Use of large population-based primary care data to model variations and trends in life expectancy

Prof Elena Kulinskaya (UEA) and Mr Nigel Wright (Aviva Life Analytics)

with contributions from Prof Nicholas Steel and Lisanne Gitsels (UEA),
and Sarah Allen (Aviva)

Content

IFoA grant on modelling longevity

- Aims
- Methodology

Case Study on Statins

- Rationale
- Study design & Methodology
- Results

Discussion

Use Of Big Health And Actuarial Data For Understanding Longevity And Morbidity Risks, IFoA 2016-2020

Consortium:

University of East Anglia:
School of Computing Sciences (CMP) and Norwich Medical School (NMS).
Aviva Life Plc.

Principal Investigator Prof Elena Kulinskaya, Aviva Chair in Statistics, CMP

UEA co-investigators: Dr Beatriz de la Iglesia, Senior Lecturer, CMP;
Prof Ruth Hancock, NMS, Prof Nick Steel, NMS.

Aviva co-investigators: Mr Nigel Wright, actuary; Ms Sarah Allen, Senior Data Analyst, the Life Risk Analytics team.

Main objectives

Development of novel statistical and actuarial methods for:

modelling mortality

modelling trends in morbidity and uptake of health interventions

assessing basis risk

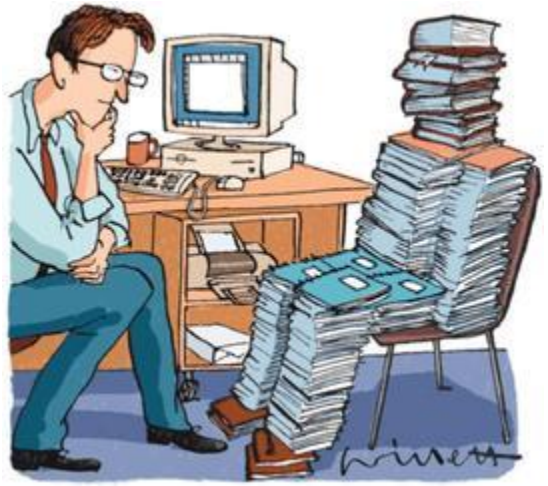
evaluating longevity improvement based on Big Health and Actuarial Data

tools to forecast longevity risk of a book

Data

The Health Improvement Network (THIN) data

- Medical records from primary care
- Representative of the UK when adjusted for deprivation



- All patients born before 1960 and followed to 01.01.2015,
  this includes 3.4 million patients

- Added various social economic status variables such as IMD and Mosaic

- The Continuing Mortality Investigation (CMI) data

Aim 1: Identification and quantification of the key factors affecting mortality/ longevity

We intend to have a target list of between 3-5 conditions or interventions.

We propose to consider statin prescription, an established longevity-improving intervention as one of the target scenarios.

Other conditions may include type 2 diabetes or heart failure.

Health interventions may include an introduction of NICE guidelines on use of particular health sustaining drugs such as calcium channel blockers, or targeted outcomes such as the blood pressure targets.

Lifestyle factors may include obesity or smoking.

Design and methods

For each of these conditions we will design a population-based prospective cohort study using an appropriate extract of the primary care data.

We intend to use a case-control design with cases matched with several controls from the same GP practice. This provides balanced and comparable cohorts of cases and controls and simplifies the study of comparatively rare conditions without loss of efficiency.

The full list of relevant confounding variables will be established from medical literature such as systematic reviews, and from expert knowledge within the team, and then the subset of these variables to be adjusted for will be found through backward elimination.

To account for the interdependence of patients from the same GP practice, multilevel modelling and multiple imputation will be used.

Aim 2. Modelling of temporal changes in the factors affecting morbidity and mortality

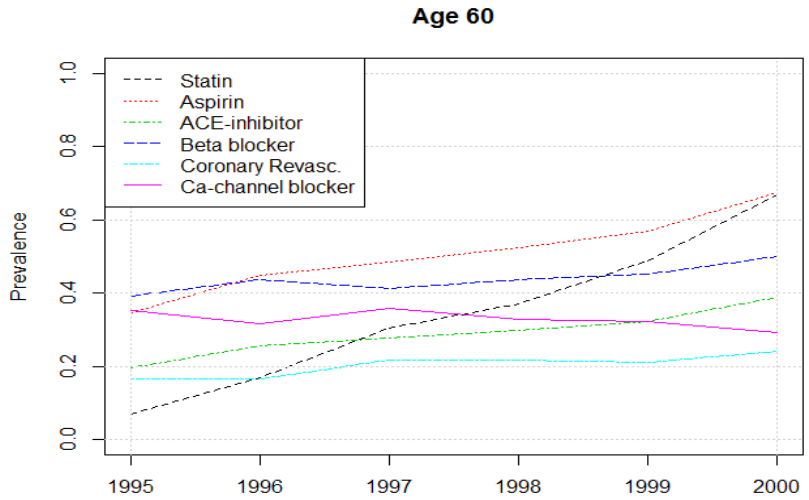
Trends in the incidence and/or prevalence of particular medical conditions and/or lifestyle factors will also be obtained from the primary care data.

