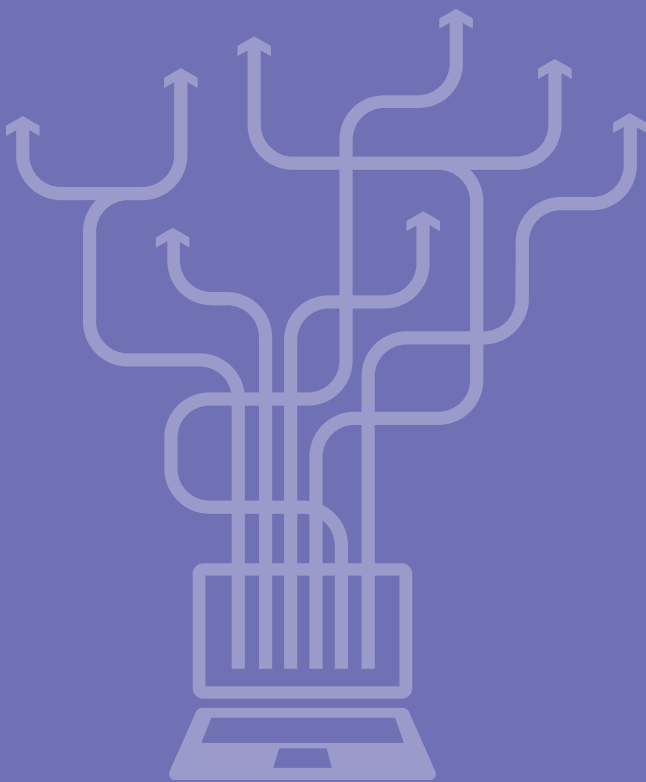


Contents

1. Introduction by our President, Colin Wilson	3
2. Foreword by Michael Tripp	4
3. Theoretical foundations of data science, Sofia Olhede and Russell Rodrigues	5
4. Opinion: Is big data just a big hype? Peter Coveney and Roger Highfield	11
5. Routes to Diagnosis for cancer, Lucy Elliss-Brookes	13
6. Understanding longevity and morbidity risk, Elena Kulinskaya and Lisanne Gitsels	16
7. Personalised risk prediction: genomics and beyond, Philippa Brice	20
8. Big data in action: wearables, Matthew Edwards	23
9. The ethical challenges of biomedical data, Luciano Floridi	26
10. Recent developments and events	29



The views expressed in this publication are those of invited contributors and not necessarily those of the Institute and Faculty of Actuaries. The Institute and Faculty of Actuaries do not endorse any of the views stated, nor any claims or representations made in this publication and accept no responsibility or liability to any person for loss or damage suffered as a consequence of their placing reliance upon any view, claim or representation made in this publication. The information and expressions of opinion contained in this publication are not intended to be a comprehensive study, nor to provide actuarial advice or advice of any nature and should not be treated as a substitute for specific advice concerning individual situations. On no account may any part of this publication be reproduced without the written permission of the Institute and Faculty of Actuaries.

6. Use of big health and actuarial data for understanding longevity and morbidity risk

Professor Elena Kulinskaya and Lisanne Gitsels, PhD candidate, University of East Anglia

Introduction

Estimating longevity risk and evaluating associated uncertainty is one of the main topics of concern to actuarial community. It is well known that longevity is increasing considerably both in developed and developing countries, including the United Kingdom. We believe that to be able to establish the drivers of this change, and to predict how they may change over time and how this would affect life expectancy, researchers need to harvest Big Health Data (Hemmingway, 2014), i.e. to access large health databases, and to use sophisticated tools for modelling the mortality experience of participating populations using individual level health data. Big Actuarial Data such as the Continuous Mortality Investigation (CMI) data are of the utmost importance in translating the results to the reference population of relevance to the actuarial community.

Contemporary evidence-based underwriting needs to account for a large number of important and time-varying determinants of health and longevity, such as demographic factors (gender, social class), lifestyle factors (smoking, obesity, alcohol usage) and medical advances, and their interactions. Many public health interventions are aimed at increasing the health of populations. These vary from offering flu vaccination to encouraging lifestyle changes to management of chronic medical conditions. However, actuarial and medical research often aim at somewhat differing objectives. While mortality is of primary interest to an actuary, exacerbation of medical conditions is often the interest of a medical researcher. For instance, not death but a cardiac event may be the primary endpoint in many medical studies of heart disease or smoking. Additionally, clinical trials while of the gold standard when studying medical interventions, deal with a selective population of patients, and usually are of short duration.

