University Library

MINERVA ACCESS

A gateway to Melbourne's research publications

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Hayes, AJ;Leal, J;Gray, AM;Holman, RR;Clarke, PM

Title:
UKPDS Outcomes Model 2: A new version of a model to simulate lifetime health outcomes of patients with type 2 diabetes mellitus using data from the 30 year united kingdom prospective diabetes Study: UKPDS 82

Date:
2013-09-01

Citation:
Hayes, A. J., Leal, J., Gray, A. M., Holman, R. R. & Clarke, P. M. (2013). UKPDS Outcomes Model 2: A new version of a model to simulate lifetime health outcomes of patients with type 2 diabetes mellitus using data from the 30 year united kingdom prospective diabetes Study: UKPDS 82. Diabetologia, 56 (9), pp.1925-1933. https://doi.org/10.1007/s00125-013-2940-y.

Persistent Link:
https://hdl.handle.net/11343/138331

Article

UKPDS Outcomes Model 2: a new version of a model to simulate lifetime health outcomes of patients with type 2 diabetes mellitus using data from the 30 year United Kingdom Prospective Diabetes Study: UKPDS 82

A. J. Hayes[1], J. Leal[2], A. M. Gray[2], R. R. Holman[3] and P. M. Clarke[4]

[1] Sydney School of Public Health, University of Sydney, NSW 2006, Australia

[2] Health Economics Research Centre, Department of Public Health, University of Oxford, Oxford, UK

[3] Diabetes Trials Unit, Oxford Centre for Diabetes, Endocrinology and Metabolism, Churchill Hospital, Oxford, UK

[4] School of Population Health, University of Melbourne, VIC, Australia

Corresponding author: A. J Hayes, Sydney School of Public Health, University of Sydney, NSW 2006, Australia
Email: alison.hayes@sydney.edu.au

Received: 2 January 2013 / Accepted: 24 April 2013

Abstract

Aims/hypothesis The aim of this project was to build a new version of the United Kingdom Prospective Diabetes Study (UKPDS) Outcomes Model (UKPDS-OM1), a patient-level simulation tool for predicting lifetime health outcomes of people with type 2 diabetes mellitus.

Methods Data from 5,102 UKPDS patients from the 20 year trial and the 4,031 survivors entering the 10 year post-trial monitoring period were used to derive parametric proportional hazards models predicting absolute risk of diabetes complications and death. We re-estimated the seven original event equations and estimated new equations for diabetic ulcer and some second events. The additional data permitted inclusion of new risk factor predictors such as estimated GFR. We also developed four new equations for all-cause mortality. Internal validation of model predictions of cumulative incidence of all events and death was carried out and a contemporary patient-level dataset was used to compare 10 year predictions from the original and the new models.

Results Model equations were based on a median 17.6 years of follow-up and up to 89,760 patient-years of data, providing double the number of events, greater precision and a larger number of significant covariates. The new model, UKPDS-OM2, is internally valid over 25 years and predicts realistic event rates for complications, which are lower than those from the existing model.

Conclusions/interpretation The new UKPDS-OM2 has significant advantages over the existing model, as it captures more outcomes, is based on longer follow-up data, and more comprehensively captures the progression of diabetes. Its use will permit detailed and reliable lifetime simulations of key health outcomes in people with type 2 diabetes mellitus.

Keywords Complications, Life expectancy, Patient-level simulation, Risk modelling, Survival, Type 2 diabetes mellitus

Abbreviations

CHF	Congestive heart failure
eGFR	Estimated GFR
IHD	Ischaemic heart disease
LDS	Lipids in Diabetes Study
MI	Myocardial infarction
OM1	Outcomes Model version 1
OM2	Outcomes Model version 2
PVD	Peripheral vascular disease
PTM	Post-trial monitoring
QALYs	Quality-adjusted life years

SBP	Systolic BP
UKPDS	United Kingdom Prospective Diabetes Study

Introduction

Computer simulation is a method of modelling the progression of type 2 diabetes mellitus and predicting long-term outcomes of the disease. Since the publication of the first United Kingdom Prospective Diabetes Study (UKPDS) Outcomes Model (UKPDS-OM1) [1], the use of simulation modelling in diabetes has increased, with at least eight models in use worldwide [2,3], many of which use the published equations from UKPDS-OM1 [3]. UKPDS-OM1 is a multi-application model and has been used in a wide variety of applications, including cost-effectiveness analyses [4, 5] and prediction of life expectancy [6].

The UKPDS-OM1 has been tested alongside several other diabetes simulation models at Mount Hood Challenge meetings. A general conclusion was that the models performed reasonably well in terms of predicting the relative risk of interventions vs control treatments, but less well in predicting absolute risk [3]. Additionally, a temporal validation study found that UKPDS-OM1 over-predicted the probability of events for high-risk subgroups [7].

Model building is an iterative process and models need to be updated as new information becomes available [8]. The UKPDS-OM1 was based on trial data collected in the UKPDS up until 1997. Additional information collected during the UKPDS 10 year post-trial monitoring (PTM) period [9] provided an opportunity to update the simulation model and to incorporate data on new risk factors and outcomes that were unavailable when the UKPDS-OM1 was constructed.

Our aim was to build a new model, Outcomes Model version 2 (UKPDS-OM2), based on the larger dataset which included additional more recent data that were also more clinically relevant, as participants were no longer in a clinical trial. This involved: (1) re-estimating, over a longer duration of follow-up, the seven original risk equations for complications (myocardial infarction [MI], ischaemic heart disease [IHD], stroke, congestive heart failure [CHF], amputation, blindness and renal failure); (2) estimating new equations, not in the original model, for diabetic ulcer and some second events; (3) developing new equations for all-cause mortality; (4) exploring the use of new risk factors such as microalbuminuria which have been shown to be predictive of diabetes-related complications.

We also present internal validation of the UKPDS-OM2 over 25 years of follow-up, carry out a sensitivity analysis and compare predictions from the original and new models using a contemporary patient-level input dataset.

