Optimisation of Health Services for Sight Threatening Diabetic Retinopathy in the UK

By Sajjad Haider

A thesis submitted to the University of Birmingham for the degree of Doctor of Philosophy

Applied Health Research
College of Medical and Dental Sciences
University of Birmingham
May 2020

Abstract

Aims: This research project aimed to investigate the future burden and risk stratification of diabetic retinopathy in the UK.

Methods: The disease burden was calculated through cross-sectional studies, and its future projections through a double exponential smoothing model using primary care Data. The prediction model was designed by a systematic review of the existing models, finalising the predictors list used therein to a shorter list using nominal group technique and evidence evaluation and then a prediction model development and validation in a retrospective cohort study.

Results: There is a trend of increasing prevalence of diabetes and diabetic retinopathy. This trend is highest and accelerating in Sight Threatening Diabetic Retinopathy in patients with type 1 Diabetes Mellitus. The forecast for 2030 is 1.6 million people with diabetic retinopathy and .65 million with Sight Threatening Retinopathy. The systematic review yielded 14 models and 78 predictors. A list of 19 candidate predictors was finalised. The new model has moderately good performance.

Conclusion: Disease burden estimation needs to be carried out periodically to capture changing trends. External validation, clinical benefit and a nomogram are also needed. Newer imaging modalities and artificial intelligence are likely to play a part in future prognostic model research.

Dedication

To my family, patients, and colleagues

Acknowledgements

I would like to begin by thanking my patients, first and foremost, for keeping my motivation focussed on positive and productive pursuits throughout my working life and now in my retirement.

Without the support, time and effort from my supervisors, Dr Krishnarajah Nirantharakumar, Dr David Moore and Dr Malcolm Price, I would not have been able to accomplish this research work and thesis. I would also like to thank Kym Snell from Keele University for her continuing support.

In addition, I would like to thank Anuradha Sabramanian, Rasiah Thayakaran, Nicola Adderley, Krishna Gokhale, Dawit Zemedikun, Deepiksana Keerthy, Kelvin Okoth and Joht Chandan based at the University of Birmingham for their help and collaboration on statistics, data science and read codes development. Thanks also to Salman Naveed Sadiq for help as a second reviewer of the systematic review and for validating the results of the evidence evaluation studies in Chapter 4. Thanks also to Harpreet Sihre, Eniya Lufumpa and Mohammad Tallouzi for help with the qualitative study design and conduct. Lastly, I am grateful to Tina Mcleod for providing editing support for formatting and language of this thesis.

Contributorship statement

The chapters of this thesis are entirely a product of my own work. However, the work has been supported by my supervisors, Dr Krishnarajah Nirantharakumar, David Moore and Malcolm Price, during the concept development phase, in study designs, data analysis and interpretation, as well as write up of the studies and the final thesis preparation.

The thesis is split into various manuscripts, each one addressing different work package and areas of expertise. Further support and advice came from other co-authors, namely Konstantinos Toulis as a diabetologist, Kym Snell as a statistician and a modelling expert, and Harpreet Sihre, Eniya Lufumpa and Mohammad Tallouzi for their expertise in qualitative studies.

My Motivation

In January 2016, after about 38 years in clinical Ophthalmology (26 years in the UK), I was getting ready for retirement. During the last 2 years of this period I had a leadership role as clinical director for head and neck specialities and operating theatres. This managerial / leadership role opened my eyes to the vast room for further improvement in the way we ran the services, and the unmet need for evidence. I was interested in contributing to an optimisation process with fresh evidence base. I did an MSc in a related field (Health Technology Assessment) back in 2006, but in busy clinical years, never really had time to use these skills. I wanted to use my retirement for furthering my knowledge and skills in this area of expertise. By 2016, I had also been running a charitable eye hospital and the first optometry graduate course in Pakistan, in my home town in Pakistan, for almost 20 years (starting from scratch in 1997). The graduate course that I founded produces over 300 graduates per year in that country. Its worst need was, and still is, a severe lack of postgraduate teaching faculty. I am in negotiation with the University of Health Sciences Lahore for an M Phil course to assure further sustainability of this course. The availability of a dedicated PhD teacher will help my application and I am happy to volunteer.

My retirement had to include these interests. I retired in May 2016 from clinical Ophthalmology to undertake this present piece of work. I saw this PhD project as an opportunity to learn how to build an evidence base behind service optimisation in hospital eye services, as well as to help build the capacity of the optometry course in Pakistan. The NHS has started an initiative in risk stratification of patients and I feel I might be able to help in this venture as well.

Table of Contents

Contents

List of tables

List of figures

List of abbreviations

BMI	Body Mass Index
CPRD	Clinical Practice Research Datalink
DESP	Diabetic Eye Screening Programme
DMO	Diabetic Macular Oedema
DR	Diabetic Retinopathy
E/O ratio	Expected / Observed Ratio
eGFR	estimated Glomerular Filtration Rate
HbA1c	Glycated Haemoglobin
HES	Hospital Eye Services
IMRD	IQVIA Medical Research Data
NGT	Nominal Group Technique
NICE	National Institute of Clinical Excellence
PDR	Proliferative Diabetic Retinopathy
PHE	Public Health England
RCOphth	Royal College of Ophthalmologists
STR	Sight threatening retinopathy
T1DM	Type 1 Diabetes Mellitus
T2DM	Type 2 Diabetes Mellitus
THIN	The Health Improvement Network
UK	United Kingdom
VEGF	Vascular endothelial growth factor

There is an additional list of abbreviations (for the systematic review) in appendix 14.

Thesis format

This thesis is presented in accordance with the Alternative Format Thesis Guidelines https://intranet.birmingham.ac.uk/as/studentservices/graduateschool/documents/public/rsa/alternative-format-thesis-guidelines.pdf

The pages of the publications / manuscripts for submission (Chapters 2 to 5) themselves may not always be included in the pagination sequence of the submission. Referencing and numbering of tables and figures will be self-contained within each chapter to maintain numerical clarity. The incorporation of publications / manuscripts for submission chapters will inevitably lead to some duplication, since each publication / manuscript for submission chapter will have self-contained components that may overlap with parts of other sections of the thesis.

This research thesis focuses on the disease burden and risk of diabetic retinopathy (DR) progression.

Publications and presentations related to the thesis / research

Publication:

Haider S, Sadiq SN, Moore D, Price MJ, Nirantharakumar K.
Prognostic prediction models for diabetic retinopathy progression: a systematic review. Eye. 2019;33(5):702-13.

Manuscripts from the thesis under review for publication:

1. **Sajjad Haider**, Rasiah Thayakaran, Anuradhaa Subramanian, Konstantinos A Toulis, David Moore, Malcolm James Price, Krishnarajah Nirantharakumar

   Disease Burden of Diabetes, Diabetic Retinopathy and their Future Projection

2. **Sajjad Haider**, Salman Naveed Sadiq, Harpreet Sihre, Eniya Lufumpa, Mohammad Tallouzi, David Moore, Krishnarajah Nirantharakumar, Malcolm James Price

   Predictors for Diabetic Retinopathy Progression – Findings from Nominal Group Technique and Evidence Review Studies

3. **Sajjad Haider**, Anuradhaa Subramanian, Nicola J. Adderley, Kym Snell, David Moore, Malcolm James Price, Krishnarajah Nirantharakumar

   Diabetic Retinopathy Progression to Treatment or Vision Failure: Development and Internal Validation of a Multivariable Prediction Model

Presentations that I have made related to the thesis:

1. **Sajjad Haider,** David Moore, Krishnarajah Nirantharakumar, Malcolm James Price
   Oral 05/09/2018 at Durham and Darlington NHS Trust

2. **Sajjad Haider,** David Moore, Krishnarajah Nirantharakumar, Malcolm James Price
   Oral 05/09/2018 at Durham and Darlington NHS Trust

3. **Sajjad Haider**, Harpreet Sihre, Eniya Lufumpa, David Moore, Krishnarajah Nirantharakumar, Malcolm James Price

Oral 10/10/2018 at Surrey and Sussex NHS Healthcare Trust

4. **Sajjad Haider,** Mohammad Tallouzi, David Moore, Krishnarajah Niranthakumar, Malcolm James Price
   Oral 25/10/2018 at Sandwell and West Birmingham NHS Trust

5. **Sajjad Haider**, Salman Naveed Sadiq, Harpreet Sihre, Eniya Lufumpa, Mohammad Tallouzi. David Moore, Krishnarajah Nirantharakumar, Malcolm James Price
   Risk (Prognostic) Factors for Diabetic Retinopathy Progression: Nominal Group Technique and a Literature Review
   Poster at 11/04/2019 MDS Festival of Graduate Research 2019

1 Chapter 1: Background

1.1 Introduction

This PhD thesis aims to produce evidence for the optimisation of diabetic retinopathy management services in the UK through disease burden and individual risk stratification studies.

This chapter provides the background and rationale for the research conducted. The chapter defines the objectives, with later chapters giving more extensive and precise details in their introduction sections. The chapter is divided into various sections giving a general introduction to diabetic retinopathy disease, its prevention and treatment, organisation of services, and the need, justification, aims and objectives of the research projects within this thesis.

1.2 Diabetes mellitus

Diabetes mellitus is a group of metabolic diseases in which patients either do not produce enough, or do not respond to, insulin, the hormone that controls the amount of sugar in the blood. In type 1 diabetes mellitus (T1DM), destruction of pancreatic beta islet cells that produce insulin leads to insulin deficiency (1). In type 2 diabetes mellitus (T2DM) there is an increasing lack of response to insulin (2). This reduction in the formation of, or response to, insulin results in persistently raised blood sugar levels, leading to macrovascular damage (coronary artery disease, peripheral arterial disease, and stroke) or microvascular damage to blood vessels in the feet, nerves, and organs such as the kidneys (3). In the back of the eye, this manifests as diabetic retinopathy (DR) (3).

1.3 Diabetic retinopathy pathogenesis and progression

DR is a major cause of blindness due to damage to the retina, which is the light sensitive layer at the back of eye. Diabetes causes damage to blood vessels through ischaemia, causing the death of retinal microvascular pericytes and endothelial cells. It also causes thickening of the basement membrane, leading to increased permeability of vessel walls (3) bringing about the following changes:

A. Increased permeability leads to the leakage of blood, fluid, lipids and proteins from vessels in the retinal layers of the eye. This results in a condition known as diabetic macular oedema (DMO), often associated with hard exudates made of lipid and protein. This localised swelling and damage to the macula / fovea (centre of the retina) can cause central sight loss (4).

B. Further ischaemic injury and cell death causes the release of vaso-proliferative factors, such as vascular endothelial growth factor (VEGF), which stimulate the development of new vessels – a process known as neovascularization. These new vessels in the retina can cause bleeding into the vitreous, (a transparent gel in the centre of the eye), impeding light reaching the retina, causing loss of vision for a prolonged period. This can also result in retinal detachment, causing permanent loss of vision. New vessels can also grow into the iris causing severe glaucoma (5), a condition with high intraocular pressure and pain.

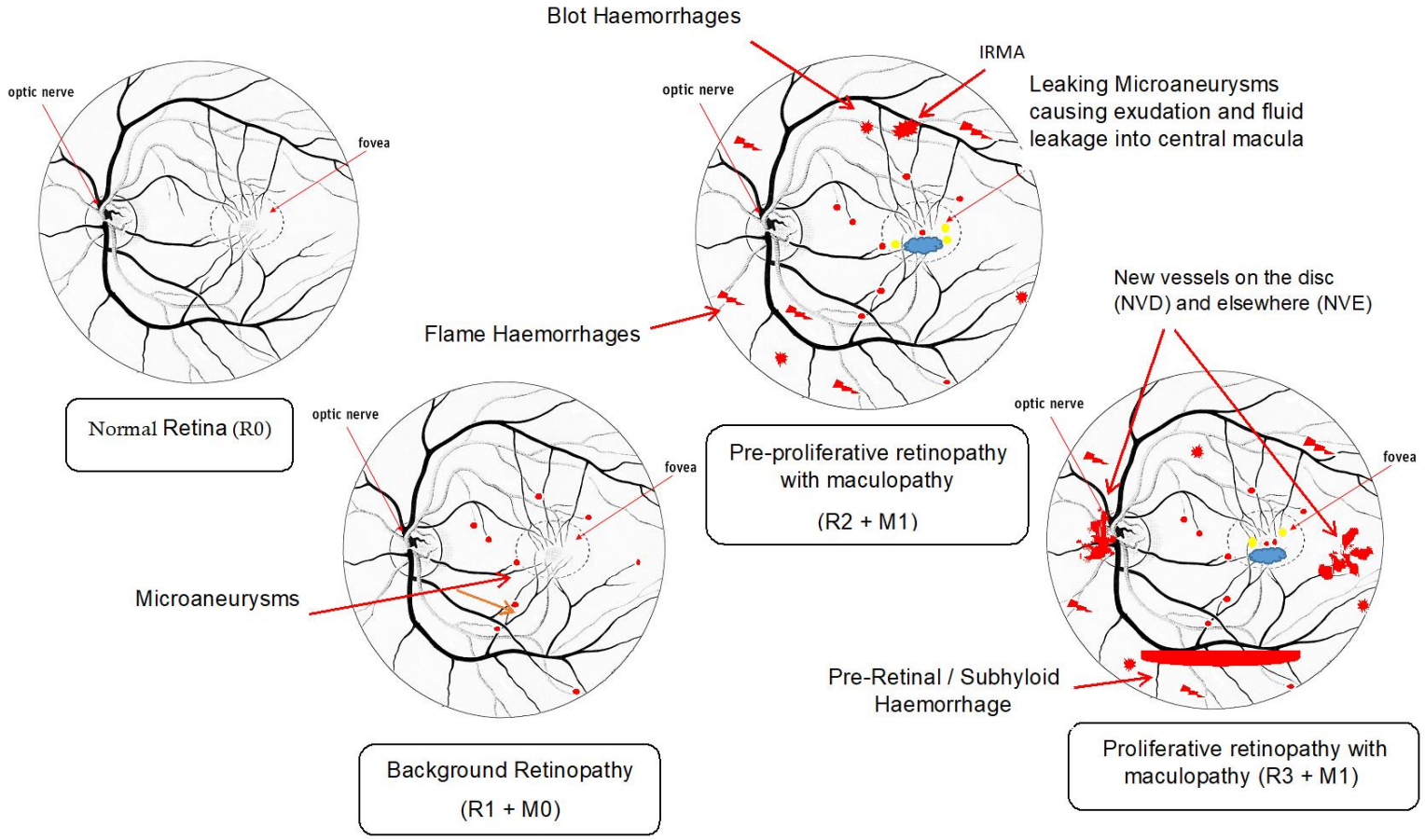


Figure 1: Diagrammatic representation of Diabetic Retinopathy pathogenesis and progression process. DR (Diabetic Retinopathy), IRMA (Intra-retinal Microvascular Anomalies), RO (No DR), R1 (Background DR), R2 (Pre-proliferative DR), R3 (Proliferative DR), M1 (Macuolopathy), M0 (No Maculopathy). Modified from "Blind Spot", © Exploratorium, www.exploratorium.edu

1.4 Classification of diabetic retinopathy

DR is classified according to the level of severity and location in the eye. There are a number of classification schemes as outlined in Table 1. In the UK, the classification adopted by the Royal College of Ophthalmologists (RCOphth) (4) is based on visible clinical signs, down a worsening scale of severity. There are alternate schemes in the college guidance (4), with small differences in the coverage of maculopathy, descriptions of the fundus view, treatments, categorization at slightly different levels, and details of retinopathy signs. In the UK, the Diabetic Eye Screening Programme (DESP) classification may be the most commonly used classification, with R denoting retinopathy and M denoting maculopathy. I have therefore used this classification in my thesis, using the categories below, as shown in brackets.

- No DR - (R0) describes no signs of retinopathy in a patient with diabetes.
- Background DR - (R1) with micro-aneurysm as the only sign.
- Pre proliferative DR - (R2) exhibiting in addition haemorrhages, cotton wool spots, venous beading and intra-retinal microvascular abnormalities (IRMA).
- Proliferative (PDR) - (R3) with new vessels on the disc or elsewhere in the retina, pre retinal / vitreous haemorrhage or consequent traction retinal detachment (advanced stage).
- No Maculopathy - (M0) describes no signs of maculopathy in a patient with diabetes.
- Diabetic Maculopathy - (M1) presenting with thickening of the macula due to fluid within retina or exudation.

Table 1: Classification of diabetic retinopathy – modified from The Royal College of Ophthalmologists (RCOphth) Guidelines (4)

Early Treatment Diabetic Retinopathy Study, 1991 (6) (7)	**EURODIAB coding system 1995 (8)**	**AAO International DR and diabetic macular oedema disease Severity scale, 2003 (9)**	**National Screening Committee UK, 2003 (10)**	**Scottish Diabetic Retinopathy Grading Scheme 2007 v1.1 (11-13)**	**RCOphth 2012 (4)**
10 none	Level 0	No apparent retinopathy	R0 none	R0 none	None
20 microaneurysms only	Level 1	Mild NPDR	R1 mild/ background	R1 mild background	Low risk
35 mild NPDR	Level 2	Moderate NPDR	[a]	[a]	[a]
43 moderate NPDR	Level 3	[a]	R2 pre proliferative	R2 – moderate BDR - observable	High risk
47 Moderately severe NPDR	[a]	[a]	[a]	[a]	[a]
53A-D severe NPDR	[a]	Severe NPDR	[a]	R3 severe BDR - referable	[a]

53E very severe NPDR					
61 mild PDR 65 Moderate PDR 71, 75 High risk PDR 81, 85 Advanced PDR	Level 5	PDR	R3 proliferative	R4 proliferative	PDR
[a]	[b] Level 4 - Scars of photocoagulation in any field.	[a]	[b] R3S. Stable treated PDR	[a]	[a]
CSME	[c] Maculopathy not gradable	[c] Diabetic Macular Oedema Apparently absent or present (Retinal thickening or exudates on the posterior pole) – • Mild DMO – Away from central macula • Moderate DMO Approaching central macula	[c] Maculopathy present (M1)	[c] • No maculopathy (M0) • Maculopathy present > 1 DD from foveal centre - observable (M1)	[a] • Focal oedema • Diffuse oedema • Ischaemic or

		• Severe DMO - involving central macula		• Maculopathy present < 1 DD from foveal centre - referable (M2)	• Mixed
[a]	[a]	[a]	[a]	[d] R5 – Enucleated eye	[a]
[a]	[a]	[a]	[a]	[e] R6 – Inadequate visibility	[a]

EDTRS - Early Treatment of Diabetic Retinopathy Study, AAO - American Academy of Ophthalmology

CSME - Clinically Significant Macular Oedema

[a] Empty cells do not correspond, therefore are not applicable, [b] Treated retinopathy, [c] Maculopathy / Diabetic Macular Oedema Disease Severity Scale, [d] signs could not be visualized

1.5 Vision loss - Low Vision and Blindness

World Health Organization (WHO) defines low vision as a visual acuity of less than 6/18, but

equal to or better than 3/60 in the better eye with best possible correction (14). The WHO standards (15) are as follows

- Moderate visual Impairment is visual acuity worse than 6/18 on the Snellen's Chart Severe visual Impairment is visual acuity worse than 6/60
- Blindness is visual acuity worse than 3/60

In the United Kingdom, patients with failing sight are certified as blind or partially sighted by their consultant ophthalmologist, which triggers referral to local social services at which point they are registered. The patients' anonymised information is used for epidemiological studies. While there have been reports of underuse of certification as blind or visually impairment (16), more recent literature indicates that following the introduction of a comprehensive retinal screening service, the incidence of visual impairment certification secondary to diabetic retinopathy appears to be decreasing (17, 18).

1.6 Global burden of diabetes and diabetic retinopathy

The diabetes burden is identified as a global epidemic (19, 20) and its rise impacts on the burden of DR. In 2015, 415 million people in the world were reported to have diabetes, and this number is expected to rise to 642 million by 2040 (21). However, the estimates for the global burden projections for 2030 vary from 439 million to 552 million (22, 23). The global diabetes prevalence is expected to grow to 7.7% by 2030 (23). World population is forecasted to grow to 8.5 billion in 2030. That brings population with diabetes forecast to 654.5 millions worldwide. Yau et al quoted one third of population with diabetes have some form of diabetic retinopathy and about one in ten have a more severe condition, sight-threatening retinopathy (STR) (24). Assuming these prevalence rates hold true in 2030, total global burden of DR will be 222.5 millions and STR requiring closer observation and treatment to be upwards of 65.5 millions

In 2015, the global burden of moderate to severe vision impairment due to diabetic retinopathy was 2.5 million world wide, is expected to rise to 3.2 million in 2020, and varies from region to region (25). The DR is one of the leading causes of blindness and visual impairment among the working age adults (26). In the developed countries, a recent decline in the incidence of blindness due to DR has been observed (27). This has been attributed to improved control of risk factors like blood sugar and blood pressure control as well as screening programmes for the early detection and then prompt treatment of severe DR (28, 29).

1.7 Incidence and prevalence of diabetes and diabetic retinopathy in the UK

An earlier forecast of UK diabetes burden significantly underestimated the figures (30). A study based on IQVIA Medical Research Data (IMRD), previously called The Health Improvement Network (THIN) noted the diabetes prevalence increase between 1996 and 2005 from 2.8% to 4.3% (31). Based on a Clinical Practice Research Datalink (CPRD) a study comparing 2004 with 2014, the prevalence of DR was higher in T1DM at 48.4% and lower in T2DM population at 28.3%. These trends for DR prevalence remained stable and STR increased (32). A study in 2019 failed to forecast the UK wide disease burden of diabetic retinopathy due to the absence of reliable data (33). The expected estimates of the rise in diabetic retinopathy vary widely from 20 % to 80 % in the next 20 years (34), creating difficulty in planning for future. The commonest cause of loss of vision among population with diabetes in the UK is centre involving macular oedema, making 10% of the DR patients under hospital care (35).

In the UK, 3.4% of people diagnosed with diabetes and attending diabetic eye screening were visually impaired and 0.39% were severely visually impaired in 2009. Blacks and South Asians, people with higher age, type 1 diabetes and those living in Yorkshire had a higher prevalence of visual impairment (36). Risk of first presentation with severe Diabetic retinopathy and vision loss was also higher in patients with deprivation (37). Disease burden of DR and vision loss is the subject under study in chapter 2.

1.8 Management of diabetic retinopathy

1.8.1 The services for diabetic retinopathy

These services are organized into Diabetic eye screening services for low risk patients, secondary care (mostly district hospitals) for high risk patients and tertiary care centres providing vitreoretinal surgical services. DR patients are referred into the hospital eye service at a stage when they develop clinical STR.

Low risk patient services - DESP

The diabetic retinopathy screening programme introduced in England and Wales in 2003, achieved population coverage across the whole of England in 2008 and has helped reduce the prevalence of visual impairment and blindness (10). The DESP sends out a letter of invitation to every diabetic over 12 years of age, except for those already under ophthalmology and the terminally ill, inviting them for a diabetic eye screening appointment once a year. Suitably trained staff deliver the programme which has an internal and external quality assurance mechanism. The screen-positive patients and those with poor-quality images are referred into ophthalmology for assessment and treatment if necessary. This photographic screening uses two-field mydriatic digital images and the uptake of the programme in 2015–16 was 82.8%. The sensitivity of the screening test is 87.8%, the specificity is 86.1% and poor-quality image rate is 3.7% (10). The screening could be moved to 2 yearly intervals for low risk of sight loss if no retinopathy is found on two consecutive screens (10). It is also likely that automated retinal image analysis can take over some of the roles of manual graders (38).