This explains why the existing medical publications and their syntheses published in numerous systematic reviews, though certainly important, are not sufficient for actuarial purposes, and the direct involvement of actuarial researchers in the modelling of health-related data is of utmost importance. In-depth actuarial longevity research should concentrate on statistical modelling of population-based individual level data collected over the long term. Some advances in this direction are already being made (Ryan et al., 2013; Lu et al., 2014).

The title of this article (Use of big health and actuarial data for understanding longevity and morbidity risks) is, in fact, the name of a research programme recently funded by Institute and Faculty of Actuaries (IFoA). This is a joint project between the School of Computing Sciences and Norwich Medical School within the UEA, and Aviva Life. This research will use the data on 3.4 million patients born before 1960 from The Health Improvement Network (THIN) primary care database, and also the CMI data. The main objectives are the development of novel statistical and actuarial methods for modelling mortality, modelling trends in morbidity, assessing basis risk and evaluating longevity improvements based on individual level big health and actuarial data.

Programme description

The first aim of the programme is the mortality modelling. This includes identification and quantification of the key factors affecting mortality/longevity such as lifestyle choices, medical conditions and/or interventions. A target list will include between 3-5 conditions or interventions. Statin prescription, an established longevity-improving intervention (Longevity Science Panel, 2014) is one of the target scenarios. The choice of the medical and social developments to be included in

research will be based on current models of disease burden in England (Newton et al., 2013), and combined with the availability of relevant information in the general practice data such as THIN.

The top causes of premature death in England are: heart disease, stroke, respiratory disease, cancer and Alzheimer's disease. Important health interventions and social developments include widening of statins prescription, possible changes in blood pressure targets, rise in obesity and type 2 diabetes, reduction in smoking, trends in diet and physical activity according to socioeconomic status. Some of the information required for tackling these diseases and interventions is not available in THIN. This includes the details on cancer severity, and this, unfortunately, takes cancer off the list.

After careful consideration of the importance of various conditions, interventions and lifestyle factors, and availability of the required information, the research team agreed on the following list which includes the main cardiovascular conditions: myocardial infarction, heart failure, atrial fibrillation, and stroke. The lifestyle factors of interest in respect to cardiovascular disease are smoking and obesity. Additionally, type 2 diabetes would contribute to all of the above conditions. Health interventions to study include statin prescriptions and a possible change in systolic blood pressure (BP) targets to 120 mm Hg. This is a very novel possible development, following the results from the just published SPRINT trial (2015). In this large trial, the lower BP target resulted in considerably lower all-cause mortality (hazard ratio, 0.73; 95% CI, 0.60 to 0.90). However this may also bring side effects, such as rise in acute kidney injury.

The second aim is the modelling of trends in morbidity and the uptake of health interventions. Trends in the incidence and/or prevalence of particular medical conditions and/or lifestyle factors will also be obtained from the primary care data, establishing 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. A new intervention may be of benefit to only the most privileged individuals, at least initially. Similarly, outcomes of a public-health campaign aimed at healthier lifestyle choices are often associated with SES and will, therefore, result in SES dependent changes in the incidence of a disease. This will lead to widening the gap in longevity between individuals from different backgrounds. Thus to be able to ascertain an effect on longevity of a population, the incidence of a condition or an uptake of an intervention needs to be modelled over time in parallel to modelling mortality.

As often happens with an existing portfolio of insured lives, the precise 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,
- 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 (Haberman et al., 2014). The data submitted to the Continuing Mortality Investigation will be used for this purpose.

Finally, an open source R package will be developed. It will incorporate the models derived from the analyses of THIN and CMI data and provide 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 behaviors and treatments. This will be an open source software available from the project website along with an accompanying manual for its use. Teaching materials for the actuarial community on the modelling techniques used in the project, and the use of the developed R package will be available from the project website.

Our programme is funded by IFoA for four years from October 2016. However, we expect to obtain the first results and to present them to actuarial community within the first year.

A case study: statins and longevity

This case study focuses on longevity improvement due to the widening guidelines on the prescription of statins to healthy patients. The results below are based on the preliminary research within Aim 1 by the second author, and are published in Gitsels et al. (2016).

Cardiovascular disease (CVD) is one of the main causes of death, accounting for 28% of all deaths in the United Kingdom. Statins are prescribed for primary and secondary prevention of CVD. For primary prevention, the risk of CVD is quantified by the so called QRISK2 score as the 10-year risk of a first cardiac event. In July 2014, the National Institute for Health and Clinical Excellence (NICE) lowered the risk threshold for which statins are prescribed at from 20% (2006 recommendation) to 10% (NICE, 2014). This translates to an increasing number of people being eligible for the drugs; that is an additional 4.5 million UK residents. From an actuarial perspective the question becomes whether the new NICE policy would materially affect mortality in the UK, and if yes then how.

