Supplemental Material

1 Supplemental Tables and Figures

Supplemental Table 1. Predictive deviance for the validation period. Scenario B (trend-based scenario) shows lower values of predictive deviance through the validation period

Year	CHD men		CHD women		STK men		STK women	
	A	B	A	B	A	B	A	B
2003	37.5	16.6	66.3	142.2	27.6	51.2	49.4	50.0
2004	277.2	38.7	359.1	42.0	62.0	18.7	169.9	52.5
2005	336.9	26.7	582.9	63.3	234.0	35.8	805.0	235.5
2006	1230.2	170.0	1920.6	208.9	650.9	119.8	2092.8	779.3
2007	2504.0	434.1	3625.7	467.3	864.1	136.4	2302.0	664.5
2008	2887.9	293.0	3985.9	339.6	1178.7	171.2	3084.3	836.5
2009	4257.7	543.1	7396.8	1099.2	1976.0	300.9	4709.2	1448.4
2010	5158.7	628.6	8470.0	1029.4	2027.1	296.3	4862.5	1269.4
2011	6245.0	744.9	10212.8	1311.5	2551.0	385.9	5869.2	1570.3
2012	7240.5	831.0	12382.4	1630.6	2699.6	419.5	6461.7	1669.4

Supplemental Table 2. Population adjusted (2012 population baseline) **mortality for Coronary Heart Disease and Stroke in 2030, stratified by age and gender with lower (LCrI) and upper (UCrI) credible intervals.** Table 1A – conventional projections, Table 1B – trend-based projections.

2A.

	CHD			**Stroke**		
Age	Mean	LCrI	UCrI	Mean	LCrI	UCrI
M 25-34	773	589	978	280	181	388
M 35-44	5032	3988	6475	931	638	1325
M 45-54	19452	15811	25701	2931	2110	4101
M 55-64	29537	24073	37876	6192	4448	8642
M 65-74	33888	27284	44607	10900	7725	15265
M 75-84	36158	29416	47811	16698	11505	24312

M 85+	35244	28754	45643	15126	10676	21356
W 25-34	241	170	320	241	179	301
W 35-44	1563	1166	2029	759	579	977
W 45-54	7233	5692	9171	2253	1802	2872
W 55-64	16189	12730	21065	4584	3702	6096
W 65-74	22349	17726	28957	9415	7391	12531
W 75-84	29729	23236	38495	18614	14896	24727
W 85+	54967	43105	71589	34476	27326	43812
Totals	292354	233743	380716	123400	93159	166705

2B.

	CHD			**Stroke**		
Age	**Mean**	**LCrI**	**UCrI**	**Mean**	**LCrI**	**UCrI**
M 25-34	776	410	1429	247	56	1484
M 35-44	4253	2545	6900	761	238	3115
M 45-54	13561	8679	20933	1935	608	8002
M 55-64	21120	13374	32884	3989	1157	15914
M 65-74	23591	14815	38354	6859	2032	26873
M 75-84	25466	16563	40277	9636	2903	38303
M 85+	24374	15607	39518	8849	2609	34621
W 25-34	261	122	634	192	70	469
W 35-44	1232	686	2610	557	252	1240
W 45-54	4679	2812	9130	1566	735	3287
W 55-64	10520	6443	20755	3203	1449	6783
W 65-74	14590	8699	29351	6395	3147	12992
W 75-84	18786	11478	37355	12715	6180	24155
W 85+	34193	21157	68329	23852	10828	45715
Totals	197403	123389	348460	80754	32266	222952

2 Methods

2.1. Bayesian Age Period Cohort model

The APC model is often regarded as a log-linear Poisson model. Here we model the logit of the probability of death from CHD in age group *i* in period *j* as a linear combination of an intercept μ, age effects $\theta_i \ (i=1,\ldots I)$, period effects $\varphi_j \ (j=1,\ldots J)$ and cohort effects $\psi_k \ (k=1,\ldots K)$:

$$log\left(\frac{p_{ij}}{1-p_{ij}}\right)=\mu+\theta_i+\varphi_j+\psi_k.$$

I is the total number of age groups and J the total number of periods. Cohorts are defined by $k=C(I-i)+j$ and the total number of cohorts is $K=C(I-1)+J$, where C is the width of the age bands.

To make the model identifiable, the so called "sum to zero" constraints $\left(i.e.\sum\theta_i=\sum\varphi_j=\sum\psi_k=0\right)$ need to be added. The model is only identifiable when it is possible to obtain a unique set of parameters after infinite linear transformations. There is another identifiability problem due to the collinear relationship of age, period and cohort effects which makes it impossible to identify and interpret the separate contributions of each individual effect [1]. However, non-linear trends or change points can be interpreted because the non-identifiability only affects linear trends [2].

Random walk of first and second order

For the Bayesian APC model, it is assumed that effects from adjacent time bands tend to be more alike [3]. There are two types of priors that can represent such relationship: a first difference prior and a second difference prior.

In a first difference prior, used for **<u>scenario A "conventional projections"</u>**, each effect is derived from the immediately preceding effect, thus preventing it from varying much with respect to adjacent estimates. Additionally, a restriction is added to bound those first differences stochastically to zero and preserve a constant trend. This is known as a random walk of first order (RW1) and for the age effect is described by:

$$\theta_i \sim N\left(\theta_{i-1},\kappa^{-1}\right),\ i=2,\ldots,I\,.$$

$$p(\theta_1)\propto const.$$

where the hyperparameter κ is also known as the precision parameter: larger values allow baseline effects to vary only slightly, while small values allow more heavy variation.