Higher risk patient services – Hospital Eye Services and digital surveillance

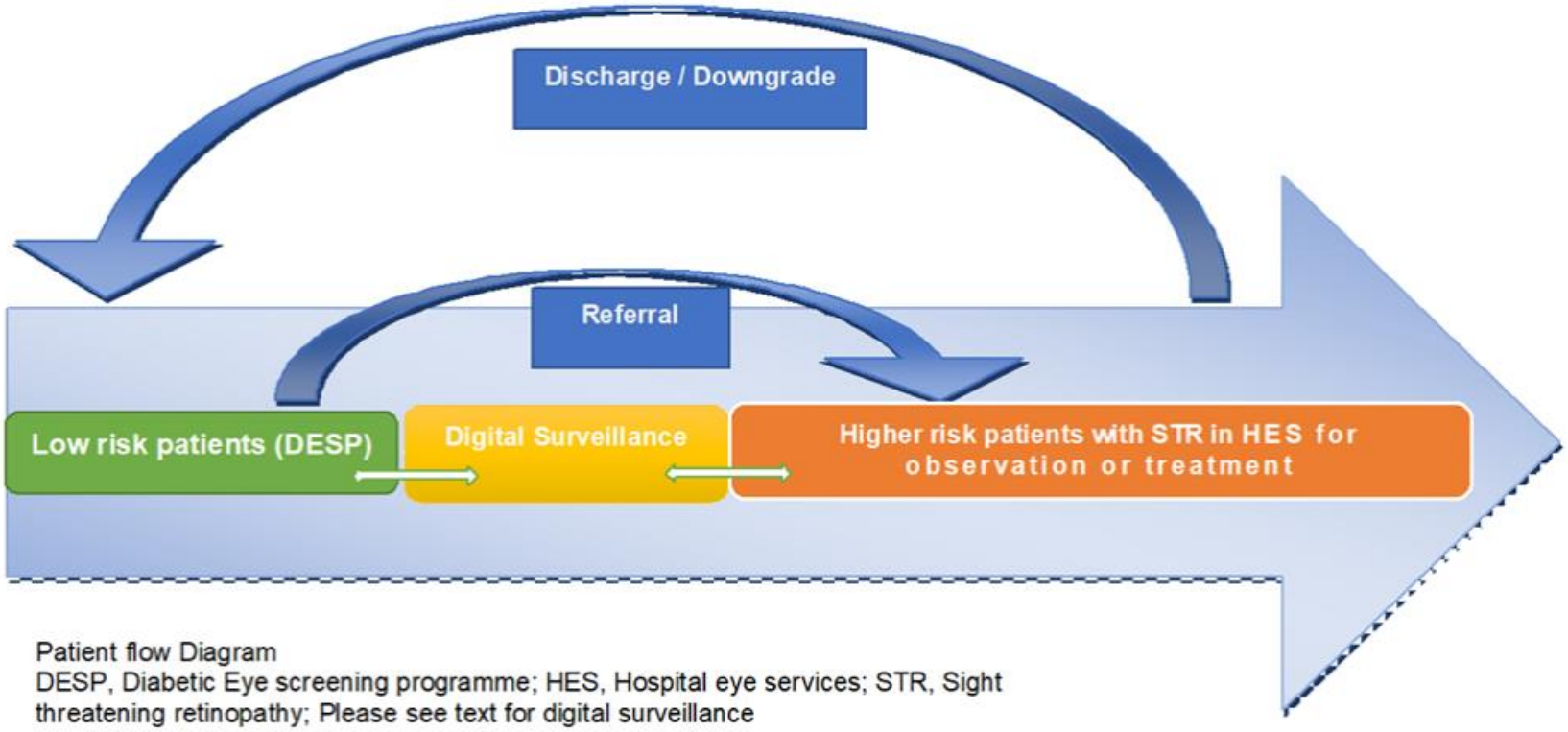
The UK diabetic retinopathy services are organized into Diabetic eye screening services in the community by DESP and Hospital Eye Services (HES) if and when they develop clinical signs of STR. Since the beginning of 2015 the digital surveillance was introduced as part of the patient pathway as an additional option for patients not likely to be treated but needing closer observation (39, 40). Under 25% of the referrals with M1 patients were found to need treatment (41). It is possible that optical coherence tomogram (OCT) may improve this and that already seems to be the direction of travel in the digital surveillance (39).

1.8.2 Early detection to prompt treatment of high risk patients

Recently published pathway standards (42) is an update document on screening standards with an aim to ensure early detection, monitoring and treatment of DR to prevent vision loss. Treatments are carried out in HES which are subject to guidance in pathway standards which gives a list of objectives to be achieved by the collaboration between local DESP and HES. This includes the expected percentage number of people with STR seen within specific periods of time seen and if needed, treated. For example, patients at pre-proliferative stage (R2) and maculopathy (M1), are referred into hospitals or digital surveillance clinic from DESP and need to be seen within 13/52. Proliferative diabetic retinopathy (R3) patients are referred into hospital from the screening programme and need to be seen and treated within 6/52. Needless to say, that these are relatively low risk R3 patients without vitreous haemorrhage

or vision loss at this stage. Patients with ungradable photographs are referred into slit lamp bio-microscopy clinics and seen within 13/52

After referral, the decision to treat is based on treatment guidelines below. The decision on the length of follow up intervals is based on the guidelines (43-45) and an impression of clinical risk based on ocular clinical features seen. Recently, there has been an initiative by RCOphth in the form of an Ophthalmic services guidance (46) on risk stratification of patients based on National Institute of Clinical Excellence and RCOphth guidance.



Patient flow Diagram
DESP, Diabetic Eye screening programme; HES, Hospital eye services; STR, Sight threatening retinopathy; Please see text for digital surveillance

Figure 2: Patient flow diagram

The international council of Ophthalmology (45) gives an explicit follow up schedule which seems to be in line with above giuidanace but will not be binding on care providers. ROCPHTH guidance (4) summarises evidence on DR treatment including systemic DM, treatment, but does not offer specific follow up advice in the care pathway in this respect.

There have been recent successes in the management of diabetic retinopathy in quite a few fields. Early Treatment of Diabetic Retinopathy Study (EDTRS) (47) initiated laser treatment and lately the diabetic retinopathy screening services have helped early detection. The new National Institute of Clinical Excellence (NICE) guidance facilitated anti Vascular Endothelial Growth Factor (anti VEGF) treatment for diabetic macular oedema have helped reduce blindness as evidenced by the fact that severe visual loss was reduced to one third when comparing figures between 1985 and 2013 (29, 48).

1.8.3 Prevention of diabetic retinopathy progression

Prevention strategies in DR involve treatment of disease modifying risk factors and early detection, so the treatment can be offered in time to prevent loss of vision. The risk of DR progression and visual failure can be significantly reduced by stricter glycaemia control (49). Evidence is conflicting on direct effect of the newer agents like in DR progression. On the whole their indications are more advanced diabetes with vascular complications. Wang and DeFronzo reported no association with an increased risk of DR. Similar results were observed for DPP4i and GLP1RA (50, 51). But Marso reported Semaglutide was associated with a worsening of diabetic retinopathy (52). Same study gave the reason for the worsening diabetic retinopathy

in SUSTAIN-6 of rapid reduction in A1c (52). Longer term studies are awaited on the subject.

A recent Royal college of Ophthalmologists (RCOphth) document mentioned a pilot survey finding of 50% patients in need of diabetes treatment regimen change and 10% needing referral to endocrinologists (34). The recommendation was to use the opportunity created by the patient's presence in the eye clinic to be capitalised on, for better control of blood sugar, blood pressure and lipids to prevent progression and also reduce the appopintment burden on patients. This requires a higher degree of integration of different disciplines caring for population with diabetes..

1.8.4 Treatment of diabetic retinopathy

The treatment options range from systemic treatment of diabetes for the control of modifiable risk factors for its progression to ocular treatment with laser (focal / grid laser for maculopathy and pan retinal photocoagulation for PDR), intravitreal injection treatment with Anti Vascular Endothelial Growth Factor (Anti VEGF) for diabetic macular oedema (DMO) but also for some PDR patients, steroid injections for DMO, and surgery for traction retinal detachments (53, 54).

1.8.5 Evidence based treatment guidelines

There is no local treatment required for background diabetic retinopathy (R1) but it is possible to significantly reduce and delay the risk of further progression by means of risk factor modification - intensive control of glycaemia (55), control of blood pressure (56) and by using Fenofibrate (lipid control) (57). Patients are kept within DESP for observation, but may be referred in for conditions other than DR like cataract etc.

In pre-proliferative stage (R2) loss of vision can result from accompanying macular oedema or rapid progression to proliferative stage. This STR stage does not generally require active treatment and patients are monitored for the development of next stage before starting treatment. However, there are exceptional cases where treatment may be considered for a patient based reason (58). Prophylactic laser treatment to vulnerable areas of non-perfusion has been shown to prevent the progression of R2 to R3 stage in a Japanese study (58). However, in routine practice, patients with this disease stage are not offered any treatment but closely monitored (4 to 6 monthly) for further progression except under high risk patient circumstances (4, 45).

Temporary loss of vision can result from bleeding and permanent loss of vision through traction retinal detachment or ischaemia of macula in PDR (R3). For just under 40 years, since the DRS study (59), laser has been the only treatment for this stage. While this treatment reduced the risk of vision loss by 50%, it does cause minimal vision loss of one line on the chart in about 5% of patients and reduces peripheral visual field in about 3% of patients. Anti VEGF agent injections have been proven to be as effective as laser panretinal photocoagulation (60).

Maculopathy (M1) classed as STR, is referred into hospitals from the Diabetic Eye screening programme. For over 30 years since EDTRS, laser has been the main treatment for diabetic macular oedema (M1). The vision improved from baseline in 57% and worsened in 28% of patients at 3 years. (47). The treatment is only offered in maculopathy in case of centre involving macular oedema, clinically significant macular oedema and exudates threatening central eye sight. VEGF is a key chemical factor in new vessels formation (R3) and increased vascular permeability leading to

diabetic macular oedema. Anti VEGF agents are being used in DMO treatment (61). The three Anti VEGF agents used are Aflibercept, Ranibizumab and Bevacizumab are the relatively new treatment modality and have for a decade been recommended as an option. They showed an improvement of a mean of 17.1, 13.6 and 12.1 letters respectively at two years follow up (62). Steroids in the form of Triamcinolone solution, Dexamethasone and Fluocinolone acetonide intravitreal implants are also injected frequently. The latter two are slow release, long acting preparations. They are almost entirely restricted to patients who have had cataract surgery to prevent the significant risk of secondary glaucoma. All these treatments are recommended by National Institute of Health and Care Excellence (NICE).

The intravitreal agents (mainly Anti VEGF but also steroids), since introduced, have quickly been adopted. They have changed the goal posts of treatment from maintenance of vision to improvement. However, one third of the patients still don't improve their eye sight (63) and new treatment strategies are being sought to overcome that problem including combining steroids with anti-VEGF agents, switching anti-VEGF agents, increasing the frequency of injections and using surgery (53).

The recommendation is for Anti VEGF to be continued till DMO continues to improve, to discontinue when oedema stabilizes, and restart when the condition worsens. Intravitreal anti-VEGF injections have now replaced laser as standard treatment for DMO, although, laser photocoagulation may still be used when a complete response is not seen (53).

1.8.6 Barriers to services

Non attendance of clinic appointments is a barrier to the services. Lack of awareness in patients about diabetes causing blindness, fear of laser treatment and guilt associated with poor control were common causes of non attendance of clinic appointments (64). The providers were not fully aware of the difficulties faced by patients to get time off work for clinic attendance (65). Only one-third of African Americans adhered to follow-up appointments, even if cost and accessibility were minimized. Provision of more complete information / counselling about DR, and making eye clinic attendance more convenient for patients were thought to help (65, 66). Frequent non-attendance of clinic appointments is associated with increased risk of developing sight threatening diabetic retinopathy (67).

1.9 Diabetic retinopathy progression and risk factors

A recent review reported (mostly Asian studies) annual progression estimates ranging between 3.4% to 12.3% (68). The progression rate was higher in patients with mild disease at baseline compared with those with no disease (68). An earlier systematic review of rates of progression in diabetic retinopathy (29) concluded that rates of progression to PDR and severe visual loss had reduced as compared to 1985. This was thought to be due to significant improvements in the medical management of diabetes, early detection and management of diabetic retinopathy. Diabetic retinopathy and maculopathy were the fourth commonest cause of severe sight Impairment registration, ranked significantly lower in 2013 as compared to in 2007/2008 (48).

The risk factors for DR progression (69) are at the core of prevention strategies and prediction of a clinical outcome. The risk factor needs to be present in subjects (individual, family, a population) before the outcome of interest appears. The Royal College of Ophthalmologists (RCOphth) diabetic retinopathy guidelines mention as non- modifiable risk factors, duration of diabetes, age of onset of diabetes, genetic factors, gender and puberty (4). The modifiable risk factors they summarised are control of HBA1c, blood pressure, body mass index (BMI), vitamin D, smoking, lipid levels and pregnancy. They also included additional factors of renal impairment and carotid artery occlusion. The modification of these risk factors is an important intervention to prevent the incidence and progression of DR and its effects on vision.

Some of these risk factors help predict prognosis. Such a prognostic factor is any measure used to predict in people with a certain health condition like diabetes, an outcome of interest, such as treatment or vision failure. This will be dealt with in detail in chapter 3,4,5.

1.10 Justification for research

Despite all the above-mentioned improvements within the UK, including the falling rates of blindness due to better diabetes control, better detection services and more effective treatments, there is an ever increasing number of patients to be screened and treated. With the high risk patients referral into hospitals to be followed up and treated, a serious mismatch between capacity and demand has occurred. New successful treatments have also added to the demand on eye clinics, accounting for 30% increase in eye clinic attendances over the last ten years (70).

During the decade before 2017-18, there was 80% growth in hospital outpatient attendances in NHS. 6 % of all outpatient attendances in 2017-18 were for eye care, highest of any other single speciality (33, 71). The injection procedures went up by 215% between 2011 to 2015. In 2014/2015, there were an average of 714 episodes per 100,000 population. There is no evidence that this trend will go down (72), putting HES under severe pressure. A recent analysis of UK hospitals' activity showed Ophthalmology being the busiest outpatient specialty in 2018 - 2019 (73). Future projections of this disease and treatment burden are needed to estimate the changing need for the services. Additionally, approximately 50-70% of referrals are false positives and are either returned back to screening services or observed in the hospital eye services (34, 41). Various service modifications are being considered, such as digital surveillance, using optical coherence tomography (OCT) (39), and virtual clinics (34) to reduce the numbers of these false positives.

The unmet need for risk stratification of the non-proliferative retinopathy was stressed in a recent literature review (74). Recent evidence (75-77) suggests that monitoring patients at intervals based on individual risks could reduce the number of screening episodes needed as well as people becoming screen positive before the allocated screening date. The number of appointments needed may be reduced significantly, thus reducing costs while still keeping patients safe. A similar approach may also help hospital eye services decide on the treatment modality chosen, the dosing regimens and follow up frequency to reduce demand and capacity mismatch. So, risk stratification of DR patients could be one mechanism for redesigning the services to seek further optimization by specifically redirecting resources to higher risk patients. This can be done by risk groupings based on a cut-off to define categories of this

score. A new individual can be assigned to a particular group based on their individual score. However, in this case, the individuals within the same group all get managed within the same group uniformly, and the difference between their individual risk scores are not considered. On the other hand, an individual prediction model can more precisely predict an individual's survival probability or outcome risk and is therefore preferable (78).

The main difficulties are increasing demand and lack of capacity. The solutions are limited mainly due to lack of convincing evidence. Understanding disease burden and individual risk stratification may be the way forward to deal with these problems.

1.10.1 Aims

The research aim in this thesis was to investigate the ways in which the services can be optimised. The sub-question were to 1) determine the size of the demand on the hospital diabetic retinopathy treatment services for the foreseeable future and 2) to design an individual risk stratification model, with diabetic retinopathy progression to a stage requiring treatment or vision failure as the outcome. This will help prioritize services within the pool of high-risk diabetic retinopathy patients under care of hospital eye services.

1.10.2 Objectives

The objectives of this PhD work are as follows:

1. To quantify and project the burden of diabetes mellitus and diabetic retinopathy in 2017 and beyond. This work is presented in chapter 2.
2. To develop an individual risk stratification model to differentiate high risk patients through following steps

a) To critically review all the existing prognostic models in the area of diabetic retinopathy progression and vision loss. This work is presented in chapter 3.
b) To identify the prognostic prediction factor variables and investigate an evidence based and parsimonious set of variables, presented in chapter 4.
c) To develop a prognostic prediction model and internally validate it for diabetic retinopathy progression to the stage of treatment or vision loss. This is presented in chapter 5.

1.11 References:

1. American Diabetes Association. Diagnosis and classification of diabetes mellitus. Diabetes care. 2009;32 Suppl 1(Suppl 1):S62-7.
2. Leahy JL. Pathogenesis of type 2 diabetes mellitus. Arch Med Res. 2005;36(3):197-209.
3. Fowler MJ. Microvascular and Macrovascular Complications of Diabetes. Clinical Diabetes. 2008;26(2):77-82.
4. The Royal College of Ophthalmologists. Diabetic Retinopathy Guidelines 2012 update July 2013 [Available from: https://www.rcophth.ac.uk/wp-content/uploads/2014/12/2013-SCI-301-FINAL-DR-GUIDELINES-DEC-2012-updated-July-2013.pdf, accessed 2020, May 11.
5. Fernández-Vigo J, Castro J, Macarro A. Diabetic iris neovascularization. Natural history and treatment. Acta Ophthalmologica Scandinavica. 1997;75(1):89-93.
6. EDTRS. Grading diabetic retinopathy from stereoscopic color fundus photographs--an extension of the modified Airlie House classification. ETDRS report number 10. Early Treatment Diabetic Retinopathy Study Research Group. Ophthalmology. 1991;98(5 Suppl):786-806.
7. DRS. Diabetic retinopathy study. Report Number 6. Design, methods, and baseline results. Report Number 7. A modification of the Airlie House classification of diabetic retinopathy. Prepared by the Diabetic Retinopathy. Invest Ophthalmol Vis Sci. 1981;21(1 Pt 2):1-226.
8. Aldington SJ, Kohner EM, Meuer S, Klein R, Sjolie AK. Methodology for retinal photography and assessment of diabetic retinopathy: the EURODIAB IDDM complications study. Diabetologia. 1995;38(4):437-44.
9. Wilkinson CP, Ferris FL, 3rd, Klein RE, Lee PP, Agardh CD, Davis M, et al. Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales. Ophthalmology. 2003;110(9):1677-82.
10. Scanlon PH. The English National Screening Programme for diabetic retinopathy 2003-2016. Acta Diabetol. 2017;54(6):515-25.
11. Scotland G, McKeigue P, Philip S, Leese GP, Olson JA, Looker HC, et al. Modelling the cost-effectiveness of adopting risk-stratified approaches to extended screening intervals in the national diabetic retinopathy screening programme in Scotland. Diabetic Medicine. 2016;33(7):886-95.
12. SDRGS. Scottish Diabetic Retinopathy Grading Scheme 2007 v1.1.
13. Scottish Diabetic Retinopathy Screening Collaborative. Diabetic Retinopathy Screening Services in Scotland: A Training handbook July 2003 [Available from: https://www.ndrs.scot.nhs.uk/?page_id=1609.

14. World Health Organization. Blindness and vision impairment prevention - Priority eye diseases [Available from: https://www.who.int/blindness/causes/priority/en/index4.html.
15. World Health Organization. Blindness and vision impairment - key facts, 2019 [Available from: https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment, accessed August 06, 2020.
16. Barry RJ, Murray PI. Unregistered visual impairment: is registration a failing system? British Journal of Ophthalmology. 2005;89(8):995-8.
17. Arora S, Kolb S, Goyder E, McKibbin M. Trends in the incidence of visual impairment certification secondary to diabetic retinopathy in the Leeds metropolitan area, 2005–2010. Diabetic Medicine. 2012;29(7):e112-e6.
18. Lin S, Gupta B, James NL, Ling RHL. Visual impairment certification due to diabetic retinopathy in North and Eastern Devon. Acta Ophthalmologica. 2017;95:e75V e762.
19. Raman R, Gella L, Srinivasan S, Sharma T. Diabetic retinopathy: An epidemic at home and around the world. Indian J Ophthalmol. 2016;64(1):69-75.
20. Scanlon PH, Aldington SJ, Stratton IM. Epidemiological issues in diabetic retinopathy. Middle East African journal of ophthalmology. 2013;20(4):293-300.
21. Ogurtsova K, da Rocha Fernandes JD, Huang Y, Linnenkamp U, Guariguata L, Cho NH, et al. IDF Diabetes Atlas: Global estimates for the prevalence of diabetes for 2015 and 2040. Diabetes research and clinical practice. 2017;128:40-50.
22. Wild S, Roglic G, Green A, Sicree R, King H. Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. Diabetes care. 2004;27(5):1047-53.
23. Shaw JE, Sicree RA, Zimmet PZ. Global estimates of the prevalence of diabetes for 2010 and 2030. Diabetes research and clinical practice. 2010;87(1):4-14.
24. Yau JW, Rogers SL, Kawasaki R, Lamoureux EL, Kowalski JW, Bek T, et al. Global prevalence and major risk factors of diabetic retinopathy. Diabetes care. 2012;35(3):556-64.
25. Flaxman SR, Bourne RRA, Resnikoff S, Ackland P, Braithwaite T, Cicinelli MV, et al. Global causes of blindness and distance vision impairment 1990-2020: a systematic review and meta-analysis. Lancet Glob Health. 2017;5(12):e1221-e34.
26. Leasher JL, Bourne RR, Flaxman SR, Jonas JB, Keeffe J, Naidoo K, et al. Global Estimates on the Number of People Blind or Visually Impaired by Diabetic Retinopathy: A Meta-analysis From 1990 to 2010. Diabetes care. 2016;39(9):1643-9.
27. Sabanayagam C, Yip W, Ting DS, Tan G, Wong TY. Ten Emerging Trends in the Epidemiology of Diabetic Retinopathy. Ophthalmic epidemiology. 2016;23(4):209-22.
28. Klein R, Klein BE. Are individuals with diabetes seeing better?: a long-term epidemiological perspective. Diabetes. 2010;59(8):1853-60.
29. Wong TY, Mwamburi M, Klein R, Larsen M, Flynn H, Hernandez-Medina M, et al. Rates of progression in diabetic retinopathy during different time periods: a systematic review and meta-analysis. Diabetes care. 2009;32(12):2307-13.
30. Bagust A, Hopkinson PK, Maslove L, Currie CJ. The projected health care burden of Type 2 diabetes in the UK from 2000 to 2060. Diabetic medicine : a journal of the British Diabetic Association. 2002;19 Suppl 4:1-5.
31. Gonzalez EL, Johansson S, Wallander MA, Rodriguez LA. Trends in the prevalence and incidence of diabetes in the UK: 1996-2005. Journal of epidemiology and community health. 2009;63(4):332-6.
32. Mathur R, Bhaskaran K, Edwards E, Lee H, Chaturvedi N, Smeeth L, et al. Population trends in the 10-year incidence and prevalence of diabetic retinopathy in the UK: a cohort study in the Clinical Practice Research Datalink 2004-2014. BMJ open. 2017;7(2):e014444.
33. Buchan JC, Norman P, Shickle D, Cassels-Brown A, MacEwen C. Failing to plan and planning to fail. Can we predict the future growth of demand on UK Eye Care Services? Eye (London, England). 2019.

