Linking survival modelling results to life expectancy differentials

Elena Kulinskaya
School of Computing Sciences,
University of East Anglia

30/10/2019

www.actuaries.org.uk/arc

Quantifying Longevity Changes

- Medical and social advances are the major drivers in the longevity increase. But how to quantify this relationship?

- In medicine, Randomized Control Trials (RCTs) are considered to be the gold standard.

- RCTs estimate the hazard or force of mortality in a (selective) sample of people and summarised over the observed (limited) time period.

- New health interventions are usually based on these estimated hazards obtained from clinical trials. A lengthy lead time would be needed to observe their effect on population longevity.

Actuarial
Research Centre
Institute and Faculty
of Actuaries

Our approach, 1

- Our research uses The Health Improvement Network (THIN) primary care data to develop statistical models of longevity.

- The advantage of using individual-level medical data is that it is possible to model both the uptake of medical treatment and the effect of that treatment on longevity conditional on the individual sociodemographic and health factors instead of the aggregated profile.

- We carefully design each observational study and match cases to controls. Survival models, usually the Cox regression, are fitted to such individual level data.

- The conclusions are generalisable to the general population.

Actuarial
Research Centre
Institute and Faculty
of Actuaries

Hazard ratio

- The type of regression model typically used in survival analysis in medicine is the Cox proportional hazards regression model.

- The Cox model estimates the hazard $\mu_i(x)$ for subject i at time x as $\mu_i(x, \beta, Z_i) = \mu_0(x)\, r_i(\beta, Z_i) = \mu_0(x)e^{\beta\, Z_i}$

- Taking a ratio of the hazard functions for two subjects i and j who differ in one risk factor z and not in the other risk factors,

$$\mu(x, \beta, Z) = \frac{\mu_i(x,\beta,Z_i)}{\mu_j(x,\beta,Z_j)} = \frac{\mu_0(x)e^{\beta\, Z_1}}{\mu_0(x)e^{\beta\, Z_0}} = \frac{e^{\beta_z\, z_1}}{e^{\beta_z\, z_0}} = e^{\beta_z\,(z_0 - z_1)}$$

From a hazard to effective age

- For simplicity, consider a binary risk factor with the reference value y = 0 and (at risk) y = 1.

- On the log scale, the log-hazards are $\lambda_1(t) = \lambda_0(t) + \beta$. This means that the log-hazard lines differ only by an increment

- For a monotone-increasing hazard, find the (unique) time increment $\Delta(t)$ such that $\lambda_1(t) = \lambda_0(t+ \Delta(t))$
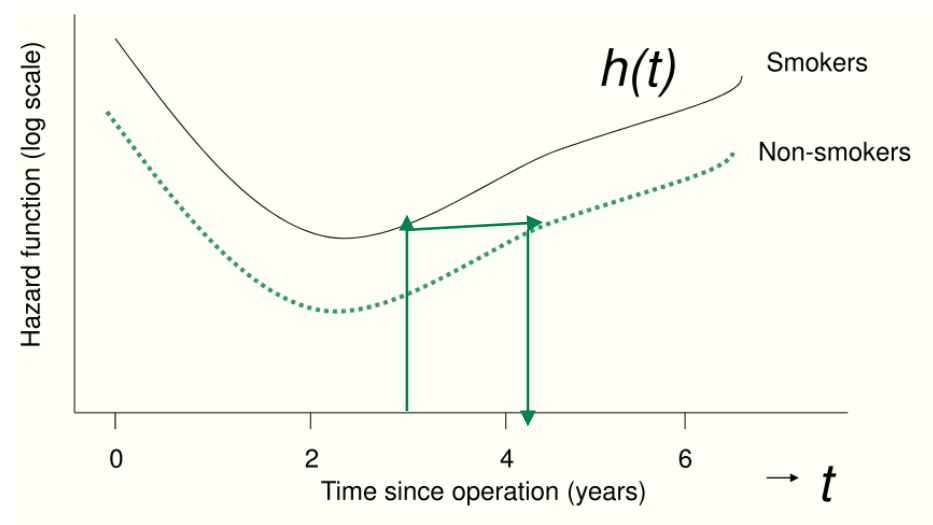
Value of t + $\Delta(t)$ is, by definition, the **effective age** of the person with risk y = 1 at chronological age t.