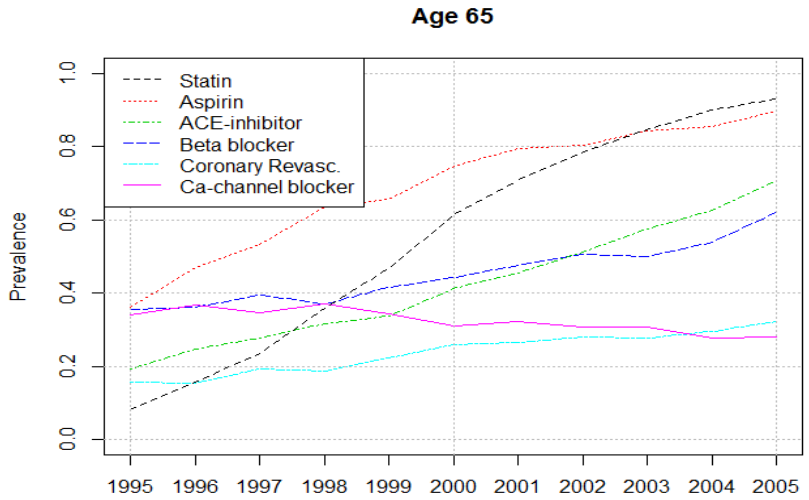
This will enable us to establish patterns due to social or geographic inequalities, such as socio-economic status (SES), age or postcode lottery.

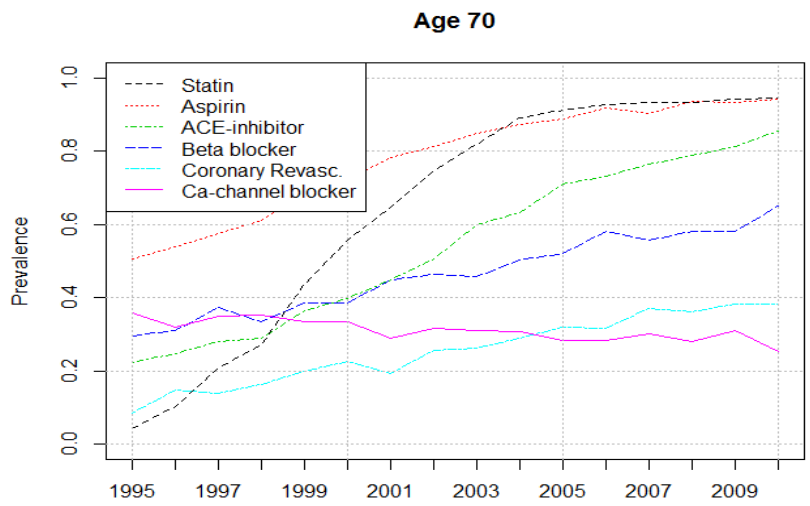
For instance, the patients in the more deprived areas may be disadvantaged in regards to the latest interventions and/or public health campaigns at least initially. This will result in widening the gap in longevity between individuals from different backgrounds.

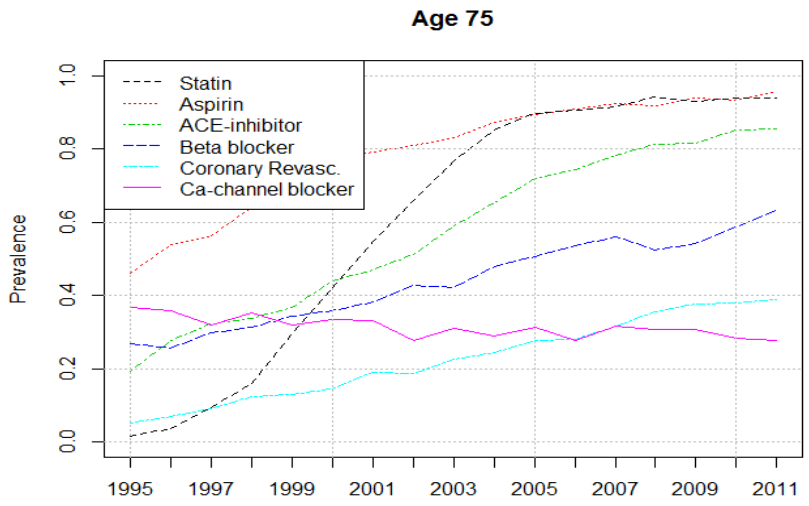
Thus to be able to ascertain an effect on longevity of a population, we need to model the incidence of a condition or an uptake of an intervention over time in parallel to modelling mortality.