34. The Royal College of Ophthalmologists. The way forward age-related macular degeneration and diabetic retinopathy 2016

[ ]. Available from: https://www.rcophth.ac.uk/wp-content/uploads/2015/10/RCOphth-The-Way-Forward-AMD-300117.pdf, Accessed 2019, July 20.
35. Keenan TD, Johnston RL, Donachie PH, Sparrow JM, Stratton IM, Scanlon P. United Kingdom National Ophthalmology Database Study: Diabetic Retinopathy; Report 1: prevalence of centre-involving diabetic macular oedema and other grades of maculopathy and retinopathy in hospital eye services. Eye (London, England). 2013;27(12):1397-404.
36. Sivaprasad S, Gupta B, Gulliford MC, Dodhia H, Mohamed M, Nagi D, et al. Ethnic variations in the prevalence of diabetic retinopathy in people with diabetes attending screening in the United Kingdom (DRIVE UK). PloS one. 2012;7(3):e32182.
37. Denniston AK, Lee AY, Lee CS, Crabb DP, Bailey C, Lip PL, et al. United Kingdom Diabetic Retinopathy Electronic Medical Record (UK DR EMR) Users Group: report 4, real-world data on the impact of deprivation on the presentation of diabetic eye disease at hospital services. The British journal of ophthalmology. 2019;103(6):837-43.
38. Tufail A, Kapetanakis VV, Salas-Vega S, Egan C, Rudisill C, Owen CG, et al. An observational study to assess if automated diabetic retinopathy image assessment software can replace one or more steps of manual imaging grading and to determine their cost-effectiveness. Health Technol Assess. 2016;20(92):1-72.
39. Leal J, Luengo-Fernandez R, Stratton IM, Dale A, Ivanova K, Scanlon PH. Cost-effectiveness of digital surveillance clinics with optical coherence tomography versus hospital eye service follow-up for patients with screen-positive maculopathy. Eye (London, England). 2019;33(4):640-7.
40. Public Health England. NHS Diabetic Eye Screening Programme, Information for health professionals, 2016 [Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/505587/DES_07_GP_information_sheet_March_2016.pdf (Accessed 05/03/2020).
41. Jyothi S, Elahi B, Srivastava A, Poole M, Nagi D, Sivaprasad S. Compliance with the quality standards of National Diabetic Retinopathy Screening Committee. Primary care diabetes. 2009;3(2):67-72.
42. Public Health England. Diabetic eye screening standards valid for data collected from 1 April 2019, Guidance, 2019 [Available from: https://www.gov.uk/government/publications/diabetic-eye-screening-programme-standards/diabetic-eye-screening-standards-valid-for-data-collected-from-1-april-2019, Accessed 2020 May 14.
43. Aiello LM. Perspectives on diabetic retinopathy. American Journal of Ophthalmology. 2003;136(1):122-35.
44. Wong TY, Sun J, Kawasaki R, Ruamviboonsuk P, Gupta N, Lansingh VC, et al. Guidelines on diabetic eye care: the International Council of Ophthalmology Recommendations for screening, follow-up, referral, and treatment based on resource settings. Ophthalmology. 2018;125(10):1608-22.
45. International Council of Ophthalmology. ICO Guidelines for Diabetic Eye Care, updated 2017 [Available from: http://www.icoph.org/downloads/ICOGuidelinesforDiabeticEyeCare.pdf, accessed 2020, May 14.
46. The Royal College of Ophthalmologists. Risk stratification coding framework March 2020 [updated 2020. Available from: https://www.rcophth.ac.uk/wp-content/uploads/2020/03/Measuring-follow-up-timeliness-and-risk-for-performance-reporting-improvement-actions-and-targeting-failsafe-procedures-in-England.pdf, accessed 2020, May 11.
47. EDTRS. Photocoagulation for diabetic macular edema. Early Treatment Diabetic Retinopathy Study report number 1. Early Treatment Diabetic Retinopathy Study research group. Archives of ophthalmology (Chicago, Ill : 1960). 1985;103(12):1796-806.

48. Quartilho A, Simkiss P, Zekite A, Xing W, Wormald R, Bunce C. Leading causes of certifiable visual loss in England and Wales during the year ending 31 March 2013. Eye (London, England). 2016;30(4):602-7.
49. Stratton IM, Kohner EM, Aldington SJ, Turner RC, Holman RR, Manley SE, 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.
50. Wang T, Hong JL, Gower EW, Pate V, Garg S, Buse JB, et al. Incretin-Based Therapies and Diabetic Retinopathy: Real-World Evidence in Older U.S. Adults. Diabetes care. 2018;41(9):1998-2009.
51. DeFronzo RA. Combination therapy with GLP-1 receptor agonist and SGLT2 inhibitor. Diabetes Obes Metab. 2017;19(10):1353-62.
52. Marso SP, Bain SC, Consoli A, Eliaschewitz FG, Jodar E, Leiter LA, et al. Semaglutide and Cardiovascular Outcomes in Patients with Type 2 Diabetes. New England Journal of Medicine. 2016;375(19):1834-44.
53. Jampol LM, Bressler NM, Glassman AR. Revolution to a new standard treatment of diabetic macular edema. Jama. 2014;311(22):2269-70.
54. Gross JG, Glassman AR, Liu D, Sun JK, Antoszyk AN, Baker CW, et al. Five-Year Outcomes of Panretinal Photocoagulation vs Intravitreous Ranibizumab for Proliferative Diabetic Retinopathy: A Randomized Clinical Trial. JAMA Ophthalmol. 2018;136(10):1138-48.
55. DCCT. The effect of intensive diabetes treatment on the progression of diabetic retinopathy in insulin-dependent diabetes mellitus. The Diabetes Control and Complications Trial. Archives of ophthalmology (Chicago, Ill : 1960). 1995;113(1):36-51.
56. Adler AI, Stratton IM, Neil HA, Yudkin JS, Matthews DR, Cull CA, 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.
57. Accord Study Group, Accord Eye Study Group, Chew EY, Ambrosius WT, Davis MD, Danis RP, et al. Effects of medical therapies on retinopathy progression in type 2 diabetes.[Erratum appears in N Engl J Med. 2011 Jan 13;364(2):190], [Erratum appears in N Engl J Med. 2012 Dec 20;367(25):2458]. New England Journal of Medicine. 2010;363(3):233-44.
58. Sato Y, Kojimahara N, Kitano S, Kato S, Ando N, Yamaguchi N, et al. Multicenter randomized clinical trial of retinal photocoagulation for preproliferative diabetic retinopathy. Jpn J Ophthalmol. 2012;56(1):52-9.
59. DRS. Photocoagulation treatment of proliferative diabetic retinopathy. Clinical application of Diabetic Retinopathy Study (DRS) findings, DRS Report Number 8. The Diabetic Retinopathy Study Research Group. Ophthalmology. 1981;88(7):583-600.
60. Gross JG, Glassman AR, Jampol LM, Inusah S, Aiello LP, Antoszyk AN, et al. Panretinal Photocoagulation vs Intravitreous Ranibizumab for Proliferative Diabetic Retinopathy: A Randomized Clinical Trial. Jama. 2015;314(20):2137-46.
61. Virgili G, Parravano M, Evans JR, Gordon I, Lucenteforte E. Anti-vascular endothelial growth factor for diabetic macular oedema: a network meta-analysis. The Cochrane database of systematic reviews. 2018;10:Cd007419.
62. Jampol LM, Glassman AR, Bressler NM, Wells JA, Ayala AR. Anti-Vascular Endothelial Growth Factor Comparative Effectiveness Trial for Diabetic Macular Edema: Additional Efficacy Post Hoc Analyses of a Randomized Clinical Trial. JAMA Ophthalmol. 2016;134(12).
63. Stewart MW. Treatment of diabetic retinopathy: Recent advances and unresolved challenges. World journal of diabetes. 2016;7(16):333-41.
64. Strutton R, Du Chemin A, Stratton IM, Forster AS. System-level and patient-level explanations for non-attendance at diabetic retinopathy screening in Sutton and Merton (London, UK): a qualitative analysis of a service evaluation. BMJ open. 2016;6(5):e010952.

65. Lewis K, Patel D, Yorston D, Charteris D. A qualitative study in the United Kingdom of factors influencing attendance by patients with diabetes at ophthalmic outpatient clinics. Ophthalmic epidemiology. 2007;14(6):375-80.
66. Keenum Z, McGwin G, Jr., Witherspoon CD, Haller JA, Clark ME, Owsley C. Patients' Adherence to Recommended Follow-up Eye Care After Diabetic Retinopathy Screening in a Publicly Funded County Clinic and Factors Associated With Follow-up Eye Care Use. JAMA Ophthalmol. 2016;134(11):1221-8.
67. Forster AS, Forbes A, Dodhia H, Connor C, Du Chemin A, Sivaprasad S, et al. Non-attendance at diabetic eye screening and risk of sight-threatening diabetic retinopathy: a population-based cohort study. Diabetologia. 2013;56(10):2187-93.
68. Sabanayagam C, Banu R, Chee ML, Lee R, Wang YX, Tan G, et al. Incidence and progression of diabetic retinopathy: a systematic review. The lancet Diabetes & endocrinology. 2019;7(2):140-9.
69. Offord DR, Kraemer HC. Risk factors and prevention. Evidence Based Mental Health. 2000;3(3):70-1.
70. Keenan TDL, Johnston RL, Donachie PHJ, Sparrow JM, Stratton IM, Scanlon P. United Kingdom national ophthalmology database study: Diabetic retinopathy; report 1: Prevalence of centre-involving diabetic macular oedema and other grades of maculopathy and retinopathy in hospital eye services. Eye (Basingstoke). 2013;27(12):1397-404.
71. NHS Digital. Hospital Outpatient Activity, 2017-18 2017-2018 [Available from: https://digital.nhs.uk/data-and-information/publications/statistical/hospital-outpatient-activity/2017-18.
72. Hollingworth W, Jones T, Reeves BC, Peto T. A longitudinal study to assess the frequency and cost of antivascular endothelial therapy, and inequalities in access, in England between 2005 and 2015. BMJ open. 2017;7(10):e018289.
73. NHS Digital. Hospital Outpatient Activity 2018-19 2019 [Available from: https://digital.nhs.uk/data-and-information/publications/statistical/hospital-outpatient-activity/2018-19.
74. Sivaprasad S, Pearce E. The unmet need for better risk stratification of non-proliferative diabetic retinopathy. Diabetic medicine : a journal of the British Diabetic Association. 2019;36(4):424-33.
75. Lund SH, Aspelund T, Kirby P, Russell G, Einarsson S, Palsson O, et al. Individualised risk assessment for diabetic retinopathy and optimisation of screening intervals: A scientific approach to reducing healthcare costs. British Journal of Ophthalmology. 2016;100(5):683-7.
76. Scanlon PH, Aldington SJ, Leal J, Luengo-Fernandez R, Oke J, Sivaprasad S, et al. Development of a cost-effectiveness model for optimisation of the screening interval in diabetic retinopathy screening. Health Technol Assess. 2015;19(74):1-116.
77. Eleuteri A, Fisher AC, Broadbent DM, Garcia-Finana M, Cheyne CP, Wang A, et al. Individualised variable-interval risk-based screening for sight-threatening diabetic retinopathy: the Liverpool Risk Calculation Engine. Diabetologia. 2017.
78. Riley RD, van der Windt D, Croft P, Moons KGM. Prognosis Research in Health Care: Concepts, Methods, and Impact: Oxford University Press; 2019.

2 Chapter 2: Disease burden projections

The first section of this chapter describes the general methods used in this research project to determine the diabetes mellitus and diabetic retinopathy disease burden in the United Kingdom. The second section presents the research study prepared as a publication for submission in a peer reviewed journal.

2.1 General methods

The aim of this section is to introduce the overall methods and areas that are not covered in detail in the manuscript presented in section 2.

2.1.1 Primary care databases

At the planning stage of these studies, the different sets of data available were considered. The important issues were 1) conditions of access for use, 2) having sufficient information for the specific studies, 3) costs, 4) data strengths and limitations and 5) feasibility of use. Retrospective primary care data named IQVIA (1) Medical Research Data (IMRD), Diabetic Eye Screening Programme (DESP) data, hospital based data (Medisoft), and prospectively collected data were all considered. When assessing the disease burden, DESP, as well as hospital data, would have been limiting to that particular setting. An alternative was to recruit for a prospective cohort, with the advantage of greater control over the data collected. However, the cost of prospectively collecting the data was prohibitive and the time this would take was not feasible when considering the time scale of this PhD. I therefore decided in favour of primary care retrospective data, where screening and hospital data are sufficiently recorded, and also since this source has the advantage of containing information on comorbidities and prescribed medications.

2.1.2 IMRD database: strengths and limitations

Clinical Practice Research Datalink (CPRD), IMRD and Q Research are the three large databases in use on a national basis in the UK. CPRD collects primary care data through Vision, and also through Educational Management Information System (EMIS) software, with 50% of practices overlapping with practices contributing data to IMRD (previously known as The Health Improvement Network (THIN) database) (1). Q Research also collects data using EMIS software (1). IMRD collects non-identified patient data from UK General Practices, mainly from the Vision system (2).

The IMRD database was chosen for this research as the data within it are accessible through a sub-licence at the University of Birmingham, and because it is updated every four months. The database covers around 6% of UK practices (3). Demographic data (such as age and sex) are collected when a patient first registers with a general practice. At the time of routine consultations at general practices, clinically relevant information is recorded as part of clinical care using Vision software (2). Lab results are also recorded in Vision. This anonymised data is collected in full from contributing general practices when they first join IMRD, and is subsequently updated on a 4 monthly basis.

Strengths:

The strength of primary care practice data is that it is routinely collected (non-interventional) and therefore likely to be applicable in routine practice. IMRD data is collected from a wide geographical area and allows analyses of regional variations and is representative of the UK population by age and gender distribution (4).

With more than 700 general practices contributing their data, it is possible to extract large samples from the IMRD database. The data accessed via the University of Birmingham license are updated every year. Because consultation data is linked with prescriptions, comprehensive and accurate (validation checked) prescribing data are available. The data provider (IQVIA) performs quality checks of IMRD data. A key parameter is the Acceptable Mortality Reporting (AMR) date. This is the year when a practice is considered to reliably report all-cause mortality. This is determined by comparing each practice with the corresponding national annual age and sex-specific death rates, to assess the year that death reporting is most complete for each practice. This provides an external standard that is used to assess completeness of mortality data in automated primary care databases and serves as a quality control measure (5). Time efficiency and low cost, as compared to prospective patient recruitment for data collection, are other benefits. There have been a large number of population-based studies carried out in the area of this research project in IMRD (6-10).

Limitations:

The accuracy of this research is dependent on the quality of the IMRD data. However, since the data are routinely collected during GP consultations, the available information may not be fit for research. On the whole, prescribing and demographic (age and sex) data are comprehensive and accurate, but lifestyle and socio-economic data are less so. The ethnicity is poorly recorded although it is improving (11). With the introduction of Quality Outcome Framework (QOF) incentives, and at times their removal (12), it must be considered that there may have been variations in data recording. A large proportion of the data may be missing, and

possibly not missing at random. General practices may vary in coding and recording information. There are internal validation checks, which can exclude implausible values, however values that are incorrect but plausible are difficult to detect (13). Information can only be recorded for individuals who are registered, and therefore cannot be as complete as prospective data. Also, because the database is so vast, studies are likely to be overpowered, causing statistical significance in all analyses, and therefore the clinical significance of data resulting from such databases should be considered.

Table 1: Information included in IMRD data

Specific data	Overall what is available	Relevant to the study
Demographics	Dates patients registered at practices	✓
	Dates patients left practices	✓
	Patient registration status	✓
	Temporary/Permanent	
	Year of birth	✓
	Sex	✓
	Ethnicity	✓
	Townsend deprivation score	✓
Diagnoses	All conditions, symptoms and examinations during consultations, such as diabetes, using the Read Codes version 2	✓
Iinformation on referrals to secondary care and received back	Details on hospital admissions, for instance for vitreoretinal surgery	✓
	Discharge diagnosis, such as diabetic retinopathy / vision loss	✓
	Outpatient consultation diagnosis	✓
	Investigation and treatment outcomes	✓
Prescribing	The drug prescribed using the drug codes, i.e. Insulin / Metformin	✓
	Acute treatments and medicines such as laser / Anti VEGF	✓
Additional Health Information	Information on lifestyle - smoking, BMI etc	✓
	Tests and laboratory results, i.e. blood sugar	✓

2.1.3 IMRD structure and data protection information

The structure of the IMRD database is mainly based on general practice files, which consist of patient files linked by practice ID and patient ID. The four main files are patient, therapy, medical, and additional health data (AHD) files, and there are another three linked files (postcode variable indicators (PVI), staff, and consult files). Patients are made aware of the IMRD data collection scheme, and of their right to withdraw at any time with previous data removed.

2.1.4 IMRD data use and defining variables

Prior to data extraction, a series of meetings were held to develop a comprehensive list of Read codes, AHD, and some drug codes, to define the population, outcomes, exposures and variables for both studies. A systematic approach, described in the IMRD data guide 1705 (14), was used to search the Read code dictionary for key words. The variables were primarily chosen from the systematic review list of 78 candidate predictors in Appendix 12 in Chapter 3.

2.1.5 Scientific and ethical approval

The IMRD database has blanket approval by the NHS South East Multi-centre Research Ethics Committee (MREC) in 2003 (15). The study protocols were submitted to both the Scientific Review Committee (SRC) and the Science, Technology, Engineering and Mathematics Ethical Review Committee at the University of Birmingham for review and approval. Approvals were granted on 12/04/2018 and 18/04/2019 respectively (Appendix 1 and Appendix 2). The approved protocol is also attached (Appendix 3).

2.1.6 Read codes and their selection methods

Read code dictionary version 2 was used in this study. Read codes cover clinical features, diagnosis, procedures, investigations and, to some extent, drugs. They have a hierarchy with increasingly specific items down the order. Selection was carried out in collaboration with Jhot Chandan and Kelvin Okoth, fellow doctoral researchers. The seven step process followed is given in detail in Appendix 4 and Appendix 5.

2.1.7 Data extraction

Data was extracted by using an automated platform developed at the University of Birmingham (DExtER). A series of cross-sectional data of patients with diabetes, was extracted, with an index date of the 1st of every year, from 1998 to 2018, with the aim of measuring the disease and treatment burden. Further details are in section 2.

2.1.8 Adjudication for diabetes and retinopathy diagnosis in IMRD studies

Patients were initially classified according to the diagnostic Read codes for the presence of type 1 or type 2 diabetes. This was further confirmed with the prescriptions of anti-diabetic medications. For a classification of type 1 diabetes, the patient needs to be on insulin, not taking any oral medications (other than metformin), and must have a type 1 diabetes diagnostic code. The most severe diabetic retinopathy read code was used. For the longitudinal data in Chapter 5, all individuals identified as having type 1 or type 2 diabetes after passing through this process were included in the final analysis. Further details are in section 2, and Chapter 5.

2.1.9 References:

1. Kontopantelis EE. Primary Care data signposting CPRD, THIN and other databases. Institute of Population Health, University of Manchester. 2014.
2. Cegedim. Vision medical software. Smarter, faster and better healthcare 2019 [Available from: https://www.visionhealth.co.uk/vision-medical-software/.
3. Kontopantelis E, Stevens RJ, Helms PJ, Edwards D, Doran T, Ashcroft DM. Spatial distribution of clinical computer systems in primary care in England in 2016 and implications for primary care electronic medical record databases: a cross-sectional population study. BMJ open. 2018;8(2):e020738.
4. Blak BT, Thompson M, Dattani H, Bourke A. Generalisability of The Health Improvement Network (THIN) database: demographics, chronic disease prevalence and mortality rates. Informatics in primary care. 2011;19(4):251-5.
5. Maguire A, Blak BT, Thompson M. The importance of defining periods of complete mortality reporting for research using automated data from primary care. Pharmacoepidemiology and drug safety. 2009;18(1):76-83.
6. Gonzalez EL, Johansson S, Wallander MA, Rodriguez LA. Trends in the prevalence and incidence of diabetes in the UK: 1996-2005. Journal of epidemiology and community health. 2009;63(4):332-6.
7. Martin-Merino E, Fortuny J, Rivero-Ferrer E, Garcia-Rodriguez LA. Incidence of retinal complications in a cohort of newly diagnosed diabetic patients. PloS one. 2014;9(6):e100283.
8. Martin-Merino E, Fortuny J, Rivero-Ferrer E, Lind M, Garcia-Rodriguez LA. Risk factors for diabetic retinopathy in people with Type 2 diabetes: A case-control study in a UK primary care setting. Primary care diabetes. 2016;10(4):300-8.
9. Martin-Merino E, Fortuny J, Rivero E, Garcia-Rodriguez LA. Validation of diabetic retinopathy and maculopathy diagnoses recorded in a U.K. primary care database. Diabetes care. 2012;35(4):762-7.
10. McEwan P, Bennett H, Ward T, Bergenheim K. Refitting of the UKPDS 68 risk equations to contemporary routine clinical practice data in the UK. Pharmacoeconomics. 2015;33(2):149-61.
11. Mathur R, Bhaskaran K, Chaturvedi N, Leon DA, vanStaa T, Grundy E, et al. Completeness and usability of ethnicity data in UK-based primary care and hospital databases. J Public Health (Oxf). 2014;36(4):684-92.
12. Minchin M, Roland M, Richardson J, Rowark S, Guthrie B. Quality of Care in the United Kingdom after Removal of Financial Incentives. The New England journal of medicine. 2018;379(10):948-57.
13. Forstmeier W, Wagenmakers EJ, Parker TH. Detecting and avoiding likely false-positive findings - a practical guide. Biol Rev Camb Philos Soc. 2017;92(4):1941-68.
14. QuintilesIMS. THIN Dataguide For Researchers. 2017.
15. NHS Health Research Authority. IQVIA Medical Research Data 2018 [Available from: https://www.hra.nhs.uk/planning-and-improving-research/application-summaries/research-summaries/the-health-improvement-network-thin-database/.