Our approach, 2: for an individual

- For an individual, the hazard ratios obtained from the survival models are translated into effective age changes.

- Effective age at y=1 is the average chronological age with the same hazard as when y=0.

- Effective ages are often used by insurers as a way of applying the correct rating to an underwritten life.

Log force of mortality for UK population based on 2010 period life table (Office for National Statistics 2017).
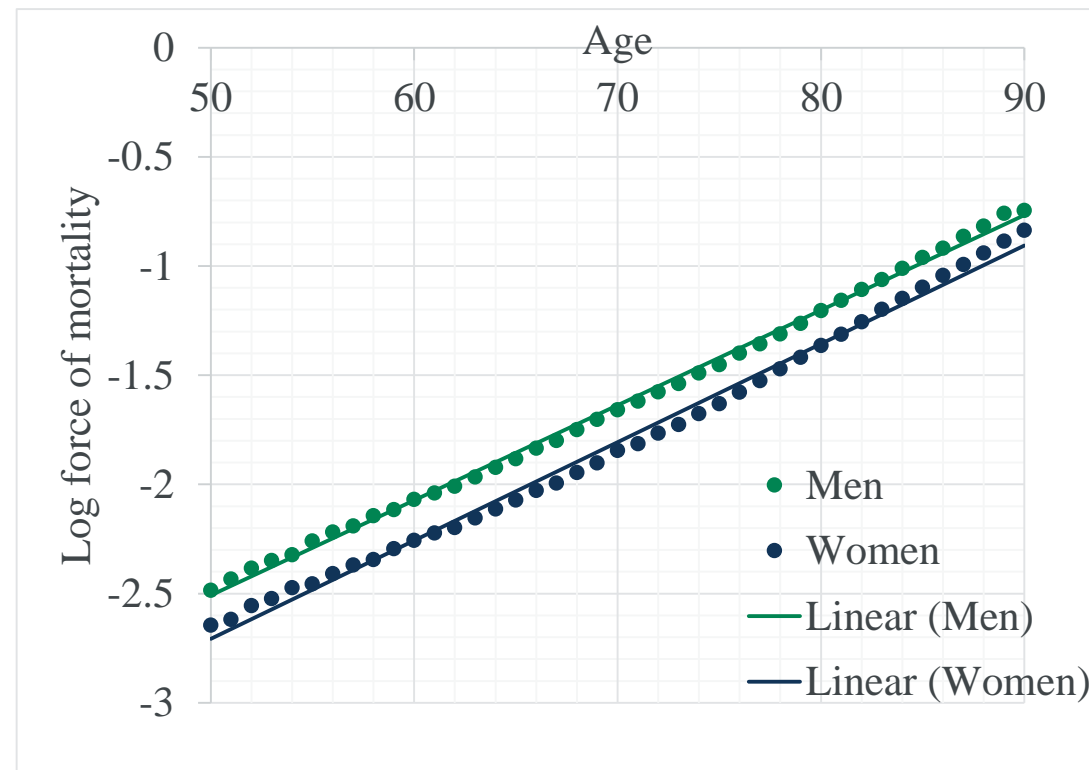
What does HR mean for an individual

- Using Gompertz law $\lambda_0(t)=a+bt$, the increase in annual hazard of mortality associated with ageing one year is approximately constant between ages 50 and 90.

   For y=1, $\lambda_1(t) = a+bt+ \beta=a+b(t+\Delta)$ ➡ $\Delta=\beta/b$

- For England and Wales in 2010-2012, the increase in the hazard between those ages was approximately 1.1 per year.