The objective of our study was to estimate the survival benefits of statins for different risk groups at various ages in the general population. Data from THIN database were used, comprising medical records from 1987 to 2011 of people born between 1920 and 1940. Four cohorts aged 60, 65, 70, or 75 years with no previous history of CVD were studied, with sample sizes 118,700, 199,574, 247,149, and 194,085, respectively.

The hazard of mortality associated with statin prescription in patients at <10%, 10-19%, or ≥20% CVD risk was calculated by a multilevel Cox proportional hazard regression, adjusted for covariates including sex, year of birth, Mosaic (lifestyle groups defined by postcode), diabetes, blood-pressure regulating drugs, high cholesterol, Body Mass Index (BMI), and smoking status.

There was low uptake of statin therapy in the eligible population as seen in Graph 2. People at <10% CVD risk did not have a mortality benefit from statin prescription at any age, whereas people at 10-19% CVD risk had a mortality benefit of 11-21% by the age of 70. Furthermore people at ≥20% CVD risk had a mortality benefit of 14-18% by the age of 65 as shown in Figure 6.

Graph 2: Statins prescriptions rates in the UK based on the THIN data

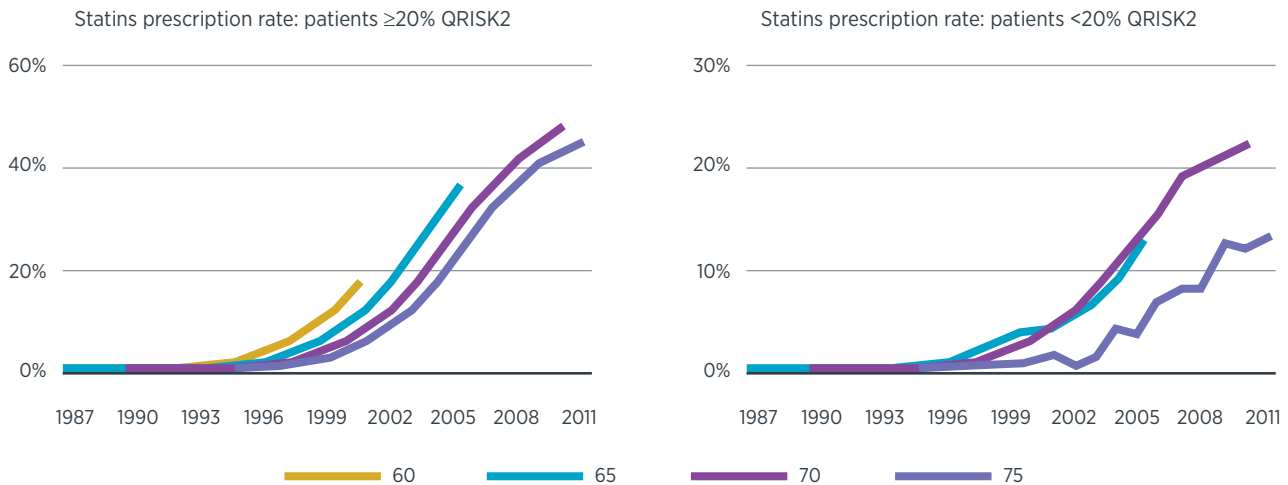
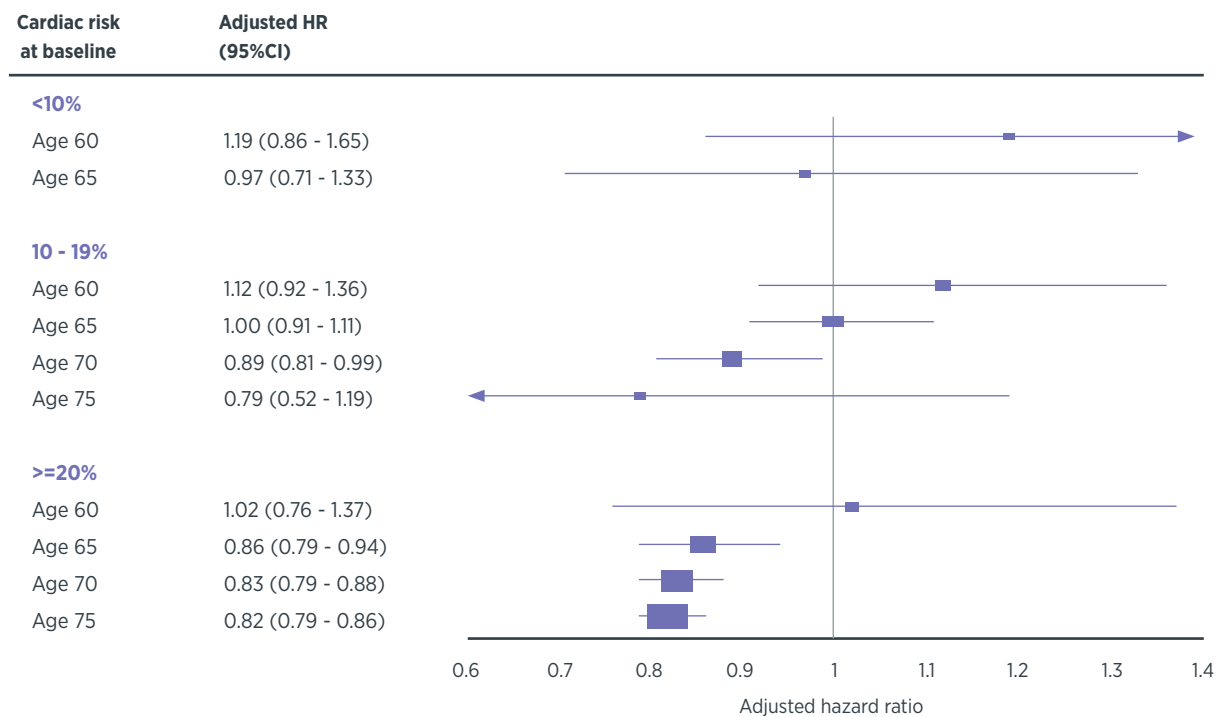