Methods

The model—UKPDS-OM2

In common with UKPDS-OM1 [1], the major objective of the new model is to simulate major diabetes-related complications over a lifetime and to calculate health outcomes such as life expectancy. It is a patient-level epidemiological model for a target population of adults aged 30 and over with any duration of diabetes. UKPDS-OM2 integrates separate risk equations for eight diabetes-related complications and death. Figure 1 shows the sequence of modelling events. The inputs to the model are individual patient characteristics (demographic, clinical risk factors and complication history); outputs include annual incidence of death or complications (MI, stroke, IHD, CHF, amputation, blindness, renal failure and ulcer), including second events for MI, stroke and amputation, life expectancy and quality-adjusted life years (QALYs). In each annual cycle the probability of death or of experiencing one or more complications is calculated for each patient according to the appropriate risk equation. Each probability is compared with a random number drawn from a uniform distribution ranging from zero to one to determine whether an event actually occurs for this patient. Equations for complications are executed in random order and if an event is predicted to occur in a given cycle it will inform the remaining set of equations still to be estimated in the same cycle. Probability of death is based on whether complications have occurred, and which complications have occurred in the current annual cycle. If the model predicts that an individual dies, their total years lived and QALYs are calculated and the individual exits the simulation; if the individual survives that cycle, the age, years of diabetes, clinical risk factor values and event histories are updated and carried forward to the next annual cycle. Clinical risk factors can either be updated from a known patient-level dataset or be modelled using risk factor time-path equations. Quality of life decrements based on the complications experienced can also be calculated; details of these decrements and new time-path equations will be reported elsewhere.

Derivation of risk model equations

Study subjects and measurement of outcomes Model equations were based on patient-level data for the 5,102 UKPDS participants with newly diagnosed type 2 diabetes mellitus, aged 25-65 years, recruited between 1977 and 1991 [10]. These patients were followed until the trial concluded in 1997. All 4,031 surviving participants entered the 10 year PTM observational study [9], during which time they returned to their community or hospital-based diabetes care providers, with no attempt made to maintain their previous allocated trial regimen. All patients provided written informed consent. Approval was obtained from the ethics committees at all 23 clinical centres, and the study conformed to the Declaration of Helsinki guidelines.

During the main trial, patients were seen three or four times each year in UKPDS clinics. During PTM, patients were seen annually for 5 years in UKPDS clinics, with continued standardised collection of outcome data plus clinical examination every 3 years. In years 6-10, patient and general practitioner questionnaires were used to follow patients remotely, since funding for clinic visits was not available. The vital status of all patients who were still living in the UK was obtained from the Office for National Statistics.

Outcomes were adjudicated exactly as in the original trial, by the UKPDS endpoint committee, which was blinded to study groups. The definition of the outcomes used in the UKPDS-OM2 match adjudicated trial endpoints and the original UKPDS-OM1 outcomes, except for vascular cardiac events, where CHF and other IHD now include both fatal and non-fatal events. Additional outcomes of diabetic ulcer of the lower limb, and second events for MI, stroke and amputation, were derived. Definitions of outcomes by ICD-9 are detailed in electronic supplementary material (ESM) Table 1.

Clinical risk factors We used a set of clinical risk factors as candidate predictor variables that were similar to those used in UKPDS-OM1 (i.e. systolic blood pressure [SBP], HbA_{1c}, lipids) but with the following modifications: HDL and LDL cholesterol were included separately; BMI, peripheral vascular disease (PVD) and atrial fibrillation were updated rather than baseline values used. We also included risk factors shown in recent studies to be potentially predictive of diabetic complications: micro- or macroalbuminuria [11], estimated GFR (eGFR) [12], heart rate [13] and white blood cell count [14]. Haemoglobin was also included, as it has been shown to be an independent predictor of mortality in patients with CHF [15].

Statistical analysis Risk equations for first occurrence of eight diabetes complications and three additional second event equations for MI, stroke and amputation were developed. Multivariate semi-parametric proportional hazards survival models were derived with time to event determined in continuous time from onset of diabetes, using the censor date of death or the date of last contact with the patient. The set of candidate covariates for each equation included time invariant factors (e.g. sex, age at diagnosis of diabetes), time varying clinical risk factors (e.g. HbA_{1c} and SBP) and time varying comorbidities (e.g. history of stroke). One year lagged values were used for clinical risk factors to avoid possible confounding of risk factors measured post complication. A full description of risk factors covariates is presented in ESM Table 2.

Risk equations for all-cause mortality were developed to take account of patients' complication status in different years. These included logistic models to capture the high mortality in the year of a complication, and Gompertz proportional hazards survival models for years in which there were no complications. Thus, in any patient-year, only one of the four mutually exclusive equations for prediction of absolute annual risk of mortality would be used. Preliminary analysis showed that all complications except blindness and ulcer were associated with mortality in the current year ($p<0.05$). Hence, logistic regression models were used to estimate the probability of death in the year of any MI, stroke, amputation (first or second), CHF, IHD or renal failure. As a result of testing to best fit and to maximise transparency, we derived two separate logistic equations for patients with and without a history of complications.

The two remaining equations for death are multivariate Gompertz proportional hazards survival models to estimate the hazard of death in years without any of the complications defined above. Time to death was determined in continuous time, with time at risk modelled by a patient's current age in order to allow extrapolation beyond the observed follow-up period [16]. The censor date for deaths was 30 September 2007, the

date of linkage to the national mortality database from the Office for National Statistics, or the date of emigration from the UK, which represents date lost to follow-up in the national statistics.

Proportional hazards models for complications and death were derived using a consistent process to select significant covariates from the candidate risk predictors. First, binary covariates were excluded from particular event equations if cross-tabulations indicated that they occurred rarely (fewer than ten occurrences). Then a multivariate model was fitted with all remaining covariates, and any not significant at $p>0.3$ were dropped. The significant covariates of the final risk model were selected in a backwards stepwise regression at $p<0.05$. The parametric form of the underlying hazard was examined graphically and models were chosen by consideration of Akaike's information criterion for exponential, Weibull and Gompertz parametric forms. The proportional hazards assumption was tested by examination of Schoenfeld residuals [17] in comparable Cox models and through Cox–Snell semi-log plots. If the effect of any covariates was identified as non-linear, they were modelled either as categorical variables or as continuous spline functions with suitable knot points. We specifically investigated any U-shaped HbA_{1c} effect [18] using continuous splines. All analyses were carried out using Stata version 12.0 software (Stata, College Station, TX, USA).

Handling uncertainty and heterogeneity

Modelled outcomes are subject to several sources of uncertainty, which are important to report [19]. Two forms of uncertainty are addressed within UKPDS-OM2: (1) Monte Carlo or 'first order' uncertainty arises as a result of comparing probabilities from risk equations against a random number to determine whether events take place at a patient level. Thus, in any model cycle, two identical patients may have different outcomes due to chance. We minimise this uncertainty by using large numbers of Monte Carlo replications until the mean of the outcome of interest is stable: (2) Parameter or 'second order' uncertainty in the estimated coefficients of the equations arises as a result of natural variation in the patient sample and limitations in the sample size for deriving the equations. Model parameters cannot be known with certainty but only within a certain parameter distribution. We captured parameter uncertainty by bootstrapping (with replacement) the UKPDS patient-level data and re-estimating all equations to derive sets of fully correlated regression coefficients. Parameter uncertainty was then propagated by using, in turn, these sets of regression coefficients to estimate different outcomes, thus providing a distribution from which CIs can be derived. This approach conforms to the American Diabetes Association guidelines on computer simulation modelling in diabetes [20].