In a second difference prior, used for **scenario B "trend-based"**, each effect is derived from its two immediate predecessors. A random walk of second order (RW2) restricts the second differences stochastically to zero, which penalizes deviations from a linear trend:

$$\theta_i \sim N\left(2\theta_{i-1} - \theta_{i-2}, \kappa^{-1}\right),\ i = 2,\ldots,I\ .$$

$$p(\theta_1) = p(\theta_2) \propto const.$$

For both types of prior, κ is assumed to follow a gamma distribution $G(a,b)$. For RW1 $a = 1$ and $b = 0.001$ are set as initial values, and for RW2 initial values are $a = 1$ and $b = 10^{-5}$. Similar priors are given to the period φ and cohort effects ψ with hyperparameters λ and ν, respectively.

Because the non-identifiability only affects linear trends, the RW1 solves the problem by keeping the age, period and cohort effects as constant as possible. However, the RW2 parameters are unidentifiable since the RW2 assumes a linear time trend. Fortunately in a Bayesian framework, it is not essential to ensure identifiability of age, period and cohort effects because $log\left(\frac{p_{ij}}{1-p_{ij}}\right)$ is fully identified [4]. Additionally, many have considered the problem of identifiability in the APC context as unsolvable [5-7].

Finally, the model can be easily extended to account for additional unstructured heterogeneity which cannot be explained by the age, period or cohort effects but by unknown and unobserved covariates:

$$log\left(\frac{p_{ij}}{1-p_{ij}}\right) = \mu + \theta_i + \varphi_j + \psi_k + z_{ij}\ .$$

$$z_{ij} \sim N\left(0, \delta^{-1}\right),$$

where the hyperparameter δ is assumed to follow a gamma distribution $G\left(a_\delta, b_\delta\right)$ with the same starting values as κ.

Estimation

The models can be implemented in BAMP [2] which uses Markov chain Monte Carlo (MCMC) simulation for the model estimation. MCMC constructs a Markov chain whose stationary distribution approximates the posterior distribution of interest using a single-component Metropolis-Hastings algorithm [8] for the age, period and cohort effects and a Gibbs sampler algorithm [9] for the hyperparameters.

After it has converged to its stationary distribution, samples from the chain can be used to calculate the parameter estimates. We fitted gamma distributions, using the function mass in R, to the samples of each hyperparameter obtained from the chain to generate new values of *a* and *b*. We used these values as new initial values and the models were recalculated.

The new samples of the posterior distribution were used to calculate credible intervals for the probability of death. A $100(1-2p)\%$ credible interval $\left[c_p - c_{1-p}\right]$ is a Bayesian interval estimate where c_p is equal to the p^{th} quantile of the scalar component and where c_{1-p} is equal to the $(1-p)^{th}$ quantile.

Forecast and comparison

Future rates $p_{ij},\ j = J+1,\ldots j+T$ can be calculated by projecting the age, period and cohort effects into the future by repeated application of the RW1 or RW2 model definitions. Similarly projected effects $z_{ij},\ j = J+1,\ldots j+T$ can be computed independently from $N\left(0,\delta^{-1}\right)$.

Finally, when different BAPC models need to be compared (e.g. RW1 versus RW2), the Deviation Information Criterion (DIC) [10] provides a simple yet powerful way of evaluating different models. DIC is the sum of two components: a term that measures goodness of fit known as posterior expectation of the deviance and a penalty term for increasing model complexity. Models with smaller DIC should be preferred.

2.2. Validation

We used observed data from 1979-2002 to project CVD mortality for 2003-2012 using both RW1 and RW2 and calculated predictive deviance. The predictive deviance at time *t* is defined as

$$D_t = 2\sum_j l(\widehat{y_{jt}}) - l(y_{jt})$$

Where $l(\widehat{y_{jt}})$ is the log-likelihood of predicted number of cases in age *j* and period *t* and $l(y_{jt})$ is the maximum log-likelihood achievable. Likewise, models with smaller predictive deviance are preferred.

2.3. Lee-Carter Model

As we discussed in the result section of the manuscript, our trend-based scenario underestimated the number of deaths during 2003-2012. Therefore, we alternatively used the well know Lee-Carter to project mortality over the same period.

Lee and Carter[11] developed a method commonly used in actuarial science, that combines a demographic model with time-series methods of forecasting. The model uses a single parameter (mortality index) governs the dynamics of a mortality trend.

Let $m_{x,t}$ be the central mortality rate for age x in year t, the model states

$$\ln(m_{x,t}) = \alpha_x + \beta_x k_t + \varepsilon_{x,t}$$

Where $k_t$the mortality index is at time t, α_x is the average pattern of mortality by age, $\beta_x$is the relative change with respect to the mortality index at age x and $\varepsilon_{x,t}$ is the residual at age x and year t.

Generally, Lee-Carter models are most suitable for forecasting mortality trends characterised by strong period effects, such as infectious respiratory disease[12]. Period effects might capture temporal change in factors associated with medical and broader societal development.

We fitted the basic Lee-Carter model to the 1979-2002 data and projected 2003-2012. Because Lee-Carter produces mortality rates m_{ij}, we transformed these into probabilities of death p_{ij} using the next formula [13] to allow for comparison with our BAPC scenarios:

$$p_{ij} = \frac{m_{ij}}{1 + \frac{m_{ij}}{2}}$$

Lee-Carter model was implemented using the R package demography [14] , using monotonic regression spines for smoothing, due to having data available only in 5-year age group bands[15]

2.4. Race CVD mortality disparities

The BAMP software allows saving the samples of the posterior distribution of the outputs generated during the estimation process. We used these samples to calculate population-adjusted and age-adjusted rate ratios to compare CVD mortality in Blacks and Hispanic with Whites.

For the age-adjusted rate ratios, we obtained the age and gender adjusted probability of death using each iteration of the BAMP software and the 2012 US standard population. We then used bootstrapping to estimate the age and gender adjusted relative risk of mortality and 95% bootstrapped CI.