- A HR can be translated to the numbers of years gained in effective age as

   $\Delta=\log(HR) / \log(1.1) \approx 10*\log(HR)$.
   [Brenner, 1993; Spiegelhalter, 2016]

- For LE at age t, $e_1(t) = e_0(t +\Delta)$.

Risk factor
Baseline
Hazard of mortality
Chronological age   Effective age
Age

Actuarial Research Centre
Institute and Faculty of Actuaries
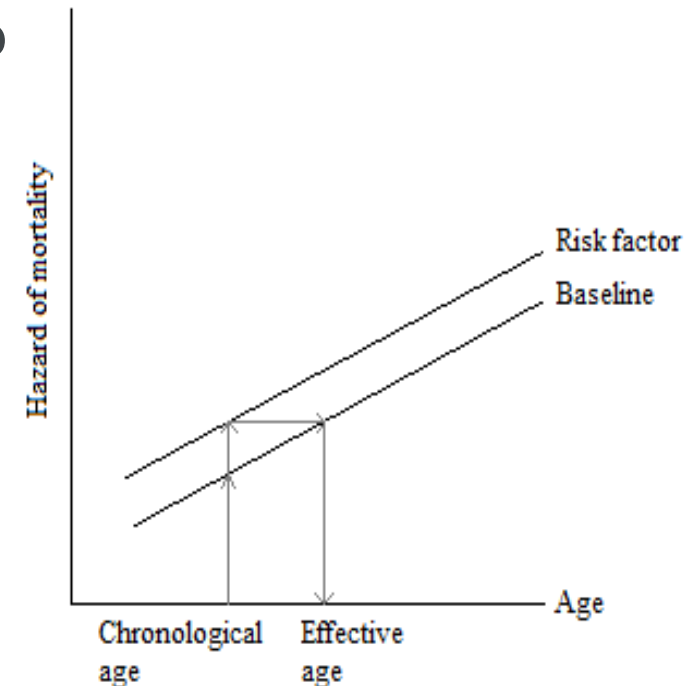
Our approach, 3: Period life expectancy

- Consider a population consisting of J risk groups of prevalence q_j, j=1,…J; $\Sigma\, q_j = 1$.

- Treatment of interest (i=0, 1) is prevalent in the population from age T but its effects vary across risk groups.

- Prevalence of the treatment of interest in group j at age T is $p_{j,I}$

- Then the overall survival function S(T) at age x=T is the weighted mean of the survival functions in the individual risk groups

$$S(T) = \left[\sum_j q_j p_{j,1} S_{1j}(T) + \sum_j q_j (1 - p_{j,1}) S_{0j}(T)\right] / \sum_j q_j.$$

The Cox model

- Assume that the hazards are proportional, so that the hazards

   $\mu_{ij} = \mu_0(x)\mu_{ij}(Y)$, where $\mu_0(x)$ is the baseline hazard at age x

and

$$\log(\mu_{ij}(Y)) = a_{ij} = a_0(T) + \alpha_i + \beta_j + \gamma_{ij} + \beta^T Y,$$

where a $_0$(T) is the baseline value which may depend on intervention time T, α_i, β_j and γ_{ij} are the main effects and interaction of risk group j and treatment I, and the other covariates Y have no interactions with the treatment or the risk of interest.

Survival function under Gompertz-Cox model

- The log-hazards in a risk group (i,j) are just a_{ij} + bx, i.e. the straight lines with the same slopes but differing intercepts.

