence- 33%. On average, people will live 6.47 years with this condition. When the threshold was changed to 190 mg/dl (not shown in the table), the entering cohort prevalence was estimated to be about 55%, and lifetime prevalence- about 83%.

Prevalence of smoking in our cohort was about 17.5% at entry, which is a bit less than the officially estimated 22% of the population of England who are current smokers. Again, our results slightly differ because we have included 1-15 age group, while the official estimates applied to those 16 years and older. Estimated lifetime prevalence of smoking of 88.3 % may seem a bit high, but it includes all people who ever smoked (even for 1 year). Also, this can be an overestimate because our model does not assume dependence of current smoking status on smoking in previous periods. For example, current non-smokers can be less likely to smoke in the future, while the model assumes that their age and gender-specific

yearly risk of smoking is the same as for current smokers. On average, people were estimated to smoke 7.16 years in their lives.

The average proportion of our cohort who were defined as alcoholics at entry (i.e., drinking 5 or 6 days a week, or more) was estimated to be 13.4%. At some point in their lives, 92% will be expected to be alcoholic. Again, this is likely to be an overestimate, given that our model does not allow dependence between drinking status in prior and current period. However, it may still give reasonably accurate estimates over the entire lives, as lifetime risk can be considerably driven by a number of people whom the model defined as "alcoholic" just for 1 or 2 years in their lives.

The prevalence of obesity at entry was estimated to be around 25.8% (almost equal to officially estimated prevalence of 24%). The lifetime prevalence of obesity was 30%, and people in this cohort were expected to live 7.93 years on average with this condition. Twenty-eight percent of people in our cohort were estimated to be inactive at entry, which is a bit lower than the official estimates of about 34% (again, our estimate can be lower because our cohort includes those aged 1-15 years). On average, people are estimated to be inactive for 13.9 years of their lives.

Prevalence of diabetes was estimated to be about 7%, which is slightly higher than the officially estimated 5%. The lifetime risk of diabetes was estimated to be about 26%, and people, on average, live for 3.43 years with this condition. In comparison, the lifetime risk of developing diabetes mellitus in the USA for people born in 2000 was estimated to be 32.8% for men, and 38.5% for women (Narayan, Boyle et al. 2003).

The prevalence of myocardial infarctions at entry was estimated to be about 0.4%- quite close to officially estimated prevalence of 0.25% in the past 12 months. In addition, the lifetime prevalence of this condition was estimated to be about 17.6% which is slightly higher than the 13.5% of respondents in official survey aged 75+ who said that they ever had heart attacks (HSE, 2006). Our figure may be slightly higher because it includes both survivors and those who have died.

The prevalence of people in the cohort who had stroke at entry was about 0.3% (increasing to 0.5% next year), which is slightly lower than the officially estimated prevalence of 0.5%. 12.7% were estimated to ever have had strokes, which is very close in comparison to about 12% of surveyed people in England aged 75+ who reported ever having had a stroke (Craig and Mindell 2008).

About 1% of people in the cohort were estimated to have COPD, with the lifetime prevalence of this condition of about 4.5%. Lung cancer prevalence was estimated to be around 0.04%, with lifetime

prevalence of 1.3%. The prevalence of breast cancer among women was about 0.8% at entry, with lifetime prevalence increasing to 9%. Colon cancer was present in 0.1% of the cohort at entry, with lifetime prevalence rising to 3.4%, which is slightly lower than the officially estimated lifetime risk of 5% (Cancer Screening NHS website http://www.cancerscreening.nhs.uk/). Finally, 1.5% of people in the cohort were estimated to have dementia, with lifetime prevalence of the condition being equal to 17.7%.

Finally, the model predicted that about 26.7% of all deaths were cardiovascular disease related, 3.3% related to lung cancer, 1.2% from colon cancer, and 0.7% from breast cancer. Note that mortality from lung cancer is significantly greater than from breast cancer, despite the fact that the lifetime prevalence from the latter condition is considerably greater than from the former one. This is not surprising, since lung cancer has remained a very deadly condition, while mortality from breast cancer has been steadily falling.