2.2 Disease burden of diabetes, diabetic retinopathy and their future projections

Disease burden of Diabetes, Diabetic retinopathy and their future projections

Sajjad Haider[1], Rasiah Thayakaran[3], Anuradhaa Subramanian[3], Konstantinos A Toulis[3], David Moore[*2], Malcolm James Price[*2], Krishnarajah Nirantharakumar[*2]

1. Doctoral researchers, Institute of Applied Health Research, University of Birmingham
2. [*]Joint Senior Authors Institute of Applied Health Research, University of Birmingham, Birmingham, United Kingdom
3. Institute of Applied Health Research, University of Birmingham

Joint Corresponding authors:

Sajjad Haider, Doctoral Researcher, Institute of Applied Health Research, University of Birmingham, Birmingham, United Kingdom

E mail: [redacted]

Krishnarajah Nirantharakumar, Institute of Applied Health Research, University of Birmingham, Birmingham, United Kingdom

E mail: [redacted]

Under submission

2.3 Abstract

Introduction / Aims

To estimate the current disease burden, trends and future projections for diabetes mellitus and diabetic retinopathy and to forecast future projections in a primary care IMRD database.

Methods

We performed a cross-sectional study of patients aged 12 and above to determine the prevalence of diabetes mellitus and diabetic retinopathy from an IMRD database in January 2017. We also carried out a series of cross-sectional studies to look into prevalence trends, and then applied a double exponential smoothing model to forecast the future burden of diabetes mellitus and diabetic retinopathy in UK.

Results

The crude diabetes mellitus prevalence in 2017 was 5.2%. The Diabetic Retinopathy, sight threatening retinopathy and diabetic maculopathy prevalence figures in 2017 were 33.78%, 12.28% and 7.86% respectively in our IMRD cross-sectional study. There are upward trends in the prevalence of diabetes mellitus, diabetic retinopathy, and sight-threatening retinopathy, most marked and accelerating in sight threatening retinopathy in type 1 diabetes mellitus but slowing in type 2 diabetes, and in the overall prevalence of diabetic retinopathy.

Conclusion

Our results suggest differential rising trends in the prevalence of diabetes mellitus and diabetic retinopathy. Preventive strategies, as well as treatment services planning, can be based on these projected prevalence estimates. Improvements that are necessary for the optimisation of care pathways, and preparations to meet demand and capacity challenges, can also be based on this information. The limitations of the study can be overcome by a future collaborative study linking diabetic retinopathy screening and hospital eye services data.

.

Summary Box

What was already known on the subject of disease burden?

- Incidence of DR 2004 to 2014, trends and prevalence of diabetic retinopathy to 2014 in the United Kingdom.
- Increasing prevalence of diabetes between the years 1996 and 2005 and up to date DM prevalence for the population over 17 years old.
- Future projections of diabetes and diabetic retinopathy, mostly underestimations

What this study adds?

- This is an up to date study to give diabetes mellitus and diabetic retinopathy prevalence trends from 1998 to 2018.
- This study forecasts the future diabetic retinopathy disease burden up to 2030 to enable preparation for impending challenges.
- Current prevalence of age 12 and over, diagnosed Diabetes Mellitus, Diabetic Retinopathy, Sight Threatening Retinopathy, Diabetic Macular Oedema disease and treatment burden in United Kingdom

2.4 Introduction

Diabetic Retinopathy (DR) is the fourth most common cause of blindness and visual impairment in high-income countries (1). Services are overburdened and optimisation requires accurate estimates of disease and the expected treatment burden (2). A recent systematic review of studies estimating the incidence of DR (3) highlighted the paucity of contemporary evidence from developed countries on the disease burden. The review also recommended that estimates should be based on populations with diabetes mellitus (rather than the general population) so as not to dilute the estimates. A recent attempt to forecast the UK-wide disease burden of DR was hindered by the need for reliable data (4).

Previous studies have been conducted on the prevalence of DR (5-9), with the most recent UK-wide study being performed in 2014 based on Clinical Practice Research Datalink (CPRD). Two of these studies also explored trends in DR incidence and prevalence (6, 9). A significant amount of heterogeneity in the populations studied, the classification of DR, and the definition of its presence and severity was present in these studies. Studies of the forecasts of the future disease burden of diabetic retinopathy would be useful both for preparing health care delivery systems for the future, and in preventing blindness in patients with DM. There is a Europe wide forecast study with UK component based on pre 2009 data dealing with DR only (10). The disease burden estimate of diabetic retinopathy will not be complete without a similar estimate for the diabetes burden. A UK wide study upto date study dealing with DM, DR and Sight Threatening Retinopathy (STR) is needed.

A previous study on future projections of diabetes mellitus in the United Kingdom was found to underestimate prevalence (11). Moreover, evidence suggests that the rate of increase is not constant or uniform across DM subtypes (namely type 2 diabetes mellitus (T2DM) and type 1 diabetes mellitus (T1DM)), especially in children (12). The incidence rate of type 1 diabetes mellitus (pooled estimate of European centres, UK included) in children is expected to continue to rise at a rate of 3.4% per annum (13). Gonzalez et al (14) reported an increasing prevalence of diabetes for the 10 years up to 2005. Public Health England figures are available for 2019, based on the Quality Outcome Framework, except in Scotland where they are based on Scottish Diabetes Survey (15). However, these figures are limited to those over 17 years old. We aimed to estimate recent trends in the disease burden of DM, and to use this as a base on which to estimate the disease burden for DR and STR in the UK. We then set out to design, train and validate a forecasting model to support future projections of these disease burdens and to assess detailed contemporary prevalence estimates for the year 2017. Since diabetic retinopathy screening is offered after age 12 only, the population of interest to us was age 12 or over only.

2.5 Methods

2.5.1 Study design and data source

To study the trend, and to forecast the future burden of diagnosed DM, DR and STR, we used the IMRD database, a primary care database, to conduct a series of yearly cross-sectional analyses on the 1st of each year from 1998 to 2018. In addition, a detailed cross sectional study was carried out on the 1st of January 2017 to estimate

the prevalence of T1DM and T2DM in the whole UK population, and of diabetic retinopathy in patients with T1DM and T2DM.

IMRD is a large UK general practice electronic database containing anonymised patient records from 787 general practices, with over 15 million patient records, of which around 3.7 million are active at a given time point (6.2% of the UK population). IMRD provides information on demographics, lifestyle, diagnoses, and prescriptions, and is quality checked (16). Based on the demographics distribution observed in IMRD, it is considered generalizable to the UK population (17). IMRD has previously been used and validated to estimate prevalence trends of DM and DR, and to identify risk factors for DR (14, 18-21).

2.5.2 Study population

To ensure that only high quality data was included, and that all important covariates were documented, general practices were eligible only if they showed acceptable mortality rates one year before the cross-sectional study date (16), and had been using the electronic medical record system for at least a year. Patients from these eligible general practices must have been registered with their practice for at least one year and must be aged 12 years or above to be included in the study (to match the DESP criteria). For estimation of the prevalence of T1DM and T2DM, the whole registered population was included as the denominator population (per 1000). For estimation of STR and DR prevalence, patients with DM served as the denominator (%). Estimates are stratified by type of diabetes.

2.5.3 Case definition of diagnoses of diabetes mellitus and diabetic retinopathy

Clinical diagnosis and symptoms in the IMRD database are recorded using the Read code classification system (22). Read codes were selected using a rigorous seven step process (Appendix 4 and 5). Read codes are given in Appendix 6. Patients with a Read code record of DM before the study entry date were identified. Patients with a record of DM specified as type 1 were categorised as type 1 if they had at least one prescription record for insulin and no record for any oral glucose-lowering medication other than metformin in the database. The remaining patients with diabetes were categorized as type 2. Prevalence estimates calculated were verified against Public Health England (PHE) estimates of DM (23).

The most severe diabetic retinopathy Read code recorded before study entry was used to classify their DR or STR status. Stages of DR among those patients identified with DM were classified using the Royal College of Ophthalmology modified classification (24). However, patients with a retinopathy record were stratified into mutually exclusive categories of 1) Pre-Sight Threatening Retinopathy (Pre-STR) including no retinopathy and background retinopathy, 2) Sight Threatening Retinopathy (STR) and 3) Retinopathy unspecified as either pre-STR (background retinopathy) or STR. Pre-STR was further categorized into mutually exclusive categories: 1) R0 or 2) R1. STR was further categorized into mutually exclusive categories of 1) STR based on diagnostic codes and 2) STR that needed treatment or resulted in vision loss. Within STR we categorised pre-proliferative DR (R2) and proliferative DR (R3) as mutually exclusive groups. STR was further stratified into overlapping categories based on the presence of sight-threatening retinopathy (R2/3)

and maculopathy (M1). Treatment and vision loss codes included: (i) laser therapy, (ii) vitreous injection and other vitreous procedures, (iii) low vision or blindness.

2.5.4 Time trend analysis and forecasting models

A double exponential smoothing model was chosen to cover the level and trend, as this was yearly cross-sectional data with no seasonal / cyclical variation expected or observed (25) not unlike Adams et al published model (26). The IMRD serial cross-sectional data for the prevalence of DM and diabetic retinopathy (STR and any retinopathy) were split into two portions - 1998 to 2013 (training data) and 2014 to 2018 (test data). The model was fitted to the training data and then prediction was carried out from 2014 to 2018. This was then compared with the test data for validation. Thereafter, the yearly prevalence of DR and STR were projected up to 2030 using the same model with 95% prediction intervals. This was done using the statistical software R (2019) (27). Prevalence rates were then converted into patient numbers, using projected population figures from the Office of National Statistics (28).

2.5.5 IMRD data analysis for annual prevalence of DM and DR.

Prevalence trends between the two decades before and after 2008 were also compared for trend analysis. Patients identified as T1DM or T2DM on or before 1st of January in each year analysed were identified as the numerators for calculating the prevalence of T1DM and T2DM. The prevalence was estimated by dividing the numerator population by the eligible registered population aged above 12 years (denominator) on 1st of January for the corresponding year. Among these patients, those diagnosed with any retinopathy and those with STR were numerators for

calculating the prevalence of DR and STR respectively. Prevalence estimates are provided for patients with T1DM and T2DM separately with 95% confidence intervals, A description of patients aged 12 or above with a diagnosis of DM is also given for the year 2017. Baseline characteristics such as age, and age at diagnosis of diabetes were summarized as the mean (SD), and as frequency (percentage) for sex, Townsend deprivation quintile and ethnicity. These characteristics were also reported as stratified by type of DM. A detailed description of the proportion of DM patients (T1DM and T2DM aged 12 or above) with DR in the year 2017 categorized by DR severity is also presented.

Estimates from IMRD were compared to estimates obtained from 2017 Diabetic Eye Screening Programme (DESP) referral data (29) and also other data from UK studies (5-7, 9, 30) for verification (Appendix 9).

Ethics:

The study protocols were submitted to both the Scientific Review Committee and the Science, Technology, Engineering and Mathematics Ethical Review Committee at the University of Birmingham for review and approval, which were granted (Appendix 1 and Appendix 2). The approved protocol is also attached (Appendix 3).

2.6 Results

Figure 1 gives the Patients flow and case selection algorithm. As of 1st January 2017, 2,813,916 people were eligible to be included in the primary cross-sectional analysis. The demography characteristics of the sample are given in Table *1**Figure 1***. The mean age of patients with T1DM and T2DM as of 1st January 2017 was 42.5 (17.2) and 66.3 (13.0) respectively. The mean age at diagnosis of T1DM and T2DM were

21.4 (14.3) and 57.0 (13.1) respectively. Nearly 80% and 55% of patients respectively had their Townsend deprivation and ethnicity recorded in IMRD.

2.6.1 Prevalence trends

Figure 3 gives the prevalence trends of STR and DR among patients aged 12 and above with T1DM and T2DM from 1998 to 2018. The results in figures 2 and 3 show an almost a global upward trend in the prevalence of both types of diabetes (T1DM and T2DM) and in diabetic retinopathy (all types of DR / STR). The highest rise was seen in STR in those with T1DM (3.7 times increase in two decades). The second highest rise was in all types of DR in the T2DM population (2.8 times). Splitting this data by the decades (1998 to 2007 versus 2009 to 2018), the end of the first decade showed a higher increase in every category (diabetes as well as diabetic retinopathy) as compared to the second decade, except in T1DM (Appendix 7), where it was higher in second decade. T2DM increased more than T1DM between 1998 and 2018, but while the increase in T2DM prevalence slowed recently, the increase in T1DM prevalence accelerated significantly in the recent decade.

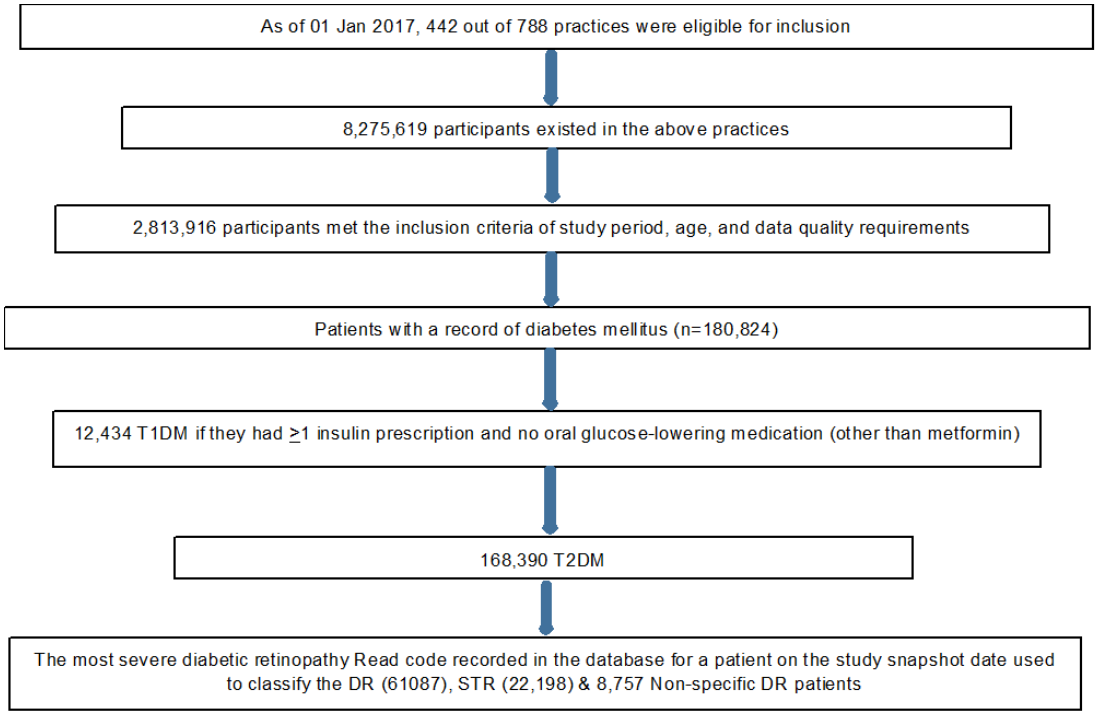


Figure 1: Patients flow and case selection algorithm

Table 1: Demography of patients with diabetes mellitus in IMRD data on 1st of January 2017

	DM (N)	% (SD)	T1DM (N)	% (SD)	T2DM (N)	% (SD)
Total	180,824	100.00%	12,434	6.88%	168,390	93.12%
Gender						
Male	101,628	56.20%	7,192	57.84%	94,436	56.08%
Female	79,196	43.80%	5,242	42.16%	73,954	43.92%
Age	180,824	64.7 (SD 14.7)	12,434	42.5 (SD 17.2)	168,390	66.3 (SD 13.0)
Age at diagnosis	180,788	54.6 (SD 16.0)	12,422	21.4 (SD 14.3)	168,366	57.0 (SD 13.1)
Townsend						
1	27,616	15.27%	2,037	16.38%	25,579	15.19%
2	30,011	16.60%	2,206	17.74%	27,805	16.51%
3	32,434	17.94%	2,222	17.87%	30,212	17.94%
4	31,332	17.33%	1,978	15.91%	29,354	17.43%
5	24,606	13.61%	1,568	12.61%	23,038	13.68%
Missing	34,825	19.26%	2,423	19.49%	32,402	19.24%
Ethnicity						
Caucasian	88,420	48.90%	6,584	52.95%	81,836	48.60%
Black afro Caribbean	2,738	1.51%	98	0.79%	2640	1.57%
Chinese/Middle eastern/ others	567	0.31%	45	0.36%	522	0.31%
South Asians	6,361	3.52%	124	1.00%	6237	3.70%
Mixed race	1243	0.69%	32	0.26%	1211	0.72%
Missing	81,495	45.07%	5551	44.64%	75944	45.10%

DM-Diabetes Mellitus; T2DM – Type 2 Diabetes Mellitus; T1DM – Type 1 Diabetes Mellitus

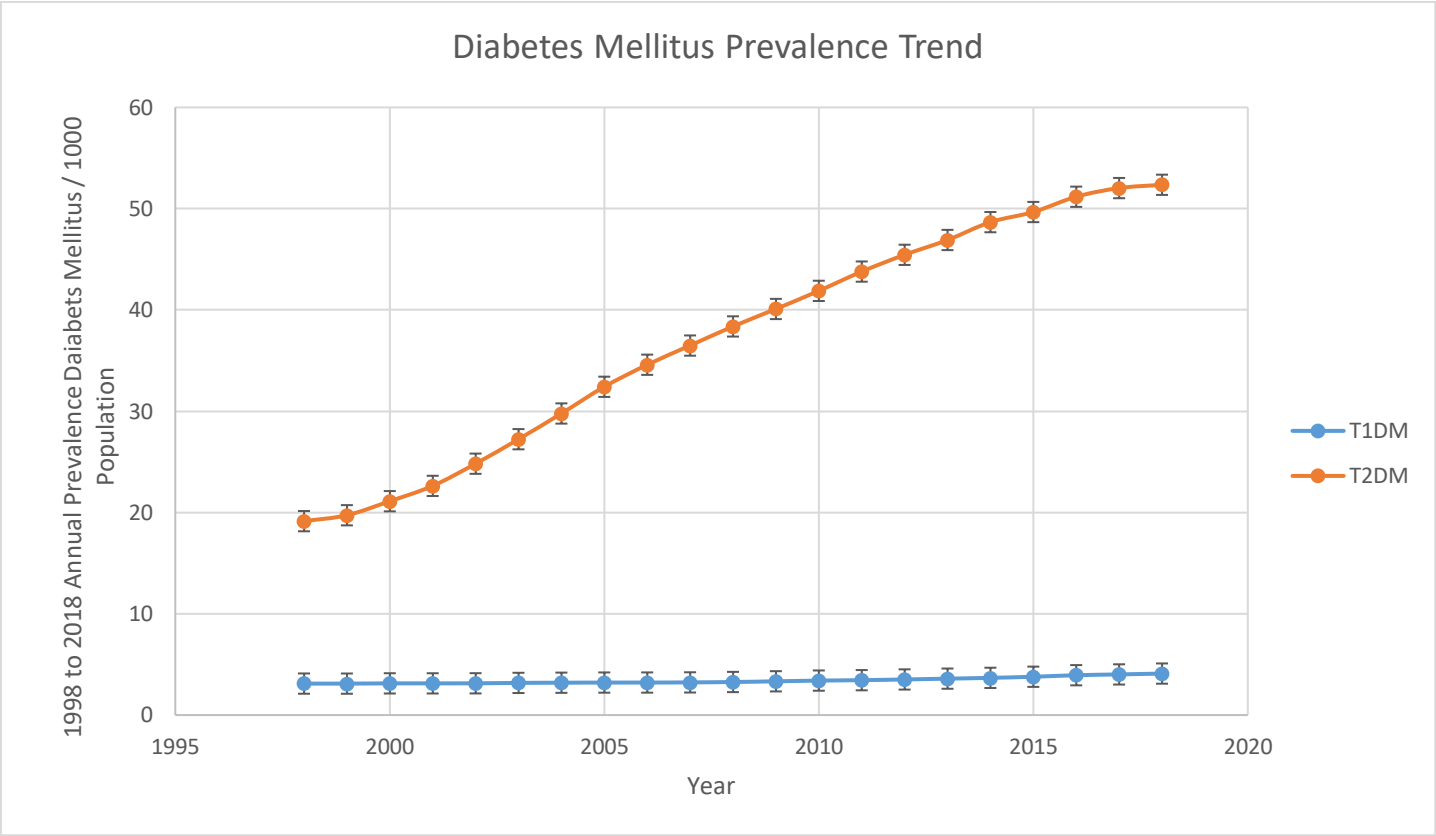


Figure 2: Prevalence trends of DM from year 1998 to year 2018

T1DM - Type 1 Diabetes Mellitus, T2DM - Type 2 Diabetes Mellitus

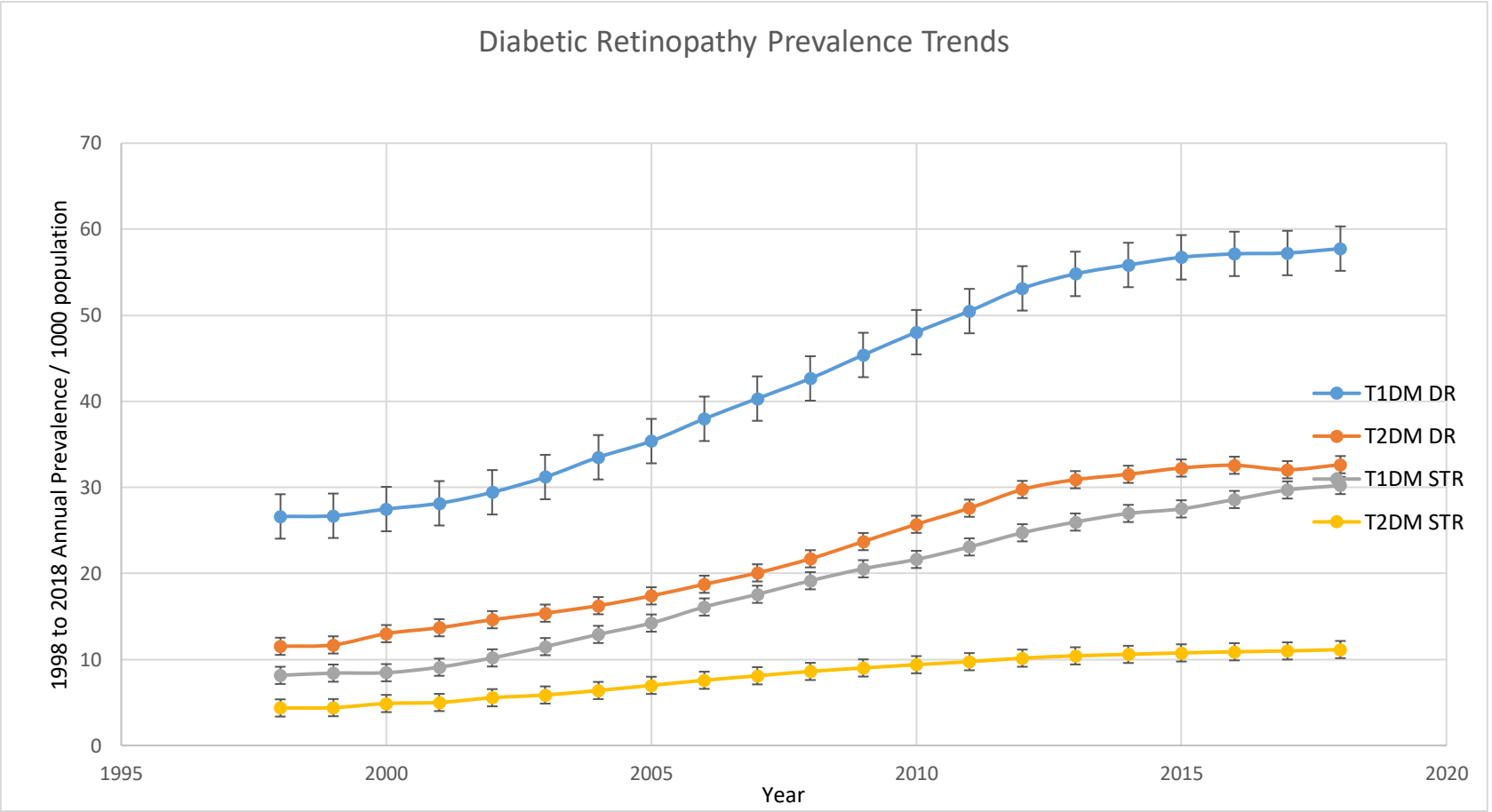


Figure 3: Annual prevalence (95% CI) of DR and STR from year 1998 to year 2018

DR - Diabetic Retinopathy, STR - Sight threatening Retinopathy

2.6.2 Forecasting model

The forecasted annual UK prevalence values of T1DM, T2DM, DR and STR, with their 95% prediction intervals (PI), are given in the Appendix 8. These suggest that the prevalence will increase by 24% (5-43%), 7.1% (-28-41%), 9% (-50-65%) and 17% (-21-54%) respectively by 2030. Corresponding estimates of the absolute numbers of people in the UK forecast to have these conditions are shown in Table 2. These correspond to 0.36 (.3 -.4), 4 (2.6 - 5.3), 1.6 (.7-2.5), and 0.64 (.42-.86) million people respectively having each condition respectively. We verified our UK forecast for 2019 and found the total figure (3,800,572) to be close to the Quality Outcome Framework provided estimate of diagnosed DM of 3,809,119.