Prevalence of treatment by cohort's age in patients with a history of acute myocardial infarction

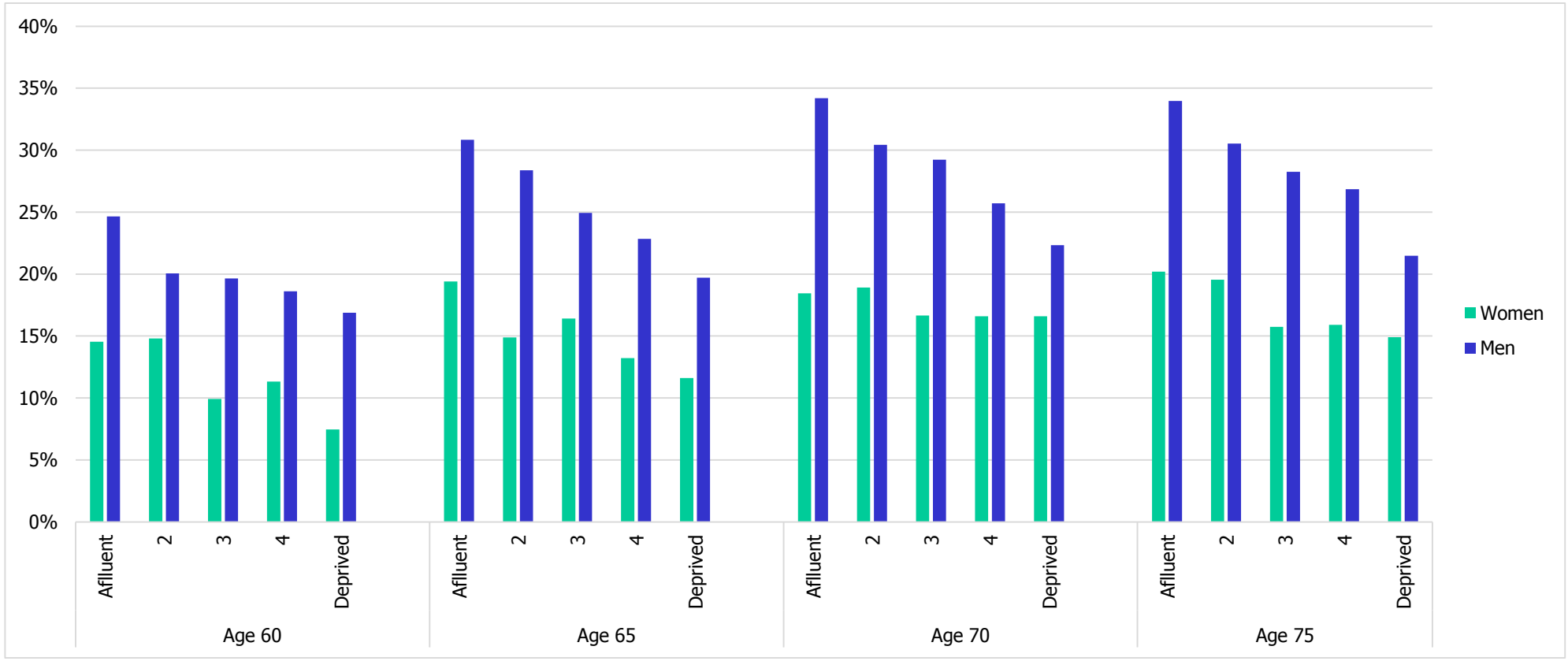








Example: Coronary Revasculisation given IHD



Aim 3. Evaluation of plausible scenarios in mortality trends due to particular medical advances or lifestyle changes on the population of insureds

As often happens with the existing portfolio of insured lives, the minute health details of a life are not available. Instead, the interest lies in the mortality trends of the whole book.

To be able to provide this information, three components are required:

- established in Aim 1 model for survival differentials associated with a particular disease or intervention;
- developed in Aim 2 model for the incidence/ prevalence of this condition or uptake of this intervention over time,
- and the sufficient knowledge of the population to which it is desired to translate trends in longevity established in general population to be able to assess the basis risk.

Aim 4. Tools to forecast longevity risk of a book

We will develop an R package incorporating our models and providing analytical and graphical means to forecast longevity of a general UK population, and also of a population of a user defined composition under a number of scenarios for changes in disease incidence, health behaviours and treatments.

This will be an open source software available from the project website along with an accompanying manual for its use.

We also intend to develop teaching materials for the actuarial community on the modelling techniques used in the project, and the use of the developed R package. These materials will be available from the project website.

Case Study:

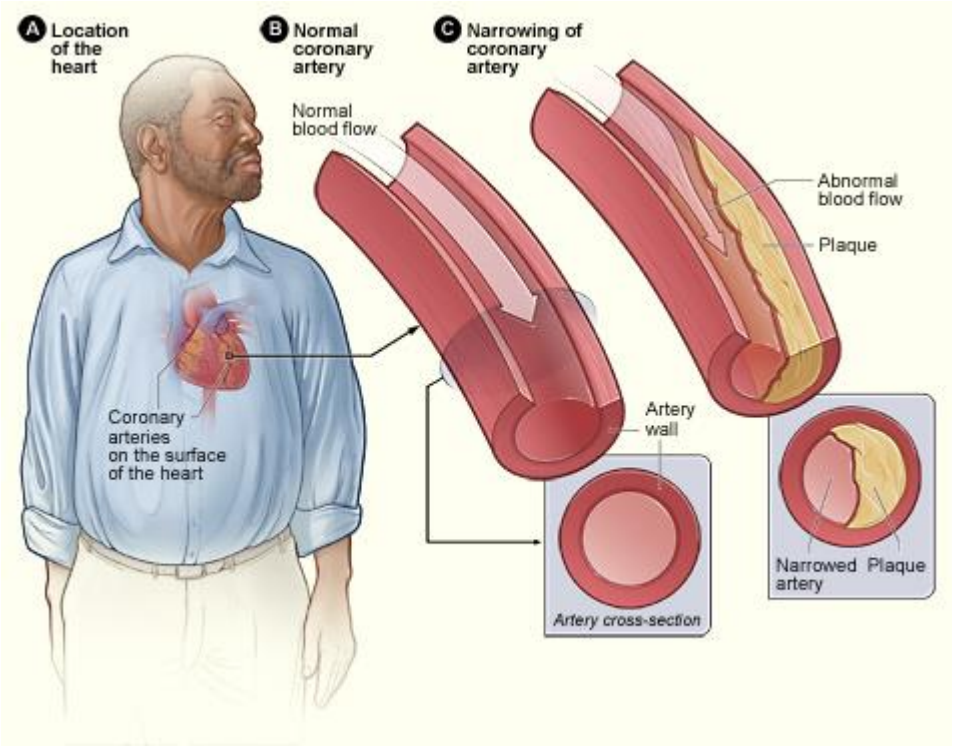
Statins and Life Expectancy

Lisanne Gitsels, PhD candidate

Elena Kulinskaya
Nick Steel
Nigel Wright (Aviva)

Cardiovascular disease (CVD)

Disease of the heart or blood vessels
Leading cause of global and UK death: **33%**



Four main types of CVD:

- coronary heart disease
- stroke
- peripheral arterial disease
- aortic disease

Risk factors for CVD:

- high blood pressure (hypertension)
- smoking
- high blood cholesterol
- diabetes
- lack of exercise
- being overweight or obese
- family history / ethnic background

Primary prevention of CVD

Primary prevention: no previous history of CVD

Example: lipid-lowering therapy - statins

National Institute of Health and Clinical Excellence (NICE):

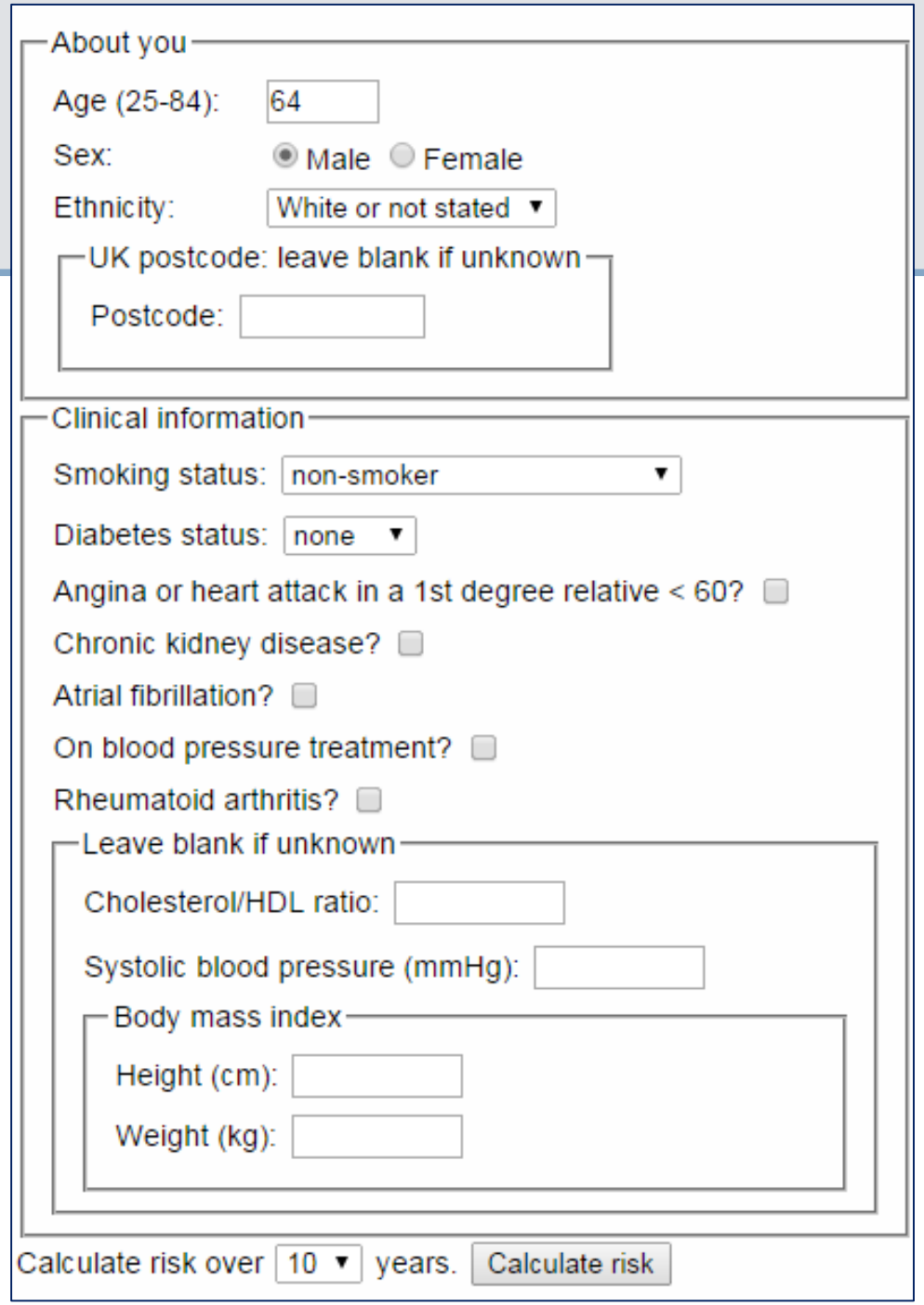
Offer atorvastatin 20 mg for the primary prevention of CVD to people who have a 10% or greater 10-year risk of developing CVD.

