Development and validation of risk prediction equations to estimate future risk of blindness and lower limb amputation in patients with diabetes: cohort study.

Journal:	*BMJ*
Manuscript ID	BMJ.2015.027121.R1
Article Type:	Research
BMJ Journal:	BMJ
Date Submitted by the Author:	11-Sep-2015
Complete List of Authors:	Hippisley-Cox, Julia; Nottingham University, General Practice Coupland, Carol; University of Nottingham, Division of Primary Care
Keywords:	risk prediction, blindness, amputation, diabetes, epidemiology, QResearch, CPRD, personalised medicine

Development and validation of risk prediction equations to estimate future risk of blindness and lower limb amputation in patients with diabetes: cohort study.

Presenting Original Research

Authors

Julia Hippisley-Cox Professor of Clinical Epidemiology & General Practice. MD, FRCP, FRCGP, MBChB

Carol Coupland Associate Professor in Medical Statistics PhD, MSc, BSc

Institutions

[1]Division of Primary Care, 13th floor, Tower Building, University Park, Nottingham, NG2 7RD, UK

Author for correspondence

Julia Hippisley-Cox

Email: Julia.hippisley-cox@nottingham.ac.uk
Carol.coupland@nottingham.ac.uk

Telephone: +44 (0)115 8466915

Fax: +44 (0)115 8466904

Word count: 3780 words

ABSTRACT

Objective To develop and externally validate risk prediction equations to estimate the 10 year risk of blindness and lower limb amputation in patients with diabetes aged 25-84 years.

Design Cohort study using routinely collected data from general practices in England (UK) contributing to the QResearch® and CPRD databases during the study period 1998 to 2014.

Setting We used 763 QResearch® practices in England to develop the equations. We validated them in 254 different QResearch® practices and 357 CPRD practices.

Participants 454,575 patients with diabetes in the derivation cohort, 142,419 in the QResearch® validation cohort, 206,050 in the CPRD validation cohort.

Measurement Incident diagnoses of blindness and amputation recorded on the patients linked electronic GP, ONS mortality or hospital record. Baseline risk factors included age, type of diabetes, diabetes duration, smoking, ethnicity, deprivation, HBA1C, systolic blood pressure, body mass index, total serum cholesterol/high density lipoprotein ratio, atrial fibrillation, congestive cardiac failure, cardiovascular disease, treated hypertension, peripheral vascular disease, chronic renal disease, rheumatoid arthritis and proliferative retinopathy.

Methods We used Cox proportional hazards models to derive separate risk equations for blindness and amputation in men and women evaluated at 10 years.

Measures of calibration, discrimination, sensitivity and specificity were determined in the two validation cohorts.

Results In the QResearch® derivation cohort, there were 4,822 new cases of lower limb amputation and 8,063 of blindness during follow-up. The risk equations were well calibrated in both validation cohorts. Discrimination was good in the external CPRD cohort for amputation (D statistic: 1.69, Harrell's C statistic: 0.77 in men) and blindness (D statistic 1.40, Harrell's C statistic 0.73 in men) with similar results in women. The CPRD validation results were marginally better than those for the QResearch® validation cohort.

Conclusions We have developed and externally validated risk prediction equations to quantify absolute risk of blindness and amputation in men and women with diabetes. They can be used to identify patients at high risk for prevention or further assessment.

What is known and what this paper adds:

- Patients with type 1 or type 2 diabetes are at increased risk of blindness and amputation but generally do not have an accurate assessment of the magnitude of their individual risk.
- We have developed and externally validated new risk prediction algorithms which calculates absolute risk of developing these complications over a 10 year period in patients with diabetes, taking account of their individual risk factors.

Web calculator Here is a web calculator to calculate the absolute risk of complications among patients with diabetes. It also has the open source software.

URL http://qdiabetes.org/amputation-blindness/index.php

1 Introduction

Diabetes is associated with macrovascular complications including increased risk of coronary heart disease or stroke and microvascular complications such as kidney failure, blindness and amputation[1-3]. Intensive control of risk factors, such as glycosylated haemoglobin and systolic blood pressure, lowers incidence of microvascular disease in type 1[2 4] and type 2 diabetes[5 6]. Tight control of blood parameters is the cornerstone of national guidance[7 8], national audits[3] and quality improvement incentives schemes[9]. However, patients need good quality information on how likely they are to develop complications and the expected risk and benefits from interventions to reduce the risk since very few patients are able to accurately quantify this[10]. Guidelines for cardiovascular disease recommend the use of calculators such as QRISK2 to estimate absolute risk of cardiovascular disease taking account of patient characteristics[7]. Whilst QRISK2 and related tools can be used to assess individualized absolute risk of cardiovascular disease[11], stroke[12] and kidney failure[13] in patients with diabetes, there are currently no tools available to calculate risk of other complications such as amputation or blindness. This is important since these are the complications which patients with diabetes fear most and which most impair quality of life[14]. They are also the complications for which patients are most likely to over-estimate their risk and over-estimate the benefits of intensive treatment.[10]

The UK Prospective Diabetes Study (UKPDS) is a source of information on the incidence of amputation and blindness, based on a cohort which originated from a trial of 5,102 patients aged 25-65 with newly diagnosed type 2 diabetes recruited between 1977 and 1991 and followed up until 1997[5]. Very few patients in the cohort however experienced blindness

(n=116) or amputation (n=45) during follow-up[16]. Also its generalisability is limited because of its historical nature and exclusion of people aged over 65 and those with various comorbidities.

We aimed to derive and externally validate risk prediction equations to quantify absolute 10 year risks of blindness and amputation in patients with diabetes using variables recorded in their primary care electronic record. Our intention was to provide a readily accessible method to quantify an individual patients' absolute risks of blindness and amputation to complete a risk profile for patients with diabetes. This information could be used to provide better information for patients and doctors and to prioritise those patients at highest levels of risk to inform treatment decisions and for closer management of modifiable risk factors.

2 Methods

2.1 Study design and data source

We undertook a cohort study to derive and validate the risk equations in a large population of primary care patients with diabetes using the UK QResearch® database (version 39, www.qresearch.org). We also carried out an external validation using the Clinical Research Practice Datalink (CPRD) database. QResearch® is a continually updated patient level pseudonymised database with data extending back to 1989. It includes clinical and demographic data from over 1,000 general practices covering a population of > 20 million patients, collected in the course of routine healthcare. The primary care data includes demographic information, diagnoses, prescriptions, referrals, laboratory results and clinical

values. Diagnoses, symptoms and clinical values are recorded using the Read code classification[15]. QResearch® has been used for a wide range of clinical research including the development and validation of risk prediction models[11 12 16]. The primary care data is linked at individual patient level to Hospital Episode Statistics (HES), and mortality records from the Office for National Statistics (ONS). HES provides details of all National Health Service (NHS) inpatient admissions since 1997 including primary and secondary causes coded using the ICD-10 classifications. ONS provides details of all deaths in England with primary and underlying causes, also coded using the ICD-10 classification. Patient records are linked using a project specific pseudonymised NHS number which is valid and complete for 99.8% of primary care patients, 99.9% for ONS mortality records and 98% for hospital admissions records[1].

We included all QResearch® practices in England who had been using their Egton Medical Information Systems (EMIS) computer system for at least a year. The EMIS computer system is the predominant commercial IT system used by 55% of family doctors in the UK for routine recording of health data for individual patients (https://www.emishealth.com/). We randomly allocated three quarters of these practices to the derivation dataset and the remaining quarter to a validation dataset. In both datasets we identified open cohorts of patients aged 25-84 years registered with eligible practices between 01 Jan 1998 and 31st July 2014. We then selected patients with diabetes if they had a Read code for diabetes or more than one prescription for insulin or oral hypoglycaemics. We classified patients as having type 1 diabetes if they had been diagnosed under the age of 35 and prescribed insulin[17], all remaining patients were classified as having type 2 diabetes. We excluded patients without a postcode related deprivation score. We determined an entry date to the cohort for each patient, which was the latest of the following dates: date of diagnosis of

diabetes, 25th birthday, date of registration with the practice plus one year, date on which the practice computer system was installed plus one year, and the beginning of the study period (01 January 1998). Patients were censored at the earliest date of the diagnosis of the relevant complication (blindness or lower limb amputation), death, de-registration with the practice, last upload of computerised data, or the study end date (1st August 2014).

We undertook an external validation using general practices in England contributing to the Clinical Research Practice Datalink (CPRD). CPRD is a similar database to QResearch® except that it is derived from practices using a different clinical computer system. We used the subset of 357 CPRD practices linked to ONS mortality and hospital admission data. We used the same definitions for selecting a validation cohort as for QResearch® except that the study end date was 1st August 2012, the latest date for which linked data were available.

2.2 Outcomes

We had two outcomes of interest

1. Lower limb amputation based on a recorded diagnosis or procedure (including above knee and below knee amputations).
2. Blindness (including blindness in one or both eyes, registered blind, severe visual impairment).

We classified patients as having the outcome if there was a record of the relevant diagnosis either in their primary care record, their linked hospital record or ONS mortality record. We used Read codes to identify recorded diagnoses from the primary care record. We used ICD-10 clinical codes and procedure codes from the 4th revision of the Office of Population, Censuses and Surveys Classification of Surgical Operations and Procedures (OPCS-4)[18] to identify incident cases of each outcome from hospital. We used ICD-10 codes to identify

cases from either the primary or underlying cause of death as recorded on the linked ONS mortality record. See appendix 1 for a list of the Read codes, OPCS-4 and ICD-10 codes used. We used the earliest recorded date of the relevant diagnosis or procedure on any of the three data sources as the index date for the diagnosis.

Patients with lower limb amputation at baseline were excluded from the cohort for the analyses of lower limb amputations during follow-up and similarly for blindness.

2.3 Predictor variables

- We examined the following predictor variables based on established risk factors for vascular disease[1 6 11 19-21]: Age at cohort entry (continuous)[22]
- Type of diabetes: type 1 or type 2[2]
- Number of years since diagnosis of diabetes (<1 year; 1-3; 4-6; 7-10; ≥ 11 years)
- Smoking status (non-smoker; ex-smoker; light(1-9 cigarettes/day); moderate(10-19 /day); heavy (20+/day) [22]
- Ethnic group (White/not recorded, Indian, Pakistani, Bangladeshi, Other Asian, Black Caribbean, Black African, Chinese, Other) [19]
- Townsend deprivation score (continuous) [11 21].
- Glycosylated haemoglobin HbA1c mmol/mol (continuous)[1 22-24]
- Systolic blood pressure (continuous) [6 22]
- Body mass index kg/m^2(continuous)
- Total serum cholesterol/high density lipoprotein cholesterol/HDL ratio (continuous)[11]
- Atrial fibrillation[11]
- Congestive cardiac failure
- Cardiovascular disease
- Treated hypertension[11]
- Peripheral vascular disease[21]
- Chronic renal disease
- Rheumatoid arthritis[11]

- Proliferative retinopathy or maculopathy

For each of the continuous clinical variables, we used the value recorded closest to the baseline cohort entry date out of all those recorded prior to the baseline date or within the 6 months after this date. All other predictor variables were based on the latest information recorded in the primary care record before entry to the cohort. The UK now uses the Standard International (SI) unit of millimoles of HbA1c per mole of Hb (mmol/mol) instead of the percentage[25]. Historical values recorded in percentages were converted to the mmol/mol[26].