Figure 6: Hazard ratio of death given statins prescription for patients stratified by QRISK2.

Hazard ratios adjusted for sex, year of birth, socioeconomic status, diabetes, hypercholesterolaemia, blood pressure regulating drugs, body mass index, smoking status, and general practice.



The mortality benefits translate to an increase in life expectancy of 1.2 to 2 years, respectively. In the course of the project we shall extend these results and combine them with a novel model for the uptake of statins over time (Aim 2) and we will develop an adjustment for the basis risk based on the CMI data (Aim 3) to provide a plausible scenario of temporal changes in longevity due to statins. Our model can also be used for the future cost-benefit analysis of the new NICE guidelines which would account for the additional drug costs and additional healthcare resource use.

References

- Haberman S. et al. (2014). *Longevity basis risk: a methodology for assessing basis risk*. Research report by Cass Business School and Hymans Robertson LLP
- Hemmingway, H. (2014). *Big Health Data*. International Mortality and Longevity Symposium, Birmingham 2014.
- Lu, J. L. C., Wong, W., Bajekal, M. (2014). *Mortality improvement by socio-economic circumstances in England (1982 to 2006)*. British Actuarial Journal 19: 1-35.
- National Institute for Health and Care Excellence (2014). *Wider use of statins could cut deaths from heart disease*.
- Newton J.N. et al. (2013). *Changes in health in England, with analysis by English regions and areas of deprivation, 1990–2013: a systematic analysis for the Global Burden of Disease Study*, The Lancet 386(10010): 2257-2274.
- Ryan, D., Rion, S., Rechfeld, F. (2013) *Disease-based models of longevity: the future of human longevity: cardiovascular health, longer lives*. 10-12 November 2013. Swiss Re.
- Longevity Science Panel (2014). *What is aging? Can we delay it?*
- The SPRINT Research Group (2015). *A randomized trial of intensive versus standard blood-pressure control*. New England Journal of Medicine , 373:2103-2116
- 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

Biographies

Elena Kulinskaya is Professor in Statistics (Aviva Chair in Statistics) at UEA. Her research interests are meta-analysis and evidence synthesis, statistical methods of big data analysis, actuarial and medical statistics and design of statistical software.

Lisanne Gitsels is a PhD candidate and Senior Research Associate at UEA. Her research interests are in analysing longitudinal data to contribute to solving current social, health problems from an epidemiological and economic point of view.



**Actuarial
Research Centre** [®]

Institute and Faculty
of Actuaries

The Actuarial Research Centre (ARC) is an international network of actuarial researchers around the world delivering cutting-edge research programmes that aim to address some of the significant, global challenges in actuarial science. To find out more about this programme and other world-class research programmes, please visit: www.actuaries.org.uk/ARC