Estimate the level of risk using the QRISK2 assessment tool

www.nice.org.uk/guidance/cg181/

www.qrisk.org/2016/

Up to 17 million UK residents eligible for statins



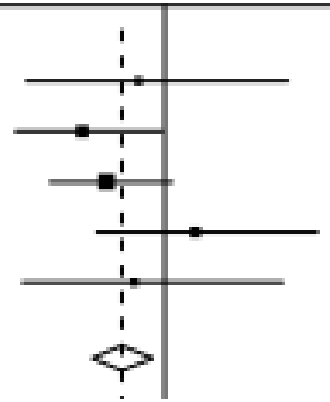
Previous research on effectiveness of statins

Meta-analysis of 27 randomised clinical trials by Cholesterol Treatment Trialists' (CTT) Collaborators, The Lancet 2015 (http://dx.doi.org/10.1016/S0140-6736(14)61368-4)

- Overall 9%relative reduction in all-cause mortality with statins
- Equivalent to absolute reduction of 1 per 1,111 patients per year per mmol – others no benefit

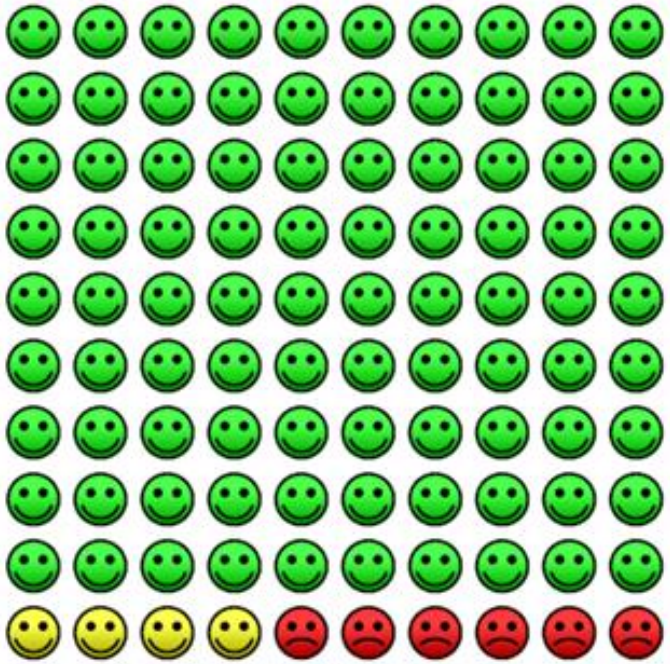
Webfigure 9: Effects on any deaths per 1.0 mmol/L reduction in LDL cholesterol at different levels of risk, by history of vascular disease and overall

5–year MVE risk at baseline	Deaths (% per annum) Statin/more	Control/less	RR (CI) per 1.0 mmol/L reduction in LDL cholesterol	Trend test
Participants without vascular disease				
< 5%	164 (0.38)	177 (0.41)	0.94 (0.71 – 1.26)	
≥ 5%,<10%	372 (0.77)	446 (0.93)	0.83 (0.69 – 0.99)	
≥ 10%,<20%	703 (1.99)	778 (2.19)	0.88 (0.76 – 1.02)	χ^2_1=1.57
≥ 20%,<30%	363 (5.13)	339 (4.73)	1.06 (0.86 – 1.32)	(p=0.2)
≥ 30%	192 (10.76)	192 (11.44)	0.94 (0.70 – 1.25)	
Subtotal	**1794 (1.33)**	**1932 (1.42)**	**0.91 (0.85 – 0.97) p= 0.007**	



Cates plot for benefits of statin treatment

Cardiovascular risk 10% over 10 years: taking atorvastatin



If all 100 people take atorvastatin for 10 years, over that time on average:

- 4 people will be saved from developing CHD or having a stroke (the yellow faces)
- 90 people will not develop CHD or have a stroke, but would not have done anyway (the green faces)
- 6 people will still develop CHD or have a stroke (the red faces).

https://www.nice.org.uk/guidance/cg181/resources/patient-decision-aid-243780157

Limitations of statin trials

- Randomised clinical trials are the 'gold standard' for evidence of effectiveness
  - Confounders randomized equally to both groups (in theory)
- Generalisability from trial participants to general population
  - Exclusion on grounds of age, comorbidity, intolerance to intervention
- Short follow-up
  - Maximum 5 years
- Commercial trial data not available for individual scrutiny
  - Lack of transparency
- Large observational datasets can fill these gaps with robust statistical analyses

Research question

What is the survival benefit associated with statin prescription as primary prevention of cardiovascular disease for different risk groups at various ages in the general population?