Patient heterogeneity is reflected through individual patient-level simulation, where it is possible to simulate whole populations, one patient at a time, and aggregate their outcomes. Each individual has a unique set of risk factors for estimation of their probability of events. Simulations presented in this manuscript use real data on 5,102 (UKPDS) and 3,984 (Lipids in Diabetes Study [LDS]) unique patients.

Internal validation of the model using the UKPDS trial population

Internal validation is a necessary step in the development of a model, providing confidence that model equations have been correctly specified and coded [8]. We carried out internal validation of the simulation model by testing its performance in replicating the incidence of complications and mortality over 25 years of follow-up. This involved using the observed clinical risk factor profiles of all 5,102 UKPDS patients over 25 years, with risk factors carried forward when missing or at the end of follow-up. We compared simulated cumulative failure of each of the major outcomes of the model with the observed (Kaplan–Meier) cumulative failure of events under the assumption adopted in many clinical studies that death as well as date of last contact are censoring events.

External validation, which tests model output against independent data, is beyond the scope of this manuscript and will be fully addressed in future publications. We present, instead, simulated outcomes using an external patient-level dataset as inputs to check on the consistency and face validity of the model.

Comparisons of outcomes from UKPDS-OM1 and UKPDS-OM2 simulations

We compared UKPDS-OM1 and UKPDS-OM2 predictions using as model inputs data on 3,984 patients with non-missing risk factors from a contemporary external dataset, the LDS [21]. We used both models to predict 10 year cumulative event rates and remaining life expectancy for selected age groups. There are no observed event rates for the LDS, due to the study being stopped early. Given the illustrative nature of these applications we assumed clinical risk factors to remain constant over the 10 years and did not apply a discount rate. For comparisons of life expectancy, we ran both models to age 100 for selected age groups: 50-54 years, 60-64 years and 70-74 years. Point estimates of life expectancy were derived from 1,000 Monte Carlo replications and 95% CIs were determined from non-parametric bootstrapping.

Sensitivity analysis

We carried out one-way deterministic sensitivity analysis to increase understanding of the relationship between model inputs and outputs [20, 22] and to determine the relative importance of patient characteristics in driving aggregate outcomes of life expectancy. Using patient-level data from the LDS as inputs, we investigated the impact on remaining life expectancy of individually changing continuous risk factors by ±1 SD of the mean and of doubling and halving the rates of binary variables such as smoking.

Results

Risk equations

The median follow-up time of patients was 17.6 years and up to 89,760 patient-years of data were available for model fitting. Numbers of events and average event rates observed during the UKPDS trial and PTM are shown in Table 1. For each outcome we had at least twice the number of events as used in the estimation of the UKPDS-OM1 risk equations (e.g. 504 stroke events compared with 157 in UKPDS-OM1).

We observed many linkages between events (e.g. having a history of IHD increases the probability of having an MI), shown schematically in ESM Fig. 1. The new model has more linkages between equations: in UKPDS-OM1 there were only five linkages across seven event equations, whilst in UKPDS-OM2 there are 15 linkages between the same seven equations (ESM Table 3).

In general, there were more significant covariates in the new set of event equations. Comparing the seven common event equations across both models, UKPDS-OM1 equations had approximately five, whereas UKPDS-OM2 equations have a mean of 11 covariates per equation (ESM Table 3). The new risk factors such as eGFR and micro- or macroalbuminuria were associated with a number of outcomes (ESM Fig. 1), including several types of vascular events (e.g. MI). White blood cell count, an indicator of inflammation, was also associated with a wide range of complications (MI, stroke, blindness, amputation and renal failure). A description of the risk equation covariates, including units, transformations and interpretation of hazard ratios, is presented in ESM Table 2. All fully specified risk equations, including constants, significant coefficients and standard errors are provided in ESM Tables 4-6, and worked examples of how to calculate the absolute risk of an event occurring are in the ESM statistical appendix.

A diagram summarising the four mutually exclusive death equations and how they relate to patients' complications status in different years is presented in ESM Fig. 2. Smoking was a significant predictor in three of the four mortality equations, but the classic risk factors HbA_{1c} and SBP were not significant predictors of mortality.

Internal validation

The results of the internal validation are presented in Fig. 2. These graphs compare the simulated and actual cumulative failure (calculated from 1 minus Kaplan–Meier survival) fcor each diabetes complication and death up to 25 years after diagnosis of diabetes. The model used actual time-paths of clinical risk factors observed in the UKPDS data, but event histories were simulated by the model. The predicted curves were within the 95% CIs of the actual cumulative failure curves for all events and death. For these simulations with an input dataset of 5,102 patients, the predictions stabilised at around 200 Monte Carlo replications.

Comparison of UKPDS-OM1 and UKPDS-OM2

Estimates of 10 year event rates for the entire LDS cohort (n=3984) and for three selected age cohorts are shown in Table 2. For the cohort as a whole, fewer events were predicted by UKPDS-OM2 than by UKPDS-OM1 (e.g. UKPDS-OM2 predicted that 10% have MI compared with 21% predicted by UKPDS-OM1), with the differences most pronounced in the 70-74 year age group for all events. Predictions of all-cause mortality were also lower at 22.5% from UKPDS-OM2 compared with 31.6% from UKPDS-OM1. The new model predicted less than half the number of renal failure events predicted by UKPDS-OM1 (0.5% vs 1.3%, respectively), but stroke, blindness and amputation were similar in both versions of the outcomes model.

Consistent with lower predictions of all-cause mortality, life expectancy predictions from UKPDS-OM2 were greater than from UKPDS-OM1 for each age group examined (Table 3). For example, remaining life expectancy for the 50-to 54-year-old cohort was predicted at 25 years from UKPDS-OM2 compared with only 20 years from UKPDS-OM1. Notably, the 95% CIs of life expectancy predictions were narrower from UKPDS-OM2.

Sensitivity analysis

A tornado plot of the results of the sensitivity analysis (Fig. 3) shows the relative importance of the classic risk factors (SBP, HbA_{1c} and lipids) in predicting life expectancy. It also shows the importance of many of the novel risk factors, in particular eGFR, micro- or macroalbuminuria, heart rate and white blood cell count. By contrast, haemoglobin and atrial fibrillation were significant in few equations and had less impact on aggregate model outcomes.