2.4 Derivation of the models

We developed risk prediction equations for lower limb amputation and blindness in the derivation cohort using established methods[11 12]. We derived separate equations for men and women. Initially we used complete case analyses to derive fractional polynomial terms[27] to model non-linear risk relationships with continuous variables if appropriate (age, body mass index, systolic blood pressure, serum cholesterol/high density lipoprotein ratio, HBA1C). We then used multiple imputation to replace missing values for continuous values and smoking status and used these values in our main analyses[28-30]. All the candidate predictor variables listed above were included in the multiple imputation models along with the log of survival time and the censoring indicator. We log transformed body mass index, HBA1C, cholesterol and HDL prior to imputation as they had skewed normal distributions. We carried out 10 imputations to improve the statistical efficiency of the estimates[31]. We used Cox's proportional hazards models to estimate the coefficients for each risk factor for both of our outcomes using the fractional polynomial terms obtained from the complete

case analyses. We used Rubin's rules to combine the regression coefficients across the imputed datasets[32]. We fitted full models initially then retained variables if they had a hazard ratio of < 0.80 or > 1.20 (for binary variables) and were statistically significant at the 0.05 level. We examined interactions between predictor variables and age and included these where they were significant, plausible (i.e. similar in direction for both men and women and consistent with the literature) and improved model fit. Model fit was assessed by measuring the AIC and BIC values for each imputed set of data.

We used the regression coefficients for each variable from the final model as weights which we combined with the baseline survivor function evaluated up to 15 years to derive risk equations over a period of 15 years of follow-up[33]. This enabled us to derive absolute risk estimates for each year of follow-up, with a specific focus on 10 year risk estimates. We estimated the baseline survivor function based on zero values of centred continuous variables, with all binary predictor values set to zero.

2.5 Validation of the models

We used multiple imputation in the two validation cohorts to replace missing values for continuous variables and smoking status. We carried out 10 imputations. We applied the risk equations for men and women obtained from the derivation cohort to the validation cohorts and calculated measures of discrimination. We calculated R^2 values (explained variation in time to diagnosis of outcome[34]), D statistics[35] (a measure of discrimination where higher values indicate better discrimination) and Harrell's C statistics[36] (an extension of the receiver operating characteristic(ROC) statistic to survival data) over 10 years and combined these model performance measures across imputed datasets using Rubin's rules. We assessed calibration, comparing the mean predicted risks at 10 years with the observed

risk by tenth of predicted risk. The observed risks were obtained using Kaplan-Meier estimates evaluated at 10 years. We applied the risk equations to the validation cohorts to define thresholds for the 10% and 20% of patients at highest estimated risk at 10 years and calculated sensitivity, specificity and observed risks for these thresholds.

We used all the available data for eligible patients on each database to maximise power and generalisability. We used STATA (version 13.1) for all analyses. We adhered to the TRIPOD statement for reporting[37].

Patient Involvement

Patients were not involved in setting the research question, the outcome measures, the design or implementation of the study. Patient representatives from the QResearch Advisory Board have written the information for patients on the QResearch website about the use of the database for research. They have also advised on dissemination including the use of lay summaries describing the research and its results.

3 Results

3.1 Overall study population

Overall, 1017 QResearch® practices in England met our practice inclusion criteria, of which 763 were randomly assigned to the derivation dataset with the remaining 254 practices assigned to the validation cohort. We identified 455,551 patients aged 25-84 years with diabetes in the derivation cohort. We excluded 976 patients (0.21%) without a recorded Townsend deprivation score leaving 454,575 for the derivation analysis. We identified 142,718 patients aged 25-84 years with diabetes in the QResearch® validation cohort. We excluded 299 patients (0.21%) without a recorded Townsend deprivation score leaving 142,419 for validation analysis.

We identified 206,050 patients aged 25-84 years with diabetes in the CPRD validation cohort from the 357 practices with linked Townsend scores and hospital admissions and mortality data.

3.2 Baseline characteristics

Table 1 shows baseline characteristics of 454,575 patients with diabetes in the derivation cohort at study entry. Of these, 94% had type 2 diabetes. Just over half had been diagnosed with diabetes for less than a year at cohort entry, 17% had been diagnosed for 1-3 years, 9% for 4-6 years, 8% for 7-10 years and 12% for 11 or more years. Smoking status was recorded in 95% of patients, ethnicity in 75%, body mass index in 90%, systolic blood pressure in 97% HBA1C in 71% and cholesterol/HDL ratio on 53%. Of the 454,575 patients in the derivation

cohort, 266,142 (58.6%) had missing data for at least one of these variables (including ethnicity).

Baseline characteristics for patients in the QResearch® validation cohort were similar to corresponding values in the derivation cohort. Of the 142,419 patients in the QResearch validation cohort, 83,403 (58.6%) had missing data for at least one variable. Baseline characteristics of the CPRD validation cohort were also similar except the recording of ethnicity (45%), cholesterol/HDL ratio (40%) and HBA1C (58%) was substantially lower in CPRD than in QResearch. Of the 206,050 patients in the CPRD validation cohort, 166,648 (80.9%) had missing data for at least one variable.

3.3 Primary outcomes of amputation and blindness

Table 2 shows the number of incident cases of each outcome during follow-up and the age standardized incidence rates in each cohort. In the QResearch® derivation cohort, there were 4,822 cases of amputation and 8063 cases of blindness. There were 1524 cases of amputation and 1524 cases of blindness in the QResearch® validation cohort and 2294 cases of amputation and 2845 of blindness in the CPRD validation cohort. The rate of blindness in men was lower in CPRD (2.33 per 1000 person years) than in both QResearch cohorts (3.03 per 1000 person years) and was also lower in women, but rates of amputation were similar.

3.4 Predictor variables

Table 3 shows the adjusted hazard ratios for variables in the final models for men and women in the derivation cohort.

3.4.1 Lower limb amputation

The final model for lower limb amputation in women included: age, systolic blood pressure, HBA1C, deprivation, duration of diabetes, smoking status, ethnicity, rheumatoid arthritis, congestive cardiac failure, peripheral vascular disease and chronic renal disease. The final model in men also included type of diabetes and atrial fibrillation. Body mass index and the serum cholesterol/high density lipoprotein ratio were not significantly associated with risk in men or women. Increasing duration of diabetes was associated with an increased risk of lower limb amputation in men and women. Increasing levels of smoking were associated with increased risk of amputation with the association being more marked amongst women than men. For heavy smokers compared with non-smokers, there was a 1.9 fold increase in risk of amputation for women and a 1.3 fold increased risk for men. South Asian ethnic groups had a lower risk compared with those whose ethnic group was either white or not recorded, Caribbean and black African men also had lower risks. Pre-existing peripheral vascular disease was associated with the highest risks (4-fold in women and 3-fold in men) followed by chronic renal disease (2.7- fold in women and 2.3 fold in men).

Figures 1-3 show adjusted hazard ratios for age, HBA1C and systolic blood pressure. Increasing values of age, systolic blood pressure and HBA1C were associated with an increased risk of lower limb amputation in men and women.

3.4.2 Blindness

The final models for blindness in men and women included age, cholesterol/HDL ratio, systolic blood pressure, HBA1C, deprivation, duration of diabetes, type of diabetes, chronic renal disease and existing proliferative retinopathy or maculopathy. Body mass index and smoking status were not significantly associated with risk. Increasing values of age, systolic blood pressure and HBA1C were associated with an increased risk of blindness (Figures 1-3).

Increasing values of the serum cholesterol/high density lipoprotein ratio were also associated with increased risk of blindness. Increasing duration of diabetes was associated with increased risk despite adjustment for age and other risk factors. There was a significant interaction between renal disease and age. Pre-existing proliferative retinopathy or maculopathy was the strongest risk factor with a 2.7 fold increase for women and a 2.9 fold increase for men.

The web calculator which implements the risk equations for the final models can be found at http://qdiabetes.org/amputation-blindness/index.php along with the open source software which includes the equations (published separately as these will be updated over time as newer data becomes available).

3.5 Validation

3.5.1 Discrimination

Table 4 shows the performance of each equation in both validation cohorts. For men in the CPRD cohort, the equations explained 40.6% of the variation in time to diagnosis of amputation and 31.9% for blindness and discrimination was good for amputation (D statistic of 1.69, Harrell's C statistic of 0.77) and blindness (D statistic of 1.40, Harrell's C statistic of 0.73). The results for women in the CPRD cohort were very similar to those for men. The results for both sexes in the CPRD cohort were similar to those for the QResearch® validation cohort although the point estimates for CPRD tended to be marginally higher.

3.5.2 Calibration

Figure 4 shows the mean predicted risks and observed risks of both outcomes at 10 years by tenth of predicted risk applying the equations to men and women in the QResearch

validation cohort. Figure 5 shows comparable results for the CPRD cohort. There was close correspondence between the mean predicted risks and the observed risks within each model tenth indicating that the equations were well calibrated across both validation cohorts.

3.5.3 Performance at threshold for the 10% and 20% of patients at highest risk

Table 5 shows the sensitivity, specificity and observed risk for the 10% and 20% of men and women at highest predicted risk of each outcome for both validation cohorts for illustrative purposes. For example, using a 10 year risk threshold of 3.2% for amputation in men in CPRD to identify the 20% at highest predicted risk, the sensitivity was 58%, the specificity was 80.5% and the observed risk was 7%.

3.6 Implementation

Figure 6 shows a clinical example of the implementation of the equations as a web calculator using http://qdiabetes.org/amputation-blindness/index.php. The example is for a woman, aged 50, non-smoker with newly diagnosed type 2 diabetes with an HBA1C of 65 mmol/mol; cholesterol/HDL ratio of 2 and a systolic blood pressure of 140 mmg Hg. Her 10 year risk of blindness is 1% and her risk of amputation is 0.5%.

Figure 7 shows the results for a man, aged 75 diagnosed with type 2 diabetes 10 years ago , who is a moderate smoker, has chronic kidney disease, HBA1C of 70 mmol/mol, cholesterol ratio of 4 and systolic blood pressure of 160 mm Hg. His 10 year risk of blindness is 14.7% and his risk of amputation is 12.1%.