Design and Data Selection

Population-based prospective cohort study

Restrictions data:

- Medical records from 1987 to 2011 of people born between 1920 and 1940

Target ages:

- 60, 65, 70, and 75

Exclusion:

- Patients with a history of cardiovascular disease

Missing data

Incomplete records in: BMI, smoking status, and risk of cardiac event

Multiple imputation

- Joint modelling
  - » Linear regression for BMI and risk of cardiac event
  - » Ordered probit regression for smoking status
- Multilevel on GP practice
- MCMC (Monte Carlo Markov Chain) 500 iterations resulting in 10 imputed datasets
- REALCOM-Imputation software

Model specification

Cox's proportional hazard regression estimates the hazard λ_{ij} for patient *i* from GP practice *j*: $\lambda_{ij} = \lambda_0(t)\, Z_j\, e^{\beta X_{ij}}$

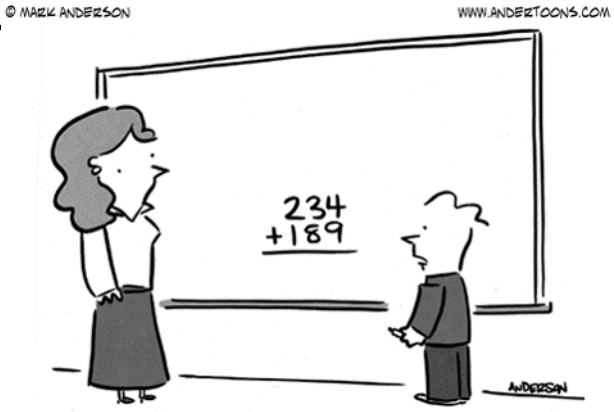
where λ_0 = baseline hazard (function of time),
Z_j = shared frailty term on GP practice,
β = coefficients (constant),
and X_{ij} = exposures, e.g. statins (constant).

Models specified:

- Ages: 60, 65, 70, and 75
- Risk groups:
  - » Low <10% risk of cardiac event
  - » Moderate 10-19% risk of cardiac event
  - » High ≥20% risk of cardiac event