Discussion

We have developed a substantially enhanced UKPDS outcomes model that uses an additional 38,000 patient-years of observational data (primarily from the PTM period), almost doubling the follow-up time used to develop the original model. During the extra follow-up many additional complications were observed, including second events, as the patients were older and had a longer duration of diabetes; this was also possibly because the patients were no longer participating in a clinical trial. The new outcomes model has a number of important enhancements: re-estimation of the event equations to include many additional risk factors; inclusion of an additional outcome of lower extremity ulcer; prediction of second events for MI, stroke and amputation; inclusion of additional linkages between complications; and substantial changes to modelling mortality. The greater number of linkages and the greater number of significant clinical risk factors in the equations compared with UKPDS-OM1 reflects the greater statistical power of a much larger dataset. For example, in the original model, being diagnosed with IHD elevated the subsequent risk of MI; but, in UKPDS-OM2, IHD also elevated the risk of stroke and blindness. We note that these interrelationships between complications may be due to other common factors not currently captured in the model.

Internal validation demonstrated a high degree of consistency between simulated and observed events over a long time period. This reliable representation of the epidemiology of diabetes complications is pertinent to future use of the model in cost-effectiveness analyses, as outcomes are usually modelled over a lifetime horizon. Sensitivity analysis confirmed the importance of the classic risk factors in driving model outcomes, but also demonstrated the importance of the new risk factors eGFR, micro- and macroalbuminuria, heart rate and white blood cells. The relative importance of clinical risk factors in predicting life expectancy depends on the number of risk equations in which they are significant and their associated hazard ratios.

In the head-to-head comparison of both simulation models, UKPDS-OM2 predicted fewer macrovascular events and higher survival over a 10 year period. There are a number of explanations. UKPDS-OM1 equations were derived from shorter durations of diabetes (up to 10 years) and represent greater out-of-sample extrapolation when evaluated over longer durations of diabetes. Also, the simplifying assumption in this example, of clinical risk factors remaining at baseline levels, confers additional reductions in risk for UKPDS-OM2 that are not captured in UKPDS-OM1, but the models diverge in their predictions in year 1 when no assumptions are made (ESM Table 7). Finally, there were some changes in definitions in some endpoints (such as the removal of vascular death from MI) that limited the degree to which results can be directly compared. Life expectancy projections were also longer using UKPDS-OM2 for all age cohorts. These are consistent with downward secular trends in cardiovascular disease and improvements in mortality and are in line with previously reported estimates of the reduction in life expectancy due to diabetes [23].

The major strength of our model is that it is based on data from the longest follow-up study of patients with type 2 diabetes, including both clinical trial and observational data in patients with a long duration of diabetes who were considerably older than usual clinical trial participants (up to age 90 years). The explicit modelling of second events will enable the model to be used more confidently in secondary prevention for patients with already complicated diabetes [7]. This comprehensive model of type 2 diabetes mellitus incorporates detailed modelling at a patient level and has the capacity to inform individualised medicine and analyses for patient subgroups.

There are a number of limitations to the simulation model in its current form. While the range of modelled outcomes has been extended, complications such as hyper- and hypoglycaemic episodes are not included. These were collected during the UKPDS, but not as true frequencies beyond a small number of deaths from hypoglycaemia. By expanding the number of outcomes and the number of input risk factors, the model has increased in complexity, but in many ways this reflects the nature of a disease that is characterised by so many different complications, the occurrence of which often is determined by interrelationships between risk factors and the patient's clinical history. Finally, as we have used stepwise regression it would be useful to explore the relationship between risk factors and outcomes in other populations to test further the associations observed here.

There are a number of areas requiring further development. In its present form, the model requires individual patient time-paths of clinical risk factors or assumptions regarding the time-paths of baseline risk factors. We are currently developing models to predict these time-paths, which can be easily integrated as additional sub-models to reflect diabetes management practices in the populations of interest. We also need to demonstrate external validity and the applicability of the model to other populations such as those in South East Asia, which are known to have a different profile of complications [24]. Further comparisons with other diabetes simulation models and stand-alone equations such as the UKPDS risk engine [25] will allow us to assess whether the new outcomes model has improved predictions in different diabetic populations. Finally, the use of this new outcomes model for cost-effectiveness analysis will require derivation of QALY weights for events including ulcer and second events, and estimation of costs associated with complications. These enhancements will be addressed in future work.

A common criticism of many computer simulations is that they are a 'black box' with users having little understanding of the underlying relationships between input values and outcomes of the model. By contrast, the UKPDS-OM2 takes a completely transparent approach [6, 26], in which we have fully reported its development, the equations that determine all outcomes and the algorithm used to bring the elements of the model together.

The model will contribute to a greater understanding of the progression of diabetes and its complications and is likely to be used widely be epidemiologists, health economists and trialists. It will play a major role in comparative effectiveness, in cost-effectiveness analyses and in the evaluation of strategies for the management of diabetes in the future.

Acknowledgements

We acknowledge the UKPDS Group (see ESM). We thank P. Kelly (University of Sydney), B. Mihalova and M. Alva (both University of Oxford) for helpful discussions, and R. Coleman and I. Kennedy (Diabetes Trials Unit) for providing information on the UKPDS clinical trial data. RRH is an NIHR senior investigator.

Funding

A. J. Hayes and P. M. Clarke acknowledge the support of Australian National Health and Medical Research Council project grant no. 512463 and capacity building grant no. 571372. A. M. Gray, J. Leal and R. R. Holman acknowledge the support of a UK Medical Research Council project grant on disease modelling (grant ID: 87386).

Duality of interest

The authors declare that there is no duality of interest associated with this manuscript.

Contribution statement

AJH carried out statistical analyses, designed and programmed the model and drafted the manuscript; JL contributed to statistical analyses, carried out UKPDS-OM1 simulations and contributed to aspects of the manuscript; AMG contributed to the design of the study and aspects of the manuscript; RRH provided all the clinical data and provided clinical input on design of the study; PMC conceived the study, gave statistical and modelling advice and drafted the manuscript. All authors critically revised earlier drafts of the manuscript. All authors approved the final version of the manuscript.

References

1. Clarke PM, Gray AM, Briggs A et al (2004) A model to estimate the lifetime health outcomes of patients with type 2 diabetes: the United Kingdom Prospective Diabetes Study (UKPDS) Outcomes Model (UKPDS no. 68). Diabetologia 47:1747-1759

2. The Mount Hood 4 Modeling Group (2007) Computer modeling of diabetes and its complications: a report on the fourth Mount Hood Challenge meeting, Diabetes Care 30:1638-1646

3. Mt Hood Challenge Group (2013) Computer modeling of diabetes and its complications: a report on the fifth Mount Hood Challenge meeting. Value Health doi:10.1016/j.val.2013.01.002

4. Clarke PM, Gray A, Briggs A, Stevens R, Matthews D, Holman R (2005). On behalf of the UK Prospective Diabetes Study (UKPDS 72). Cost utility analyses of intensive blood-glucose and tight blood-pressure control in Type 2 diabetes. Diabetologia 48:868-877.