4 Discussion

4.1 Key findings

We have developed and externally validated risk prediction equations to quantify the absolute risks of blindness and lower limb amputation over 10 years in men and women with type 1 and type 2 diabetes. The equations are well calibrated and have good discrimination with C statistic values of at least 0.72 in the external CPRD validation cohort. To our knowledge, these are the first tools for predicting both the 10 year risk of blindness and amputation – two of the complications which most concern patients with diabetes and affect quality of life.

4.2 Clinical implications

These algorithms are designed to provide better information on the absolute risks of blindness and amputation for patients and doctors to inform management decisions. Patients with diabetes tend to over-estimate their risk of complications and also over-estimate the benefits of treatment[10]. For example, in one study, patients believed they were 1.5 times more likely to become blind and 13 times more likely to have a lower leg amputation than estimates of absolute risk based on the DCCT trial[2 10]. Some may argue that over-estimating risk of complications might result in patients being more likely to take intensive treatment. However, from a holistic and ethical point of view, more accurate individualised information on risk of complications may help patients to make more informed decisions about the balance of risks and benefits of treatment options reflecting their own values and choices. Over-estimation of the risk of complications might lead to increased levels of anxiety and depression which could negatively affect quality of life. This

is especially important since patients with diabetes are more likely to experience anxiety and depression than the general population[38].

For clinicians and the health service, more accurate methods for stratifying patients according to their absolute risk of complications could enable screening programmes more tailored to an individual's level of risk and support the more rational use of scarce resources. For example, blindness can be prevented by screening and treatment of retinopathy[39] and patients at high risk of blindness might need retinal screening more often than once a year. Those at higher risk of amputation might benefit from a proactive targeted program to prevent lower-extremity amputation (including more frequent checks, tailored patient education, specially designed protective footwear, early reporting of foot injuries) since this has been shown to substantially reduce risk of emergency admissions, use of antibiotics, foot operations and lower limb amputation compared with usual practice[40 41]. Better information on absolute risk of individual complications could also prompt more intensive treatment of modifiable risk factors - such as lowering of HBA1C and tighter blood pressure control - which are generally considered to lower risk of microvascular complications such as blindness[2 5 42].

4.3 Comparisons with the literature

The incidence rates of amputation and blindness are comparable to the amputation rate of 1.6 per 1000 patient years and blindness rate of 3.5 per 1000 patient years reported by UKPDS[5]. However our study is approximately 100-fold larger than UKPDS with almost 5,000 incident amputations and over 8,000 cases of recorded blindness and is 10 times larger than the US hospital based cohort study reported by Zhao et al[24] . Our study is also more recent than the UKPDS study which started almost 40 years ago and ended almost 20 years ago [5].

Our study included patients with prevalent type 1 and type 2 diabetes as well as those with a new diagnosis, enabling us to account for the important contribution of duration of diabetes to risk and ensure that the results can be applied to patients with either newly diagnosed or prevalent diabetes.

We included established risk factors in our equations and report hazard ratios similar in both magnitude and direction to those reported elsewhere for lower limb amputation[1], progression of retinopathy and blindness[1 20] which increases the clinical face validity of the equations. As in UKPDS[6], increased systolic blood pressure was associated with increased risks of blindness and lower limb amputation[20] and increased levels of HBAC1 were associated with increased risk of blindness and amputation when compared over equivalent ranges [1 24]. Deprivation and smoking were associated with increased risk of amputation in our study and others[21]. However, smoking was not associated with an increased risk of blindness in our study which is consistent with other research[20]. Non-white ethnic groups had lower risks of lower limb amputation compared with the white group. This contrasts with a US study where Black Africans had a higher risk of amputation[19] .

There are three economic models based on the DCCT[2] and UKPDS[5] studies. The CORE diabetes[43 44] and the Sheffield diabetes models[45] are based on equations derived from the DCCT trial and the UKPDS study. The EAGLE model[46] is based on equations derived from UKPDS, the DCCT as well as the Wisconsin Epidemiological Study of Diabetic Retinopathy. The CORE model predicts risk of amputation[46] whilst the CORE, EAGLE and Sheffield models predict retinopathy rather than blindness.

4.4 Methodological considerations

The methods used to derive and validate these models are very similar to those for other

risk prediction tools derived from the QResearch® database, the strengths and limitations of which have been discussed in detail[11 12]. In summary, key strengths include cohort size, duration of follow up, representativeness, and lack of selection, recall and respondent bias. UK general practices have good levels of accuracy and completeness in recording clinical diagnoses and prescribed medications[47]. The QResearch® database has linked hospital and mortality records for nearly all patients and is therefore likely to have picked up the majority of cases lower limb amputation thereby minimising ascertainment bias. The QResearch database is updated regularly allowing us to update the algorithms over time which can reflect changes in data quality, population characteristics or requirements thereby keeping the tools up to date. We undertook two validations, one using a separate set of practices and patients contributing to QResearch® and the other using a fully external set of practices contributing to CPRD. The results of both validations were extremely similar which is consistent with previous validation studies showing comparable performance using different practice populations[48 49]. Whilst we have derived and validated the equations using UK datasets, the equations could be used internationally by using alternative deprivation scores relevant to the setting (which would need to be scaled to conform with the Townsend score). Local validation should be done to ensure good calibration and discrimination in the applicable population since patients from different countries may have different rates of complications or distributions of risk factors .

Limitations of our study include the lack of formal adjudication of diagnoses, and potential for bias due to missing data which we have addressed using multiple imputation. Whilst we have provided analysis of several thresholds for illustrative purposes, we have not provided definite comment on what threshold of absolute risk should be used to define a “high risk”

group as that would require (a) consideration of the balance of risks and benefits for individuals and (b) cost-effectiveness analyses which are outside the scope of this study.

5 Conclusion

We have developed and validated new risk prediction equations to quantify the absolute risks of blindness and lower limb amputation in patients with diabetes. They can be used to identify patients with diabetes at high risk of these complications for further assessment. Further research is needed to evaluate the clinical outcomes and cost effectiveness of using these risk equations in primary care.

6 Other information

6.1.1 Funding:

There was no external funding for this project.

6.1.2 Acknowledgements

We acknowledge the contribution of EMIS practices who contribute to the QResearch® and EMIS for expertise in establishing, developing and supporting the database.

6.1.3 Approvals:

The project was reviewed in accordance with the QResearch® agreement with NRES Committee East Midlands - Derby [reference 03/4/021]. The project was reviewed by the independent scientific committee of the Clinical Research Practice Datalink [reference 13_079].

6.1.4 Contributorship

JHC initiated the study, undertook the literature review, data extraction, data manipulation and primary data analysis and wrote the first draft of the paper. CC contributed to the design, analysis, interpretation and drafting of the paper.

6.1.5 Competing Interests

All authors have completed the Unified Competing Interest form at

www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare: JHC is professor of clinical epidemiology at the University of Nottingham and co-director of QResearch® – a not-for-profit organisation which is a joint partnership between the University of Nottingham and Egton Medical Information Systems (leading commercial

supplier of IT for 60% of general practices in the UK). JHC is also a paid director of ClinRisk Ltd which produces open and closed source software to ensure the reliable and updatable implementation of clinical risk equations within clinical computer systems to help improve patient care. CC is associate professor of Medical Statistics at the University of Nottingham and a paid consultant statistician for ClinRisk Ltd. This work and any views expressed within it are solely those of the co-authors and not of any affiliated bodies or organisations.

6.1.6 Copyright

The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, an exclusive licence (or non-exclusive for government employees) on a worldwide basis to the BMJ Publishing Group Ltd, and its Licensees to permit this article (if accepted) to be published in BMJ editions and any other BMJPGL products and to exploit all subsidiary rights, as set out in our licence (bmj.com/advice/copyright.shtml).

6.1.7 Data Sharing

The equations presented in this paper will be released as Open Source Software under the GNU lesser GPL v3. The open source software allows use without charge under the terms of the GNU lesser public license version 3. Closed source software can be licensed at a fee.

6.1.8 Transparency statement

JHC is the manuscript's guarantor and affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies are disclosed.

7 References

1. Stratton IM, Adler AI, Neil HA, et al. Association of glycaemia with macrovascular and microvascular complications of type 2 diabetes (UKPDS 35): prospective observational study. BMJ 2000;**321**(7258):405-12.
2. The Diabetes Control and Complications Trial Research Group. The Effect of Intensive Treatment of Diabetes on the Development and Progression of Long-Term Complications in Insulin-Dependent Diabetes Mellitus. New England Journal of Medicine 1993;**329**(14):977-86.
3. Health and Social Care Information Centre. National Diabetes Audit 2012-2013 Report 2 Complications and Mortality, 2014:37.
4. Group DER, Aiello LP, Sun W, et al. Intensive diabetes therapy and ocular surgery in type 1 diabetes. N Engl J Med 2015;**372**(18):1722-33.
5. UK Prospective Diabetes Study (UKPDS) Group. Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). UK Prospective Diabetes Study (UKPDS) Group. Lancet 1998;**352**(9131):837-53.
6. Adler AI, Stratton IM, Neil HA, et al. Association of systolic blood pressure with macrovascular and microvascular complications of type 2 diabetes (UKPDS 36): prospective observational study. BMJ 2000;**321**(7258):412-9.
7. National Clinical Guideline Centre. Lipid modification: cardiovascular risk assessment and the modification of blood lipids for the primary and secondary prevention of cardiovascular disease. London, 2014:286.
8. Goff DC, Lloyd-Jones DM, Bennett G, et al. 2013 ACC/AHA Guideline on the Assessment of Cardiovascular Risk: A Report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. Circulation 2013.
9. Centre NI. The Quality and Outcomes Framework. http://www.ic.nhs.uk/services/qof Accessed 25th February 2015, 2015.
10. Meltzer D, Egleston B. How patients with diabetes perceive their risk for major complications. Eff Clin Pract 2000;**3**(1):7-15.
11. Hippisley-Cox J, Coupland C, Vinogradova Y, et al. Predicting cardiovascular risk in England and Wales: prospective derivation and validation of QRISK2. BMJ 2008:bmj.39609.449676.25.
12. Hippisley-Cox J, Coupland C, Brindle P. Derivation and validation of QStroke score for predicting risk of ischaemic stroke in primary care and comparison with other risk scores: a prospective open cohort study. BMJ 2013;**346**:f2573.
13. Hippisley-Cox J, Coupland C. Predicting the risk of Chronic Kidney Disease in Men and Women in England and Wales: prospective derivation and external validation of the QKidney(R) Scores. BMC Family Practice 2010;**11**:49.
14. Clarke P, Gray A, Holman R. Estimating utility values for health states of type 2 diabetic patients using the EQ-5D (UKPDS 62). Med Decis Making 2002;**22**(4):340-9.
15. Wikipedia. Readcodes. Secondary Readcodes 2015. http://en.wikipedia.org/wiki/Read_code.
16. Hippisley-Cox J, Coupland C, Robson J, et al. Predicting risk of type 2 diabetes in England and Wales: prospective derivation and validation of QDScore. BMJ 2009;**338**:b880-.
17. Hippisley-Cox J, Pringle M. Prevalence, Care and Outcomes for patients with diet controlled diabetes in general practice: cross sectional survey. Lancet 2004;**364**:423-28.
18. Health and Social Care Information Centre. OPCS-4 Classification. Secondary OPCS-4 Classification. http://systems.hscic.gov.uk/data/clinicalcoding/codingstandards/opcs4/.