Table 2: Future Projections of Diabetes and Diabetic Retinopathy Burden

Year	Projected Population	*T1DM	*T2DM	*Total DM	#DR	#STR
2019	66,435,550	278083	3504300	3782382	1305586	479242
2020	66,832,812	285715	3547631	3833347	1333906	493320
2021	67,195,769	293270	3589396	3882666	1361933	507386
2022	67,530,759	300764	3629901	3930665	1389770	521473
2023	67,844,183	308220	3669464	3977684	1417526	535618
2024	68,138,263	315643	3708183	4023827	1445230	549831
2025	68,413,827	323031	3746086	4069117	1472884	564109
2026	68,671,302	330381	3783177	4113558	1500482	578447
2027	6,892,124	337740	3820023	4157763	1528242	592929
2028	69,163,320	345105	3856598	4201702	1556151	607548
2029	69,397,439	352472	3892888	4245360	1584200	622301
2030	69,624,055	359842	3928912	4288754	1612395	637187

*The DR and STR forecast is actual IMRD based figures projected for the UK population (26). Formula used is Affected Population = Projected Prevalence X Projected Population. # In calculating projections for diabetic retinopathy we have applied the retinopathy rates of those aged 12 and above for the whole diabetes population rather than for those above 12 years old (age at which retinopathy screening commences and was one of our inclusion criteria). This approximately gives the projected total population, as the breakdown for those over 12 years is not available but the number of patients with DM below 12 years is negligibly small.

2.6.3 2017 Cross-sectional analysis

In the 2017 data analysis, 180,824 patients had a code for diabetes prior to this date of which 12,434 (6.9%) were identified as T1DM and 168,390 (93.1%) were identified as T2DM. Patients with DM were more likely to be men (56.2% vs 43.8%). The prevalence of diabetic retinopathy in different stages of progression is given in Appendix 7 3. Prevalence of any DR and STR among patients with DM aged 12 and above was 33.8% and 12.3% respectively. When stratified by diabetes type, a higher proportion of patients with T1DM had a more severe form of retinopathy than patients with T2DM (prevalence of STR was 29.7% vs 11%), while prevalence of pre-STR (R0/R1 & M0) was higher among patients with T2DM (31.8% in T1DM vs 37.8% in T2DM). Each subcategory among STR population (R2 / R3 / M1 and their combinations), was present in higher proportion of patients with T1DM as compared to T2DM (R2: 3.7% vs 1.2%; R3: 12.1% vs 1.9%; and M1: 19.6% vs 7.0% respectively)]. A higher proportion of patients with T1DM compared to T2DM also received treatment procedures (Laser: 7.1% vs 1.3%; Vitreous injection and procedures: 5.1% vs 1.1%). There was also a higher proportion of documented cases of visual impairment or vision loss among T1DM [3.1% vs 2.8%].

Table 3: Diabetic Retinopathy in patients with diabetes mellitus in IMRD data on 1st of January 2017

	DM		T1DM		T2DM	
Diabetes (N)	180,824	**%**	12,434	**%**	168,390	**%**
No Retinopathy coding available	82,119	45.41%	3,846	30.93%	78,273	46.48%
Retinopathy Coding available	98,705	54.59%	8,588	69.07%	90,117	53.52%
Pre-STR	**67750**	37.47%	**3951**	31.78%	**63699**	37.83%
No DR (R0M0)	37,618	20.80%	1,472	11.84%	36,146	21.47%
R1	30,132	16.66%	2,479	19.94%	27,553	16.36%
STR	**22,198**	12.28%	**3,693**	29.70%	**18,505**	10.99%
STR without Rx or vision loss	13,165	7.28%	2,271	18.26%	10,894	6.47%
R2	2,487	1.38%	454	3.65%	2,033	1.21%
R3	4,729	2.62%	1,505	12.10%	3,224	1.91%
M1	14,206	7.86%	2,440	19.62%	11,766	6.99%
STR with Rx and vision loss	9,033	5.00%	1,422	11.44%	7,611	4.52%
Laser	3,092	1.71%	885	7.12%	2,207	1.31%
Vitreous injections / procedures	2,536	1.40%	637	5.12%	1,899	1.13%
Vision loss / blindness	5,050	2.79%	384	3.09%	4,666	2.77%
None specific for STR or Pre-STR	**8,757**	4.84%	**844**	6.79%	**7913**	4.70%
Any retinopathy	61087	33.78%	7016	56.43%	53971	32.05%

DR – Diabetic retinopathy, R0 – no retinopathy, M0 – no maculopathy, R1, Background retinopathy, Pre-STR is combination of no diabetic retinopathy and background retinopathy, R2 is pre-proliferative diabetic retinopathy, R3 is proliferative diabetic retinopathy, M1 is diabetic maculopathy, STR is sight-threatening retinopathy which is a combination of R2, R3 and M1, Non-specific retinopathy is where it cannot be categorised into R1 or STR. Where colour codes are assigned, the same colour indicates that they are mutually exclusive. Where colour codes are not assigned they overlap within that category. For example, patients with M1 can have either R2 or R3, likewise patients who received laser treatment could have received vitreous injection. . The WHO standards (31) were used for vision loss. Here all categories were combined into a single category.

2.7 Discussion

2.7.1 Principal findings

We explored the disease burden associated with diabetic retinopathy in the UK from the past, present and future perspectives. Our study followed a tripartite structure, comprising of 1) a series of epidemiological studies throughout a 20-year span to document disease-specific trends, 2) training a forecasting model to predict the future disease burden to guide clinical practice and service development and 3) a detailed descriptive cross-sectional analysis in 2017 using a study population of 180,824 people with diabetes to explore contemporary prevalence estimates of different forms of DR.

Between 1998 and 2018, the prevalence of DR and STR increased. The prevalence of all DR in T2DM nearly tripled and STR almost quadrupled among patients with T1DM aged 12 and above. There was a parallel increase in the overall prevalence of DM. While the growth in the numbers of T1DM patients was less than that for patients with T2DM, stratifying the calculations by two decades showed a marked rise in the rate of increase in T1DM prevalence in the latter half of the whole period between 1998 and 2018. This was in sharp contrast to the trends in T2DM, STR and DR prevalence, which showed a higher rise in the decade between 1998 and 2007 but slowed down in the later decade between 2009 and 2018.

Our forecasting model showed that, in less than ten years, over 1.5 million people with diabetes will have some DR, almost two thirds of a million of whom will have STR. With T1DM expected to rise faster and higher, it is also likely to correspond to a comparatively higher rise in STR, forcing a further increase in demand on services.

A key parameter when calculating the current and future prevalence of DR is the accuracy of estimates of the trend of the underlying condition, i.e. the presence of diabetes mellitus. T1DM showed a smaller increase in the period starting from 1998, but this has accelerated since 2009. This is the most concerning recent trend considering that these are younger patients (mean age of diagnosis of 21.4 vs 57), having to live with the condition and its complications for more life years, and also suffering from the more severe form of diabetic retinopathy, with the consequent disability, treatment burden and treatment costs. There is a recent report of a 3.4% annual increase in the incidence rate of T1DM in children (13). Although there is an association between T1DM and obesity (32), it is believed that the cause may be multifactorial, including hygiene, viral factors and vitamin D deficiency amongst others (33).

The diagnosed DM prevalence based on the 2017 IMRD cross-sectional survey is 5.2%. The detailed descriptive analysis in 2017 showed that, out of 180,824 people with diabetes, 33.8% had any DR as a complication, 12.3% had sight threatening retinopathy and, importantly, 2.8% had blindness or vision loss. STR was 52% of total DR in T1DM and 34% of total DR in T2DM. In 2017, nearly one third of all patients with T1DM were affected by a sight threatening form of DR. This analysis also confirmed the notion that, from the health care perspective, neither DM type is “benign” with regards to DR risk, since DR severity is graver in T1DM, and absolute numbers of affected individuals are higher in T2DM.

Diabetic complications are mainly macrovascular damage (coronary artery disease, peripheral arterial disease, and stroke) or microvascular damage to blood vessels in organs like kidney, foot and nerves (34). Tackling the first reduces mortality rate and

might mean these patients living longer and consequently a higher prevalence of diabetic retinopathy among higher risk patients. With greater efficacy and a rapid reduction of HBA1C, the new agents might induce progression of DR (early worsening) (35). So with increased prevalence there may be a disproportionately more high risk diabetic retinopathy cases. There are conflicting reports on direct effect of newer medical treatments like Incretin based Therapies on diabetic retinopathy (35, 36). But the follow-up and event rates are limited in these studies. Therefore, further studies with longer follow-up are needed.

2.7.2 Strengths and weaknesses of this study

This study reports up to date prevalence figures of DM, DR and STR, as well as trends from 1998 to 2018, in a clinically relevant form, which clinicians and managers leading hospital eye services can use in the management of services for diabetes and diabetic retinopathy. Our work is based on a cross-sectional analysis of primary care data and is therefore closer to routine practice. Our findings have also been verified against PHE, DESP, and other previously published figures (5-7, 14, 30, 37, 38). This is also the first observational IMRD based study to forecast the diabetes mellitus, diabetic retinopathy and sight-threatening retinopathy disease burden in the UK all together. While incorporating current evidence on the trend of underlying condition (DM), this study portrays a comprehensive analysis of the recent DR disease burden.

This study has not however been adjusted for the risk factors for the incidence/prevalence of DM or DR. Other limitations are possible coding errors, difficulties of missing data, and the potential risk of an overestimation of vision failure. The findings of this study should be interpreted within that context. Firstly, the

possible impact of coding errors, as well as subjectivity in documentation across a retrospective nationwide database involving several practices in different areas, cannot be precluded. This potential risk was minimised through a strict Read code selection process. The prevalence of severe DR was higher for those of South Asian and mixed ethnicity (9), therefore could have implications for local variations in its prevalence, and estimates could differ depending on the local ethnic mix. The potential impact of several concomitant medications on the course of DR was not captured in this study design. For the sake of future projections, estimates from individuals over 12 years old were applied to the whole population to calculate the final values, assuming that the number of DM patients under 12 is very low. Finally, we acknowledge that these projections are subject to the assumption that factors affecting the incidence, course and progression of the disease will remain stable over the next few years.

2.7.3 Strengths and weaknesses in relation to other studies, discussing particularly any differences in results

We wanted to verify our figures against data from Diabetic Eye Screening Programme which screens everyone from age 12 (39) and Mathur et al. work (9). Both these research studies used a cut off of over 12 years for their estimates. We also wanted our findings to be generalizable to the whole UK populations with diabetes including those under care of Diabetic Eye Screening Programme and Hospital Eye Services and generalisable internationally where majority of world population with diabetes is within one pool, without access to screening services. Limitations are that 2017 figures are not easily verifiable against PHE figures 2017

being over 17 years of age. So verification against that estimate is difficult and thus adds uncertainty to our UK forecast estimates

Gonzalez et al (14) reported an increasing prevalence of diabetes between the years 1996 and 2005 (10 years) based on THIN data analysis of patients aged 10 to 79 years old. They reported an overall increase of 54%. Our estimate between 1998 and 2005 (our data did not match the years) was 60%. In a Clinical Practice Research Datalink (CPRD) based study, Zghebi et al (38) found an overall increase of 64% in the patient population between 2004 to 2014, but this was limited to patients over 16 years old with T2DM. Our corresponding figures are 63%. Thus, our estimates fall midway between these two studies. Bagust is a future forecast for UK, but is limited to T2DM and is an underestimation (11). It projected T2DM figures for 2036 to be 1.1 million.

The PHE estimate for prevalence of diabetes in UK in 2017 arrived at by Quality and Outcome Framework figures was 3.7 million (5.6%) in those aged 17 years and over (37) and included diagnosed patients with diabetes. Our estimate of diagnosed patients with diabetes in 2017 of 3.4 million (crude prevalence of 5.2%) in over 12 years old population contrasts with the 2017 PHE figures. However previous analysis by our team has shown comparable figures with that of PHE when age restriction was 17 years and above. Similarly, PHE predicted the diabetes burden for 2025 to be 4.9 million for people aged over 16 years (40). It is not possible to make a direct comparison with our forecast of just under 4.3 million for 2025 because of our estimate being for people over 12 years of age. But given one would expect more it could mean the present study to be an underestimation. Alternatively, PHE figures could be an overestimation for 2017, because of the inbuilt assumptions in that

model. Our estimate for 2019 matches the quality and outcome estimate of 3.8 million.

IDF (41) estimated total diabetes prevalent cases (20 to 79 years old) to be 2.7 million in 2017, which is an underestimation when compared to PHE and our study.

A recent DR prevalence study focussed on lower risk patients with diabetes under screening services (9). The DR period prevalence in the Mathur's study (2004 to 2014) was found to be 48.4% for patients with T1DM and 28.3% for patients with T2DM, contrasting with point prevalence (2017) of 56.4% and 32.0% for patients with T1DM and T2DM respectively in our study. They also did not split the pathology into maculopathy and pre-proliferative categories, and did not include treatment and vision failure. Li et al 2019 (10) is the only study so far, that has projected DR till 2050. They estimated that 8.6 million people with diabetes (DR in 25% of the European population with T2DM and 50% with T1DM) will have diabetic eye disease inn 2050. The British studies included within this systematic review were based on diabetic screening services from pre-2009 (42) and pre-2003 data (7). Case definitions and patient pathways have since changed. Consequently their figures are a significant underestimation as compared to ours (710,510 vs 1,612,395 in 2030)

Other prevalence studies from the UK (5-7, 30) are compared with estimates from our study in detail for completeness in Appendix 10 and

Appendix 11. The majority of these UK studies are quite old, come from the screening programme setting, and do not deal with all of the categories of DR due to changed case definitions. Keenen et al (43) is a study based on work between 2007 and 2009 on hospital patients. They based their estimates of prevalence in eyes rather than patients, therefore, due to this heterogeneity, cannot be directly compared with our figures.

2.7.4 Meaning of the study: possible mechanisms and implications for clinicians or policymakers

Consecutive analyses over the course of over two decades provides information regarding the trend and severity of diabetic disease, and by a detailed analysis of different forms and severity groups, it captures the implications for the public health system. With the use of relevant outcomes, coupled with a prerequisite validation, the study provides a forecasting model which will be of use for clinicians and managers leading the professional services in planning the capacity to meet the increasing demand, and will guide public health strategy. Local demand can be calculated with the help of national figures provided by taking local factors into account.

Out of the 33.8% of total diabetic retinopathy in all patients with diabetes, 12.3% was made up of the STR. Those STR patients that actually needed treatment or experienced vision failure constituted a total of 5%. These figures reflect a high false-positive rate of referrals (50 - 70% as reported earlier (2, 44) and needs to be considered in the future relationship between diabetic retinopathy screening services and overburdened hospital eye services. Our estimated prevalence figures, in a clinically relevant form will help the clinicians and managers leading hospital eye services to optimise capacity planning for the increased demand.

2.7.5 Unanswered questions and future research

PHE used a prevalence model to predict the disease burden of diabetes in 2016 (45). The predictive factors they used were age, ethnicity, gender and deprivation index. To accommodate local variation in populations and practices, final calculations can be made using these predictive factors. The above-mentioned limitations of the study can be overcome by a future collaborative study linking diabetic retinopathy screening and hospital eye services data, with figures based on patient numbers and not eyes, to prevent heterogeneity among studies. Forecasting capacity needs is an area that should be repeated periodically with the help of the forecasting model presented.

2.8 Conclusion

In our study, the estimates suggested a trend of differential rise in prevalence rates in T1DM and T2DM. Overall, there is a continuing rise in the numbers of patients with DM and DR needing care. Preventive strategies and service planning can be based on these projected prevalence estimates to meet demand over the next ten years. Future forecasting will need repeating periodically to capture any external factors causing a change in the present trend.

Acknowledgements / Contributions: SH and KN designed the main study. AS performed data extraction. SH made the main contribution to the interpretation of data and wrote up the draft. RT designed and carried out the future projection model. AS, MP, DM, KT and KN made a contribution verbally or by critical revision of the draft. All authors approved the final version of the manuscript.

Conflict of Interest: All authors declare no conflict of interest.

Funding: There was no funding / financial support for this research.

2.9 References:

1. Bourne RRA, Jonas JB, Bron AM, Cicinelli MV, Das A, Flaxman SR, et al. Prevalence and causes of vision loss in high-income countries and in Eastern and Central Europe in 2015: magnitude, temporal trends and projections. The British journal of ophthalmology. 2018;102(5):575-85.
2. The Royal College of Ophthalmologists. The way forward age-related macular degeneration and diabetic retinopathy 2016

[ ]. Available from: https://www.rcophth.ac.uk/wp-content/uploads/2015/10/RCOphth-The-Way-Forward-AMD-300117.pdf, Accessed 2019, July 20.
3. Sabanayagam C, Banu R, Chee ML, Lee R, Wang YX, Tan G, et al. Incidence and progression of diabetic retinopathy: a systematic review. The Lancet Diabetes & Endocrinology. 2019;7(2):140-9.
4. Buchan JC, Norman P, Shickle D, Cassels-Brown A, MacEwen C. Failing to plan and planning to fail. Can we predict the future growth of demand on UK Eye Care Services? Eye (London, England). 2019;33(7):1029-31.
5. Thomas RL, Dunstan FD, Luzio SD, Chowdhury SR, North RV, Hale SL, et al. Prevalence of diabetic retinopathy within a national diabetic retinopathy screening service. The British journal of ophthalmology. 2015;99(1):64-8.
6. Misra A, Bachmann MO, Greenwood RH, Jenkins C, Shaw A, Barakat O, et al. Trends in yield and effects of screening intervals during 17 years of a large UK community-based diabetic retinopathy screening programme. Diabetic medicine : a journal of the British Diabetic Association. 2009;26(10):1040-7.
7. Younis N, Broadbent DM, Harding SP, Vora JP. Incidence of sight-threatening retinopathy in Type 1 diabetes in a systematic screening programme. Diabetic Medicine. 2003;20(9):758-65.
8. Looker H, Nyangoma S, Cromie D, Olson J, Leese G, Black M, et al. Diabetic retinopathy at diagnosis of type 2 diabetes in Scotland. Diabetologia. 2012;55(9):2335-42.
9. Mathur R, Bhaskaran K, Edwards E, Lee H, Chaturvedi N, Smeeth L, et al. Population trends in the 10-year incidence and prevalence of diabetic retinopathy in the UK: a cohort study in the Clinical Practice Research Datalink 2004-2014. BMJ open. 2017;7(2):e014444.
10. Li JQ, Welchowski T, Schmid M, Letow J, Wolpers C, Pascual-Camps I, et al. Prevalence, incidence and future projection of diabetic eye disease in Europe: a systematic review and meta-analysis. Eur J Epidemiol. 2019.
11. Bagust A, Hopkinson PK, Maslove L, Currie CJ. The projected health care burden of Type 2 diabetes in the UK from 2000 to 2060. Diabetic medicine : a journal of the British Diabetic Association. 2002;19 Suppl 4:1-5.
12. Candler T, Mahmoud O, Lynn R, Majbar A, Barrett T, Shield J. Continuing rise of type 2 diabetes incidence in children and young people in the UK. Diabetic Medicine. 2018;35(6):737-44.
13. Patterson CC, Harjutsalo V, Rosenbauer J, Neu A, Cinek O, Skrivarhaug T, et al. Trends and cyclical variation in the incidence of childhood type 1 diabetes in 26 European centres in the 25 year period 1989–2013: a multicentre prospective registration study. Diabetologia. 2019;62(3):408-17.
14. Gonzalez EL, Johansson S, Wallander MA, Rodriguez LA. Trends in the prevalence and incidence of diabetes in the UK: 1996-2005. Journal of epidemiology and community health. 2009;63(4):332-6.
15. Public Health England. Diabetes Prevalence 2019 2019 [Available from: https://www.diabetes.org.uk/professionals/position-statements-reports/statistics/diabetes-prevalence-2019, Accessed 2020, May 12.
16. Maguire A, Blak BT, Thompson M. The importance of defining periods of complete mortality reporting for research using automated data from primary care. Pharmacoepidemiology and drug safety. 2009;18(1):76-83.