For the population-adjusted rate ratios, we obtained probability of death standardised by gender, age and race group for each iteration of the BAMP software in years 2012 and 2030. We then used bootstrapping to estimate the relative risk of mortality and 95% bootstrapped CI.

2.5. Age-Period and Cohort Analysis

Random walks models of first order are useful to explore the contributions of age, period and cohort effects on CHD and stroke mortality. More specifically, the hyperparameters of the model gives an idea of the size of the effects. Larger values of the hyperparameters allow baseline effects to vary only slightly, while small values allow more heavy variation.

Hyperparameter	CHD men	Stroke men	CHD women	Stroke women
Age	0.893258	1.12454	0.720158	1.04239
Period	1373.6	637.683	1097.82	840.124
Cohort	1404.33	2346.25	1338.14	4148.42

Supplemental Table 3: Hyperparameter estimates

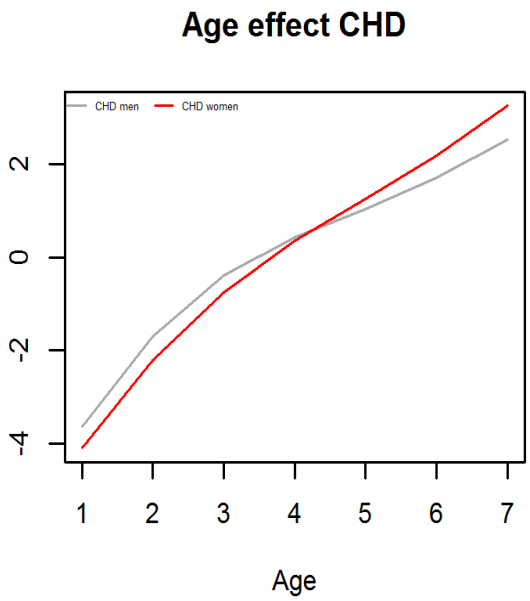
By far, the most important factor for both sexes in CHD and stroke was age, followed by period and cohort. However for CHD, cohort effects were as important as period effects, especially for men. This is an interesting finding since there is contradictory evidence whether CHD mortality has a cohort effect. APC models built for Australia [16]and New Zealand [17]and other descriptive analyses conducted for Poland and Hungary [18] suggested the absence of any cohort effect. However cohort effects were found in Singapore[19] and Norway[20]. While a descriptive study in Hong Kong [21] found cohort effects for women but not for men

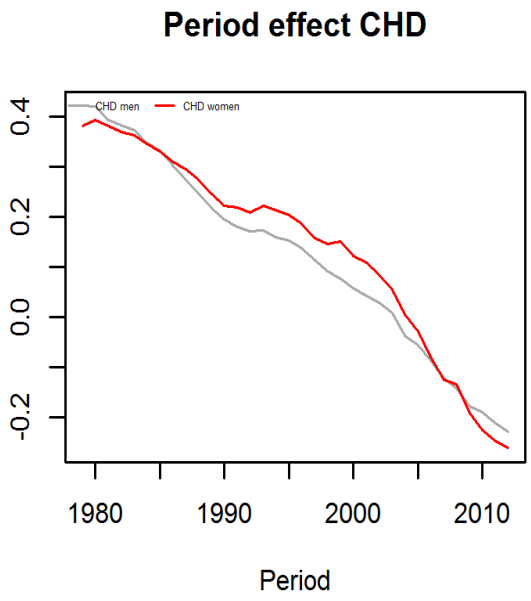
Looking at the analysis by ethnicity (supplemental table 3), we found only for white men and women the same phenomena. For the rest of the ethnicities cohort has a very small effect in comparison with age and period.

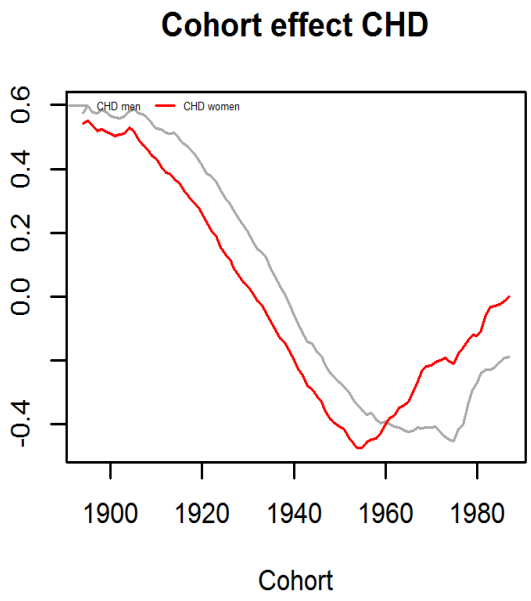
The next graphs show the age, period and cohort effects. Because of some assumptions imposed to the model, we can only interpret non-linear trends or change points. For example, for CHD and for both men and women there is a distinctive peak at the beginning of the 20th century. Then, we can observe a less steep slope for men born around 1955, which can be interpreted as a slowing of the effect. Finally, women and men born after 1955 and 1978 respectively experience an increasing trend.