19. Resnick HE, Valsania P, Phillips CL. Diabetes mellitus and nontraumatic lower extremity amputation in black and white Americans: the National Health and Nutrition Examination Survey Epidemiologic Follow-up Study, 1971-1992. Archives of Internal Medicine 1999;**159**(20):2470-5.
20. Stratton IM, Kohner EM, Aldington SJ, et al. UKPDS 50: risk factors for incidence and progression of retinopathy in Type II diabetes over 6 years from diagnosis. Diabetologia 2001;**44**(2):156-63.
21. Leggetter S, Chaturvedi N, Fuller JH, et al. Ethnicity and risk of diabetes-related lower extremity amputation: a population-based, case-control study of African Caribbeans and Europeans in the United kingdom. Archives of Internal Medicine 2002;**162**(1):73-8.
22. Moss SE, Klein R, Klein BE. The 14-year incidence of lower-extremity amputations in a diabetic population. The Wisconsin Epidemiologic Study of Diabetic Retinopathy. Diabetes Care 1999;**22**(6):951-9.
23. Klein R. Hyperglycemia and microvascular and macrovascular disease in diabetes. Diabetes Care 1995;**18**(2):258-68.
24. Zhao W, Katzmarzyk PT, Horswell R, et al. HbA(1c) and Lower-Extremity Amputation Risk in Low-Income Patients With Diabetes. Diabetes Care 2013;**36**(11):3591-98.
25. Barth JH, Marshall SM, Watson ID. Consensus meeting on reporting glycated haemoglobin (HbA1c) and estimated average glucose (eAG) in the UK: report to the National Director for Diabetes, Department of Health. Diabet Med 2008;**25**(4):381-2.
26. Drugs and Therapuetics Bulletin. Change in units for HbA1c Secondary Change in units for HbA1c 2010. http://dtb.bmj.com/site/about/HBA1C_chart_Feb_10.pdf.
27. Royston P, Ambler G, Sauerbrei W. The use of fractional polynomials to model continuous risk variables in epidemiology. Int J Epidemiol 1999;**28**:964-74.
28. Steyerberg EW, van Veen M. Imputation is beneficial for handling missing data in predictive models. J Epidemiol Community Health 2007;**60**:979.
29. Moons KGM, Donders RART, Stijnen T, et al. Using the outcome for imputation of missing predictor values was preferred. J Epidemiol Community Health 2006;**59**:1092.
30. Schafer J, Graham J. Missing data: our view of the state of the art. Psychological Methods 2002;**7**:147 - 77.
31. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. Stat Med 2011;**30**(4):377-99.
32. Rubin DB. *Multiple Imputation for Non-response in Surveys.* New York: John Wiley, 1987.
33. Hosmer D, Lemeshow S, May S. *Applied Survival Analysis: Regression Modelling of Time to Event data*. US: Wiley, 2007.
34. Royston P. Explained variation for survival models. Stata J 2006;**6**:1-14.
35. Royston P, Sauerbrei W. A new measure of prognostic separation in survival data. Stat Med 2004;**23**:723-48.
36. Harrell FE, Jr., Califf RM, Pryor DB, et al. Evaluating the yield of medical tests. JAMA 1982;**247**(18):2543-6.
37. Collins GS, Reitsma JB, Altman DG, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): The TRIPOD StatementThe TRIPOD Statement. Annals of Internal Medicine 2015;**162**(1):55-63.
38. Gavard JA, Lustman PJ, Clouse RE. Prevalence of depression in adults with diabetes. An epidemiological evaluation. Diabetes Care 1993;**16**(8):1167-78.
39. Bachmann MO, Nelson SJ. Impact of diabetic retinopathy screening on a British district population: case detection and blindness prevention in an evidence- based model. Journal of Epidemiology and Community Health 1998;**52**(1):45-52.
40. Patout CA, Jr., Birke JA, Horswell R, et al. Effectiveness of a comprehensive diabetes lower-extremity amputation prevention program in a predominantly low-income African-American population. Diabetes Care 2000;**23**(9):1339-42.

41. Rith-Najarian S, Branchaud C, Beaulieu O, et al. Reducing lower-extremity amputations due to diabetes. Application of the staged diabetes management approach in a primary care setting. J Fam Pract 1998;**47**(2):127-32.
42. Mohamed Q, Gillies MC, Wong TY. Management of diabetic retinopathy: a systematic review. JAMA 2007;**298**(8):902-16.
43. Palmer AJ, Roze S, Valentine WJ, et al. The CORE Diabetes Model: Projecting Long-term Clinical Outcomes, Costs and Costeffectiveness of Interventions in Diabetes Mellitus (Types 1 and 2) to Support Clinical and Reimbursement Decision-making. Current Medical Research and Opinion 2004;**20**(sup1):S5-S26.
44. Foos V, Palmer JL, Grant D, et al. PRM58 Long-Term Validation of the IMS CORE Diabetes Model in Type 1 and Type 2 Diabetes. Value in Health;**15**(7):A470.
45. Waugh N, Scotland G, McNamee P, et al. Screening for type 2 diabetes: literature review and economic modelling. Health Technol Assessment. London: HTA, 2007:125.
46. Mueller E, Maxion-Bergemann S, Gultyaev D, et al. Development and validation of the Economic Assessment of Glycemic Control and Long-Term Effects of diabetes (EAGLE) model. Diabetes Technol Ther 2006;**8**(2):219-36.
47. Majeed A. Sources, uses, strengths and limitations of data collected in primary care in England. Health Statistics Quarterly 2004(21):5-14.
48. Collins GS, Altman DG. Predicting the 10 year risk of cardiovascular disease in the United Kingdom: independent and external validation of an updated version of QRISK2. BMJ 2012;**344**:e4181.
49. Collins GS, Altman DG. External validation of the QDScore for predicting the 10-year risk of developing Type 2 diabetes. Diabetic Medicine 2011;**28**:599-607.

Figure 1 Adjusted hazard ratios for blindness and lower limb amputation for age in the derivation cohort.

Figure 2 Adjusted hazard ratios for blindness and lower limb amputation for HBA1C in the derivation cohort.

Figure 3 adjusted hazard ratios for blindness and lower limb amputation for systolic blood pressure in the derivation cohort.

Figure 4 Mean predicted risks and observed risks of blindness and lower limb amputation at 10 years by tenth of predicted risk applying the equations to all men and women in the QResearch® validation cohort.

Figure 5 mean predicted risks and observed risks of blindness and lower limb amputation at 10 years by tenth of predicted risk applying the equations to all men and women in the CPRD validation cohort.

Figure 6 web calculator applied to an example female patient

Figure 7 web calculator applied to an example male patient

Table 1 Baseline characteristics of patients with diabetes aged 25-84 years in the QResearch® derivation cohort and both validation cohorts. Values are numbers (percentages) unless stated otherwise.

	QResearch® derivation cohort		*QResearch® validation cohort*		*CPRD validation cohort*	
	women	**men**	**women**	**men**	**women**	**men**
total patients	199679	254896	62407	80012	90280	115770
Type 2 diabetes	188086 (94.2)	241058 (94.6)	58852 (94.3)	75717 (94.6)	85361 (94.6)	109540 (94.6)
Type 1 diabetes	11593 (5.8)	13838 (5.4)	3555 (5.7)	4295 (5.4)	4919 (5.4)	6230 (5.4)
Years since diagnosis						
newly diagnosed (<1 year)	108040 (54.1)	137725 (54.0)	34900 (55.9)	44412 (55.5)	48913 (54.2)	62922 (54.4)
1-3 years	33256 (16.7)	43790 (17.2)	9819 (15.7)	12902 (16.1)	14912 (16.5)	19345 (16.7)
4-6 years	18826 (9.4)	23855 (9.4)	5552 (8.9)	7159 (8.9)	8283 (9.2)	10535 (9.1)
7-10 years	15895 (8.0)	19950 (7.8)	4824 (7.7)	6256 (7.8)	7285 (8.1)	9255 (8.0)
>10 years	23662 (11.9)	29576 (11.6)	7312 (11.7)	9283 (11.6)	10887 (12.1)	13713 (11.8)
mean age (SD)	61.5 (14.1)	59.5 (13.4)	62 (14.0)	59.9 (13.3)	62.7 (13.7)	60.4 (12.9)
mean Townsend score (SD)	.8 (3.4)	.5 (3.4)	.4 (3.3)	.1 (3.2)	0 (3.3)	-.4 (3.2)
Ethnicity recorded	150526 (75.4)	191204 (75.0)	46575 (74.6)	59394 (74.2)	40151 (44.5)	51522 (44.5)
White/not recorded	164366 (82.3)	214557 (84.2)	53760 (86.1)	70000 (87.5)	83962 (93.0)	108518 (93.7)
Indian	6836 (3.4)	9027 (3.5)	1928 (3.1)	2606 (3.3)	1503 (1.7)	2036 (1.8)
Pakistani	5011 (2.5)	5744 (2.3)	854 (1.4)	1071 (1.3)	778 (0.9)	801 (0.7)
Bangladeshi	5979 (3.0)	6731 (2.6)	956 (1.5)	1028 (1.3)	268 (0.3)	321 (0.3)
Other Asian	3134 (1.6)	4017 (1.6)	1005 (1.6)	1393 (1.7)	865 (1.0)	1083 (0.9)
Caribbean	5614 (2.8)	4653 (1.8)	1578 (2.5)	1291 (1.6)	919 (1.0)	768 (0.7)
Black African	3831 (1.9)	4654 (1.8)	1004 (1.6)	1102 (1.4)	838 (0.9)	891 (0.8)
Chinese	693 (0.3)	719 (0.3)	193 (0.3)	222 (0.3)	138 (0.2)	168 (0.1)
Other	4215 (2.1)	4794 (1.9)	1129 (1.8)	1299 (1.6)	1009 (1.1)	1184 (1.0)