Table 2. Prevalence of risk factors and diseases, the whole cohort

	Entry prevalence	**1st year**	**Lifetime**	**YLC**	**Mortality proportion**
Age	40.28	-	-		
Males	0.493	-	-		
Age at death	74.9	-	-		
HBP	0.25	0.24	0.65	10.1	
High cholesterol	0.16	0.16	0.327	6.41	
Smoking	0.175	0.178	0.883	7.16	
Obesity	0.258	0.255	0.30	7.93	
Alcoholism	0.134	0.134	0.92	7.00	
Inactivity	0.28	0.272	0.97	13.9	
DM	0.07	0.07	0.26	3.43	
MI	0.004	0.0076	0.176	2.43	0.267
Stroke	0.003	0.005	0.127	1.34	
LD	0.01	0.01	0.045	0.83	
LC	0.004	0.005	0.013	0.13	0.033
Breast cancer	0.008	0.01	0.09	1.47	0.007
Colon cancer	0.001	0.001	0.034	0.35	0.012
Dementia	0.015	0.016	0.177	1.38	

Effect of risk factors on disease burden

From table 3, we can see what will happen to several outcomes when certain risk factors are assumed to be eliminated in the population. The simulation has been run for the highest and lowest socio-economic class. We see that eliminating hypertension will have the greatest effect on increasing the age at death for those in the highest socioeconomic status whereas eliminating smoking will have the highest impact on

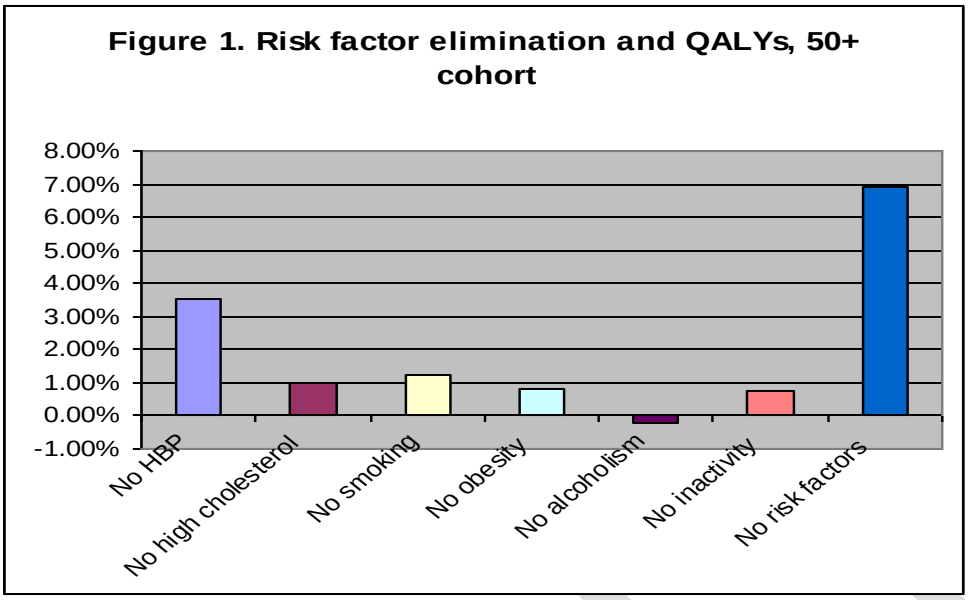
those in the lowest socioeconomic group. Since alcoholism was not assumed to be a risk factor linked to any disease in the model, its effect on the change in age at death was predictably small. Eliminating all risk factors at the same time predictably had the largest effect on the increase in the number of years at death for both groups.

The effect of eliminating risk factors on another outcome, QALYs, gives a somewhat different picture. Again, eliminating hypertension had the largest effect on the increase QALYs for the highest socioeconomic group, and also had a small positive impact on the lowest group. Eliminating alcoholism also had a relatively large positive effect on the highest group. Interestingly, eliminating individual risk factors actually seems to have a small detrimental effect on the QALYs of the lowest socioeconomic group, however eliminating all risk factors seems to improve quality of life.