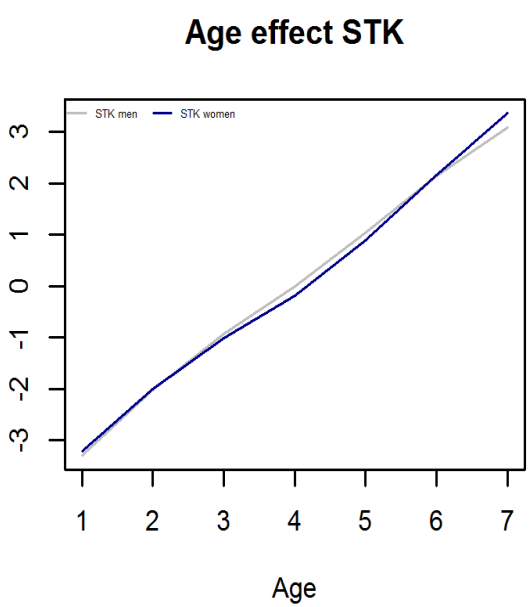
For stroke, we can see series of peaks before and around the First World War. This distinctive peak was also reported by Doll et al. [22], who reported that those birth cohorts were associated with an historical excess mortality associated with cigarette smoking. Women born around 1955-1960 experience an increasing slope. But, generations born just before 1970-1980 experiment a sharply

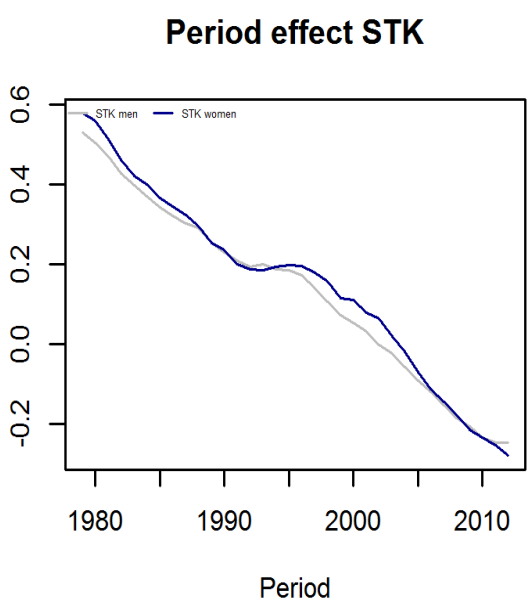
decreasing slope. New generations born after 1975 (for men) and 1980 (for women) experience an increasing trend

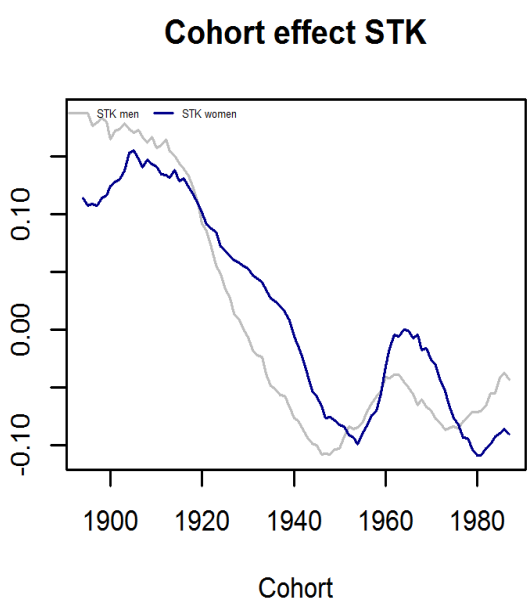

Age effect CHD
CHD men
CHD women
-4
-2
0
2
1
2
3
4
5
6
7
Age

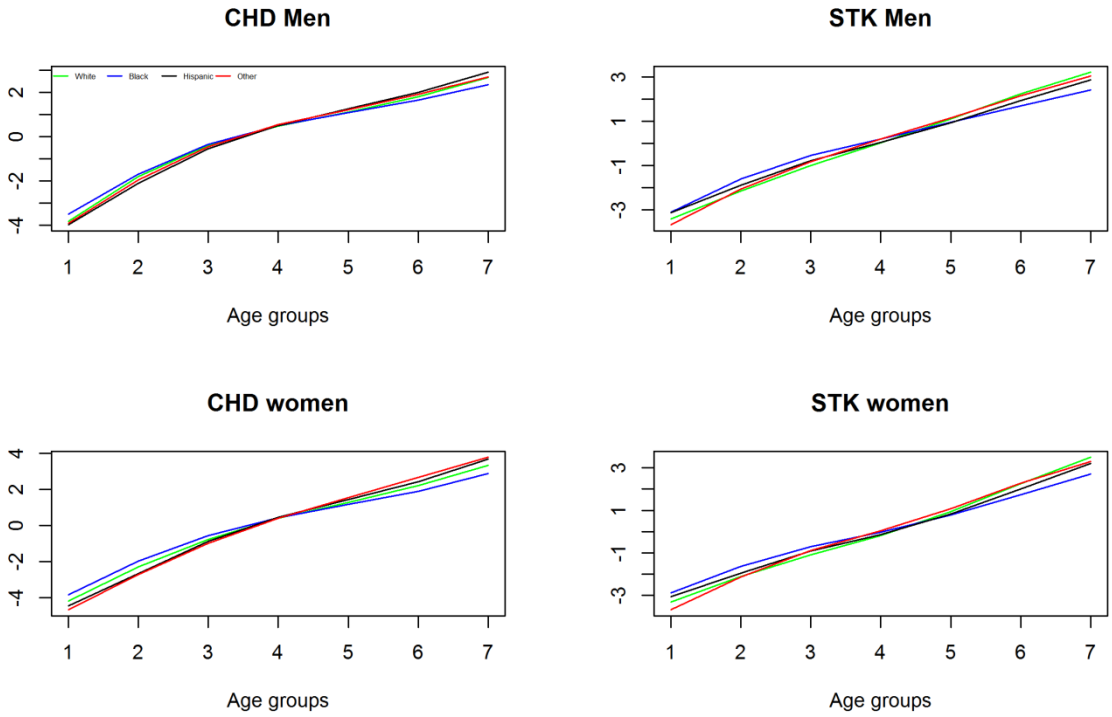

Period effect CHD
CHD men
CHD women
-0.2
0.0
0.2
0.4
1980
1990
2000
2010
Period


Cohort effect CHD
CHD men
CHD women
-0.4
0.0
0.2
0.4
0.6
1900
1920
1940
1960
1980
Cohort