- The survival functions are $\quad S_{ij}(x) = \exp(-e^{a_{ij}}b^{-1}(e^{bx}-1))\quad$.
Substituting the a_{ij}, the survival functions at age x >T are

$$S_{ij}(x|Y) = \exp(-e^{a_0(T)+\alpha_i+\beta_j+\gamma_{ij}+\beta^TY}b^{-1}(1-e^{bx})).$$

- Assuming that the prevalences do not depend on Y, Y can be integrated out to obtain

$$S(x|T) = \sum_j q_j p_{j,1} \exp(-e^{a_0(T)+\alpha_1+\beta_j+\gamma_{1j}}b^{-1}(1-e^{bx})) +$$

$$\sum_j q_j(1-p_{j,1}) \exp(-e^{a_0(T)+\beta_j}b^{-1}(1-e^{bx})).$$

Actuarial Research Centre

Institute and Faculty of Actuaries

Finding component survival functions

$$S(x|T) = \sum_j q_j p_{j,1} \exp(-e^{a_0(T)+\alpha_1+\beta_j+\gamma_{1j}} b^{-1}(1-e^{bx}))+$$

$$\sum_j q_j(1-p_{j,1}) \exp(-e^{a_0(T)+\beta_j} b^{-1}(1-e^{bx})).$$

- This is a non-linear equation with one unknown, a_0. The left-hand side is given by the period life-table, and the slope b should be determined for a particular population of interest. As S(x) is a decreasing function of a_0, it has a unique solution.

- After solving for a_0(T), we can find component survival functions S_{ij}(x) for any set of prevalences {q_j} and {$p_{j,1}$}.

Estimating changes in life expectancy

- For a Gompertz distribution G(a,b), the LE is

$$e_{G(a,b)}(z) = \frac{b^{-1}\exp(b^{-1}e^a)E_1(b^{-1}e^{a+bx})}{\exp(-e^a b^{-1}(e^{bx}-1))}.$$

So we can find component LEs $e_{ij}(z)$ for each component distribution $G(a_{ij},b)$. Then, the life expectancy at age z is

$$e(z) = \frac{\int_z^\infty S(x)dx}{S(z)} = \frac{\sum_k w_k S_k(z)\int_z^\infty S_k(x)dx/S_k(z)}{\sum w_k S_k(z)} = \frac{\sum_k w_k S_k(z)e_k(z)}{\sum w_k S_k(z)}.$$

- Taking all $p_{j,1} = 0$, we obtain a hypothetical life expectancy $e_0(z)$ if there were no intervention of interest, and, for all $p_{j,1} = 1$, a hypothetical life expectancy $e_1(z)$ with full uptake of the intervention.

Actuarial Research Centre
Institute and Faculty of Actuaries

Case study: survival benefits of statins

- We used the data for QRISK2 groups 10−19% and ≥ 20% at ages 70 and 75 (yob 1920-1940, observed 1987-2011) from Gitsels et al. [2016]

- We fitted the same Cox models after adding the QRISK2 group to the predictors. The final models adjusted for sex, birth cohort, socioeconomic status, diabetes, hyper-cholesterolaemia, blood pressure regulating drugs, body mass index, and smoking status. The models included a random effect on general practice.

- Interactions between statins, QRISK2 groups and the other risk factors were tested, but none was significant.

- We also used the adjusted HRs for all-cause mortality of heart attack survivors, from Gitsels et al. [2017], as a substitute for HRs for CVD sufferers. The HRs are given in Table 1.

Table 1. Population characteristics, statins study by Gitsels et al. 2016

Cohort	Cardiac risk	Women % (Statins %)	Men % (Statins %)
Age 70	QRISK2 10-19%	80 (9.5)	17 (5.4)
N=247,149, FU=7 years	QRISK2>20%	20 (28.2)	83 (17.4)
Age 75	QRISK2 10-19%	15 (4.6)	0 (0.0)
N=194,085, FU=6 years	QRISK2>20%	85 (19.6)	100 (19.1)