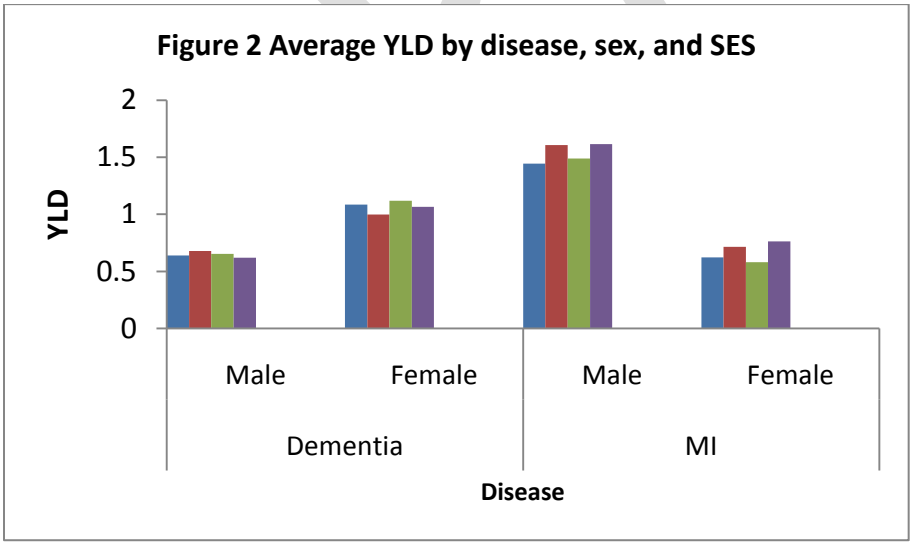
Table 3 Risk factors and basic indicators (SCI/II versus SCIV/V cohorts)

	Age at death		**QALYs, discounted**	
	SCI/II	**SCIV/V**	**SCI/II**	**SCIV/V**
Baseline	74.84	74.26	16.48	17.41
No HBP	76.62	74.18	17.67	17.45
No high cholesterol	75.34	74.2	17.43	17.2
No smoking	75.54	75.76	17.3	17.36
No obesity	75.21	75.49	16.96	17.28
No alcohol	74.88	74.27	17.55	17.37
No risk factors	76.9	76.41	17.62	17.92

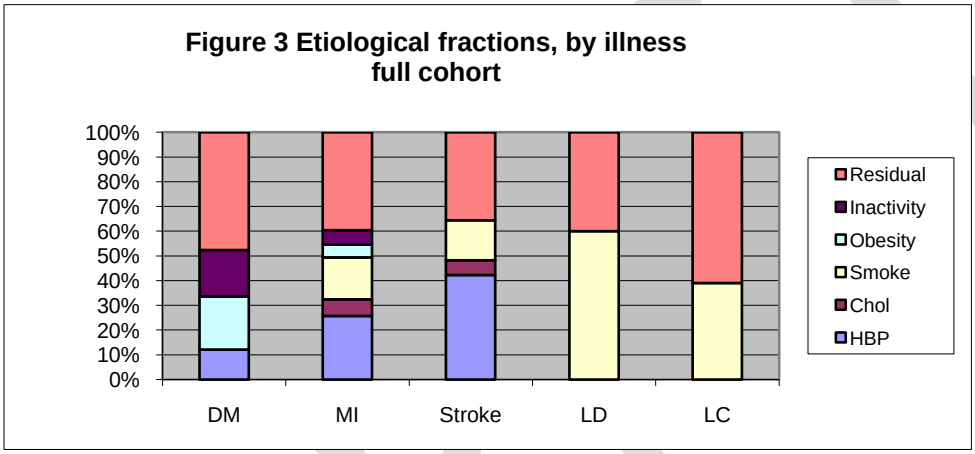
For another cohort of the population, the over 50's, figure 1 shows a graphical illustration of the effect of risk factor elimination on the change in QALYs. Also, now we are looking at this change not in absolute terms, but as a proportional change relative to the baseline case of no risk factor elimination. For this cohort, it is striking how many QALYs can be gained if all modifiable risk factors are controlled.