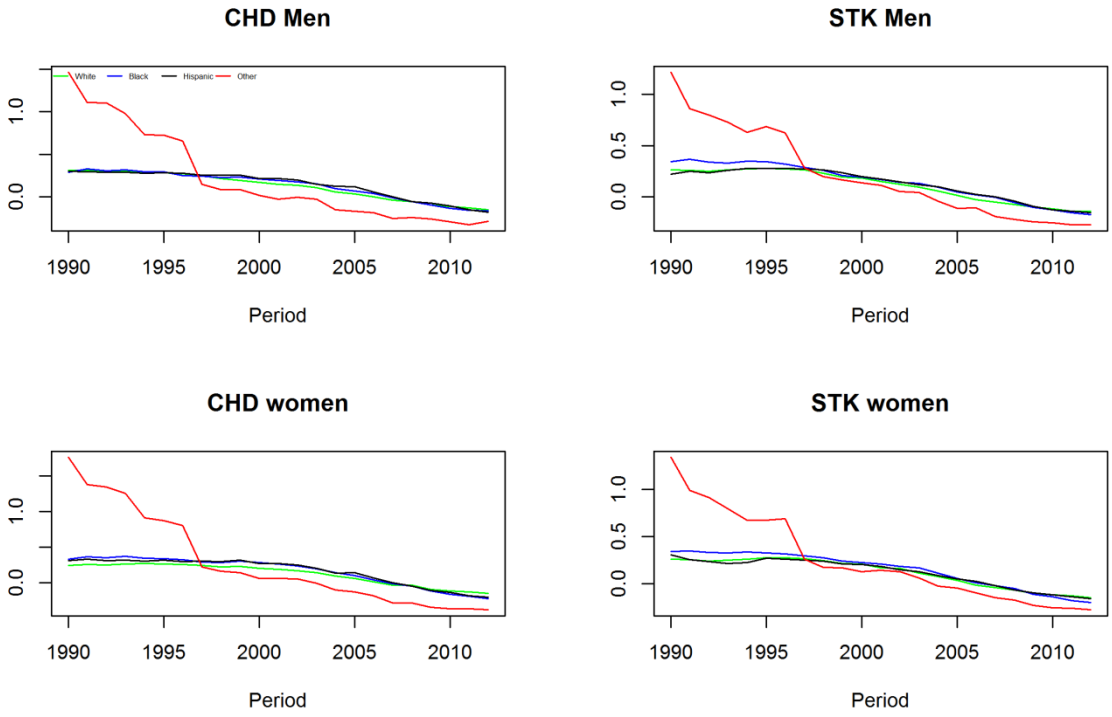

Age effect STK
STK men
STK women
-3
-2
-1
0
1
2
3
1
2
3
4
5
6
7
Age


Period effect STK
STK men
STK women
-0.2
0.0
0.2
0.4
0.6
1980
1990
2000
2010
Period

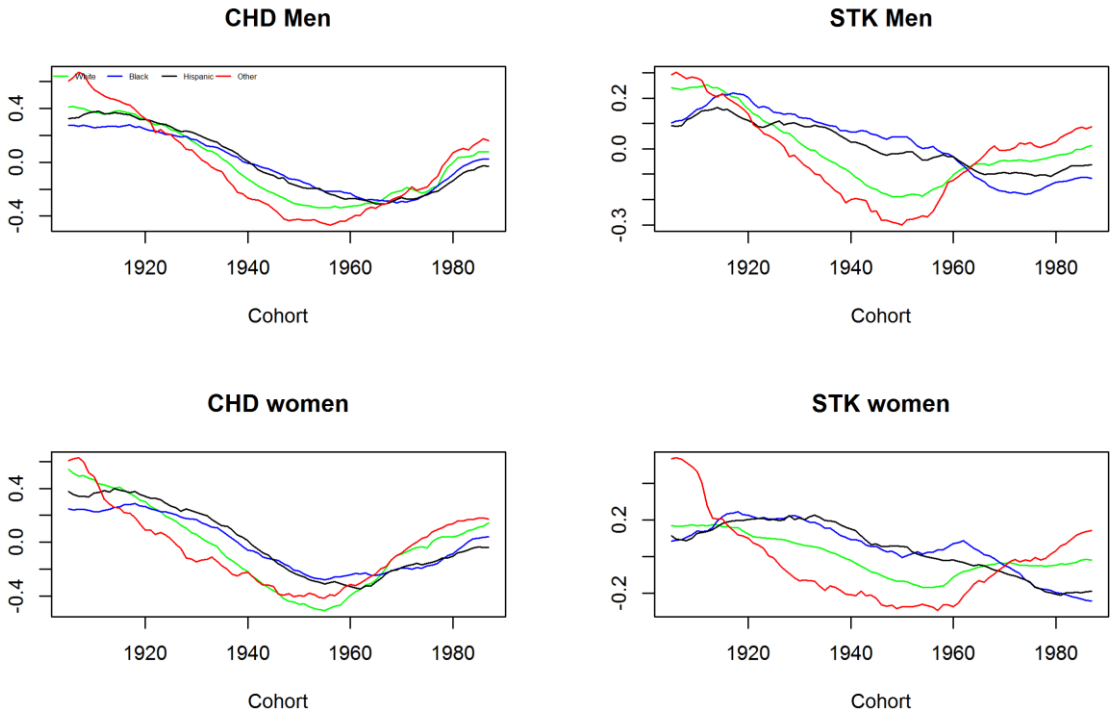

Cohort effect STK
STK men
STK women
-0.10
0.00
0.10
1900
1920
1940
1960
1980
Cohort