* FU= average follow-up

Actuarial Research Centre
Institute and Faculty of Actuaries

Table 2. Hazard ratios of statins and of cardiac risk groups

Cardiac risk	Age	HR statins (vs no statins)	Changes in effective age (men) [1]	Changes in effective age (women) [1]
No heart attack	70	0.84(0.80, 0.88)	-1.69 (-2.16,-1.24)	-1.57 (-2.01, -1.15)
	75	0.82 (0.79, 0.86)	-1.92 (-3.18,-1.46)	-1.79 (-2.97, -1.36)
Heart attack	70	0.74 (0.70, 0.78)	-2.91 (-3.45, -2.40)	-2.72 (-3.22, -2.24)
	75	0.77 (0.74, 0.81)	-2.53 (-2.91, -2.04)	-2.36 (-2.72, -1.90)
Cardiac risk	Age	HR Cardiac Risk	Changes in effective age (men)	Changes in effective age (women)
QRISK2 10-19%	70	0.80 (0.77, 0.83)	-2.16 (-2.53, -1.80)	-2.01 (-2.36, -1.68)
	75	0.87 (0.80, 0.94)	-1.35 (-2.16, -0.60)	-1.26 (-2.01, -0.56)
QRISK2>20%	70	1	0	0
	75	1	0	0
Heart attack	70	1.50 (1.42,1.59)	3.92 (3.39, 4.48)	3.66 (3.17, 4.19)
	75	1.45 (1.38,1.53)	3.59 (3.11, 4.11)	3.35 (2.91, 3.84)

[1] based on Gompertz distributions with b=0.1034 for men and 0.1108 for women.

Institute and Faculty of Actuaries

Info on the prevalences of the risk groups and the treatment

- Prevalence of risk groups (q's on slide 9): QRISK2 score increases with age and by age 70, there were practically no patients with a QRISK2 score of < 10% and by age 75, there were no male patients with a QRISK2 score of < 20%.

- Prevalence of treatment(p's on slide 9): At the end of study period in 2010, statins were prescribed in 20% of patients with a QRISK2 score of < 20%, in 45% of patients with a QRISK2 score of ≥ 20%, and in 90% of patients with CVD.