On figure 2, we can see how the burden of disease for dementia and MI is measured by YLDs compared by gender and SES (blue is SCI/II, red SCIIIN, green SCIIIM, purple SCIV/V). Not surprisingly, males have considerably more life years lived with disability caused by MIs and the lowest SES group has more YLDs associated with MI regardless of sex than the highest SES. Females spent longer with dementia than males, but there was not a strong class gradient.

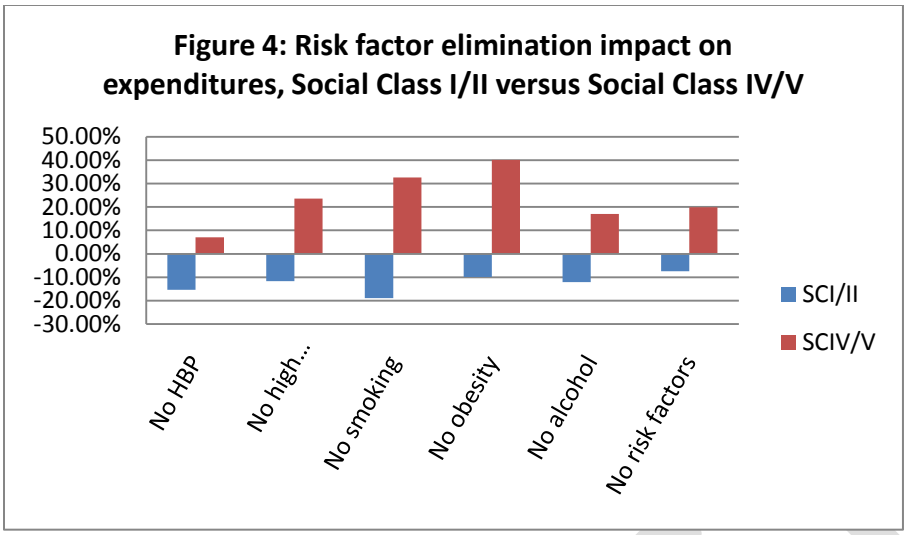


Next, on Figure 3, we can see how various risk factors contribute to disease burden (as measured by years lived with disease[2]). Only those diseases linked to risk factors in the model were chosen for analysis. As a rule, we see that the majority of the disease burden was usually accounted for by the risk factors rather than the residual, with the exception of lung cancer, where smoking accounted for only about 40% of the burden. For diabetes, the main determinant of disease burden is obesity, followed by inactivity and high blood pressure. For MIs, hypertension is now the most significant risk factor, accounting for about 25% of the burden. Smoking is the second most important determinant of disease burden, followed by high cholesterol, inactivity and obesity. The most significant predictors of stroke burden were high blood pressure and smoking. Finally, smoking accounts for roughly 60% of disease burden of COPD, and for 40% of lung cancer.