smoking status recorded	189827	(95.1)	243379	(95.5)	59409	(95.2)	76617	(95.8)	89107	(98.7)	114577	(99.0)
non smoker	118807	(59.5)	108368	(42.5)	36291	(58.2)	33839	(42.3)	43414	(48.1)	40977	(35.4)
ex-smoker	41073	(20.6)	83683	(32.8)	13572	(21.7)	27231	(34.0)	14002	(15.5)	28100	(24.3)
light smoker	16090	(8.1)	30116	(11.8)	5112	(8.2)	9080	(11.3)	4879	(5.4)	7799	(6.7)
moderate smoker	7720	(3.9)	10684	(4.2)	2512	(4.0)	3196	(4.0)	9772	(10.8)	12756	(11.0)
heavy smoker	6137	(3.1)	10528	(4.1)	1922	(3.1)	3271	(4.1)	5931	(6.6)	11363	(9.8)
Smoker amount not recorded		n/a		n/a		n/a		n/a	11109	(12.3)	13582	(11.7)
Medical conditions at baseline												
atrial fibrillation	7995	(4.0)	11009	(4.3)	2684	(4.3)	3626	(4.5)	3952	(4.4)	5273	(4.6)
congestive cardiac failure	6783	(3.4)	9986	(3.9)	2255	(3.6)	3136	(3.9)	3504	(3.9)	4641	(4.0)
cardiovascular disease	31729	(15.9)	55262	(21.7)	10170	(16.3)	17453	(21.8)	16188	(17.9)	26826	(23.2)
treated hypertension	78323	(39.2)	85634	(33.6)	24451	(39.2)	26721	(33.4)	31477	(34.9)	32465	(28.0)
peripheral vascular disease	5242	(2.6)	10380	(4.1)	1692	(2.7)	3257	(4.1)	2846	(3.2)	5344	(4.6)
chronic renal disease	2325	(1.2)	2857	(1.1)	718	(1.2)	905	(1.1)	930	(1.0)	1185	(1.0)
rheumatoid arthritis	7458	(3.7)	4651	(1.8)	2204	(3.5)	1477	(1.8)	1976	(2.2)	1206	(1.0)
proliferative retinopathy or maculopathy	5531	(2.8)	7657	(3.0)	1653	(2.6)	2162	(2.7)	1319	(1.5)	1913	(1.7)
Existing blindness	3416	(1.7)	3701	(1.5)	1126	(1.8)	1169	(1.5)	1789	(2.0)	1656	(1.4)
Existing lower limb amputation	1010	(0.5)	2073	(0.8)	346	(0.6)	728	(0.9)	455	(0.5)	1013	(0.9)
clinical values at baseline												
HBA1C recorded	141005	(70.6)	180594	(70.9)	43575	(69.8)	56107	(70.1)	51725	(57.3)	67013	(57.9)
mean HBA1C (SD)	61.4	(20.8)	63	(22.0)	61.1	(20.8)	62.9	(21.9)	60.8	(21.1)	62.6	(22.0)
BMI recorded	179818	(90.1)	232298	(91.1)	55892	(89.6)	72979	(91.2)	82814	(91.7)	107778	(93.1)
mean BMI (SD)	31.1	(6.3)	29.8	(5.3)	31.2	(6.4)	29.9	(5.3)	30.9	(6.3)	29.7	(5.3)
cholesterol ratio recorded	105436	(52.8)	138385	(54.3)	33392	(53.5)	43988	(55.0)	35174	(39.0)	46530	(40.2)
mean cholesterol/HDL ratio (SD)	4.1	(1.4)	4.5	(1.5)	4.1	(1.4)	4.5	(1.5)	4.2	(1.5)	4.5	(1.6)
systolic blood pressure recorded	194001	(97.2)	246991	(96.9)	60728	(97.3)	77707	(97.1)	88792	(98.4)	113582	(98.1)
mean SBP (SD)	139.3	(20.0)	138.4	(18.6)	139.8	(20.0)	138.6	(18.6)	141.4	(20.6)	140	(19.0)

Table 2 Numbers of incident cases of blindness and lower limb amputation during follow-up and age standardised incidence rates per 1,000 person years in men and women with diabetes aged 25-84 years in the derivation cohort and validation cohorts

	QResearch® derivation cohort		QResearch® validation cohort		CPRD validation cohort	
	cases	rate per 1,000 person years (95% CI)	cases	rate per 1,000 person years (95% CI)	cases	rate per 1,000 person years (95% CI)
Women						
amputation	1,541	1.34 (1.27 to 1.41)	482	1.32 (1.20 to 1.44)	675	1.32 (1.22 to 1.42)
blindness	4,074	3.43 (3.33 to 3.54)	1,365	3.59 (3.40 to 3.79)	1,487	2.78 (2.64 to 2.93)
Men						
amputation	3,281	2.36 (2.28 to 2.44)	1,042	2.33 (2.19 to 2.47)	1,619	2.66 (2.53 to 2.79)
blindness	3,989	3.03 (2.93 to 3.12)	1,286	3.04 (2.88 to 3.21)	1,358	2.33 (2.20 to 2.45)

Notes:
Patients with existing diagnoses of each complication at baseline were dropped from the relevant cohort.
Rates were directly age standardised to the overall age distribution of patients aged 25 to 84 within the QResearch® derivation cohort in 5-year age bands

Table 3 Adjusted hazard ratios with 95% confidence intervals for blindness and lower limb amputation in men and women in the derivation cohort. For fractional polynomial terms see footnotes and figures 1 to 3.

		adjusted hazard ratio women	*adjusted hazard ratio men*
Amputation[1]	Townsend deprivation score[3]	1.10 (1.01 to 1.19)	1.29 (1.22 to 1.36)
	duration of diabetes		
	newly diagnosed (< 1year)	1	1
	1-3 years	1.59 (1.36 to 1.85)	1.68 (1.51 to 1.87)
	4-6 years	1.69 (1.42 to 2.01)	2.03 (1.81 to 2.28)
	7-10 years	2.37 (2.01 to 2.79)	2.67 (2.39 to 3.00)
	>10 years	3.30 (2.89 to 3.78)	3.49 (3.15 to 3.86)
	smoking status		
	non smoker	1	1
	ex-smoker	1.08 (0.94 to 1.24)	0.94 (0.87 to 1.03)
	light smoker	1.59 (1.34 to 1.88)	1.28 (1.14 to 1.43)
	moderate smoker	1.58 (1.25 to 1.99)	1.15 (0.96 to 1.37)
	heavy smoker	1.89 (1.49 to 2.41)	1.26 (1.06 to 1.49)
	ethnicity		
	white/not recorded	1	1
	Indian	0.44 (0.28 to 0.68)	0.42 (0.32 to 0.55)
	Pakistani	0.72 (0.47 to 1.12)	0.40 (0.28 to 0.58)
	Bangladeshi	0.29 (0.15 to 0.56)	0.12 (0.07 to 0.22)
	Other Asian	0.70 (0.39 to 1.27)	0.42 (0.26 to 0.67)
	Caribbean	0.87 (0.65 to 1.18)	0.49 (0.36 to 0.66)
	Black African	0.92 (0.55 to 1.54)	0.38 (0.23 to 0.61)
	Chinese	0.50 (0.12 to 1.99)	0.35 (0.11 to 1.09)
	Other	0.70 (0.44 to 1.10)	0.63 (0.45 to 0.87)
	co-morbidity		
	type 1 diabetes (vs type 2)	NS	1.26 (1.09 to 1.45)
	rheumatoid arthritis	1.50 (1.19 to 1.90)	1.39 (1.11 to 1.75)
	atrial fibrillation	NS	1.26 (1.07 to 1.49)
	congestive cardiac failure	1.79 (1.44 to 2.22)	1.34 (1.14 to 1.58)
	peripheral vascular disease	4.26 (3.63 to 4.99)	3.16 (2.84 to 3.51)
	chronic renal disease	2.68 (1.96 to 3.66)	2.26 (1.80 to 2.85)
Blindness[2]	Cholesterol/HDL ratio[4]	1.06 (1.03 to 1.09)	1.03 (1.00 to 1.06)
	Townsend deprivation score[3]	1.21 (1.15 to 1.27)	1.33 (1.27 to 1.39)
	duration of diabetes		
	newly diagnosed (<1 year)	1	1
	1-3 years	1.36 (1.25 to 1.49)	1.40 (1.28 to 1.54)

	4-6 years	1.51 (1.36 to 1.67)	1.42 (1.28 to 1.58)
	7-10 years	1.72 (1.55 to 1.91)	1.57 (1.41 to 1.76)
	>10 years	2.17 (1.97 to 2.38)	2.09 (1.90 to 2.29)
	co-morbidity		
	type 1 diabetes (vs type 2)	1.50 (1.26 to 1.78)	1.44 (1.22 to 1.70)
	chronic renal disease	1.49 (1.17 to 1.89)	2.57 (1.88 to 3.52)
	Proliferative retinopathy/maculopathy	2.67 (2.37 to 3.02)	2.93 (2.61 to 3.29)

Notes

[1] amputation model in women also included terms for: age (linear), systolic blood pressure (2 FP terms:, -1 -0.5), hba1c (2 FP terms, 3 3); amputation model in men included terms for: age (linear), systolic blood pressure (2 FP terms -2 0.5) hba1c (2 FP terms 2 2).

[2] blindness model in women also included terms for: age (2 FP terms 2 2), systolic blood pressure (linear), hba1c (2 FP terms 2 2); the model in men also included terms for: age (2 FP terms 2 2), systolic blood pressure (2 FP terms 1 2), hba1c (2 FP terms -2 -2). There was an interaction between age and renal disease in men.

[3] the Townsend deprivation score ranges between -7 (most affluent) and +11 (most deprived). Adjusted hazard ratio is per 5 unit increase.

[4] adjusted hazard ratio is per unit increase.