17. Blak BT, Thompson M, Dattani H, Bourke A. Generalisability of The Health Improvement Network (THIN) database: demographics, chronic disease prevalence and mortality rates. Informatics in primary care. 2011;19(4):251-5.
18. Martin-Merino E, Fortuny J, Rivero E, Garcia-Rodriguez LA. Validation of diabetic retinopathy and maculopathy diagnoses recorded in a U.K. primary care database. Diabetes care. 2012;35(4):762-7.
19. Martin-Merino E, Fortuny J, Rivero-Ferrer E, Garcia-Rodriguez LA. Incidence of retinal complications in a cohort of newly diagnosed diabetic patients. PloS one. 2014;9(6):e100283.
20. Martin-Merino E, Fortuny J, Rivero-Ferrer E, Lind M, Garcia-Rodriguez LA. Risk factors for diabetic retinopathy in people with Type 2 diabetes: A case-control study in a UK primary care setting. Primary care diabetes. 2016;10(4):300-8.
21. McEwan P, Bennett H, Ward T, Bergenheim K. Refitting of the UKPDS 68 risk equations to contemporary routine clinical practice data in the UK. Pharmacoeconomics. 2015;33(2):149-61.
22. Dave S, Petersen I. Creating medical and drug code lists to identify cases in primary care databases. Pharmacoepidemiology and drug safety. 2009;18(8):704-7.
23. Public Health England. Diabetes Prevalence Model 2016 [Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/612306/Diabetesprevalencemodelbriefing.pdf, Accessed 2019, July 12.
24. Haider S, Sadiq SN, Moore D, Price MJ, Nirantharakumar K. Prognostic prediction models for diabetic retinopathy progression: a systematic review. Eye. 2019;33(5):702-13.
25. Rob J Hyndman and George Athanasopoulos. Forecasting: Principles and Practice, [Available from: https://otexts.com/fpp2/tspatterns.html.
26. Adams RJ, Tucker G, Hugo G, Hill CL, Wilson DH. Projected future trends of hospital service use for selected obesity-related conditions. Obesity research & clinical practice. 2008;2(2):133-41.
27. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2019 [Available from: http://www.R-project.org/.
28. Office For national Statistics. National population projections: 2018-based, [Figure 1: UK population projected to rise to 69.4 million by mid 2028 and to 72.4 million by mid 2043]. 2019 [Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationprojections/bulletins/nationalpopulationprojections/2018based, Accessed 2019, July 12.
29. Public Health England. Diabetic Eye Screening: 2016 to 2017 Data 2017 [Available from: https://www.gov.uk/government/publications/diabetic-eye-screening-2016-to-2017-data, Accessed 2019, July 12.
30. Minassian DC, Owens DR, Reidy A. Prevalence of diabetic macular oedema and related health and social care resource use in England. The British journal of ophthalmology. 2012;96(3):345-9.
31. World Health Organization. Blindness and vision impairment - key facts, 2019 [Available from: https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment, accessed August 06, 2020.
32. Łuczyński W, Szypowska A, Głowińska-Olszewska B, Bossowski A. Overweight, obesity and features of metabolic syndrome in children with diabetes treated with insulin pump therapy. European journal of pediatrics. 2011;170(7):891-8.
33. Egro FM. Why is type 1 diabetes increasing? Journal of molecular endocrinology. 2013;51(1):R1-13.
34. Fowler MJ. Microvascular and Macrovascular Complications of Diabetes. Clinical Diabetes. 2008;26(2):77-82.
35. DeFronzo RA. Combination therapy with GLP-1 receptor agonist and SGLT2 inhibitor. Diabetes Obes Metab. 2017;19(10):1353-62.

36. Wang T, Hong JL, Gower EW, Pate V, Garg S, Buse JB, et al. Incretin-Based Therapies and Diabetic Retinopathy: Real-World Evidence in Older U.S. Adults. Diabetes care. 2018;41(9):1998-2009.
37. Diabetes UK. Diabetes Prevalence 2017, Quality and Outcomes Framework 2018 [Available from: https://www.diabetes.org.uk/professionals/position-statements-reports/statistics/diabetes-prevalence-2017, Accessed 2019, July 12.
38. Zghebi SS, Steinke DT, Carr MJ, Rutter MK, Emsley RA, Ashcroft DM. Examining trends in type 2 diabetes incidence, prevalence and mortality in the UK between 2004 and 2014. Diabetes Obes Metab. 2017;19(11):1537-45.
39. Scanlon PH. The English National Screening Programme for diabetic retinopathy 2003-2016. Acta Diabetol. 2017;54(6):515-25.
40. England PH. Research and analysis - Diabetic eye screening: 2016 to 2017 data.
41. Cho NH, Shaw JE, Karuranga S, Huang Y, da Rocha Fernandes JD, Ohlrogge AW, et al. IDF Diabetes Atlas: Global estimates of diabetes prevalence for 2017 and projections for 2045. Diabetes research and clinical practice. 2018;138:271-81.
42. Thomas RL, Dunstan F, Luzio SD, Roy Chowdury S, Hale SL, North RV, et al. Incidence of diabetic retinopathy in people with type 2 diabetes mellitus attending the Diabetic Retinopathy Screening Service for Wales: retrospective analysis. BMJ. 2012;344:e874.
43. Keenan TD, Johnston RL, Donachie PH, Sparrow JM, Stratton IM, Scanlon P. United Kingdom National Ophthalmology Database Study: Diabetic Retinopathy; Report 1: prevalence of centre-involving diabetic macular oedema and other grades of maculopathy and retinopathy in hospital eye services. Eye (London, England). 2013;27(12):1397-404.
44. Jyothi S, Elahi B, Srivastava A, Poole M, Nagi D, Sivaprasad S. Compliance with the quality standards of National Diabetic Retinopathy Screening Committee. Primary care diabetes. 2009;3(2):67-72.
45. Public Health England. Diabetes Prevalnce Estimates for CCG's by GP registered populations 2016 [Available from: https://www.gov.uk/government/publications/diabetes-prevalence-estimates-for-local-populations, Accessed 2020, May 12.

2.10 Appendices for Chapter 2

3 Chapter 3: Systematic review

This systematic review aimed to summarize the characteristics and performance of existing models in predicting the progression of diabetic retinopathy and their applicability for higher-risk diabetic retinopathy patients under hospital care to predict progression. The aims were fulfilled through the following objectives

1. A systematic review of the literature to produce a comprehensive database of existing prognostic prediction models able to predict diabetic retinopathy progression to need for treatment stage and /or loss of vision and any of their external validation studies.
2. Critical appraisal of the modelling studies using high-quality, rigorous methods.
3. To apply validated tools to assess the models' applicability to high-risk hospital diabetic retinopathy patients

3.1 Systematic review

Chapter 3 is presented in the publication format as published in the “Eye” journal.

Haider S, Sadiq SN, Moore D, Price MJ, Nirantharakumar K. Prognostic prediction models for diabetic retinopathy progression: a systematic review. Eye (Lond). 2019 May;33(5):702-713. DOI: 10.1038/s41433-018-0322-x. Epub 2019 Jan 16. Review. PubMed PMID: 30651592.

Eye
https://doi.org/10.1038/s41433-018-0322-x

The ROYAL COLLEGE of OPHTHALMOLOGISTS

REVIEW ARTICLE

Prognostic prediction models for diabetic retinopathy progression: a systematic review

Sajjad Haider[1] · Salman Naveed Sadiq[2] · David Moore[1] · Malcolm James Price[1] · Krishnarajah Nirantharakumar[1]

Received: 1 September 2018 / Revised: 10 November 2018 / Accepted: 19 November 2018

Abstract

With the increasing incidence of diabetic retinopathy and its improved detection, there is increased demand for diabetic retinopathy treatment services. Prognostic prediction models have been used to optimise services but these were intended for early detection of sight-threatening retinopathy and are mostly used in diabetic retinopathy screening services. We wanted to look into the predictive ability and applicability of the existing models for the higher-risk patients referred into hospitals. We searched MEDLINE, EMBASE, COCHRANE CENTRAL, conference abstracts and reference lists of included publications for studies of any design using search terms related to diabetes, diabetic retinopathy and prognostic models. Search results were screened for relevance to the review question. Included studies had data extracted on model characteristics, predictive ability and validation. They were assessed for quality using criteria specified by PROBAST and CHARMS checklists, independently by two reviewers. Twenty-two articles reporting on 14 prognostic models (including four updates) met the selection criteria. Eleven models had internal validation, eight had external validation and one had neither. Discriminative ability with c-statistics ranged from 0.57 to 0.91. Studies ranged from low to high risk of bias, mostly due to the need for external validation or missing data. Participants, outcomes, predictors handling and modelling methods varied. Most models focussed on lower-risk patients, the majority had high risk of bias and doubtful applicability, but three models had some applicability for higher-risk patients. However, these models will also need updating and external validation in multiple hospital settings before being implemented into clinical practice.

Introduction

There has been a global increase in the number of people with diabetes, rising from 108 million in 1980 to 422 million in 2014 [1]. The detection of retinopathy has also increased through better population screening [2]. While services may be organised differently from country to country, the care pathways are likely to be similar with patients at higher risk being provided closer monitoring and care. In the United Kingdom, diabetic retinopathy (DR) services are organised into diabetic eye screening programmes (DESP) and hospital eye services. DESP provides annual diabetic retinopathy screening to all patients with diabetes above 12 years of age. Screening uptake in the year 2015/16 was 82.5% [3]. If the screening findings indicate low risk (retinopathy stage R0, R1 and M0, see Fig. 1 and Table A1) they are retained within the DESP and reviewed yearly. When the disease progresses to sight-threatening diabetic retinopathy (STR) stage—(R2, R3 or M1), they are referred to the hospital eye services for closer observation and treatment (Fig. 1). In the United States, yearly screening is recommended to all type 2 diabetes mellitus (T2DM) patients at diagnosis and afterwards, type 1 diabetes mellitus (T1DM) patients are recommended to have screening on an annual basis, commencing 5 years after diagnosis [4, 5].

In the United Kingdom, ~50% of referrals with STR do not need intervention and are observed in the hospital eye

These authors are Joint Senior Authors: David Moore, Malcolm James Price, Krishnarajah Nirantharakumar

Supplementary information The online version of this article (https://doi.org/10.1038/s41433-018-0322-x) contains supplementary material, which is available to authorized users.

✉ Sajjad Haider
drsajjadhaider@yahoo.co.uk

✉ Malcolm James Price
M.Price.2@bham.ac.uk

[1] Institute of Applied Health Research, University of Birmingham, Birmingham, UK

[2] York Teaching Hospital NHS Foundation Trust, York, UK

Published online: 16 January 2019

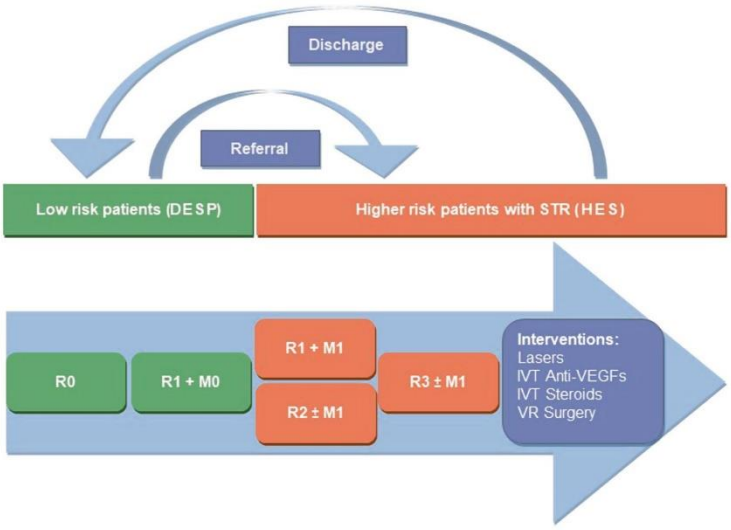


Fig. 1 Patient flow diagram[a]. [a]See Table A1 for more details on classification of diabetic retinopathy. R0, no retinopathy; R1, background retinopathy; R2, pre-proliferative retinopathy; R3, proliferative retinopathy; M0, no maculopathy; M1, maculopathy present; DESP, diabetic eye screening programme; HES, hospital eye services; STR, sight-threatening retinopathy; IVT, intra-vitreal therapy; VEGF, vascular endothelial growth factor; VR, vitreoretinal

service for a variable period of time [6], placing extra burden on these services. Various modifications to improve the service have been proposed, like digital surveillance using optical coherence tomography (OCT), and virtual clinics within DESP [6].

Prognostic modelling/nomograms can aid decision-making [7]. There have been successful attempts at optimising diabetic screening services through stratification of patients by risk of progression of DR using a prognostic prediction model [8, 9]. A similar approach will also help optimise hospital eye services. Such a model combines multiple prognostic factors to predict which patients are at higher risk of progression to visual loss and thus need closer observation or treatment to prevent loss of vision [10–12]. The aim of this systematic review is to summarise the characteristics and performance of existing models in predicting progression of retinopathy and their applicability for higher-risk DR patients under hospital care to predict need for treatment or loss of vision.

Methods/design

Patient group, selection criteria/study design/inclusion criteria

Studies were included in the review of any design that developed, updated, validated, compared or evaluated a prognostic prediction model/tool, using multiple prognostic factors to predict the risk of progression of diabetic retinopathy and/or vision loss. The searches covered studies reporting the development of a model, validation of a model and impact of a model in practice.

There was no restriction on the age of participants or type of diabetes. The primary outcomes of our review were predictive accuracy and applicability of the prognostic prediction models/tools in relation to progression of diabetic retinopathy from stages that required hospital referral (R2, R3 or M1) to treatment requiring stage or vision loss.

Search strategy and selection criteria

As prognostic model studies can be difficult to identify, several approaches were used. MEDLINE, EMBASE and COCHRANE CENTRAL (up to March 5, 2017) were searched using index and free text terms for diabetes, retinopathy and prognostic models. A sample search strategy for MEDLINE is shown in Table A2. We did not apply any restrictions on language or dates of publication. We also searched abstracts from the following national and international conferences from 2014 to March 5, 2017.

- The Royal College of Ophthalmologists, American Academy of Ophthalmology, European Society of Retina Specialists (EURETINA), European Society of Ophthalmology (SOE) and Association for Research in Vision and Ophthalmology (ARVO)

- Diabetes conferences. American Diabetic Association (ADA), Diabetes UK and International Diabetes Federation

Reference lists of included studies were screened for additional studies and authors of relevant conference abstracts were identified and their publication lists checked for additional relevant studies using Pubmed, Google Scholar and Scopus.

Search results were recorded in Endnote (version x7.4Clarivate Analytics) and duplicate entries removed. Titles (and abstracts where available) were screened for relevance using predefined screening criteria. Full texts of all potentially relevant articles were obtained and assessed against the selection criteria. Reviewer decisions, including the reason why studies were excluded from the review were recorded.

Data extraction and quality assessment

The information extracted from each study included study characteristics, source of data, study design, participants characteristics, candidate predictors, their handling, outcomes assessed, sample size, missing data and its handling, modelling methods, methods for selection of final predictors, model performance measures (discrimination, calibration and classification measures), model validation and presentation of the final prediction models. Authors were contacted where necessary (mostly for reporting deficiencies). Some models have dealt with multiple outcomes other than of interest to our review. We only considered the ocular outcomes.

Critical appraisal was carried out using a risk assessment form by combining PROBAST [10] and CHARMS [13] checklists. Risk of bias and applicability was assessed mainly by using PROBAST tool [10], however, since the tool was being piloted, CHARM checklist [13] was also used to further refine assessment of the studies.

Study selection, data extraction and risk of bias assessment were carried out by two reviewers independently with disagreements resolved by mutual discussion or with a third reviewer where required.

Analysis

As some models were the subject of more than one study, they were grouped by specific models and organised by whether describing model development, internal validation or external validation. Details on each model are presented in evidence tables and narrative summaries are given on key model features (population, samples, predictors and performance).

This review was registered prospectively with PROSPERO (registration number CRD42017057767) [14] and is reported here in accordance with the PRISMA guidelines [15].

Results

Volume of the research literature available

The searches yielded 12,118 records of which 4893 were duplicates. After screening titles and abstracts, 62 relevant articles were identified of which 22 met all selection criteria and are included in the review (Fig. 2).

Overview

This systematic review analysed 14 prognostic prediction modelling studies, including four updates. Six models had both internal and external validation, five models performed only internal validation and two were only validated in external datasets. One model lacked both internal and external validation. No studies assessing the impact of a model were identified. All studies were published within the last two decades. Figure 3 shows how the studies fit into the evolution of the modelling process. The characteristics of individual studies are summarised in Table 1.

Population

The majority of studies have been conducted post year 2000 ($n = 19$). The latest data used were from 2009 to 2014 in ISDR study [16], and the oldest data were part of the UKPDS OM1 study (1977–1997) [17], though this model has been validated in recent data [18, 19]. Two studies were in T1DM [20, 21], eight in T2DM [17–19, 22–26] and 11 in mixed populations [8, 16, 27–34], with one unspecified [35]. Population studied ranged from newly diagnosed patients [17, 22, 35] to those with a relatively severe form of the disease [19]. The included studies have used large routinely available databases such as The Health Improvement Network (THIN) [19], Q Research and Clinical Practice Research Datalink (CPRD) [28], US claims databases [35], hospital databases [21, 26], diabetic screening data and clinical trials research data [17, 20, 22, 24].

Sample size, events per variable (EPV) and follow-up

Sample size ranged from 1441 [27] to 454,575 [28] in primary development studies and from 200 [21] to 206,050 [28] in validation studies. EPVs ranged from as low as 0.86 [20] to 424.27 [28]. External validation samples were generally small and yielded low EPV. Lack of reporting

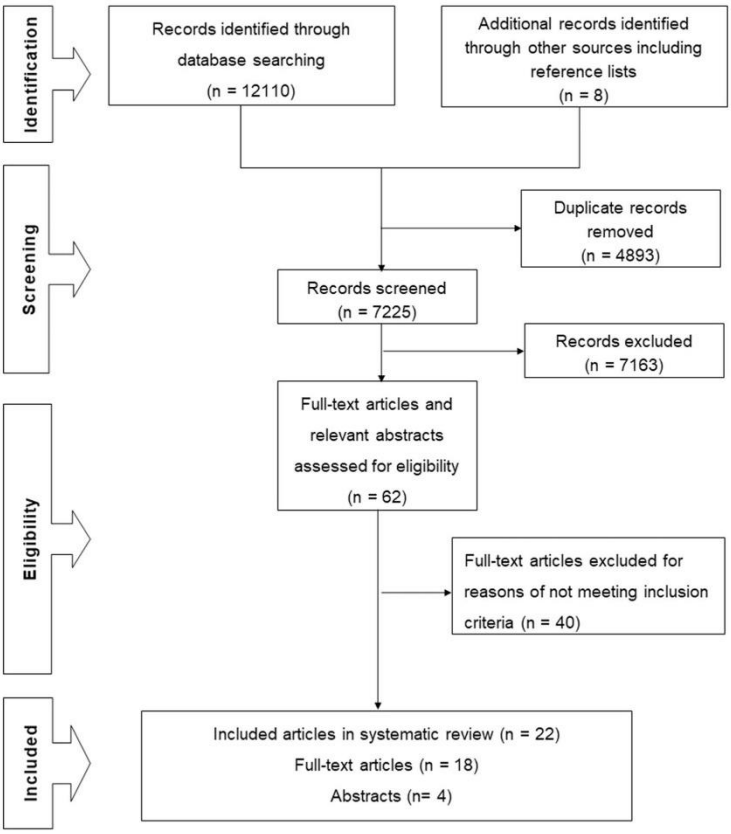


Fig. 2 PRISMA flow diagram showing the search process from the volume of the research literature available

affected assessment of EPV in two models [20, 23]. Median duration of follow-up time was 5.1 years, ranging from 1.1 [30] to 17.6 years [22].

Predictors

The models contained 78 different candidate predictors, which can be grouped into nine broad groups (Table A3 in supplementary material). The number of candidate predictors ranged from 3 [32, 33] to 51 [27] in any one model and their selection was mostly based on literature reviews or clinical intuition. Forty nine different predictors appeared in the final models. The median number of final predictors used in a given model was 5 with a range of 2 [17] to 14 [35].

Standardised definitions and measurement methods were generally used for predictors. Predictor values were recorded or measured at baseline cohort entry, or soon after. Categorisation of the continuous predictors was mostly avoided, but was not always reported [20, 24, 27, 34]. The method of selection of the final predictors was reported in 9 out of 14 primary development studies and was typically performed using backward elimination [16, 20, 22, 24, 28]. Four models [29, 31, 32, 35] used the full model approach by using all candidate predictors.

Biochemical predictors were the most common (Table A3 in supplementary material). HbA1c was the most popular predictor appearing in all studies except one [32]. It was followed by duration of diabetes ($n = 7$). Age was another important predictor used in various forms—as age ($n = 4$), age at diabetes diagnosis ($n = 4$) and age at DR diagnosis ($n = 1$). Half of the models used local predictors/ocular signs. One model [32] only used ocular signs as the sole set of predictors. They categorised baseline DR into three predictors (R0 in eyes, R1 one eye or R1 both eyes). This trend continued in subsequent related studies [34, 36],

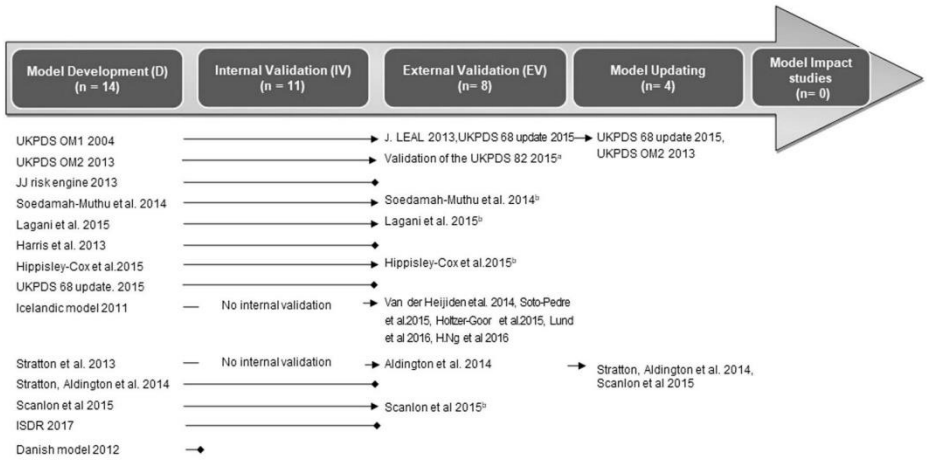


Fig. 3 Flow diagram showing the modelling process and studies in each part of this process (the arrow indicates continuation of the process; the box implies cessation of further progress). [a]Performed Internal Validation on UKPDS OM2 2013. [b]External Validation reported in same study

except R0 was later omitted and biochemical predictors were added.

Outcomes

Outcome definition varied from blindness only [17–20, 22, 23, 28], to STR [8, 25, 26, 29, 30, 32–34, 36], retinopathy progression [8, 24, 27, 35], need for treatment [31] and referable DR [16]. The definition of blindness was mostly defined as best corrected vision of less than 6/60. In some models, STR was used interchangeably with treatment [29, 31]. Validation studies defined the outcomes in a similar way to their respective model development studies.

Statistical model

The final models were based on Cox proportional hazards model in most studies [22, 24, 27, 28, 32, 34–36]. Weibull proportional hazards model was used in four studies [17, 19, 20, 29]. Logistic regression [31] and continuous-time Markov mathematical models [16] were used by one study each.

Model performance/evaluation

Performance measures were poorly reported in almost half of the studies. The remainder reported either Harrell's concordance index [20, 24, 28], area under the curve or both in the case of the ISDR model [16]. Model calibration where reported was mainly in the form of calibration plots. Five primary modelling studies [17, 22, 31, 32, 35] and one validation study [26] failed to report any performance measures. For discriminatory power, one model [16] reported a c-statistic of >0.80 (0.91) and the remainder reported moderate discrimination of 0.614 to 0.79, except for one poor c-statistic value of 0.57 [19].

Missing data

The amount of missing data varied in most studies [16, 18–20, 24, 27–29, 36], with some reporting more than 50% of participants missing at least one predictor value [19, 28]. Proportions of missing data have not always been reported [16, 18, 24, 27]. The mechanisms used to handle missing data included the last observation carried forward (average or mode) [18, 27, 29, 36], complete case analysis [36] and multiple imputations [16, 19, 20, 24, 28].