Supplemental Figure 1: Age effects by ethnicity, sex and cause of death



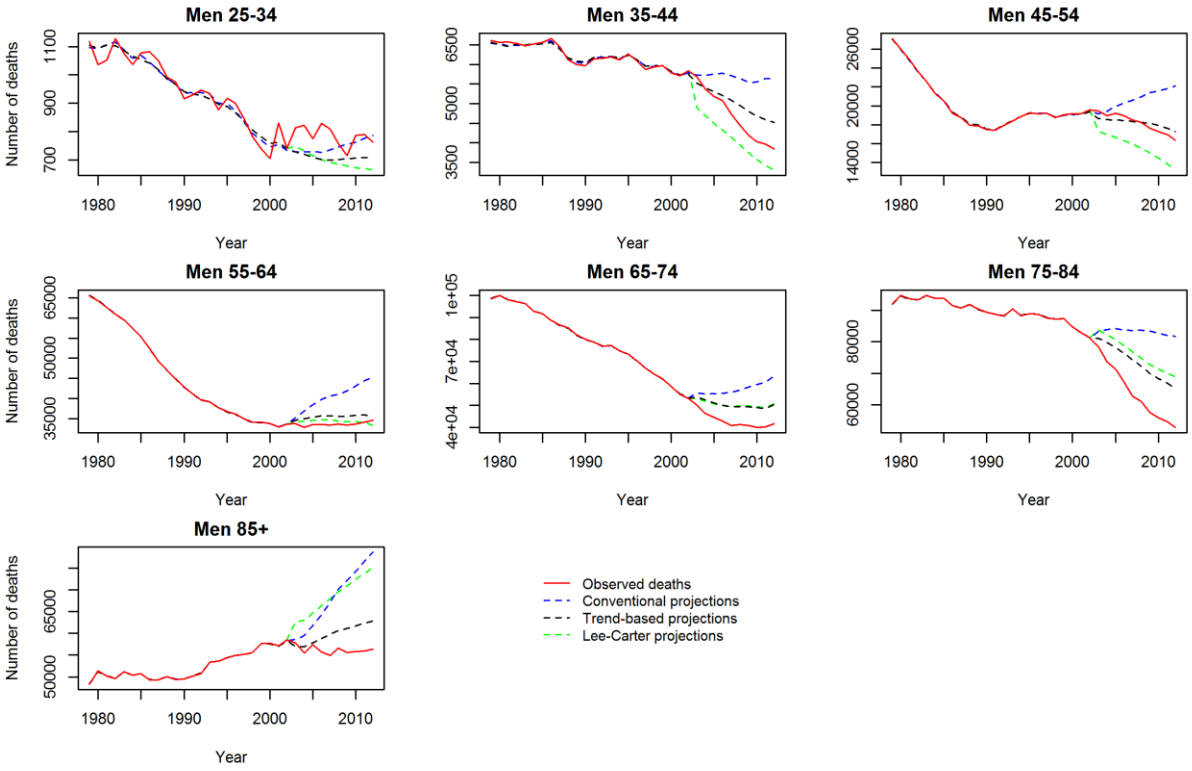
Supplemental Figure 2: Period effects by ethnicity, sex and cause of death



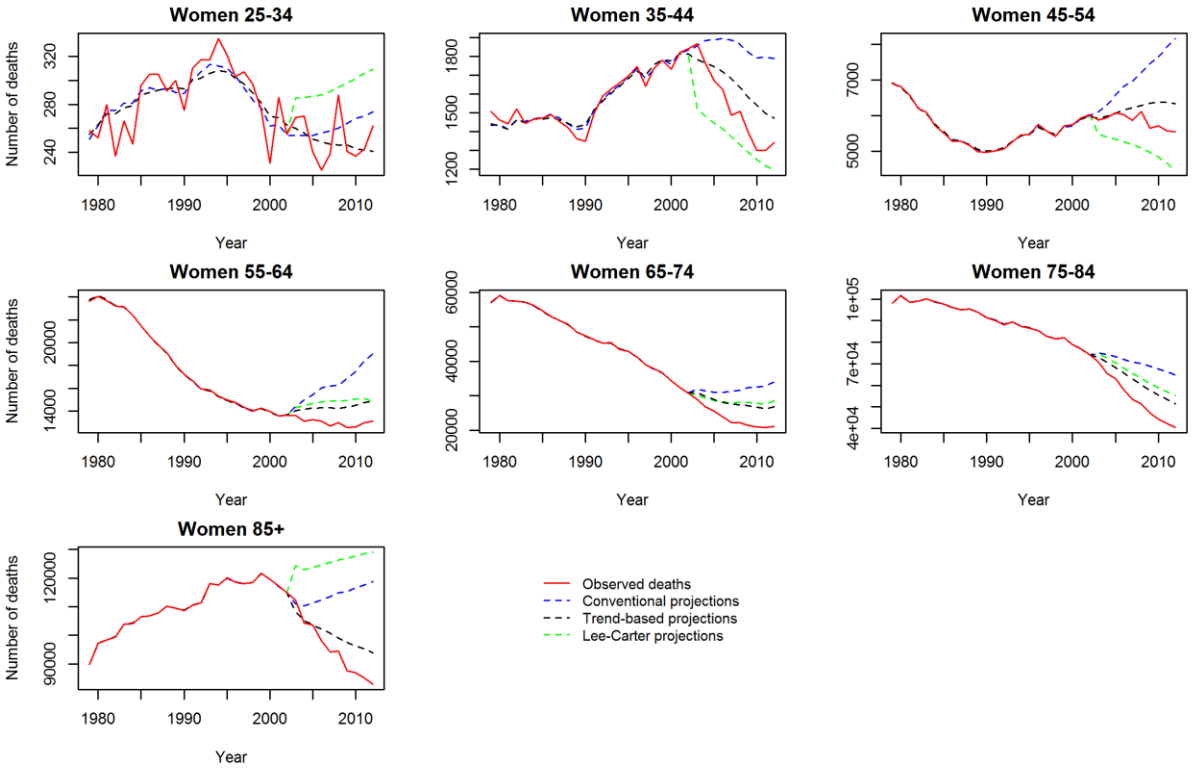
Supplemental Figure 3; Cohort effects by ethnicity, sex and cause of death