- Given cardiac risk group, statins were prescribed more in women, in younger patients, and in patients from less deprived areas.
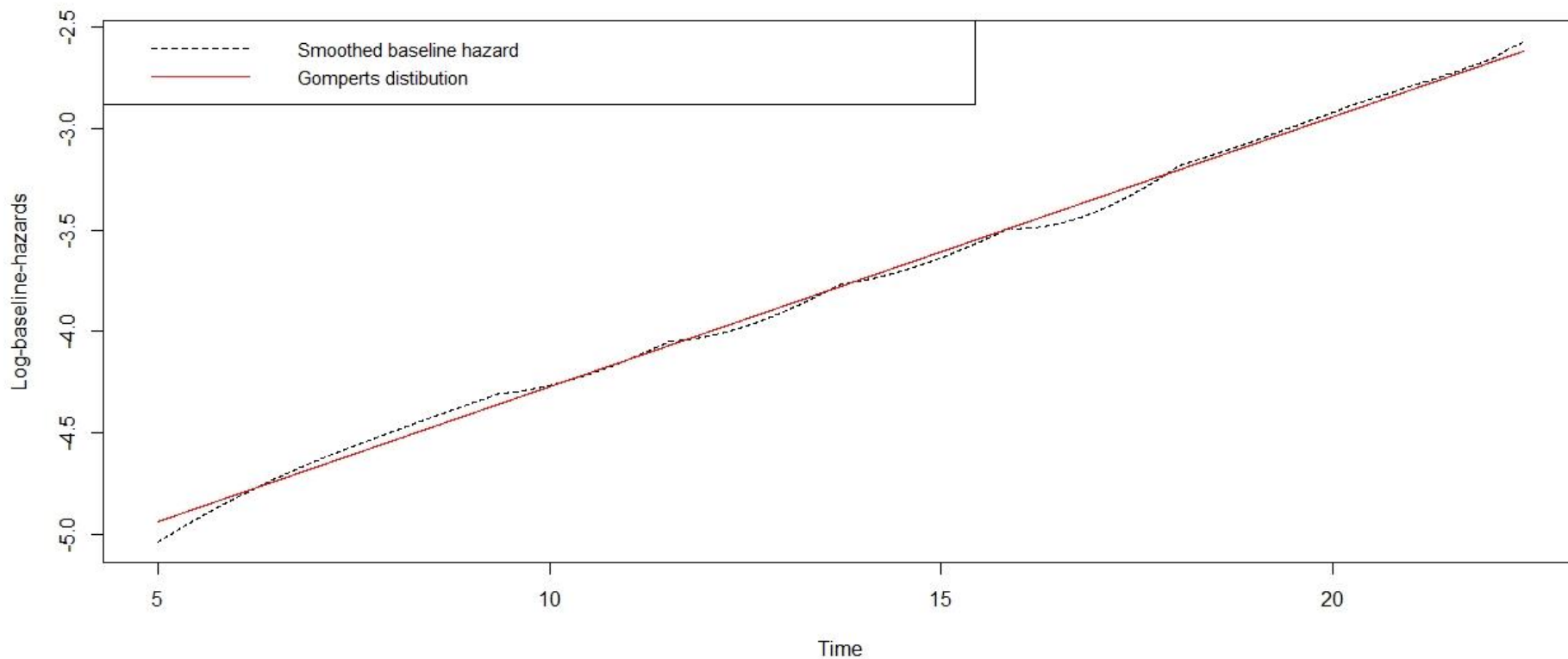
Fitting Gompertz distribution to period life tables
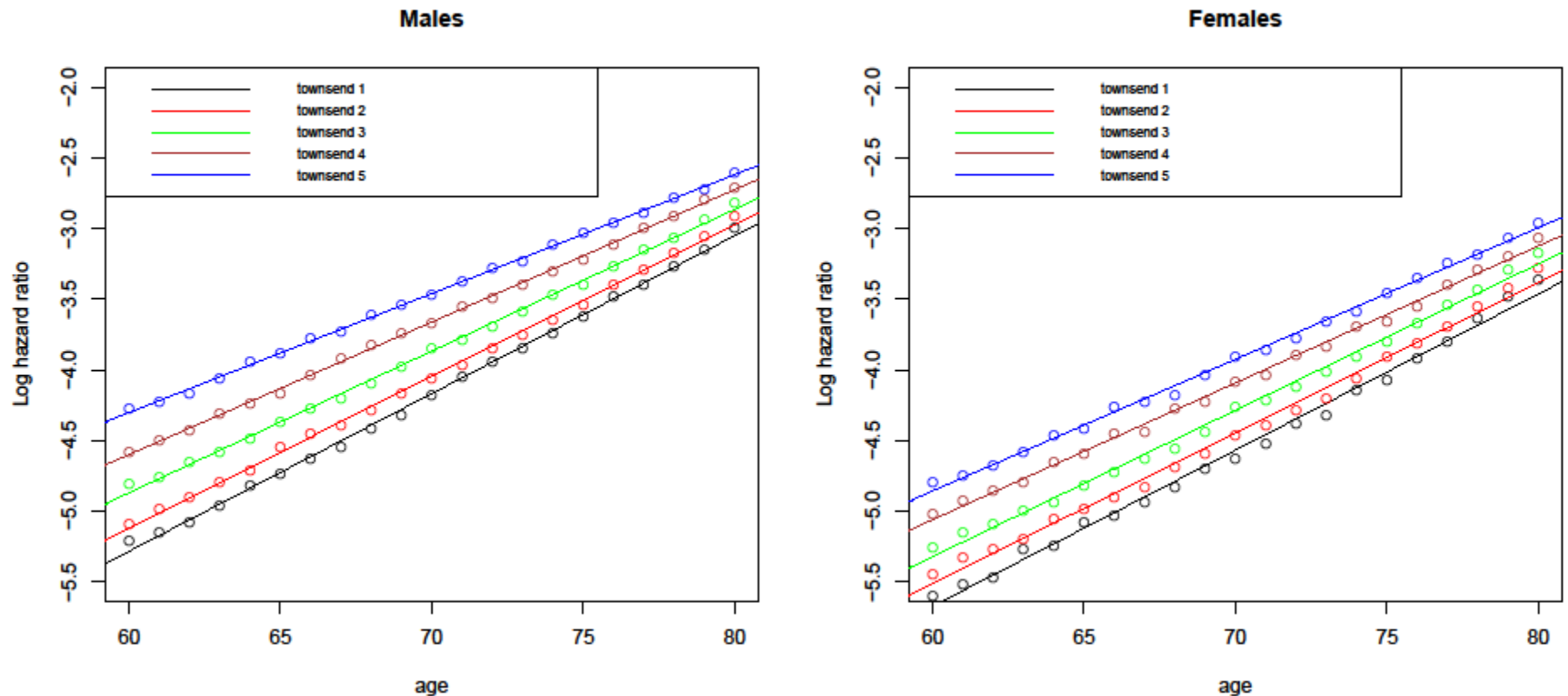
Log-hazards between the ages 70- 90 from the ONS period life table centered at 2010 (circles) and fitted regression lines by Townsend score quintiles and sex.