Summary of bias

Four modelling studies had moderate [16, 19, 28, 36] and one had low risk of bias [29] (Table 2). Risk of bias was unclear in one modelling study [34] taken from conference abstracts (lack of reporting). The remaining studies were at high risk of bias. High risk of bias was mostly due to low number of outcomes per variable (EPV < 10), lack of information on missing data, absence of external validation and lack of reporting of relevant performance measures. All validation studies have shown good discrimination ability, but were mostly at high risk of bias due to small numbers (<100) of outcomes [13] or lack of reporting [33].

Table 1 Characteristics of model development and validation studies included in the systematic review

Study ID	Study dates	Participants DM-type source	Predictors (*n*) candidate final	Outcome follow-up (years)	Sample size (*n*)	Events (*n*)	Events per variable	Model performance /evaluation	Presentation
Models development studies with external validation									
UKPDS OM1 2004 [17] (D, IV)	1977–1997	T2DM UKPDS trial	14 2	One eye blindness median 10.3y	3642	104	7	Not reported	Equation
Leal et al., 2013 [18] (EV of OM1)	1997–2007	T2DM PTM data	Same		4031	101		C-statistics (95% CI) 0.60 (0.55–0.65)	NA
UKPDS 68 update 2015 [19] (update, IV and EV of OM1)	2000–2009	T2DM (two risk groups) THIN database	14 12	Blindness low risk (LR) 5.1y Intermediate risk (IR) 6.4y	LR 54,169 IR 68,990	LR 176 IR 607	56	C-statistics LR: IV 0.59 and EV 0.57 IR: IV 0.59 and EV 0.58 Calibration plots	Risk groups (low and intermediate)
UKPDS OM2 2013 [22] (D, IV)	1997–2007	T2DM PTM data	24 7	One eye blindness median 17.6y	5102	271	11	Predicted curves in IV	Hazard ratios
UKPDS 82 2015 [23] (EV of OM2)	Not reported	T2DM	Same		Not reported	Not reported		MAPE (%) = 20.31 and $R^2 = 0.96$ Calibration plots	NA
Soedamah-Muthu et al., 2014 [20] (D, IV) EV in the same study three cohorts	1989–1999 1998–1999 1994–2009 2000–2002	T1DM EURODIAB trial data EDC FinnDiane CACTI	15 5 Same	Blindness median 7.4y Median 8.1y Median 7.5y Median 7.3y	1973 554 2999 580	12 29 Not reported 5	<1 (0.8)	Harrell's C-index: IV 0.74 C-statistics 0.79 0.82 0.73 Calibration plots	Equation score charts Kaplan–Meier plots for high, intermediate and low risk score groups
Lagani et al.,[a] 2015 [27] (D, IV) EV in the same study	1983–1993 2004–2014	T1DM DCCT, EDIC T1DM and T2DM Chorleywood, UK	51 5 Same	Retinopathy event[b] Mean 6.5y	1441(IV) T1DM 36 T2DM 294	969(IV) T1DM 17 T2DM 70	Not applicable	Concordance index: 0.66 ($p < 0.001$) T1DM 0.72 ($p = 0.002$) T2DM 0.55 ($p = 0.119$)	Equation recalibration NA
Hippisley-Cox et al., 2015 [28] (D, IV) EV in the same study	1998–2014	T1, T2DM Q Research CPRD, UK	18 9 Same	Blindness in one or both eyes Total 15y	454,575 (D) 142,419 (IV) 206,050 (EV)	8063 2651 2845	448	Harrell's C-index Women 0.73 (0.71–0.74) and men 0.75 (0.73–0.77) (IV) 0.73 (0.72–0.75) in both gender (EV) Calibration plots	Web calculator NA
Icelandic model 2011 [29] (D, EV) Five separate EV studies for Icelandic model*	Not reported	T1, T2DM Icelandic eye screening /Aarhus diabetic database	7 7	STR f/u not reported	5199	149	21	C-statistics (95% CI) 0.76 (0.74–0.78) and Calibration graph	Equation nomogram mobile app
*van der Heijden 2014 [25]	1998–2010	T2DM, Dutch	Same	STR Mean 4.4y	3319	76		0.83 (0.74–0.92)	NA
*Soto-Pedre et al., 2015 [30]	Not reported	T1, T2DM, Spain	Same	STR Median 1.1y	508	16		0.74 (0.62–0.85)	NA
*Holtzer-Goor, 2015 [26]	Not reported	T2DM, Dutch Hospital	Same	STR Mean 3.3y	888	47		Not reported	NA
*Lund et al., 2016 [8]	2010–2012	T1, T2DM, DESP, UK	Same	STR Total 2y	9687	531		T1DM 0.70 (0.67–0.73) and T2DM 0.80 (0.78–0.81)	NA

Table 1 (continued)

Study ID	Study dates	Participants DM-type source	Predictors (*n*) candidate final	Outcome follow-up (years)	Sample size (*n*)	Events (*n*)	Events per variable	Model performance /evaluation	Presentation
Models development studies with external validation									
*Ng et al., 2016 [21]	Not reported	T1DM, Dutch Hospital	Same	STR Mean 6.9y	200	22		0.75 (0.65–0.86)	NA
Stratton et al., 2013 [32] (D)	2005–2010	T1, T2DM DESP, UK	3 3	STR median 2.8y	14,554	803	268	Not reported	Kaplan–Meier plot
Aldington et al., 2014 [33] (EV)	Not reported	Two screening programmes, UK	Same	Total 4y	24,509 32,987	2328 3991		C-statistics 0.76 0.80	NA
Scanlon et al., 2015 [36] (D, IV) EV in the same study three programmes	2005–2011	T1, T2DM DESP, Gloucestershire (D)(IV) East Anglia South London Nottinghamshire	13 4 Same	STR total 5y Median 2.7y 3.8y 4.2y	7012(D) 5778(IV) 17,634 1223 1083	606(D) 490(IV) 845 94 81	47	C-statistics (95% CI) 0.79 (0.76–0.81) IV 0.78 (0.76–0.80) 0.84 (0.78–0.88) 0.82 (0.71–0.90)	Risk score NA
Models development studies with internal validation but no external validation									
JJ risk engine 2013 [24] (D, IV)	Not reported	T2DM, Japan Trial data JDCS, and J-EDIT	16 5	Progression of retinopathy Mean 7.2y	1748	178	11	Harrell's C-index (95% CI) 0.61 (0.52–0.70) Hosmer–Lemeshow $p = 0.13$	Equation web app
Harris et al., 2013 [35] (D, IV)	2001–2009	DM type not specified Claims database, US	14 14	Progression to PDR Median 1.7y	4617	307	22	Not reported	Equation risk score
Stratton et al., 2014 [34] (D, IV)	Not reported	T1, T2DM DESP, Gloucestershire (D)(IV)	6 5	STR Total 4y	6449(D) 5460 (IV)	555(D) 496(IV)	93	C-statistics (95% CI) 0.79 (0.78–0.81)	Not reported
ISDR 2017 [16] (D, IV)	2009–2014	T1, T2DM DESP, Liverpool and primary care data were combined	19 5	Referable DR Total 5y	11,806	388	20	Harrell's concordance index: 0.69 at 24 months AUC (95% CI): 0.91 (0.87–0.94) Calibration curve	Equation
Models development studies with no clear validation (performance testing only)									
Danish model 2012 [31] (D) Unclear if validated	1994–2007	T1, T2DM Arhus database, Denmark	7 7	Treatment requiring DR f/u not reported	5311 1372	559 208	80	Not reported	Equation

PTM post-trial monitoring data, *DESP* diabetes eye screening programme, *T1DM* type 1 diabetes mellitus, *T2DM* type 2 diabetes mellitus, *D* development, *IV* internal validation, *EV* external validation, *U* model update, *STR* sight-threatening retinopathy, *PDR* proliferative diabetic retinopathy, *MAPE* mean absolute percentage error, R^2 coefficient of determination/validation, *THIN* The Health Improvement Network, *CPRD* Clinical Practice Research Datalink, *DCS* Diabetes Care System West-Friesland, *Sn* sensitivity, *Sp* specificity, *LR* low risk, *IR* intermediate risk

[a]Figures from DCCT trial and in electronic supplementary material (not given in the publication)

[b]Any worsening in the retina condition that lasted at least 6 months

Low risk of bias: The Icelandic model [29] used good sample size ($n = 5199$) and EPV of 21·29 (>10). Moreover, it is the only model to have multiple external validation studies ($n = 5$). Development of the model was a bit unusual, based on hazard ratios from UKPDS and Wisconsin studies and prevalence figures from Icelandic diabetic eye screening programme. They 'empirically' tested this model in Aarhus diabetic database, which could be interpreted as an external validation.

Moderate risk of bias: Hippisley-Cox et al. [28] used a large cohort with a long follow-up and good reporting. However, there was a lack of formal adjudication of outcomes and a high number of missing variable values (up to 80.9%), though dealt by multiple imputation. The study by Scanlon et al. [36] has sound methodology, logical process of evolution and pragmatic decisions about predictors. However, there is lack of multiple imputations for missing data. The recently published ISDR model [16] has no external validation yet (which according to PROBAST is high risk of bias). We assigned it a moderate risk of bias because of a large sample ($n = 11,806$) and a robust internal validation. UKPDS 2015 [19] lacks external validation but was judged to have moderate risk of bias as it had a large sample size and used multiple imputation.

Applicability of the models

The target population and the outcomes of interest in our review were progression of diabetic retinopathy reaching the treatment requiring stage or visual loss (blindness or low vision) in patients under close observation in a hospital setting (Table 2). In the models identified, the context was broadly the early, low-risk part of the disease pathway. The population was largely a mixed diabetic population of all ages and any duration of diabetes but mainly early-onset diabetic retinopathy. None of the models were fully applicable to our review question with regard to population characteristics or outcomes used.

Even though five out of 14 models used blindness as the outcome (defined as corrected vision of 6/60 or less) [17, 19, 20, 22, 28], these models were mostly designed for multiple outcomes. Therefore, their predictor sets were also less specific for the outcomes needed for DESP or hospital eye service population. Despite having blindness as their outcome, these limitations make them high risk for applicability. Lagani et al. [27] model also focussed on multiple outcomes and defined outcome as any diabetic retinopathy event. Therefore, this was also classified as high risk. JJ risk Engine [24] model used predictors and outcomes more relevant to DESP or hospital eye service population but their participants were only patients with type 2 diabetes. We therefore assigned it as medium risk for applicability.

There was good applicability to low-risk diabetic screening patients and partial applicability for higher-risk hospital patient population in the remaining seven models [16, 29, 31, 32, 34–36]. Mehlsen et al. [31] used the outcome of DR progression to treatment-requiring stage and another study [35] used the outcome of DR progression. The remainder used STR/referable DR [16, 29, 32, 34, 36]. Only three models had a moderate to low risk of bias and also have low risk for applicability [16, 29, 36].

There were 11 different types of final predictors in these three models. Duration of diabetes and HBA1c were common among all three, and systolic blood pressure was used by two models [16, 29]. Other predictors included in these three models were presence, grade of diabetic retinopathy [29], presence of background diabetic retinopathy in one or both eyes [36], gender [29], type of diabetes [29], age at diagnosis [16] and total serum cholesterol [16].

Because of the heterogeneity introduced due to differences in populations, outcome measurements and the context in which they were studied, a meta-analysis could not be carried out.

Discussion

This systematic review summarises the details of 14 predictive model development studies, including four updates. Most of the studies dealt with the question of diabetic retinopathy progression up to the level of referral to the hospital for treatment or closer observation (lower-risk part of the disease pathway). The perspective has been largely individual patients' risk stratification in diabetic screening services using diabetic screening databases (12 out of 22 studies). Only very limited evidence was fully applicable to high-risk patients. Five modelling studies had moderate to low risk of bias and out of them, only three studies also had potential for applicability [16, 29, 36].

To our knowledge, no systematic review looking at the predictive accuracy and applicability of predictive models for patients with DR beyond referable sight-threatening diabetic retinopathy has been published. Our review considered all levels of the modelling process from model development through to validation studies. We have included all diabetic populations (both type 1 and type 2 diabetic patients). Thorough electronic and manual searches were conducted. Following up authors of relevant conference abstracts identified two studies that were not found by other means, one of which was published prior to the searches of bibliographic databases. Models from Europe, United States and Japan are included giving a global picture.

Some studies failed to report important information such as sample size and EPV [23], predictor handling, mathematical algorithm/equation (22, 28), follow-up period

Table 2 Summary of bias assessment and applicability as per PROBAST tool mainly and adding CHARMS input where necessary

Study ID	Participants	Predictors	Outcome	Sample size and participant flow	Model development and evaluation	External validation	Overall bias High Moderate Low	Applicability High concern Moderate concern Low concern	Comments for bias
UKPDS OM1 [17] 2004	L	L	L	H	H	Yes	H	H	Low EPV Did not report relevant performance measures[a]
J. Leal et al. [18] 2013	L	L	L	H	L	NA	H	H	Low EPV
UKPDS 68 update [19] 2015	L	L	L	L	L	No	M	H	High number of missing values (up to 57%) but dealt with multiple imputations. No EV after update
UKPDS OM2 [22] 2013	L	L	L	H	H	Yes	H	H	Missing data dealt with LOCF, no multiple imputations
Validation of the UKPDS 82 [23] 2015	L	L	L	H	L	NA	H	H	No. of outcomes and sample size not reported
JJ risk engine [24] 2013	L	H	L	L	L	No	H	M	Categorised continuous predictor (BMI) lacks external validation
Soedamah-Muthu et al. [20] 2014	L	L	L	H	L	Yes	H	H	Very low EPV and fewer number of outcomes
Lagani et al. [27] 2015	L	L	L	H	L	Yes	H	H	Low EPV
Harris et al. [35] 2013	L	H	L	L	H	No	H	L	Did not report relevant performance measures[a] Lacks external validation
Hippisley-Cox et al. [28] 2015	L	L	L	L	L	Yes	M[b]	H	High number of missing values, but dealt with multiple imputations
Icelandic model [29] 2011	L	L	L	L	L	Yes	L[b]	L	
Van der Heijden [25] 2014	L	L	L	H	L	NA	H	L	Low number of outcomes (<100)
Soto-Pedre et al. [30] 2015	L	L	L	H	L	NA	H	L	Low number of outcomes (<100)
Holtzer-Goor [26] 2015	L	L	L	H	L	NA	H	L	Low number of outcomes (<100)
Lund et al. [8] 2016	L	L	L	L	L	NA	L[b]	L	
H. Ng Keunen et al. [21] 2016	L	L	L	H	L	NA	H	L	Low number of outcomes (<100)
Danish model [31] 2012	L	L	L	L	H	No	H	L	Did not report relevant performance measures[a] No validation

Table 2 (continued)

Study ID	Participants	Predictors	Outcome	Sample size and participant flow	Model development and evaluation	External validation	Overall bias High Moderate Low	Applicability High concern Moderate concern Low concern	Comments for bias
Stratton et al. [32] 2013	L	L	L	L	H	Yes	H	L	Did not report relevant performance measures[a]
Aldington et al. [33] 2014	L	L	L	L	L	NA	Unclear[b]	L	Lack of information as meeting abstract only
Stratton Aldington et al. [34] 2014	L	L	L	L	L	No	Unclear[b]	L	Lack of information as meeting abstract only Lacks external validation
Scanlon et al. [36] 2015	L	L	L	H	L	Yes	M[b]	L	High amount of missing information and did CCA
ISDR [16] 2017	L	L	L	L	L	No	M[b]	L	Lacks external validation

EPV events per candidate variables, *EV* external validation, *LOCF* last observation carried forward, *NA* not applicable, *CCA* complete case analysis, *L* low, *M* moderate, *H* high

[a]Calibration, discrimination and/or classification measures

[b]See text

[29, 31] and model performance measures [17, 22, 26, 31, 32, 35]. Out of 14 primary modelling studies, three had no internal validation, and six lacked an external validation. On the whole, external validation studies had smaller samples. The studies are heterogeneous principally due to differing population characteristics, disease classification, outcomes, predictors, their handling/numbers and type of statistical models.

There have been three systematic reviews on topics related to diabetic retinopathy. Lagani et al. [37] examined the probability of complications developing in diabetic patients which included the incidence of diabetic retinopathy. However, they did not consider models for progression of diabetic retinopathy or vision loss (our precise review question). Van der Heijden's systematic review [38] is a conference abstract, so insufficient information was available to make comparison. The population of interest was only T2DM (our review includes T1DM as well). The context was screening and detection of DR (earlier low-risk part of the disease pathway), rather than progression to vision loss and treatment (higher-risk patients, the population of interest in our review). Taylor-Phillips et al. [39] investigated annual against longer screening intervals and concluded that there is insufficient evidence to support extending screening of diabetic patients for STR beyond 1 year. They based their conclusions on the lack of quality and use of different definitions of the low-risk group of patients. Their question was different from ours.

The three models identified as moderate to low risk of bias and low risk for applicability have already shown some impact in diabetic screening in lower-risk patients. Recent work by Scanlon et al. [36], Aspelund et al. [29] and ISDR [16] has clearly shown that individual patient's risk assessment and prediction can be safely and effectively achieved through the use of routine data in pre-STR patients. The evidence from Scanlon et al. [36] is also expected to have an impact on DESP, to risk-stratify patients into those suitable for 2-yearly screening and the relatively higher-risk group for yearly screening. Aspelund et al. [29] have the largest number of external validation studies and the model is being used in practice in Aarhus University. ISDR model is already the subject of an impact study [40]. One of these models could also be updated and tested on a higher-risk hospital patient population as well.

In conclusion, in countries with developed DR screening, for patients who have been referred for treatment or closer observation, a model is needed to determine their individual risk of progression to treatment stage/loss of vision, to direct the resources appropriately and further optimise the services especially for higher-risk patients. This review highlights some of the useful models available for the said purpose. Scanlon et al. [36], Aspelund et al. [29] and ISDR model [16] seem to be appropriate in terms of contemporary

participant data, accessible predictors and sound methodology, though they do not directly address the outcome of our interest. They need further external validation in diverse high-risk settings before being implemented into clinical practice. In addition to these three models, we have listed the predictors and performance of all other models. This means anyone, dependent on their own particular datasets, could use one that suits their needs such as to define screening intervals, or to target screening in poorly resourced countries.

More primary modelling studies particularly for use in hospital eye services will be useful as long as they are well-structured with good reporting. Studies should also ensure to present the final model in a simplified way, to make it easier for clinicians or policy makers to implement. The model's integration into electronic medical records can help decision-making and needs to be the goal in future models.

Acknowledgements We thank Sue Bayliss (SB) for help with the search strategy.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. NCD Risk Factor Collaboration (NCD-RisC)). Worldwide trends in diabetes since 1980: a pooled analysis of 751 population-based studies with 4.4 million participants. Lancet. 2016;387:1513–30.
2. Mathur R, Bhaskaran K, Edwards E, Lee H, Chaturvedi N, Smeeth L, et al. Population trends in the 10-year incidence and prevalence of diabetic retinopathy in the UK: a cohort study in the Clinical Practice Research Datalink 2004–2014. BMJ Open. 2017;7:e014444.
3. NHS Diabetic Eye Screening Programme Summary statistics for England. In: England PH, editor. Gov.UK; 2016.
4. Klein R, Klein BE, Moss SE, Davis MD, DeMets DL. The Wisconsin Epidemiologic Study of Diabetic Retinopathy: II. Prevalence and risk of diabetic retinopathy when age at diagnosis is less than 30 years. Arch Ophthalmol. 1984;102:520–6.
5. Klein R, Klein BE, Moss SE, Davis MD, DeMets DL. The Wisconsin Epidemiologic Study of Diabetic Retinopathy: III. Prevalence and risk of diabetic retinopathy when age at diagnosis is 30 or more years. Arch Ophthalmol. 1984;102:527–32.
6. RCOPHTH. The Way Forward Executive Summary 2017 20 Jan. 2018; (300117), p. 6–7. https://www.rcophth.ac.uk/wp-content/uploads/2015/10/RCOphth-The-Way-Forward-Executive-Summary-300117.pdf.
7. Ross PL, Gerigk C, Gonen M, Yossepowitch O, Cagiannos I, Sogani PC, et al. Comparisons of nomograms and urologists' predictions in prostate cancer. Semin Urol Oncol. 2002;20:82–8.
8. Lund SH, Aspelund T, Kirby P, Russell G, Einarsson S, Palsson O, et al. Individualised risk assessment for diabetic retinopathy and optimisation of screening intervals: a scientific approach to reducing healthcare costs. Br J Ophthalmol. 2016;100:683–7.
9. Echouffo-Tcheugui JB, Ali MK, Roglic G, Hayward RA, Narayan KM. Screening intervals for diabetic retinopathy and incidence of visual loss: a systematic review. Diabet Med: a J Br Diabet Assoc. 2013;30:1272–92.
10. Wolff R, Whiting P, Mallett S, editors. PROBAST: a risk of bias tool for prediction modelling studies. Cochrane Colloquium Vienna; 2015.
11. Moons KA, Wolff R, Whiting P. Prediction modelling: Where are we now and where do we need to go? 2018. http://www.meduniwien.ac.at/wbs/Moons170126.pdf.
12. Steyerberg EW, Moons KG, van der Windt DA, Hayden JA, Perel P, Schroter S, et al. Prognosis Research Strategy (PROGRESS) 3: prognostic model research. PLoS Med. 2013;10:e1001381.
13. Moons KG, de Groot JA, Bouwmeester W, Vergouwe Y, Mallett S, Altman DG, et al. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. PLoS Med. 2014;11:e1001744.
14. PROSPERO. Prognostic prediction models for the progression of diabetic retinopathy (DR) and vision loss in patients with sight-threatening diabetic retinopathy (STDR): protocol for a systematic review; 2017. http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42017057767.
15. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med. 2009;6:e1000100.
16. Eleuteri A, Fisher AC, Broadbent DM, Garcia-Finana M, Cheyne CP, Wang A, et al. Individualised variable-interval risk-based screening for sight-threatening diabetic retinopathy: the Liverpool Risk Calculation Engine. Diabetologia; 2017;60:2174–2182.
17. Clarke PM, Gray AM, Briggs A, Farmer AJ, Fenn P, Stevens RJ, et al. 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. 2004;47:1747–59.
18. Leal J, Hayes AJ, Gray AM, Holman RR, Clarke PM. Temporal validation of the UKPDS outcomes model using 10-year posttrial monitoring data. Diabetes Care. 2013;36:1541–6.
19. McEwan P, Bennett H, Ward T, Bergenheim K. Refitting of the UKPDS 68 risk equations to contemporary routine clinical practice data in the UK. Pharmacoeconomics. 2015;33:149–61.
20. Soedamah-Muthu SS, Vergouwe Y, Costacou T, Miller RG, Zgibor J, Chaturvedi N, et al. Predicting major outcomes in type 1 diabetes: a model development and validation study. Diabetologia. 2014;57:2304–14.
21. Ng H, Keunen J, Tack C, Nijpels G, Van Der Heijden AAW. Validation of a prediction model to optimise retinopathy screening in type 1 diabetes. Diabetologia. 2016;59(1Supplement 1):S476–7.
22. Hayes AJ, Leal J, Gray AM, Holman RR, Clarke PM. 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. 2013;56:1925–33.
23. McEwan P, Ward T, Bennett H, Bergenheim K. Validation of the UKPDS 82 risk equations within the Cardiff Diabetes Model. Cost Eff Resour Alloc. 2015;13:12.
24. Tanaka S, Tanaka S, Iimuro S, Yamashita H, Katayama S, Akanuma Y, et al. Predicting macro- and microvascular complications in type 2 diabetes: the Japan Diabetes Complications Study/the Japanese Elderly Diabetes Intervention Trial risk engine. Diabetes Care. 2013;36:1193–9.
25. van der Heijden AA, Walraven I, van 't Riet E, Aspelund T, Lund SH, Elders P, et al. Validation of a model to estimate personalised