5. Simon J., Gray A.M, Clarke P., Wade A., Neil A., Farmer A (2006). Cost-effectiveness of self-monitoring of blood glucose in the management of patients with non-insulin treated type 2 diabetes: economic evaluation of data from the randomised controlled DiGEM trial. British Medical Journal 336:1177-1180

6. Leal J, Gray AM, Clarke PM (2009) Development of life expectancy tables for people with type 2 diabetes. European Heart Journal 30:834-839.

7. Leal J, Hayes AJ, Gray AM, Holman RR, Clarke PM (2012) Temporal validation of the UKPDS outcomes model using 10-year posttrial monitoring data. Diabetes Care doi: 10.2337/dc12-1120

8. Eddy DM, Hollingworth W, Caro JJ, Tsevat J, McDonald KM, Wong JB; ISPOR-SMDM Modeling Good Research Practices Task Force (2012). Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-7. Value Health 15:843-850

9. Holman RR, Paul SK, Bethel MA, Matthews DR, Neil HA (2008).10-year follow-up of intensive glucose control in type 2 diabetes (UKPDS 80) N Engl J Med 359:1577-1589

10. UK Prospective Diabetes Study (UKPDS) Group (1998) Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). Lancet 352:837–853

11. Stehouwer CD, Smulders YM (2006) Microalbuminuria and risk for cardiovascular disease: analysis of potential mechanisms. J Am Soc Nephrol 17:2106–2111.

12.Heerspink HJ, Holtkamp FA, de Zeeuw D, Ravid M. Monitoring kidney function and albuminuria in patients with diabetes. Diabetes Care. 2011 May;34 Suppl 2:S325-329.

13. Hillis GS, Woodward M, Rodgers A et al (2012) Resting heart rate and the risk of death and cardiovascular complications in patients with type 2 diabetes mellitus. Diabetologia 55:1283-1290.

14. Tong PC, Lee KF, So WY et al (2004) White blood cell count is associated with macro- and microvascular complications in Chinese patients with type 2 diabetes. Diabetes Care 27:216-222.

15. Go AS, Yang J, Ackerson LM et al (2006) Hemoglobin level, chronic kidney disease, and the risks of death and hospitalization in adults with chronic heart failure: the Anemia in Chronic Heart Failure: Outcomes and Resource Utilization (ANCHOR) Study. Circulation 13:2713-2723

16. Nelson CL, Sun JL, Tsiatis AA, et al Empirical estimation of life expectancy from large clinical trials: use of left-truncated, right-censored survival analysis methodology. Statist. Med. 2008; 27:5525–5555.

17. Schoenfeld D. Partial residuals for the proportional hazards regression model. Biometrika 1982; 69:239-241.

18. Currie CJ, Peters JR, Tynan A, et al (2010) Survival as a function of HbA(1c) in people with type 2 diabetes: a retrospective cohort study. Lancet.375:481-489.

19. Briggs AH, Weinstein MC, Fenwick EA, Karnon J, Sculpher MJ, Paltiel AD (2012) ISPOR-SMDM Modeling Good Research Practices Task Force. Model parameter estimation and uncertainty: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-6. Value Health 15:835-842.

20. American Diabetes Association Consensus Panel (2004) Guidelines for computer modeling of diabetes and its complications, Diabetes Care 27:2262-2265.

21. Clarke PM, Simon J, Cull CA, Holman RR (2006) Assessing the impact of visual acuity on quality of life in individuals with type 2 diabetes using the short form-36. Diabetes Care. 29:1506-1511.

22. Kopec JA, Finès P, Manuel DG et al (2010) Validation of population-based disease simulation models: a review of concepts and methods. BMC Public Health 10:710.

23. Emerging Risk Factors Collaboration, Seshasai SR, Kaptoge S, Thompson A et al (2011). Diabetes mellitus, fasting glucose, and risk of cause-specific death. N Engl J Med. 364:829-841.

24. Chan JCN, Malik V, Jia W et al (2009) Diabetes in Asia: epidemiology, risk factors, and pathophysiology, JAMA:301:2129-2140.

25. Stevens RJ, Kothari V, Adler AI, Stratton IM; United Kingdom Prospective Diabetes Study (UKPDS) Group (2001). The UKPDS risk engine: a model for the risk of coronary heart disease in Type II diabetes (UKPDS 56). Clin Sci 101:671-679.

26. Smith-Spangler CM (2012). Transparency and reproducible research in modeling: why we need it and how to get there. Med Decis Making 32:733-743.

Table 1 Number of events and average event rates observed during UKPDS and PTM for 5,102 participants with newly diagnosed type 2 diabetes mellitus

Event	Total number of events	Annual event rate[a]
Death	2,260	0.0252
First MI	1,014	0.0113
Second MI	169	0.0019
First stroke	504	0.0056
Second stroke	78	0.0009
CHF	351	0.0039
IHD	749	0.0083
First amputation	171	0.0019
Second amputation	58	0.0006
Blindness	271	0.003
Renal failure	113	0.0013
Ulcer	97	0.0011

[a]Calculated from total number of events/total patient-years

Table 2 Comparisons of simulated percentage[a] of patients with events at 10 years from UKPDS-OM1 and UKPDS-OM2

Event type	50–54 years (*n*=555)		60–64 years (*n*=819)		70–74 years (*n*=578)		All ages (*n*=3,984)	
	OM1	OM2	OM1	OM2	OM1	OM2	OM1	OM2
First MI (%)	14.9	7.5	22.5	10.3	29.6	13.3	21.0	9.9
Second MI (%)	n/a	0.9	n/a	1.0	n/a	1.1	n/a	1.0
Ulcer (%)	n/a	1.5	n/a	1.9	n/a	2.2	n/a	1.8
Blindness (%)	2.2	2.2	3.5	3.1	4.9	3.95	3.3	2.9
IHD (%)	8.6	6.9	10.3	8.3	10.5	9	9.5	7.8
First stroke (%)	3.3	3.3	7.9	6.4	14.2	10.7	7.6	6.2
Second stroke (%)	n/a	0.3	n/a	0.7	n/a	1.5	n/a	0.71
Renal failure (%)	0.9	0.3	1.4	0.6	1.6	0.75	1.3	0.5
First amputation (%)	1.7	1.3	2.0	1.6	1.7	1.8	1.8	1.5
Second amputation (%)	n/a	0.4	n/a	0.6	n/a	0.4	n/a	0.44
Heart failure (%)	3.0	2.5	5.9	4.3	9.9	6.4	5.7	4.0
Death (%)	14.5	11.1	32.1	22.3	58.8	43.3	31.6	22.5