Baseline hazard in the statins survival model

The baseline hazard is well approximated by the Gompertz hazard.

Calculating component life expectancies

- Since the mortality rates, the cardiac risk distribution and the statin prescription rates differ by gender and by socio-economic status, we analysed the life tables separately for each Townsend score quintile- by-gender combination.

- For each life table, we substitute the S(x) at age x=70 or 75 (obtained from the fitted Gompertz distribution G(a,b)) into the left-hand side of the equation on top of slide 13, and solve for the value of a_0(T).

- These values were used to calculate period life tables for component cardiac risk by statin prescription subpopulations for each (i,j) combination.
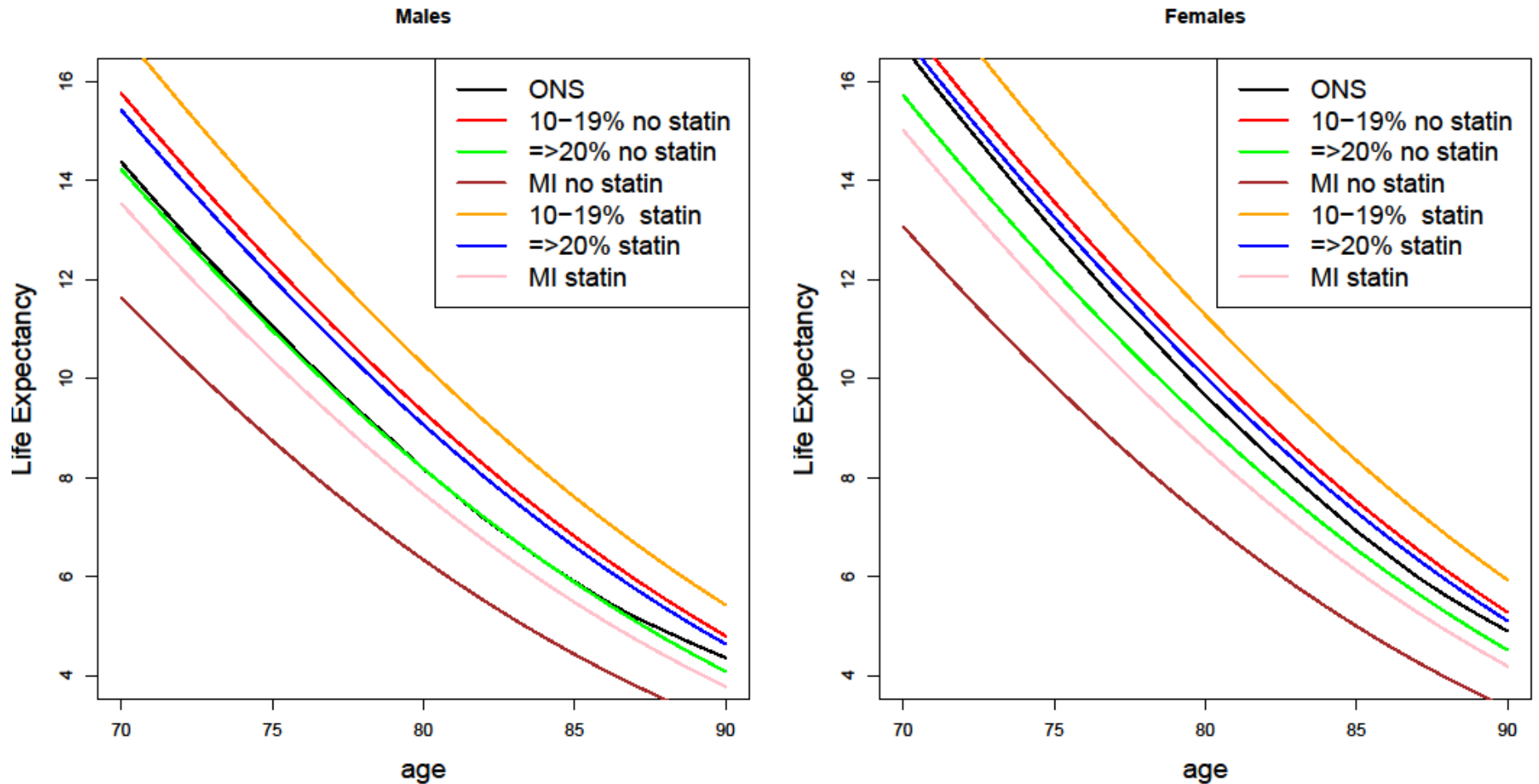
Results, 1