Table 4 Performance of the equations in men and women in CPRD validation cohort and QResearch® validation cohort

	statistic	CPRD validation cohort	QResearch® validation cohort
		Mean (95% CI)	Mean (95% CI)
women			
amputation	D statistic	1.61 (1.45 to 1.77)	1.30 (1.14 to 1.47)
	R^2 (%)	38.22 (33.61 to 42.83)	28.90 (23.70 to 34.10)
	Harrell's C statistic	0.762 (0.735 to 0.789)	0.700 (0.670 to 0.731)
blindness	D statistic	1.36 (1.27 to 1.46)	1.32 (1.23 to 1.42)
	R^2 (%)	30.78 (27.94 to 33.63)	29.44 (26.50 to 32.39)
	Harrell's C statistic	0.733 (0.719 to 0.747)	0.725 (0.709 to 0.741)
men			
amputation	D statistic	1.69 (1.59 to 1.79)	1.48 (1.38 to 1.59)
	R^2 (%)	40.57 (37.70 to 43.44)	34.42 (31.14 to 37.70)
	Harrell's C statistic	0.770 (0.755 to 0.784)	0.748 (0.730 to 0.767)
blindness	D statistic	1.40 (1.31 to 1.49)	1.33 (1.23 to 1.42)
	R^2 (%)	31.93 (29.04 to 34.82)	29.57 (26.53 to 32.62)
	Harrell's C statistic	0.732 (0.716 to 0.747)	0.714 (0.696 to 0.731)

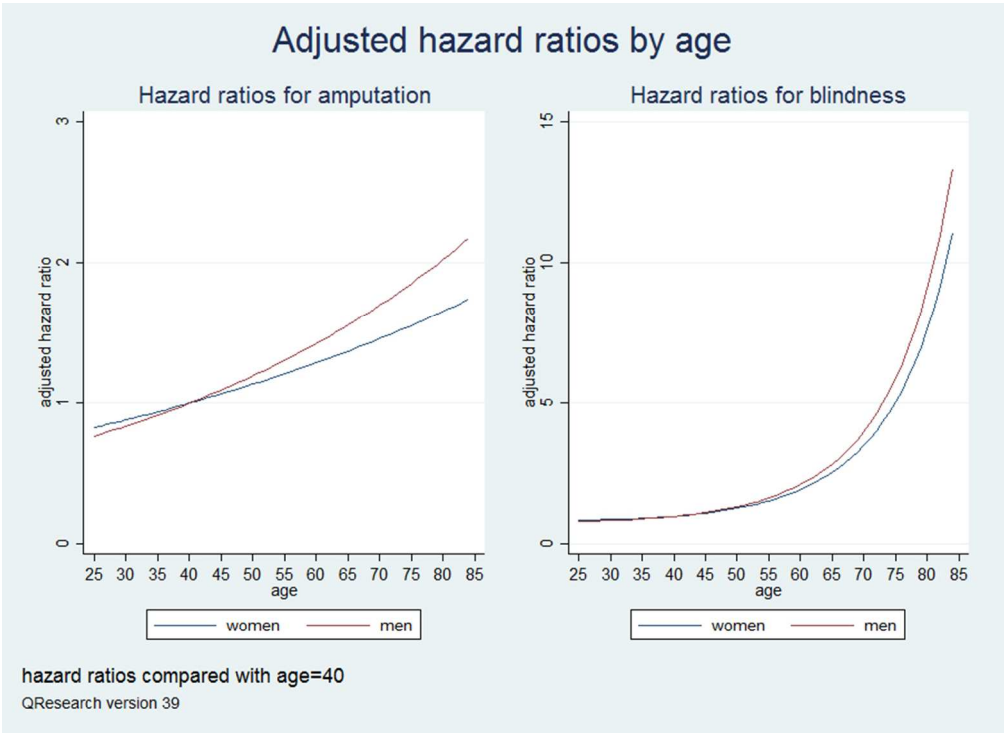
Notes on understanding validation statistics:

Harrell's C statistic is a measure of discrimination where higher values indicate better discrimination. The D statistic is also a measure of discrimination which is specific to censored survival data where higher values indicate better discrimination. R^2 measures explained variation in time to diagnosis of the outcome and higher values indicate more variation is explained.

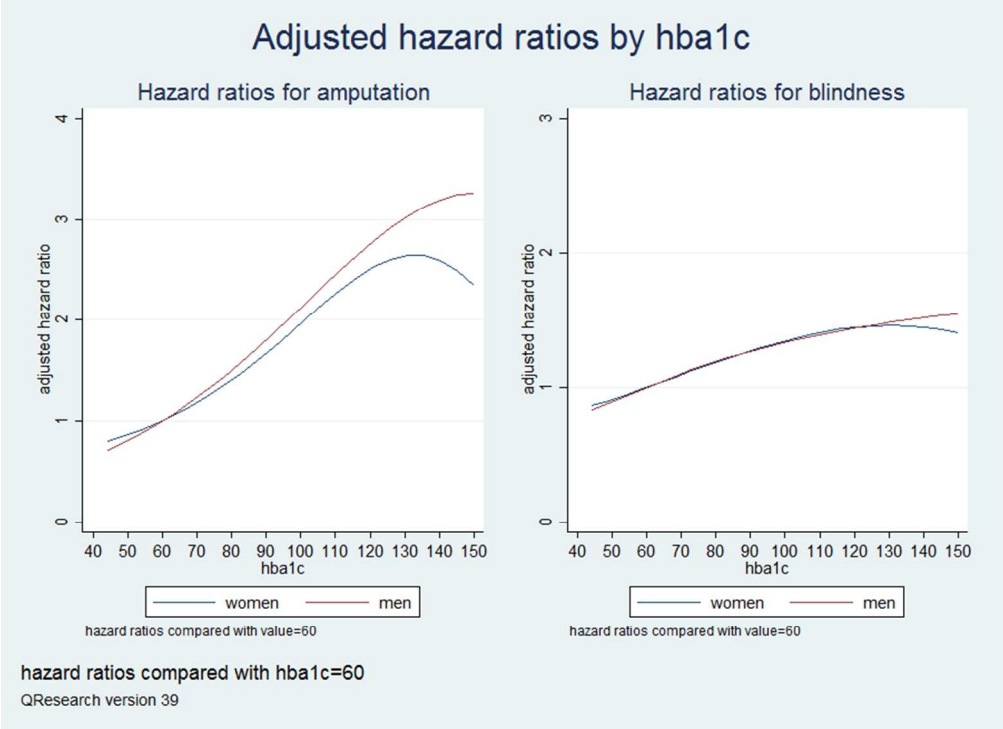
Table 5 Performance of each model in both the QResearch® and CPRD validation cohorts based on the 10% and 20% of patients at highest predicted risk

	QResearch® cohort				CPRD cohort			
	Cut off (%) for 10 year risk[1]	*sensitivity (%)*	*Specificity * (%)*	*observed risk (%)*	*Cut off (%) for 10 year risk*[1]	*sensitivity (%)*	*Specificity* (%)*	*observed risk (%)*
Women								
Amputation (top 10%)	2.6	33.2	90.2	4.6	2.9	39.4	90.2	4.9
Amputation (top 20%)	1.8	48.1	80.2	3.2	2.0	59.8	80.3	3.7
Blindness (top 10%)	8.1	27.9	90.4	12.8	8.0	25.7	90.2	8.7
Blindness (top 20%)	5.6	45.1	80.5	9.6	5.6	44.3	80.4	7.2
Men								
Amputation (top 10%)	4.5	37.5	90.3	7.9	4.8	41.9	90.4	10.2
Amputation (top 20%)	3.0	53.5	80.4	5.7	3.2	58.0	80.5	7.0
Blindness (top 10%)	6.2	27.6	90.2	9.5	6.0	31.5	90.2	8.3
Blindness (top 20%)	4.1	45.9	80.4	7.2	4.1	49.1	80.3	6.1

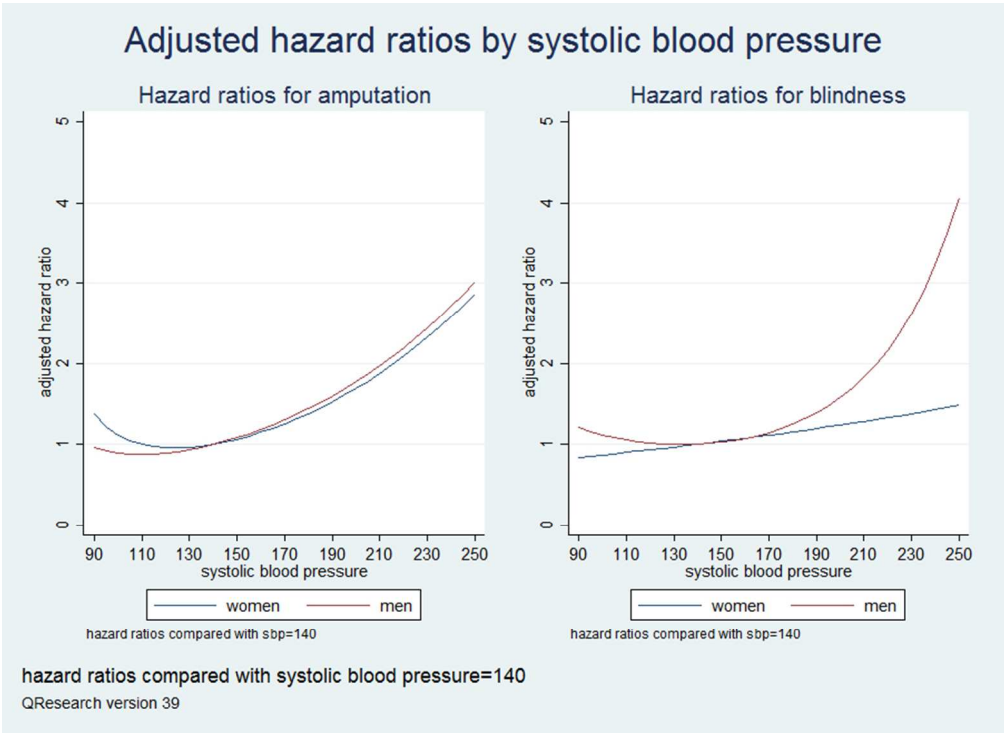
[1] This is the risk threshold for the 10% or 20% of patients at highest predicted risk of the outcome over 10 years



317x231mm (72 x 72 DPI)



317x231mm (72 x 72 DPI)



317x231mm (72 x 72 DPI)