[a]Calculated from total patients having event by 10 years/number patients at baseline

n/a, not applicable

Table 3 Simulated life expectancy (95% CI) for three age cohorts using patient-level data from the LDS cohort

Patient group	*N*	Duration of diabetes Mean (SE)	Remaining life expectancy (95% CI) in years	
			UKPDS-OM1[a]	UKPDS-OM2[b]
50-54 years	554	6.4 (0.21)	20.0 (17.7-23.0)	25.1 (24.5-25.7)
60-64 years	819	8.1 (0.21)	13.9 (12.7-16.2)	17.7 (17.1-18.3)
70-74 years	578	9.3 (0.28)	9.1 (8.4-10.7)	11.7 (11.2-12.2)

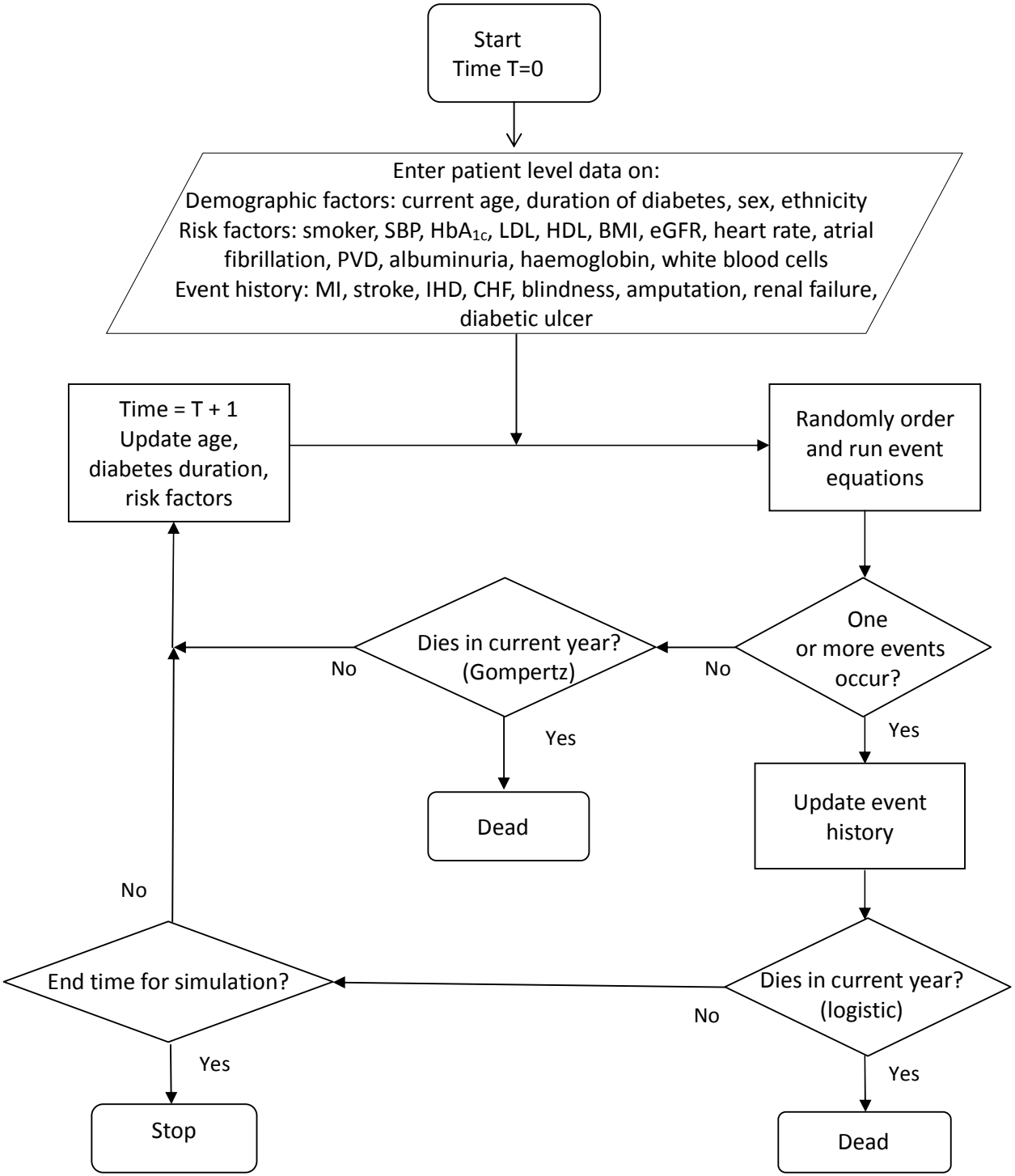
[a]Based on simulations for 40 years (maximum possible with this model)

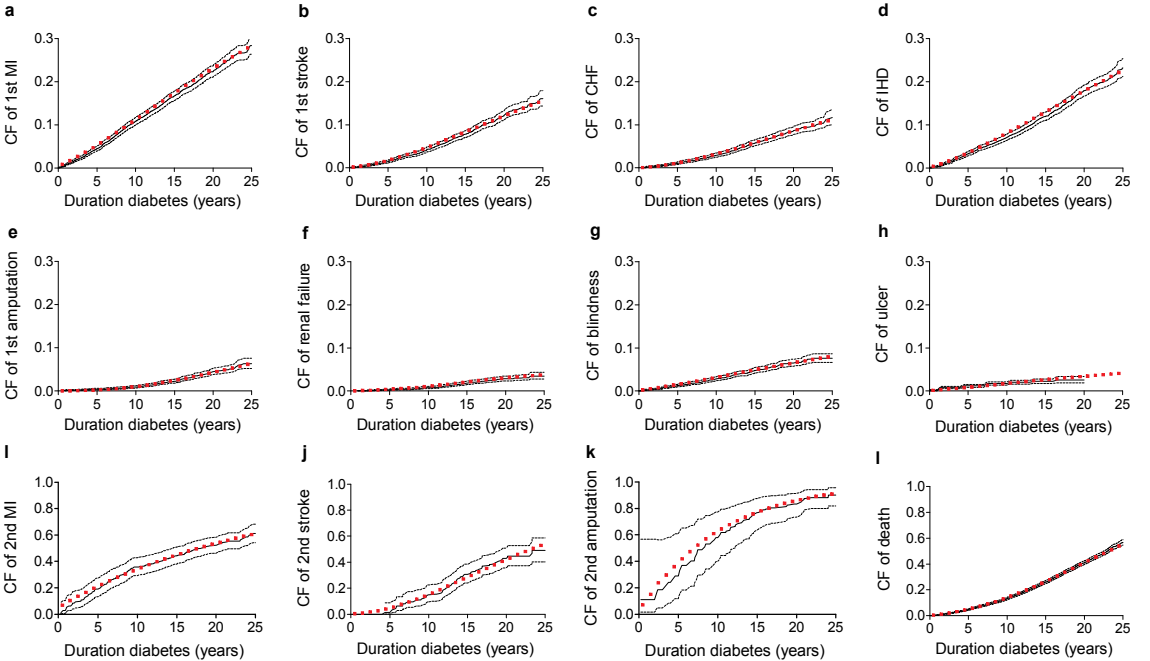
[b]Based on simulations to age 100 years