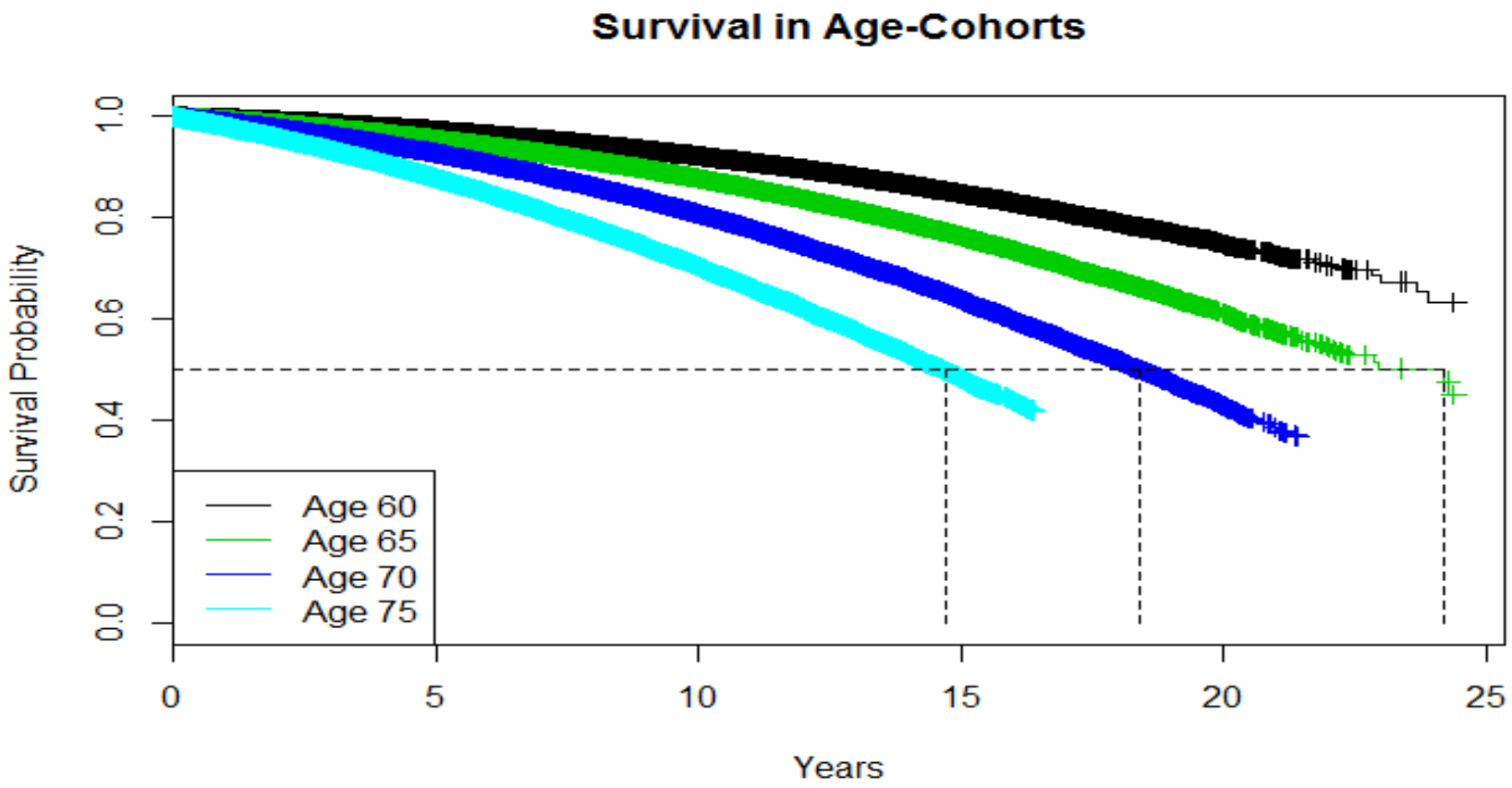
Cohorts' characteristics

Cohort	Number of patients	Number of deaths	Average follow-up time	Maximum follow-up time
Age 60	118,700	15,296 (12.8%)	12 years	24 years
Age 65	199,574	28,848 (14.5%)	10 years	24 years
Age 70	247,149	40,699 (16.5%)	7 years	21 years
Age 75	194,085	37,356 (19.2%)	6 years	16 years



"Does this count as big data?"