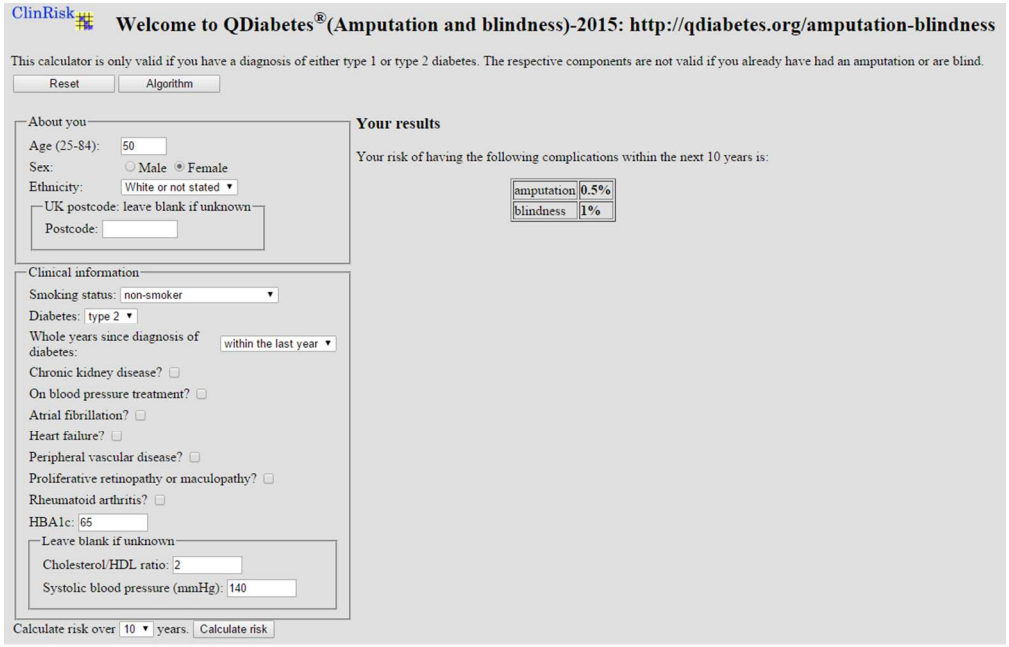
Group Name	clincial cod	clinical term
Blindness (Read code)	2B6A-1	O/E - blind R-eye
Blindness (Read code)	2B6B	O/E - R-eye completely blind
Blindness (Read code)	2B6S	O/E - pinhole R-eye completely blind
Blindness (Read code)	2B7A-1	O/E - blind L-eye
Blindness (Read code)	2B7B	O/E - L-eye completely blind
Blindness (Read code)	2B7S	O/E - pinhole L-eye completely blind
Blindness (Read code)	6688	Registered partially sighted
Blindness (Read code)	6688-1	Registered partially blind
Blindness (Read code)	6689	Registered blind
Blindness (Read code)	6689-1	Registered severely sight impaired
Blindness (Read code)	668C	Certificate of vision impairment
Blindness (Read code)	8F6-1	Blind rehabilitation
Blindness (Read code)	8F61	Blind rehabilitation
Blindness (Read code)	8F62	Blind lead dog rehabilitation
Blindness (Read code)	9m08	Excluded from diabetic retinopathy screening as blind
Blindness (Read code)	EGTONBL2	Blind (subjectively)
Blindness (Read code)	EGTONVI8	Vision - blind despite any aid
Blindness (Read code)	F49	Blindness and low vision
Blindness (Read code)	F49-1	Impaired vision
Blindness (Read code)	F49-2	Low vision
Blindness (Read code)	F49-3	Partial sight
Blindness (Read code)	F49-4	Sight impaired
Blindness (Read code)	F490	Blindness, both eyes
Blindness (Read code)	F490-98	Blind/low vision - both eyes
Blindness (Read code)	F490-99	Blind - both eyes
Blindness (Read code)	F4900	Unspecified blindness both eyes
Blindness (Read code)	F4902	Better eye: near total VI, Lesser eye: unspecified
Blindness (Read code)	F4903	Better eye: near total VI, Lesser eye: total VI
Blindness (Read code)	F4904	Better eye: near total VI, Lesser eye: near total VI
Blindness (Read code)	F4905	Better eye: profound VI, Lesser eye: unspecified
Blindness (Read code)	F4906	Better eye: profound VI, Lesser eye: total VI
Blindness (Read code)	F4907	Better eye: profound VI, Lesser eye: near total VI
Blindness (Read code)	F4908	Better eye: profound VI, Lesser eye: profound VI
Blindness (Read code)	F4909	Acquired blindness, both eyes
Blindness (Read code)	F490z	Blindness both eyes NOS
Blindness (Read code)	F491	Better eye: low vision, Lesser eye: profound VI
Blindness (Read code)	F4910	One eye blind, one eye low vision
Blindness (Read code)	F4911	Better eye: severe VI, Lesser eye: blind, unspecified
Blindness (Read code)	F4912	Better eye: severe VI, Lesser eye: total VI
Blindness (Read code)	F4913	Better eye: severe VI, Lesser eye: near total VI
Blindness (Read code)	F4914	Better eye: severe VI, Lesser eye: profound VI
Blindness (Read code)	F4915	Better eye: moderate VI, Lesser eye: blind, unspecified
Blindness (Read code)	F4916	Better eye: moderate VI, Lesser eye: total VI
Blindness (Read code)	F4917	Better eye: moderate VI, Lesser eye: near total VI
Blindness (Read code)	F4918	Better eye: moderate VI, Lesser eye: profound VI
Blindness (Read code)	F491z	One eye blind, one eye low vision NOS
Blindness (Read code)	F492	Low vision, both eyes
Blindness (Read code)	F4920	Low vision, both eyes unspecified
Blindness (Read code)	F4921	Better eye: severe VI, Lesser eye: low vision unspecified

Blindness (Read code)	F4922	Better eye: severe VI, Lesser eye: severe VI
Blindness (Read code)	F4923	Better eye: moderate VI, Lesser eye: low vision unspecified
Blindness (Read code)	F4924	Better eye: moderate VI, Lesser eye: severe VI
Blindness (Read code)	F4925	Better eye: moderate VI, Lesser eye: moderate VI
Blindness (Read code)	F492z	Low vision, both eyes NOS
Blindness (Read code)	F493	Visual loss, both eyes unqualified
Blindness (Read code)	F493-99	Blind/low vision - both eyes
Blindness (Read code)	F494	Legal blindness USA
Blindness (Read code)	F495	Profound impairment, one eye
Blindness (Read code)	F4950	Blindness, one eye, unspecified
Blindness (Read code)	F4951	Lesser eye: total visual impairment, Better eye: unspecified
Blindness (Read code)	F4952	Lesser eye: total VI, Better eye: near normal vision
Blindness (Read code)	F4953	Lesser eye: total VI, Better eye: normal vision
Blindness (Read code)	F4954	Lesser eye: near total VI, Better eye: unspecified
Blindness (Read code)	F4955	Lesser eye: near total VI, Better eye: near normal vision
Blindness (Read code)	F4956	Lesser eye: near total VI, Better eye: normal vision
Blindness (Read code)	F4957	Lesser eye: profound VI, Better eye: unspecified
Blindness (Read code)	F4958	Lesser eye: profound VI, Better eye: near normal vision
Blindness (Read code)	F4959	Lesser eye: profound VI, Better eye: normal vision
Blindness (Read code)	F495A	Acquired blindness, one eye
Blindness (Read code)	F495z	Profound impairment one eye NOS
Blindness (Read code)	F496	Low vision, one eye
Blindness (Read code)	F496-99	Blind/low vision -one eye only
Blindness (Read code)	F4960	Low vision, one eye, unspecified
Blindness (Read code)	F4961	Lesser eye: severe VI, Better eye: unspecified
Blindness (Read code)	F4962	Lesser eye: severe VI, Better eye: near normal vision
Blindness (Read code)	F4963	Lesser eye: severe VI, Better eye: normal vision
Blindness (Read code)	F4964	Lesser eye: moderate VI, Better eye: unspecified
Blindness (Read code)	F4965	Lesser eye: moderate VI, Better eye: near normal vision
Blindness (Read code)	F4966	Lesser eye: moderate VI, Better eye: normal vision
Blindness (Read code)	F496z	Low vision, one eye NOS
Blindness (Read code)	F49y	Visual loss, one eye, unqualified
Blindness (Read code)	F49y-99	Blind/low vision -one eye only
Blindness (Read code)	F49z	Visual loss NOS
Blindness (Read code)	F49z-1	Acquired blindness
Blindness (Read code)	F49z-99	Blindness/low vision NOS
Blindness (Read code)	F49z0	Charles Bonnet syndrome
Blindness (ICD10)	H54	H54 - Visual impairment including blindness (binocular or mon
Blindness (ICD10)	H540	H540 - Blindness, binocular
Blindness (ICD10)	H541	H541 - Severe visual impairment, binocular
Blindness (ICD10)	H544	H544 - Blindness, monocular
Blindness (ICD10)	H545	H545 - Severe visual impairment, monocular
Blindness (ICD10)	H549	H549 - Unspecified visual impairment (binocular)
Group Name	**clincial cod**	**clinical term**
Amputation (Read code)	14N4	H/O: limb amputation
Amputation (Read code)	14N4-1	Amputee - limb
Amputation (Read code)	14N41	H/O: lower limb amputation
Amputation (Read code)	14N4Z	H/O: limb amputation NOS
Amputation (Read code)	2G42	O/E - Amputated right leg
Amputation (Read code)	2G43	O/E - Amputated left leg