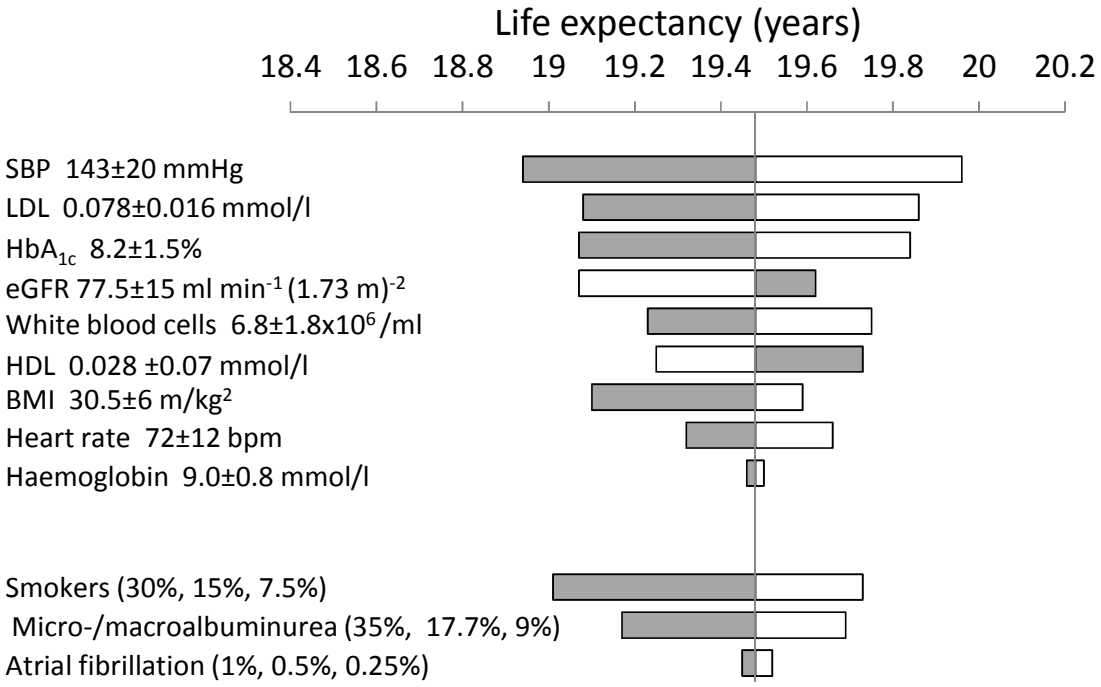
Fig. 1 Flowchart showing structure of simulation model

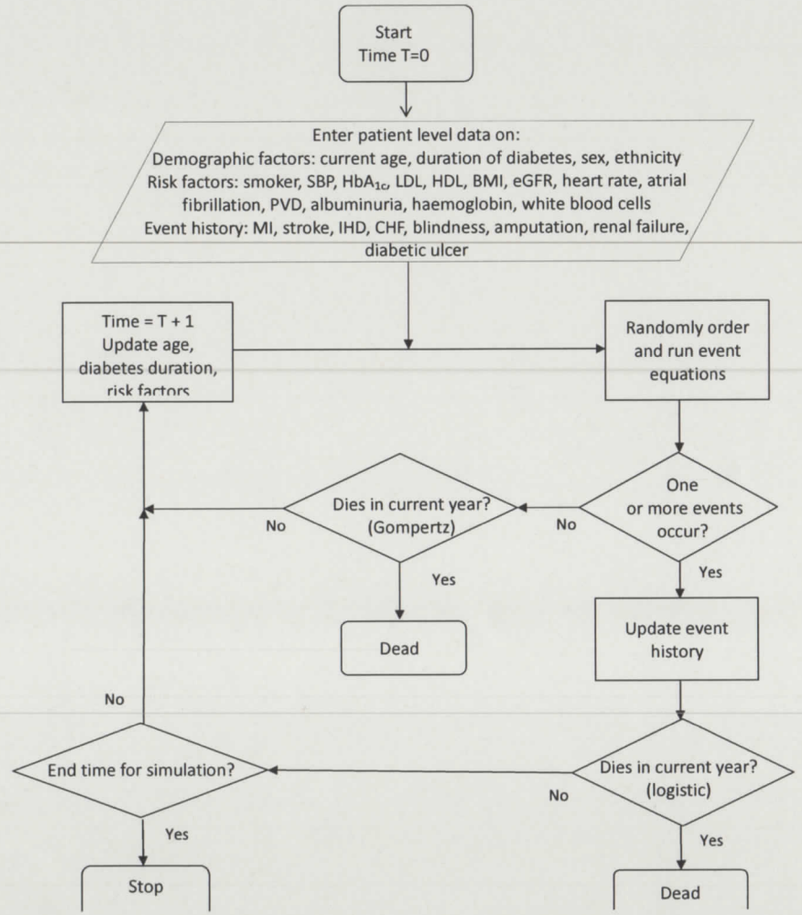
Fig. 2 Comparison of simulated (red squares) and actual (solid black lines) Kaplan–Meier cumulative failure (CF) and 95% CI (dashed black lines) of diabetes complications and death over 25 years. (**a**) First MI; (**b**) first stroke; (**c**) CHF; (**d**) IHD; (**e**) first amputation; (**f**) renal failure; (**g**) blindness; (**h**) ulcer; (**i**) second MI; (**j**) second stroke; (**k**) second amputation; (**l**) death from any cause. Graphs (**i**–**l**) have different *y*-axis scales. Large CIs for second events reflect the low numbers of patients at risk, particularly in the early years of diabetes. Ulcer data only available to 20 years.