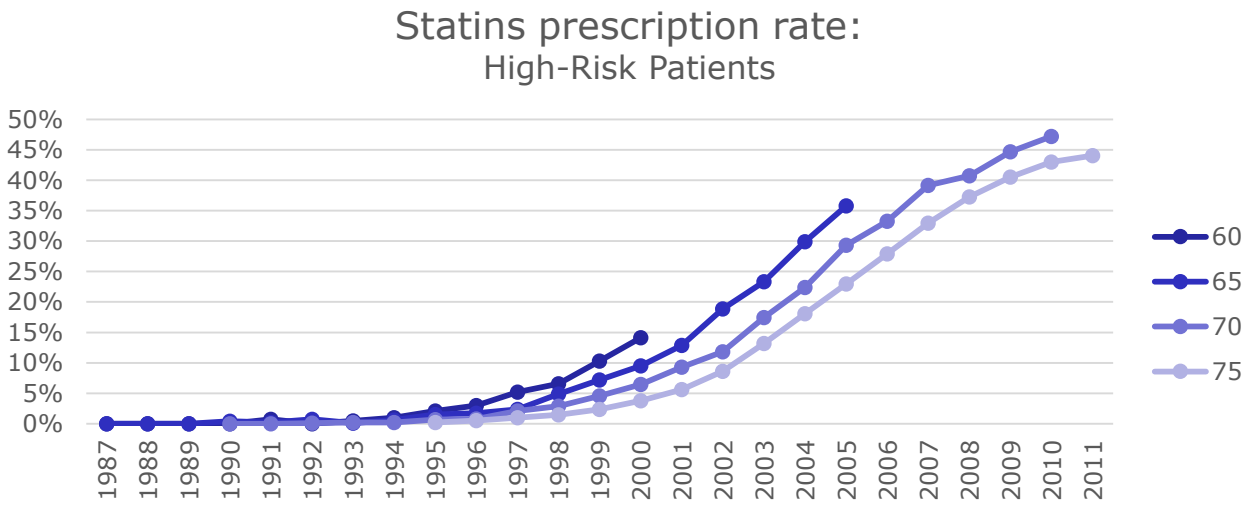
Kaplan-Meier plots

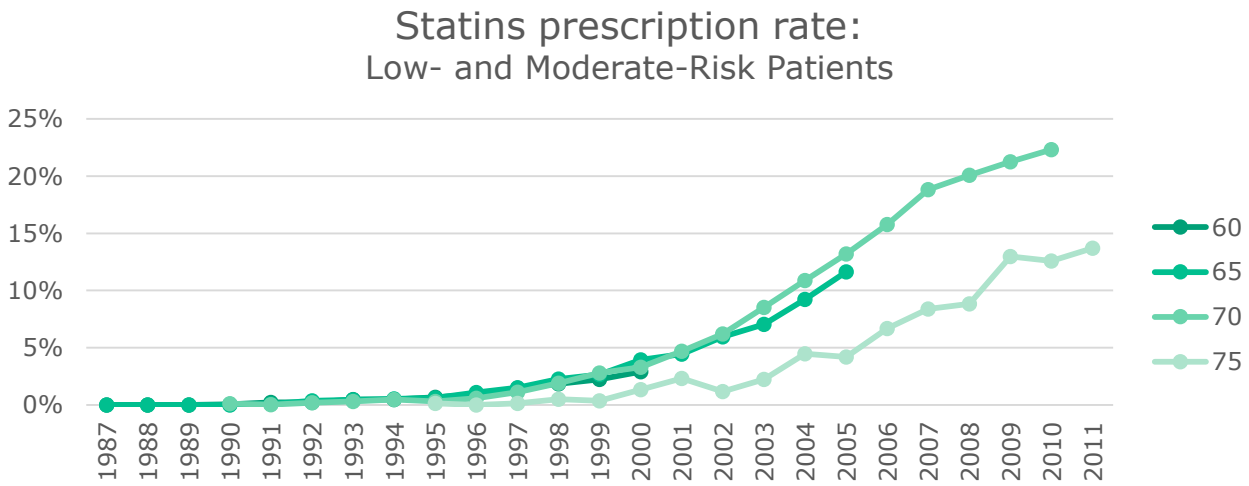


Distribution men and women across risk groups

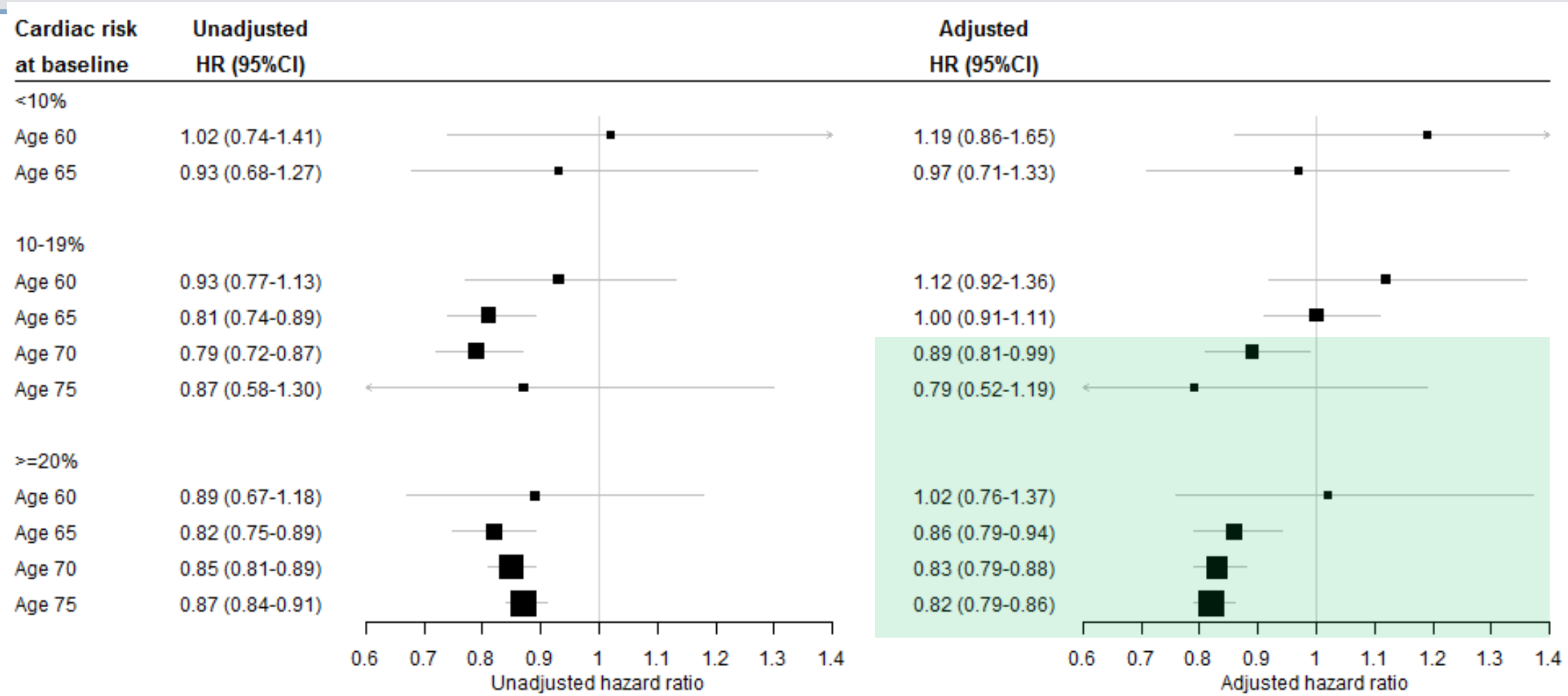
Cohort	Cardio risk	Women % (Statins %)	Men % (Statins %)
Age 60	Low	83 (1.2)	16 (0.4)
	Moderate	16 (3.7)	78 (1.3)
	High	1 (11.9)	6 (5.2)
Age 65	Low	40 (2.2)	.
	Moderate	55 (7.4)	72 (3.2)
	High	5 (26.9)	28 (12.4)
Age 70	Moderate	80 (9.5)	17 (5.4)
	High	20 (28.2)	83 (17.4)
Age 75	Moderate	15 (4.6)	.
	High	85 (19.6)	100 (19.1)

Uptake of statins by risk group





Hazard of mortality from statin prescription by age and risk group



HRs adjusted for gender, year of birth, postcode, diabetes, high cholesterol level, blood pressure regulating drugs, BMI, smoking status, general practice

Impact

Medicine and Public Health:

Are current guideline thresholds for statin therapy for primary CVD prevention too low?

- Overtreatment of people under 60 and at <10% risk?
- Recent extension of guidelines to younger and lower risk groups may need to be reconsidered?
- Clinicians discuss risks and benefits of statin initiation with their patients.

Further research needed on statins for primary prevention:

- People under 65
- People at <10% 10 year risk
- Individual data on low risk patients in trials

Insurance and Government:

- Pricing and reserving for longevity risk (annuities, pension liabilities, etc.) and morbidity and mortality risk
- Predicting volumes for coverage of medical procedures
- Predicting changes in population life expectancy

Personal:

- Information on average life expectancy (and confidence limits)
- How to structure retirement funds?
- Lifestyle changes can be made (e.g. stop smoking)
- Potential benefits of statins after age 70 at population level