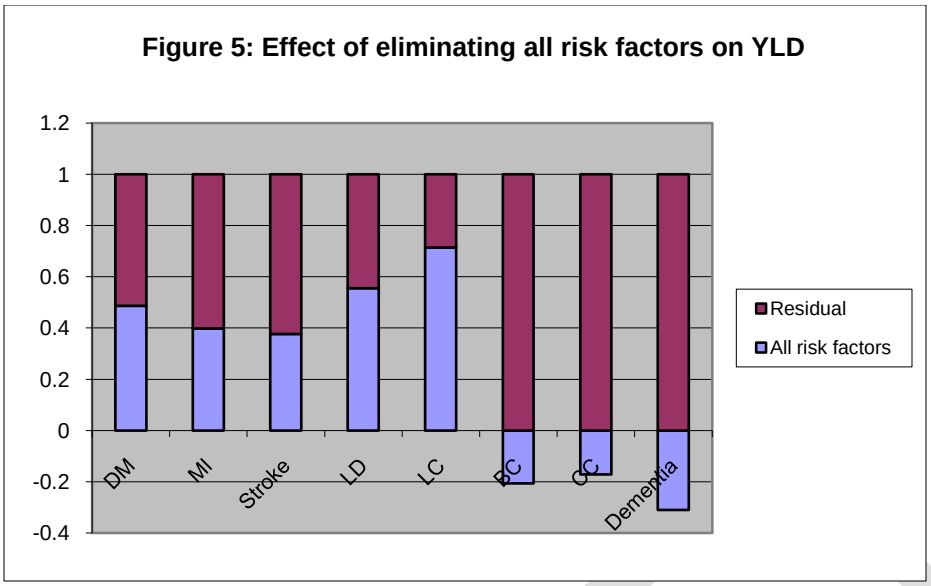


Although it is clear that most risk factors play an important role as health determinants, the question remains on whether controlling for them will allow saving money. While it is true that controlling for hypertension can prevent stokes and heart attacks and therefore these costs will not be incurred in the future, it is also possible that by bringing greater longevity they will lead to increased expenditures for other diseases. There is also the possibility of a differential between the socio-economic groups given the types of risk factors impacting on them. Figure 4 suggests this may indeed be the case. This figure shows that for the highest socioeconomic group, reducing risk factors will indeed reduce lifetime expenditure, especially when smoking is eliminated. However, for the lowest socio-economic group, expenditures increase with the elimination of certain risk factors. This may be due to the types of diseases that are common in the higher socioeconomic groups (i.e. MI)

[2] It is not necessary to adjust years lived with disease by disability weight- this will make no difference to calculating etiological fraction



A similar question can be asked in relation to another outcome- years lived with disease. What will happen to disease burden as measured by YLDs if all risk factors are eliminated at the same time? Is it possible that it may actually increase if longevity pushes disease burden for certain conditions into the future? As figure 5 shows, when all risk factors are eliminated, the model predicts that there will be a reduction in the prevalence of some diseases, but not all. Thus, eliminating all risk factors will lead to a substantial reduction in the years of life lived with DM, myocardial infarctions, strokes, COPD and lung cancers. At the same time, there will be an *increase* in the number of years lived with breast cancer, colon cancer and especially dementia. Although this may look somewhat counterintuitive, a plausible explanation for this is what can possibly be called "failure of success" (Niessen et al, 2010). In other words, the greater longevity caused by the elimination of risk factors may contribute to higher disease prevalence in later years of life for certain conditions, especially those weakly related to modifiable risk factors.



Conclusions/Discussion

The model has shown that eliminating risk factors is effective in reducing the burden of some diseases, and that eliminating certain risk factors can have a variable impact on the cost of care depending on which socio-economic group. There model does provide the ability to look at specific groups within simulated population, which potentially could help target public health interventions.

As with any model, there are several limitations. One of the main limitations of the model is that there is a simplified link between certain risk factors and diseases, as well as risk of death (eg smoking and COPD doesn't take into account amount of previous exposure). Related to this is that the risk of getting certain risk factors is independent of prior history. For example, each period, risk of smoking does not depend on whether a person smoked or not. The same for alcoholism and inactivity (this may overestimate prevalence of these factors). The binary nature of these risk factors may not be a realistic assumption and therefore may not indicate the true amount certain risk factors are contributing to disease. Although capable, the model has not been extensively tested in terms of reducing the risk factors by a certain proportion (e.g. reducing smoking in the population by 25%). Another limitation of the current model is that it cannot model dynamically into future. We can estimate all the parameters and follow the entering cohort only; however, in reality new people would be entering the population who may face different risk factors.