Fig. 3 Tornado plot showing one-way sensitivity analysis of change in life expectancy of a cohort of 3,984 LDS patients arising from +1 SD (grey) and −1 SD (white) change in continuous clinical risk factors and a doubling and halving of binary risk factors. To convert values for HbA_{1c} in % into mmol/mol, subtract 2.15 and multiply by 10.929 or use the conversion calculator at www.hba1c.nu/eng/


Start
Time T=0
Enter patient level data on:
Demographic factors: current age, duration of diabetes, sex, ethnicity
Risk factors: smoker, SBP, HbA1c, LDL, HDL, BMI, eGFR, heart rate, atrial fibrillation, PVD, albuminuria, haemoglobin, white blood cells
Event history: MI, stroke, IHD, CHF, blindness, amputation, renal failure, diabetic ulcer
Time = T + 1
Update age, diabetes duration, risk factors
Randomly order and run event equations
One or more events occur?
No
Yes
Dies in current year? (Gompertz)
No
Yes
Dead
Update event history
Dies in current year? (logistic)
No
Yes
Dead
End time for simulation?
No
Yes
Stop


a
CF of 1st MI
Duration diabetes (years)
b
CF of 1st stroke
Duration diabetes (years)
c
CF of CHF
Duration diabetes (years)
d
CF of IHD
Duration diabetes (years)
e
CF of 1st amputation
Duration diabetes (years)
f
CF of renal failure
Duration diabetes (years)
g
CF of blindness
Duration diabetes (years)
h
CF of ulcer
Duration diabetes (years)
i
CF of 2nd MI
Duration diabetes (years)
j
CF of 2nd stroke
Duration diabetes (years)
k
CF of 2nd amputation
Duration diabetes (years)
l
CF of death
Duration diabetes (years)


Life expectancy (years)
18.4
18.6
18.8
19
19.2
19.4
19.6
19.8
20
20.2
SBP 143±20 mmHg
LDL 0.078±0.016 mmol/l
HbA1c 8.2±1.5%
eGFR 77.5±15 ml min-1 (1.73 m)-2
White blood cells 6.8±1.8x10^6 /ml
HDL 0.028 ±0.07 mmol/l
BMI 30.5±6 m/kg2
Heart rate 72±12 bpm
Haemoglobin 9.0±0.8 mmol/l
Smokers (30%, 15%, 7.5%)
Micro-/macroalbuminurea (35%, 17.7%, 9%)
Atrial fibrillation (1%, 0.5%, 0.25%)



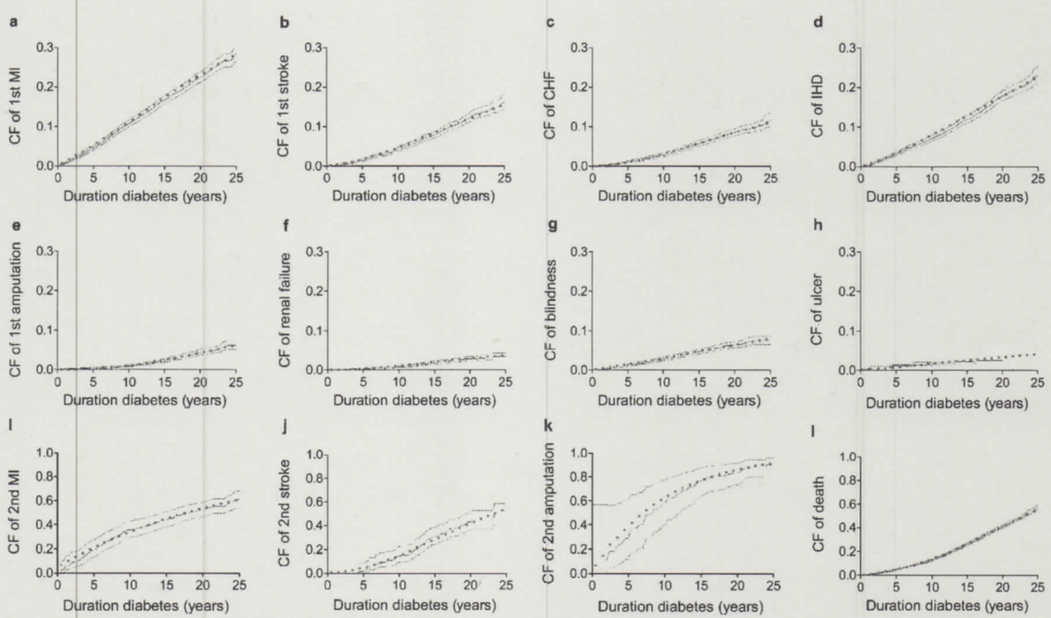
13003
Hayes
Fig. 1

Please set in 11.5 cm

No changes

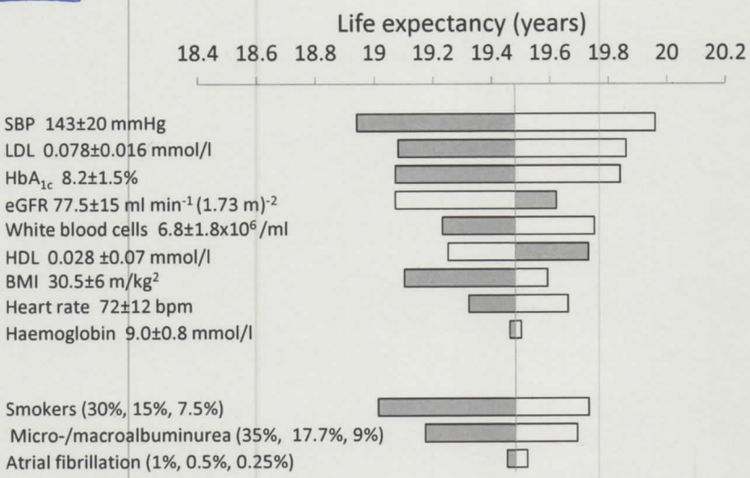
130003
Fig. 2

Please make sure weight of all axis lines is the same



Set in 15cm

13 0003
Fig 3



No changes

Set in full column width

ESM File Legends

ESM Figure 1. The event equations showing significant explanatory variables, hazard ratios, event numbers and linkages

ESM Figure 2. Schematic diagram showing the four mutually exclusive death equations and how they relate to patients' complications status in different years. Death equation 1 applies to patients with no pre-existing or current year complications. The elevated mortality in the first year of any MI, stroke, IHD, CHF, renal failure or amputation event is modelled with a logistic equation (equation 2). Death equation 3 applies when there are pre-existing but no current year complications and equation 4 for subsequent years in which there are both current and pre-existing complications. The graphs, based on real data pertain to a male, with a non-fatal stroke event at the age of 60 and amputation at the age of 79