Supplemental Figure 4. Validation of mortality projections 2003-2012: Observed rates, Lee-Carter, 'trend-based' and conventional projections. 1A CHD Men, 1B, CHD Women, 1C Stroke Men, 1D Stroke Women

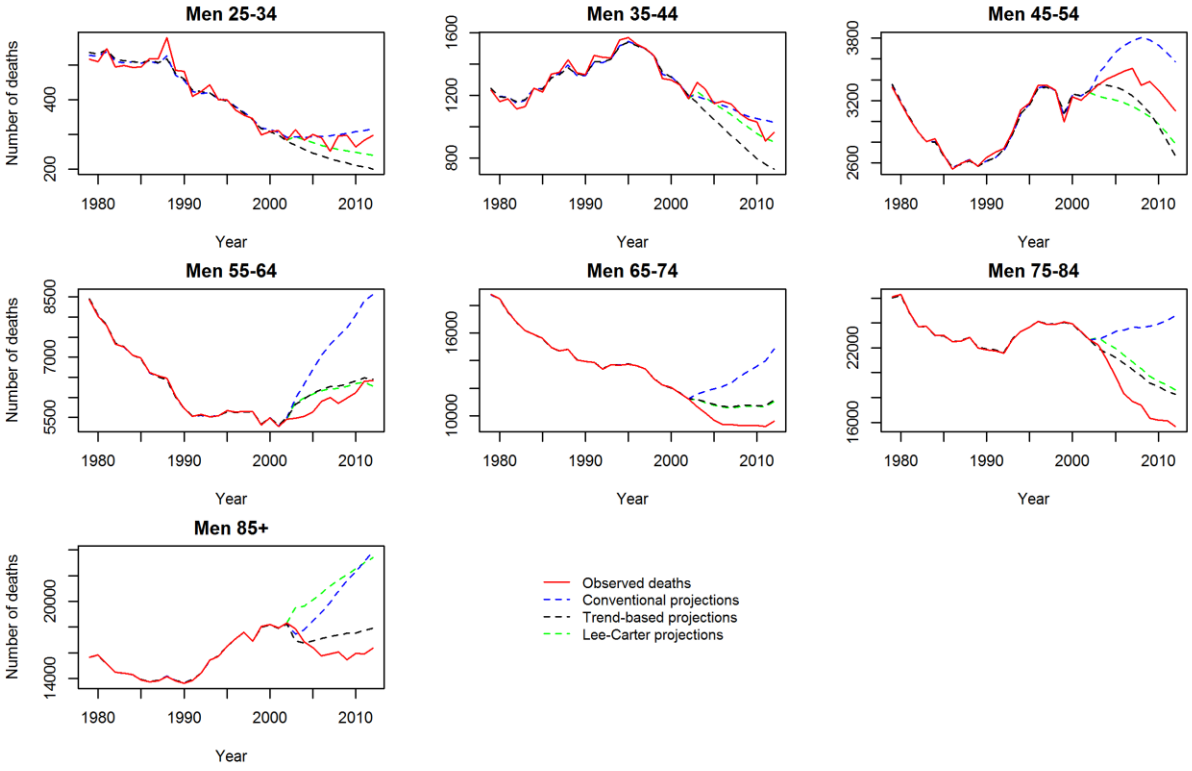
1A.



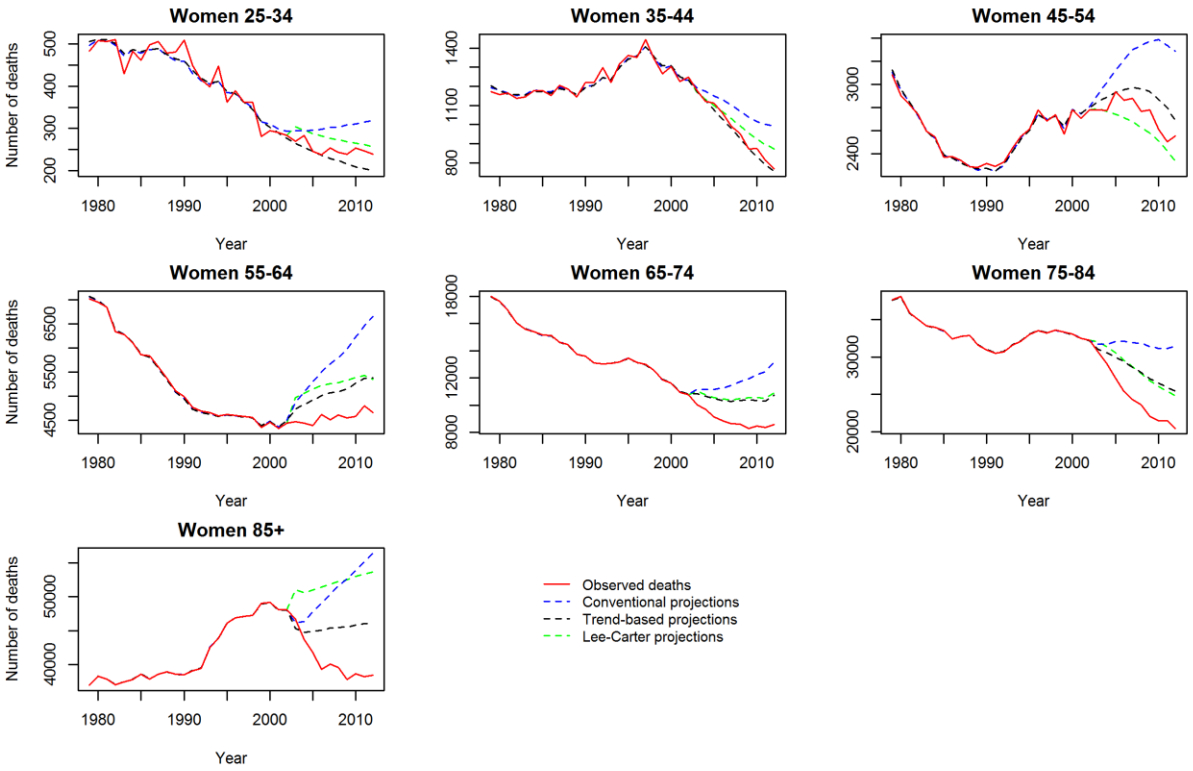
1B.