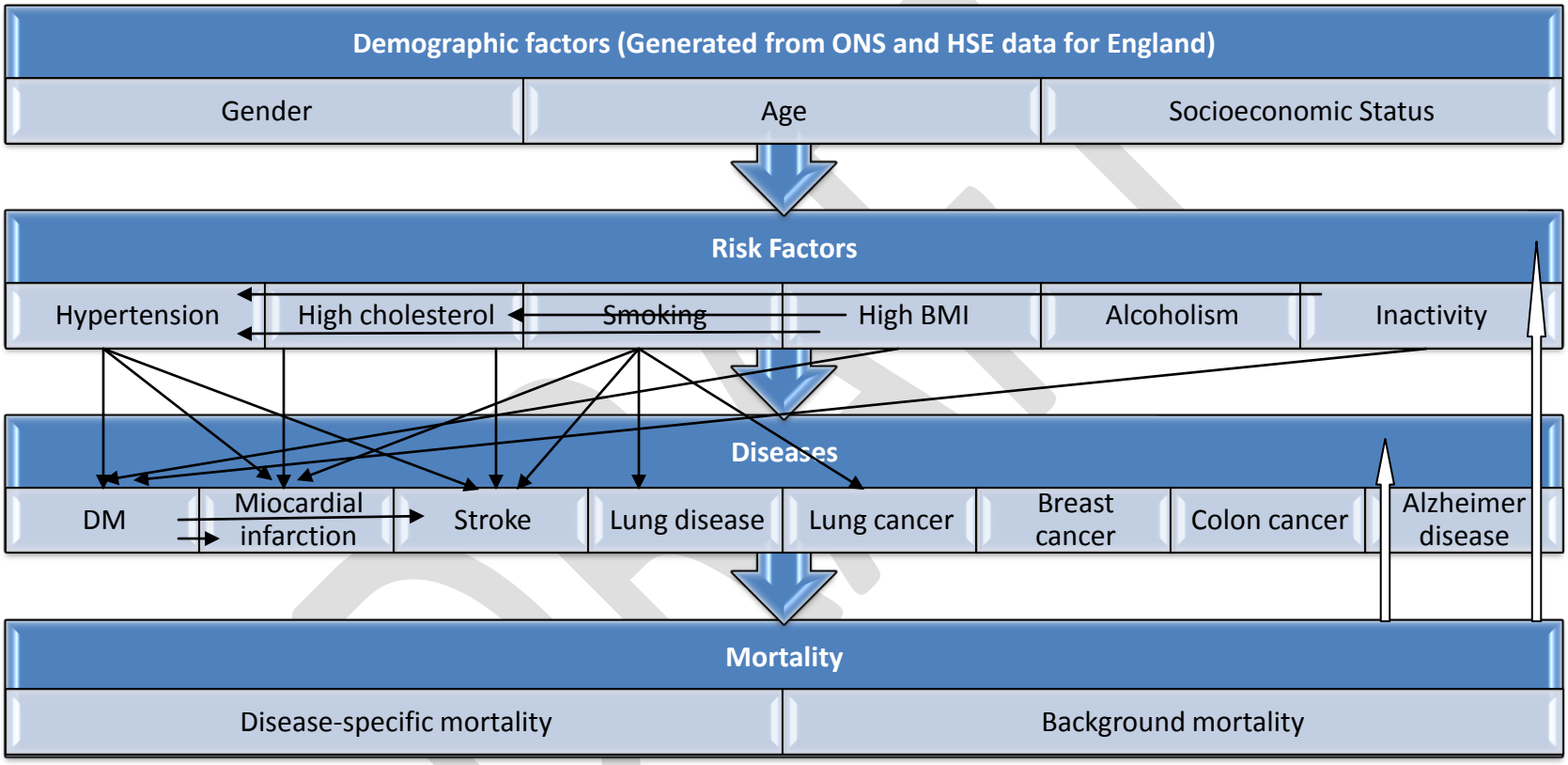
To address these limitations, we are proposing to change the variables we use to measure the risk factors (especially smoking, alcoholism, and inactivity) to reflect past history and volume (eg. someone who drinks one small glass of wine 5 or more nights a week may not be considered "alcoholic" when

compared with someone who drinks a bottle of wine a night 5 nights a week). In terms of predicting future costs and benefits, the model could use scenario analysis to predict future cohorts entering the model, but that would greatly increase the computing time and power necessary to run the model.

Another limitation is the comparability of the cost inputs into the model. As indicated, the annual cost of disease for each disease were measured in different ways, with some taking into account many more factors of the cost of care than others. Therefore, a direct comparison must be approached with caution. Another possible limitation of model inputs is the use of socioeconomic class as an indication of socio-economic status. The model has not yet been run using different measures (such as other classification systems, income, or education level).