- Increase in individual LE due to statins depends on cardiac risk, and is highest for heart attack survivors (1.41−2.02 years), and is comparable in the two QRISK2 groups (1.14−1.35 years across ages 70 and 75). The effect of statins increases with deprivation.

- We also calculated the period LE and its increase due to statins in each cardiac risk group for the total England and Wales population by averaging the LE across all TS quintiles, a and plotted the results.

Actuarial
Research Centre

Institute and Faculty
of Actuaries

Life expectancy by cardiac risk group with and without statins for ages 70 − 90 based on the ONS period life table centered at 2010

Males

Females

Results, 2

- We also calculated national life expectancy with and without statins, by averaging the LE across cardiac risk groups, taking p = 0 (for no statins) and p = 1 (for statins).

- The national life expectancy for women aged 70 or 75 would be increased by up to 0.91 or 0.79 years, respectively, if all eligible women under the current guideline of primary and secondary prevention of CVD were prescribed statins.

- Similarly, the national life expectancy for men aged 70 or 75 would be increased by up to 0.79 or 0.63 years. The most improvement would come from the areas of medium deprivation.

Actuarial
Research Centre
Institute and Faculty
of Actuaries

An app for general public

References:

1. Brenner H, Gefeller O, Greenland S. (1993) Risk and rate advancement periods as measures of exposure impact on the occurrence of chronic diseases. *Epidemiol Camb Mass*. 4(3):229–36.

2. Gitsels L.A., Kulinskaya E., Steel N. (2016) Survival Benefits of Statins for Primary Prevention: A Cohort Study. PLoS ONE **11**(11): e0166847. doi:10.1371/journal.pone.0166847

3. Gitsels LA, Kulinskaya E, Steel N Survival prospects after acute myocardial infarction in the UK: a matched cohort study 1987–2011 BMJ Open 2017;7:e013570. doi:10.1136/bmjopen-2016-013570.

4. Kulinskaya E., Gitsels, LA. and Bakbergenuly, I., 2019. Calculation of changes in individual and period life expectancy based on proportional hazards model of an intervention. *Insurance Mathematics and Economics,* under review.

5. Spiegelhalter (2016) How old are you, really? Communicating chronic risk through 'effective age' of your body and organs. *BMC Medical Informatics and Decision Making,* 16:104

Actuarial
Research Centre
Institute and Faculty
of Actuaries