Women 25-34
Women 35-44
Women 45-54
Women 55-64
Women 65-74
Women 75-84
Women 85+
Number of deaths
Year
Observed deaths
Conventional projections
Trend-based projections
Lee-Carter projections

1C.


Men 25-34
Men 35-44
Men 45-54
Men 55-64
Men 65-74
Men 75-84
Men 85+
Number of deaths
Year
Observed deaths
Conventional projections
Trend-based projections
Lee-Carter projections

1D.


Women 25-34
Women 35-44
Women 45-54
Women 55-64
Women 65-74
Women 75-84
Women 85+
Number of deaths
Year
1980
1990
2000
2010
Observed deaths
Conventional projections
Trend-based projections
Lee-Carter projections

	Hyperparameter	CHD men	STK men	CHD women	STK women
White	**Age**	0.86	0.93	0.72	0.90
	Period	1228	1040	1158	1017
	Cohort	1371	2382	1005	2980
Black	**Age**	1.10	1.37	0.85	1.37
	Period	852	752	662	764
	Cohort	1711	2118	1688	1772
Hispanic	**Age**	0.85	1.08	0.61	1.11
	Period	763	835	548	720
	Cohort	1388	2421	1033	1812
Other	**Age**	0.87	0.94	0.52	0.89
	Period	43	72	34	56
	Cohort	527	875	379	445

Supplemental Table 4: Hyperparameter estimates by ethnicity

References

1. Holford T. Age-period-cohort analysis. In: Armitage P, Colton T, eds. Encyclopedia of Biostatistics. 2nd ed. West Sussex: John Wiley and Sons; 2005: 105-23.

2. Schmid VJ, Held L. Bayesian age-period-cohort modeling and prediction-BAMP. *J Stat Software* 2007; **21**(8): 1-15.

3. Clayton D. Generalized Linear Mixed Models. In: Gilks WR, Richardson S, Spiegelhalter DJ, eds. Markov chain Monte Carlo in practice. London: Chapman & Hall; 1996: 274-301.

4. Knorr-Held L, Rainer E. Projections of lung cancer mortality in West Germany: a case study in Bayesian prediction. *Biostatistics* 2001; **2**(1): 109-29.

5. Glenn ND. Cohort analysts' futile quest: Statistical attempts to separate age, period and cohort effects. *American sociological review* 1976; **41**(5): 900-4.

6. Goldstein H. Age, period and cohort effects: A confounded confusion. *Journal of Applied Statistics* 1979; **6**(1): 19-24.

7. Glenn ND. Cohort analysis: SAGE Publications, Incorporated; 2005.

8. Metropolis N, Rosenbluth A, Rosenbluth M, Teller A, Teller E. Equation of state calculations by fast computing machines. *The Journal of chemical physics* 1953; **21**: 1087.

9. Geman S, Geman D. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 1984; (6): 721-41.

10. Spiegelhalter DJ, Best NG, Carlin BP, Van Der Linde A. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 2002; **64**(4): 583-639.

11. Carter LR, Lee RD. Modeling and forecasting U.S. sex differentials in mortality. *Int J Forecast* 1992; **8**(3): 393-411.

12. Di Cesare M MM. Forecasting mortality, different approaches for

different cause of deaths? The cases of lung cancer; influenza, pneumonia, and

bronchitis; and motor vehicle accidents. British Actuarial Journal; 2009.

13. Preston S, Heuveline P, Guillot M. Demography: Measuring and modeling population processes. Oxford: Blackwell; 2001.

14. Hyndman R. demography: Forecasting mortality, fertility, migration and population data. R package version 1.16.; 2013.

15. SN W. Monotonic smoothing splines fitted by cross validation. SIAM Journal on Scientific Computing. p. 1126-33.

16. Taylor R, Page A, Danquah J. The Australian epidemic of cardiovascular mortality 1935–2005: effects of period and birth cohort. *Journal of Epidemiology and Community Health* 2012; **66**(7): e18-e.

17. Tobias M, Sexton K, Mann S, Sharpe N. How low can it go? Projecting ischaemic heart disease mortality in New Zealand to 2015. *Special Series* 2006.

18. Bobak M, Jarvis MJ, Skodova Z, Marmot M. Smoke intake among smokers is higher in lower socioeconomic groups. *Tobacco Control* 2000; **9**(3): 310-2.

19. Hughes K. Trends in mortality from ischaemic heart disease in Singapore, 1959 to 1983. *International journal of epidemiology* 1986; **15**(1): 44-50.

20. Sverre JM. Secular Trends in Coronary Heart Disease Mortality in Norway, 1966–1986. *American Journal of Epidemiology* 1993; **137**(3): 301-10.

21. Yu TS, Wong SL, Lloyd OL, Wong TW. Ischaemic heart disease: trends in mortality in Hong Kong, 1970-89. *Journal of Epidemiology and Community Health* 1995; **49**(1): 16-21.

22. Doll R, Peto R, Boreham J, Sutherland I. Mortality in relation to smoking: 50 years' observations on male British doctors. *Journal of Epidemiology and Community Health* 2004; **58**(11): 930-.