An alternative approach that we are also looking into is to populate the model with individual data from the Health Survey for England and modelling that cohort over time, therefore allowing individuals' own "real" data to determine their risk factors over time.

Appendix A. Conceptual model



References

(2004). Appendix D: Economics costs of COPD to the NHS. Thorax, The National Collaborating Centre for Chronic Conditions. 59: i192-i194.

(2010). CancerStats Key facts: Lung Cancer and Smoking, Cancer Research UK.

Allender, S., R. Balakrishnan, et al. (2009). "The burden of smoking-related ill health in the UK." Tobacco control 18(4): 262.

Anderson, K. M., P. M. Odell, et al. (1991). "Cardiovascular disease risk profiles." American heart journal 121(1): 293-298.

Brown, J. B., A. Russell, et al. (2000). "The global diabetes model: user friendly version 3.0." Diabetes research and clinical practice 50: S15-S46.

Brownson, R. C., J. C. Chang, et al. (1992). "Gender and histologic type variations in smoking-related risk of lung cancer." Epidemiology 3(1): 61.

Commons, H. o. (2006). "Reducing brain damage: faster access to better stroke care. Fifty–second Report of Session 2005–06."

CPC. (2010). from http://www.cpc.unc.edu/projects/rlms-hse.

Craig, R. and J. Mindell (2008). "Health Survey for England 2006." London, United Kingdom: The Information Centre.

Dolan, P., D. J. Torgerson, et al. (1999). "Costs of breast cancer treatment in the United Kingdom." The Breast 8(4): 205-207.

Forman, D., D. Stockton, et al. (2003). "Cancer prevalence in the UK: results from the EUROPREVAL study." Annals of Oncology 14(4): 648.

Hansell, A., J. Hollowell, et al. (1999). "Use of the General Practice Research Database (GPRD) for respiratory epidemiology: a comparison with the 4th Morbidity Survey in General Practice (MSGP4)." Thorax 54(5): 413.

Johannessen, A., E. Omenaas, et al. (2005). "Incidence of GOLD-defined chronic obstructive pulmonary disease in a general adult population." The International Journal of Tuberculosis and Lung Disease 9(8): 926-932.

Kauf, T. L., E. J. Velazquez, et al. (2006). "The cost of acute myocardial infarction in the new millennium: evidence from a multinational registry." American heart journal 151(1): 206-212.

Knapp, M., M. Prince, et al. (2007). "Dementia UK: The full report." London: Alzheimer's Society.

Luengo-Fernandez, R., J. Leal, et al. (2010). Dementia 2010. The economic burden of dementia and associated research funding in the United Kingdom.

Mathers, C., D. M. Fat, et al. (2008). The global burden of disease: 2004 update, World Health Organization.

Matthews, F. and C. Brayne (2005). "The incidence of dementia in England and Wales: findings from the five identical sites of the MRC CFA Study." PLoS Medicine 2(8).

Narayan, K. M., J. P. Boyle, et al. (2003). "Lifetime risk for diabetes mellitus in the United States." Jama 290(14): 1884.

NHS (2007). Diabetes in the NHS: Commissioning and providing specialist services, National Diabetes Support Team.

NHS (2009). Prescribing for Diabetes in England - An Update: 2002-2008. An analysis of volume, expenditure and trends. D. H. Intelligence.

ONS. (2010). "Mid-2009 Population Estimates: England; estimated resident population by single year of age and sex " Retrieved November 2010, from http://www.statistics.gov.uk/statbase/Product.asp?vlnk=15106.

Rothman, K. J., S. Greenland, et al. (2008). Modern epidemiology, Lippincott Williams & Wilkins.

Ryan, R., A. Newnham, et al. (2005). "New cases of diabetes mellitus in England and Wales, 1994-1998: Database study." Public health 119(10): 892-899.
Simpson, C. R., J. Hippisley-Cox, et al. (2010). "Trends in the epidemiology of chronic obstructive pulmonary disease in England." British Journal of General Practice: e277.
Trueman, P., Lowson, K. , Bending, M., Ganderton M. (2007). "Bowel Cancer Services: Costs and Benefits. Final Report to the Department of Health."