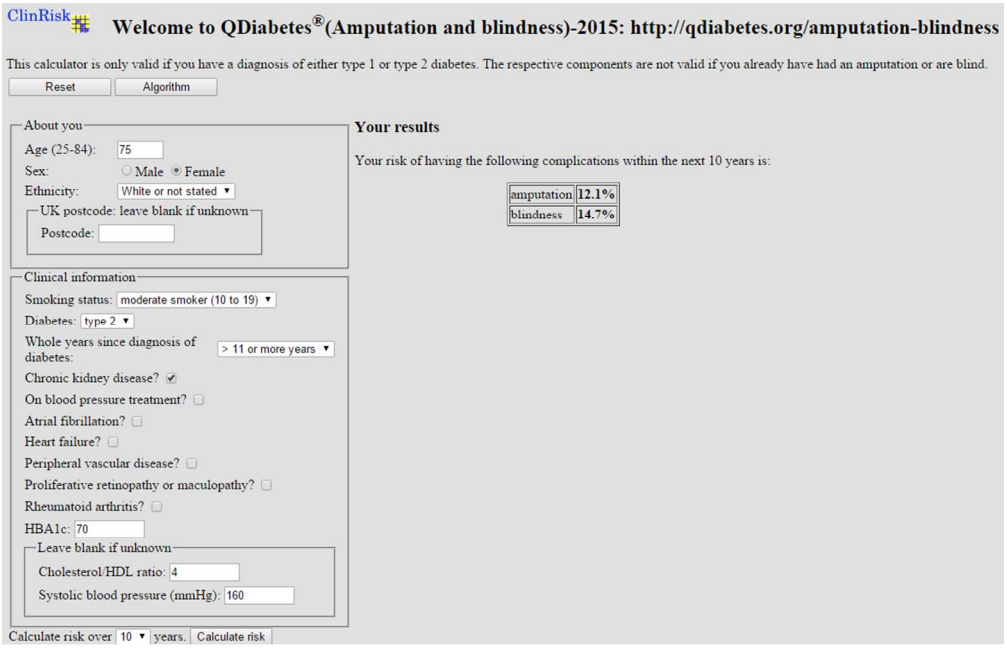
Amputation (Read code)	2G44	O/E - Amputated right above knee
Amputation (Read code)	2G45	O/E - Amputated left above knee
Amputation (Read code)	2G46	O/E - Amputated right below knee
Amputation (Read code)	2G47	O/E - Amputated left below knee
Amputation (Read code)	2G4A	O/E - amputated left midfoot
Amputation (Read code)	2G4B	O/E - amputated right midfoot
Amputation (Read code)	7L06	Amputation of leg
Amputation (Read code)	7L06-99	Amputation - lower limb
Amputation (Read code)	7L060	Hindquarter amputation
Amputation (Read code)	7L060-1	Ferre hindquarter amputation
Amputation (Read code)	7L060-2	Gordon - Taylor hindquarter amputation
Amputation (Read code)	7L060-3	Jaboulay hindquarter amputation
Amputation (Read code)	7L060-4	King hindquarter amputation
Amputation (Read code)	7L060-5	Sorrondo hindquarter amputation
Amputation (Read code)	7L060-6	Steelquist hindquarter amputation
Amputation (Read code)	7L060-7	Taylor hindquarter amputation
Amputation (Read code)	7L060-99	Hind quarter amputation
Amputation (Read code)	7L061	Disarticulation of hip
Amputation (Read code)	7L061-1	Boyd disarticulation of hip
Amputation (Read code)	7L061-2	Fitzmaurice - Kelly disarticulation of hip
Amputation (Read code)	7L061-99	Hip disarticulation
Amputation (Read code)	7L062	Amputation above knee
Amputation (Read code)	7L062-1	Kirk amputation of leg through thigh
Amputation (Read code)	7L062-2	Amputation of leg through thigh
Amputation (Read code)	7L062-99	Above knee amputation
Amputation (Read code)	7L063	Amputation through knee
Amputation (Read code)	7L063-1	Batch disarticulation of knee
Amputation (Read code)	7L063-2	Callander disarticulation of knee
Amputation (Read code)	7L063-3	Disarticulation of knee
Amputation (Read code)	7L063-4	Gritti-Stokes disarticulation of knee
Amputation (Read code)	7L063-5	Kirk disarticulation of knee
Amputation (Read code)	7L063-6	Mazet disarticulation of knee
Amputation (Read code)	7L063-7	McFaddin disarticulation of knee
Amputation (Read code)	7L063-8	Slocum disarticulation of knee
Amputation (Read code)	7L063-9	Spittler disarticulation of knee
Amputation (Read code)	7L063-99	Knee disarticulation
Amputation (Read code)	7L064	Amputation below knee
Amputation (Read code)	7L064-1	Boyd amputation of leg below knee
Amputation (Read code)	7L064-2	Burgess amputation of leg below knee
Amputation (Read code)	7L064-3	Guyon amputation of leg below knee
Amputation (Read code)	7L064-98	Below knee amputation
Amputation (Read code)	7L064-99	Supramalleolar ankle amputat.
Amputation (Read code)	7L06y	Other specified amputation of leg
Amputation (Read code)	7L06z	Amputation of leg NOS
Amputation (Read code)	7L06z-99	Amputation lower limb NOS
Amputation (Read code)	7L07	Amputation of foot
Amputation (Read code)	7L070	Amputation through ankle
Amputation (Read code)	7L070-1	Pirogoff amputation of foot through ankle
Amputation (Read code)	7L070-2	Syme amputation of foot through ankle
Amputation (Read code)	7L070-99	Disarticulation of foot

Amputation (Read code)	7L071	Disarticulation of tarsal bones
Amputation (Read code)	7L071-1	Boyd amputation of hindfoot
Amputation (Read code)	7L071-99	Amputation foot: mid-tarsal
Amputation (Read code)	7L072	Disarticulation tarsometatarsal joint
Amputation (Read code)	7L072-1	Lisfranc tarsometatarsal amputation
Amputation (Read code)	7L072-2	Tarsometatarsal amputation
Amputation (Read code)	7L072-3	Disarticulation of metatarsal bones
Amputation (Read code)	7L072-99	Amputation foot:tarsal-metatar
Amputation (Read code)	7L073	Amputation through metatarsal bones
Amputation (Read code)	7L073-1	Chopart midtarsal amputation
Amputation (Read code)	7L073-2	Ray transmetatarsal amputation
Amputation (Read code)	7L07y	Other specified amputation of foot
Amputation (Read code)	7L07z	Amputation of foot NOS
Amputation (Read code)	7L07z-1	Hey amputation of foot
Amputation (Read code)	7L08	Amputation of toe
Amputation (Read code)	7L080	Amputation hallux
Amputation (Read code)	7L080-1	Amputation great toe
Amputation (Read code)	7L081	Amputation of phalanx of toe
Amputation (Read code)	7L082	Proximal hemiphalangectomy of toe
Amputation (Read code)	7L083	Amputation lesser toe
Amputation (Read code)	7L084	Terminalisation of hallux
Amputation (Read code)	7L085	Terminalisation of lesser toe
Amputation (Read code)	7L08y	Other specified amputation of toe
Amputation (Read code)	7L08z	Amputation of toe NOS
Amputation (Read code)	7L08z-1	Disarticulation of toe NOS
Amputation (OPCS-4 code)	X093	X093 - Amputation of leg above knee
Amputation (OPCS-4 code)	X094	X094 - Amputation of leg through knee
Amputation (OPCS-4 code)	X095	X095 - Amputation of leg below knee
Amputation (OPCS-4 code)	X098	X098 - Other specified amputation of leg
Amputation (OPCS-4 code)	X099	X099 - Unspecified amputation of leg
Amputation (OPCS-4 code)	X10	X10 - Amputation of foot
Amputation (OPCS-4 code)	X101	X101 - Amputation of foot through ankle
Amputation (OPCS-4 code)	X104	X104 - Amputation through metatarsal bones
Amputation (OPCS-4 code)	X108	X108 - Other specified amputation of foot
Amputation (OPCS-4 code)	X109	X109 - Unspecified amputation of foot
Amputation (OPCS-4 code)	X11	X11 - Amputation of toe
Amputation (OPCS-4 code)	X111	X111 - Amputation of great toe
Amputation (OPCS-4 code)	X112	X112 - Amputation of phalanx of toe
Amputation (OPCS-4 code)	X118	X118 - Other specified amputation of toe
Amputation (OPCS-4 code)	X119	X119 - Unspecified amputation of toe

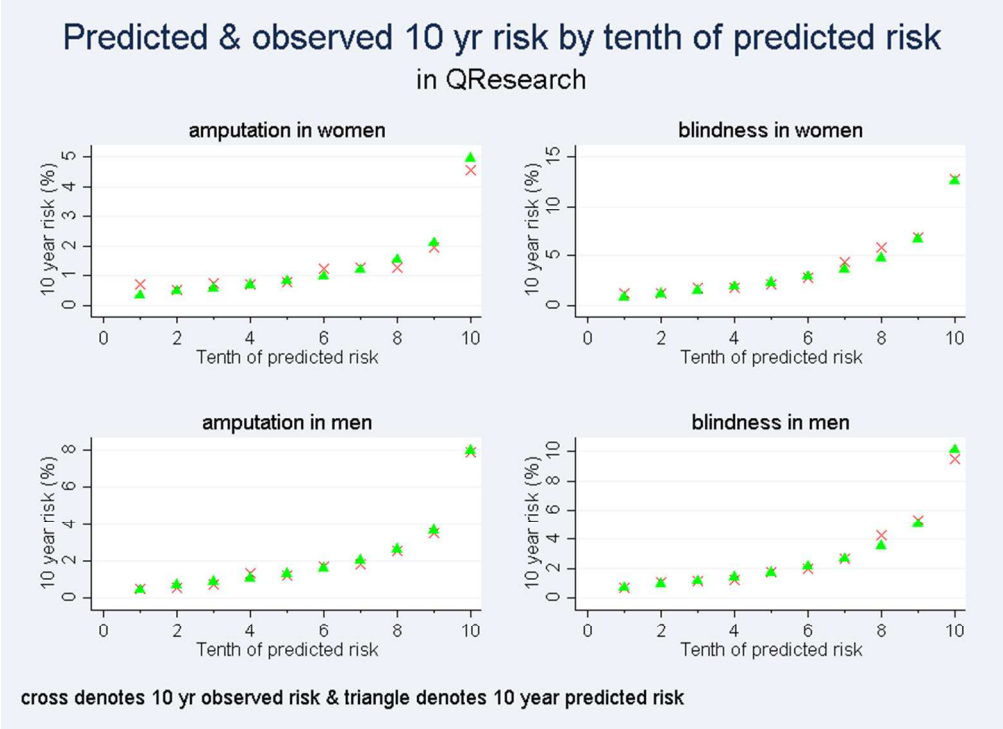
ocular)



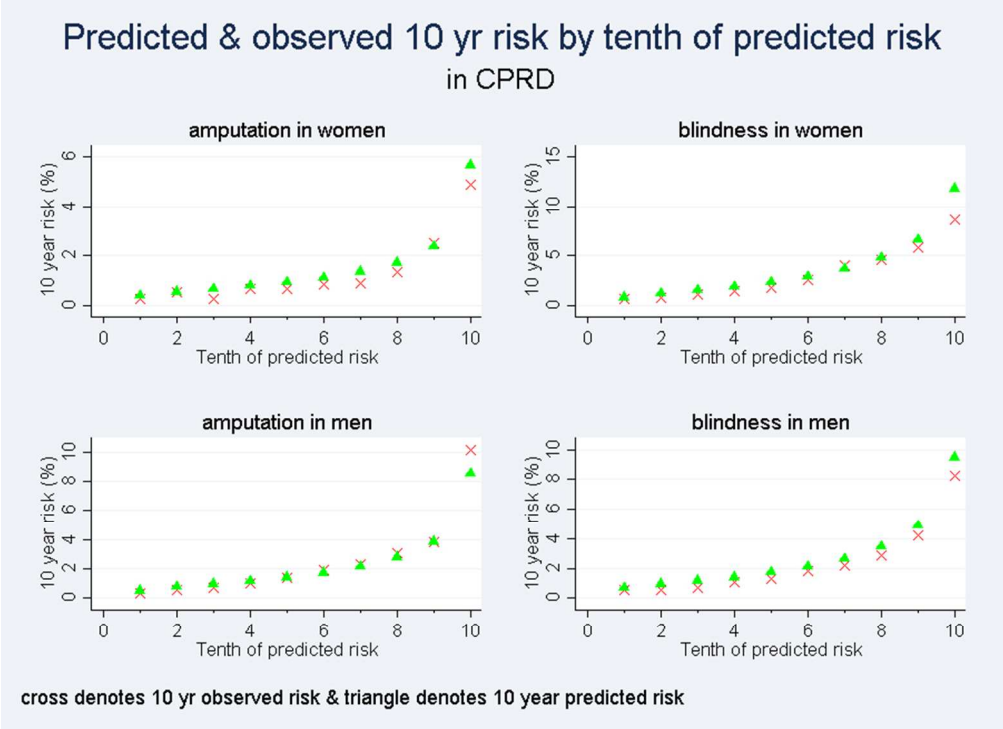
312x199mm (96 x 96 DPI)



309x198mm (96 x 96 DPI)



317x231mm (72 x 72 DPI)



317x231mm (72 x 72 DPI)