screening frequency to monitor diabetic retinopathy. Diabetologia. 2014;57:1332–8.
26. Holtzer-Goor KM, Van Der Heijden AA, Jonker M, Stolk E, Nijpels G. Validation of an algorithm to predict the risk of sight threatening retinopathy in a multi-ethnic patient group treated in a Dutch hospital. Diabetologia. 2015;1:S525–6.
27. Lagani V, Chiarugi F, Thomson S, Fursse J, Lakasing E, Jones RW, et al. Development and validation of risk assessment models for diabetes-related complications based on the DCCT/EDIC data. J Diabetes its Complicat. 2015;29:479–87.
28. Hippisley-Cox J, Coupland C. Development and validation of risk prediction equations to estimate future risk of blindness and lower limb amputation in patients with diabetes: Cohort study. BMJ (Online). 2015;351:h5441.
29. Aspelund T, Porisdottir O, Olafsdottir E, Gudmundsdottir A, Einarsdottir AB, Mehlsen J, et al. Individual risk assessment and information technology to optimise screening frequency for diabetic retinopathy. Diabetologia 2011;54:2525–32.
30. Soto-Pedre E, Pinies JA, Hernaez-Ortega MC. External validation of a risk assessment model to adjust the frequency of eye-screening visits in patients with diabetes mellitus. J Diabetes its Complicat. 2015;29:508–11.
31. Mehlsen J, Erlandsen M, Poulsen PL, Bek T. Individualized optimization of the screening interval for diabetic retinopathy: a new model. Acta Opthalmologica. 2012;90:109–14.
32. Stratton IM, Aldington SJ, Taylor DJ, Adler AI, Scanlon PH. A simple risk stratification for time to development of sight-threatening diabetic retinopathy. Diabetes care. 2013;36:580–5.
33. Aldington SJ, Stratton IM, Scanlon PH. Validation of a simple stratification algorithm for progression to sight threatening diabetic retinopathy. Diabetologia. 2014;1:S474–5.
34. Stratton IM, Aldington SJ, Farmer AJ, Scanlon PH. Personalised risk estimation for progression to sight-threatening diabetic retinopathy: How much does clinical information add to screening data? Diabet Med. 2014;31:23–4.
35. Harris Nwanyanwu K, Talwar N, Gardner TW, Wrobel JS, Herman WH, Stein JD. Predicting development of proliferative diabetic retinopathy. Diabetes care. 2013;36:1562–8.
36. Scanlon PH, Aldington SJ, Leal J, Luengo-Fernandez R, Oke J, Sivaprasad S, et al. Development of a cost-effectiveness model for optimisation of the screening interval in diabetic retinopathy screening. Health Technol Assess. 2015;19:1–116.
37. Lagani V, Koumakis L, Chiarugi F, Lakasing E, Tsamardinos I. A systematic review of predictive risk models for diabetes complications based on large scale clinical studies. J Diabetes its Complicat. 2013;27:407–13.
38. Van der Heijden AA, Badloe F, Nijpels G, Beulens JW, editors. Prediction models for the risk of Retinopathy in people with type 2 Diabetes. A Systematic Review. 27th European Association for the Study of Diabetes Eye Complications Study Group; May 2017.
39. Taylor-Phillips S, Mistry H, Leslie R, Todkill D, Tsertsvadze A, Connock M, et al. Extending the diabetic retinopathy screening interval beyond 1 year: Systematic review. Br J Ophthalmol. 2016;100:105–14.
40. ISDR. http://www.isdrprojectcouk/rcthtml; 2018.

3.2 Appendices for Chapter 3

4 Chapter 4: Prognostic factors

This chapter is presented in two sections. Section 1 is general methods and Section 2) is submitted as a publication with the associated appendices. Appendix 26 is in the form of a poster that was presented at the University of Birmingham research festival.

The 78 predictors for diabetic retinopathy progression that were present in the literature are listed in the list of candidate predictors in Appendix 12. The overall aim of this study was to arrive at a shorter, more practical list of candidate predictors that were clinically relevant to high risk patients and are based on evidence.

Objectives

1. To seek clinicians' opinions about the most important predictors in the list for high risk patients, and to identify any additional predictors not yet described.
2. To assess the evidence base for each predictor identified in the first objective from the published literature.

4.1 Section 1: General methods

4.1.1 Candidate predictors list

The candidate predictors list is subjected to a multivariate regression analysis to select a final set for prediction modelling. This final list of predictors is then used by the end user to input their measured values to give a hazard ratio, that helps predict the outcome of interest in a patient.

Recently, there have been useful reviews on predictors (1-3), however, these were mainly about relatively new predictors, not all of which are presently in clinical practice, with some still in the research domain, or were in context of eyes and not patients. If these predictors were to be used, to ensure that there was not too much missing data, the research had to be prospective. This would have been too time consuming and expensive. The systematic review list was used as the list had already been tested and the predictors were already being recorded in various data sources, including primary care databases (4). Hence the list collated from the systematic review was used.

The higher the number of the predictors, the more information there is to be collected, increasing the time and effort required for the end user, as opposed to the aim of improving efficiency. The available list of 78 predictors was therefore not practical. There was also duplication and overlap between some of the predictors.

4.1.2 Need for clinical opinion and NGT

Predictor selection is preferably based on clinical knowledge and previous studies rather than being based on statistical methods alone (5).

Against the background of the need for expert opinion for the decision, various techniques, including Nominal Group Technique (NGT), brainstorming, focus groups and Delphi Methods were considered (6). Their technical strengths and limitations were taken into account, along with time requirements and time available. A qualitative study design (Delphi) has been used with the aim to prioritise the list of already identified predictors (7). NGT was found to be favourable after a feasibility exercise based on a Gallagher et al publication (6). NGT scored higher as compared to other techniques given that it enables equal participation, detailed consideration of ideas, and group cohesiveness (Given in table 1 below).

Table 1: The relative strengths of NGT, brainstorming, Delphi and focus groups, Modified from table 1 in Gallagher et al (6)

Advantages	**Comments**	**NGT**	**Brainstorming**	**Delphi**	**Focus**
Difficult for dominant participants to control	But possible, with round Robin	Yes	No	Yes	Possibly
Encourages minority concerns/options to be voiced	Equity	Yes	No	Yes	Possibly
Avoids 'quick decision making'	Well considered ideas	Yes	No	Yes	Possibly
High degree of task completion	Effectiveness	Yes	No	Yes	Possibly
Generates a high number of comments/ideas	New predictors	Yes	Possibly	Yes	Possibly
Provides support to allow identification of personal problems and self disclosure	Personal experience	Yes	No	No	Yes
Allows measurement of importance of ideas/items to individuals	Ranking	Yes	No	Yes	Possibly

Avoids pursuit of a single train of thought ('focus-effect')	Can cope with multiple items in fairly quick succession	Yes	No	Yes	Yes
Participants value social interaction i.e. group cohesiveness	Time wastage	Possibly	Yes	No	Yes
Ease of administration	Could cope with some support	No	Yes	Yes	Possibly
Need for experienced leader	Could provide with some help	Yes	No	No	Yes

4.1.3 Need for evidence review

PROGRESS guidance was used for predictor selection. Predictors have an important place in clinical practice and PROGRESS recommends ranking a predictor as “exploration”, “confirmation” or ”replication”. Exploration refers to a predictor being mentioned in a primary study as a risk factor for being part of the causal pathway. Confirmation is established if a predictor retains prognostic value, even after adjustment for other predictors. Replication is the assessment of the predictor in multiple independent studies. Publication and reporting biases are common in predictor research (8). I reviewed the literature and established the strength of evidence for the predictors to be considered for a prognostic model.

4.1.4 References:

1. Sivaprasad S, Pearce E. The unmet need for better risk stratification of non-proliferative diabetic retinopathy. Diabetic medicine : a journal of the British Diabetic Association. 2019;36(4):424-33.
2. Jenkins AJ, Joglekar MV, Hardikar AA, Keech AC, O'Neal DN, Januszewski AS. Biomarkers in Diabetic Retinopathy. The Review of Diabetic Studies. 2015;12(1-2):159-95.
3. Lee CS, Lee AY, Baughman D, Sim D, Akelere T, Brand C, et al. The United Kingdom Diabetic Retinopathy Electronic Medical Record Users Group: Report 3: Baseline Retinopathy and Clinical Features Predict Progression of Diabetic Retinopathy. American Journal of Ophthalmology. 2017;180:64-71.
4. McEwan P, Bennett H, Ward T, Bergenheim K. Refitting of the UKPDS 68 risk equations to contemporary routine clinical practice data in the UK. Pharmacoeconomics. 2015;33(2):149-61.
5. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. European heart journal. 2014;35(29):1925-31.
6. Gallagher M, Hares T, Spencer J, Bradshaw C, Webb I. The nominal group technique: a research tool for general practice? Fam Pract. 1993;10(1):76-81.
7. Thangaratinam S, Ismail K, Sharp S, Coomarasamy A, O'Mahony F, Khan K, et al. Prioritisation of tests for the prediction of preeclampsia complications: a Delphi survey. Hypertension in pregnancy. 2007;26(1):131-8.
8. Riley RD, Hayden JA, Steyerberg EW, Moons KG, Abrams K, Kyzas PA, et al. Prognosis Research Strategy (PROGRESS) 2: prognostic factor research. PLoS medicine. 2013;10(2):e1001380.

4.2 Predictors for Diabetic Retinopathy Progression – Findings from Nominal Group Technique and Evidence Review

Accepted for publication

Sajjad Haider[1], Salman Naveed Sadiq[2], Harpreet Sihre[1], Eniya Lufumpa[1] and Mohammad Tallouzi[1] David Moore[3*], Krishnarajah Nirantharakumar[3*], Malcolm James Price[3*]

[1] Doctoral researchers, Institute of Applied Health Research, University of Birmingham

[2] Royal Victoria Infirmary, Newcastle, United Kingdom

[3*] Joint Senior Authors Institute of Applied Health Research, University of Birmingham, Birmingham, United Kingdom

Joint Corresponding authors:

Sajjad Haider, Doctoral Researcher, Institute of Applied Health Research, University of Birmingham, Birmingham, United Kingdom

[redacted]

David J. Moore, Institute of Applied Health Research, University of Birmingham, Birmingham, United Kingdom

[redacted]

4.3 Abstract

Objectives: Risk stratification is needed for patients referred to hospital eye services by Diabetic Eye Screening Programme UK. This requires a set of candidate predictors. The literature contains a large number of predictors. The objective of this research was to arrive at a small set of clinically important predictors for the outcome of the progression of diabetic retinopathy. They need to be evidence based and readily available during the clinical consultation.

Methods and Analysis: Initial list of predictors was obtained from a systematic review of prediction models. We sought the clinical expert opinion using a formal qualitative study design. A series of nominal group technique meetings to shorten the list and to rank the predictors for importance by voting were held with NHS hospital-based clinicians involved in caring for diabetic retinopathy patients in the United Kingdom. We then evaluated the evidence base for the selected predictors by critically appraising the evidence.

Results: The source list was presented at nominal group meetings (n=4), attended by 44 clinicians. 25 predictors from the original list were ranked as important predictors and 8 new predictors were proposed. Two additional predictors were retained after evidence check. Of these 35, 21 had robust supporting evidence in the literature condensed into a set of 19 predictors by categorising diabetic retinopathy.

Conclusion: We identified a set of 19 clinically meaningful predictors of diabetic retinopathy progression that can help stratify higher risk patients referred to hospital eye services and should be considered in the development of an individual risk stratification model.

Key Words: Risk Factors, Prognostic factors, Predictors, Diabetic retinopathy.

Study Design: A qualitative study and evidence review

Setting: Secondary eye care centres in North East, Midlands and South of England

Key messages

What is already known about this subject? A large list of already identified predictors, for diabetic retinopathy progression was available but had duplications and overlap.

What are the new findings? With the clinical experts' opinion using Nominal Group Technique and evidence evaluation we identified a shorter, evidence based and pragmatic list of 19 predictors, which was also ranked by clinical importance.

How might these results change the focus of research or clinical practice? This set of candidate predictors can help stratify patients with referable diabetic retinopathy under hospital ophthalmology services through risk stratification.

4.4 Introduction

Diabetic retinopathy (DR) can develop in anyone with Diabetes Mellitus and is a major cause of blindness due to damage to, and disruption of the retina (the light-sensitive layer at the back of the eye) leading to loss of sight. DR is a consequence of changes to the blood vessels in this part of the eye leading to leakage of blood and fluid and the formation of abnormal blood vessels (1).

There has been a global increase in the number of adults with Diabetes Mellitus from an estimated 108 million in 1980 to 422 million in 2014 (2). The rise in diabetes prevalence coupled with early detection of DR through better population eye screening coverage has increased the burden of patients with DR to health-care systems (3).

In the United Kingdom, DR services are organised into the Diabetic Eye Screening Programme (DESP) for lower-risk patients and the hospital eye services (HES) for higher-risk patients with referable diabetic retinopathy, with the HES providing treatment, and closer observation of patients. While services may be organised differently in other countries, the care pathways are likely to be similar for these patients. DR patients are referred to the HES when they develop clinical signs of the so called sight-threatening retinopathy (STR). The clinical signs based on photographs are the only differentiating features used for this risk stratification. However, approximately 50 - 70% of referrals (4, 5) do not need intervention and are observed in the HES for varying periods of time. This is one reason for a demand and capacity mismatch in HES.

There have been successful attempts to optimize the diabetic screening services through risk stratification of patients with the help of a prognostic prediction model (6). A similar approach could be used to improve safety and efficiency of diabetic retinopathy services in the hospital environment. However, a recent systematic review of prediction models for DR found that none of the 14 models identified were directly applicable for the higher risk patients in the hospital setting (7).

Predictors (8) are at the core of prevention and prediction of a clinical event. Predictors could be, for example, an individual attribute, a clinical feature, a

physiological, psychosocial or an environmental factor. Predictor research ranges from “exploration” to “confirmation” to ”replication”. Exploration refers to a predictor being mentioned in a primary study as a risk factor for being part of the causal pathway. Confirmation is established if a predictor retains prognostic value even after adjustment for other predictors. Replication is the assessment of the predictor in multiple independent studies. Publication and reporting biases are common in predictors’ research (9).

A prediction model combines two or more predictors (also called prognostic factors) to predict the likelihood of an outcome, for example, diabetic retinopathy progression to a treatment requiring stage or loss of vision, before it occurs (10). Previous prediction models in the systematic review (7) have between them used 78 different candidate predictors. However, there are problems with the direct use of these predictors for the HES setting. Firstly, a set of predictors this large is not feasible for use in clinical practice. Second, many of these predictors have not been confirmed. Third, a number of predictors were extrapolated from evidence for macrovascular outcomes like stroke to retinopathy progression inappropriately, for example, “smoking” and “ethnicity” (11). Fourth, there was significant duplication and overlap between the predictors, for example, diabetic nephropathy and chronic kidney disease. Such highly correlated predictors are unlikely to be independently predictive in the same model (7). Finally, there may be predictors for higher-risk patients, not reported in the prediction model studies as they were primarily for low-risk patients.

There were two comprehensive and up to date evidence reviews on the subject of predictors. However, their perspective was quite different from ours. They were a rich resource of evidence and very useful for horizon scanning (12, 13), but most of the

predictors they suggested are still in research domain and not being recorded in the patient notes, so not possible at the moment to be used within prognostic research. We instead used the modelling studies included within the systematic review assuming moats of them will be evidence based and for the practical reason that they will have been measured and thus can be extracted from the data.

4.4.1 Aims

In this qualitative, study we aimed to de-duplicate the list of predictors and through clinicians input and evidence assessment to arrive at a shorter list, more clinically relevant to high risk patients.

4.4.2 Objectives

- To identify all the predictors for diabetic retinopathy progression in the literature
- To seek clinicians' opinion about the most important predictors among them for high risk patients and to identify any new predictors not yet described.
- To assess the evidence base from published literature for each predictor identified in objective 2.
- To finalize a potentially parsimonious set, i.e; a minimal number of predictors in a future multivariate model able to give the highest predictive performance (14, 15)

4.5 Methods

4.5.1 Study design and methods

Our recent systematic review of prognostic modelling studies for the development of diabetic retinopathy of any type, including maculopathy, in patients with diabetes (7) was the primary source of the list of candidate predictors. The text and the reference lists of the modelling studies were searched for candidate predictors. The Nominal Group Technique (NGT) was then used to shortlist the list of candidate predictors and primary studies were then evaluated for evidence. In case no reference of the primary study was found in the modelling study, searches were made in Pubmed and Scopus. Subsequently, the evidence behind the shortlisted predictors was evaluated using the Quality in Prognosis Studies (QUIP) tool (16). To comply with the PROGRESS research framework for predictors (9), we used the criterion of at least confirmation within primary studies before including them into the list of our candidate predictors for future models. This ensures that using those predictors can calculate the risk of the outcomes of interest more precisely in a future model (10) Our recent systematic review of prognostic modelling studies for the development of diabetic retinopathy of any type including maculopathy in patients with diabetes (7) was the primary source of a list of candidate predictors. The text and the reference lists of the modelling studies were searched for the candidate predictors. A qualitative study design (Delphi) has been used with the aim to prioritise the list of already identified predictors (16). Techniques other than NGT such as brainstorming, focus group and Delphi methods were considered (17) along with their technical strengths and limitations. Nominal group technique (NGT) was used to shrink the list of candidate predictors. Primary studies were then evaluated for evidence. In case no reference of

the primary study was found in the modelling study, searches were made in Pubmed and Scopus. Subsequently, the evidence behind the shortlisted predictors was evaluated using the Quality in Prognosis Studies (QUIP) tool (18). To comply with the PROGRESS (Prognosis research strategy) framework for predictors (9), we used a criterion of at least “exploration” (exploring predictor’s relation to prognosis) within primary studies before including them into the list of our candidate predictors for future models, to ensure that the risk of the outcomes of interest can be calculated more precisely by using those predictors by a future model (10).

4.5.2 Nominal group technique

NGT is a qualitative research methodology, where every meeting is a structured small group exercise allowing judgements by individuals to be pooled to arrive at a decision in an uncertain situation (17, 19). NGT has been used frequently in ophthalmology and medicine (20-28) and is a highly adaptable method (29).

An information pack with the predictors list from the systematic review was sent to participating unit before the meeting (7) (Appendices 14,19 and 20). Informed consent forms were provided and signed by participants in Appendix 20. NGT was performed through a series of four meetings, each lasting around 1.5 hours. Meetings were accommodated within clinical governance/audit meetings of four NHS trusts with the permission of respective research and development directorates. NGT was chosen, because of its modifiable nature allowing brainstorming to decide on importance of predictors, a round robin for an equal opportunity for all participants, a discussion for clarifications, voting for final decision on ranking and for being time-efficient. Each meeting was conducted in 4 stages in Appendix 21 from providing

background information, round-robin recording of ideas, discussion of the list of ideas and ending in voting for ranking of the ideas generated.

During the round-robin recording, the facilitator went around the table inviting one item from each participant at a time (to give equal opportunity to all the participants) and recording them on the flip chart. It was left open to participants to choose any number of predictors from the list provided or name their own predictors using their personal experience and insight. We requested participants to restrict their final choices to a total of 15 - 20 predictors.

The round robin cycles (within a meeting and between and then the meetings) were repeated until saturation had been achieved - a point where no new predictors were being added and thus all new information had been obtained (30, 31). The centres were recruited sequentially and from the second meeting onward, after every meeting cumulative results were examined for any new predictors suggested and thus monitoring for saturation (when no further predictors were being added) before stopping to recruit.

Participants

To ensure a maximum range of views and opinions were collected, keeping purposeful sampling in mind i.e; selection of information-rich resource for the research question (32) as a priority, we approached medical retina team leaders for four NHS trusts (consultants with much longer training and experience) (33) with their teams of middle grades and registrars, nurses and optometrists, directly involved in caring for diabetic retinopathy patients, for these meetings. We aimed to over-recruit allowing for likely no-shows/dropouts on the day for each group.

Data collection

Reflexivity (34) of the authors was considered when designing the project. SH is an ophthalmologist and thus shared the participants' experience. He moderated the meetings, but one or more of the qualitative study researchers (HS, EL or MT) were also present to reduce the chances of bias. Clinicians were asked for their written consent (Appendix 20) information collected to be used for further research and publication. Flip chart and marked lists were collected at the end of discussion while voting and all other information was collated on a spreadsheet. All participant data were anonymised, with no direct quotes mentioned in any dissemination.

Data extraction in spreadsheet and collation after every meeting were carried out by one of the co-authors involved in that meeting and validated with the paper forms with the help of a second researcher. We collated all-new predictors suggested by clinical colleagues participating in these meetings, added them to the list provided and helped participants rank them for importance in prognosis prediction.

Analyses of NGT

Voting frequency was calculated for voting scores per predictor (summing of scores) as well as frequency of voting percentage (score achieved for the item / maximum possible score X 100) and tabulated according to their rank (29).

4.5.3 Evidence review

In this step, we searched the reference lists of the modelling studies identified in the systematic review (7) for primary studies that had investigated the predictors selected by NGT and subjected them to critical appraisal using the QUIP tool (18). A basic literature search was then performed in Pubmed and Scopus to identify primary studies

for predictors where no primary evidence source was referenced in any modelling study. Using the identified information we summarised the status of each predictor as explored, confirmed or confirmed and replicated in more than one study.

Patients and Public Involvement:

While there was no direct patients and public involvement required in this research, we included all the predictors from ISDR model which were chosen by an expert patient panel (35) to reflect patient input.

Ethics: The protocol was evaluated by the four NHS trusts where the NGT was carried out. Since the research was essentially a decision exercise based on clinicians' expert opinion and did not involve patient data, no ethical approval was required.

4.6 Results

The list of 78 candidate predictors from the systematic review are given in Figure (shows their breakdown and frequency of their use in the pool of 14 models) and in Appendix 14. Biochemical predictors were used most commonly. Figure 2 shows the predictor items flow during the processes of NGT and evidence evaluation.

4.6.1 NGT

We conducted four NGT meetings, within secondary care eye centres in National Health Service trusts in the following areas in the UK: (a, b) Two Midland trusts, (c) North East England, and (d) South of England. After the fourth NGT meeting, saturation was reached. The full set of 33 predictors selected by the clinicians are shown in Table . Participants roles varied from consultants, specialist registrars, middle-grade doctors and allied health professionals (one optometrist and one nurse). A total of 44 clinicians participated with numbers per session ranging from 6

to 16 (Appendix 22). 11 out of 19 participating consultants were medical retina trained. After multiple round robin cycles (within a meeting and different meetings) had been repeated to ensure saturation had been achieved and no new predictors were being added (Appendix 23), the final set of candidate predictors as decided by NGT is given in Table 1.

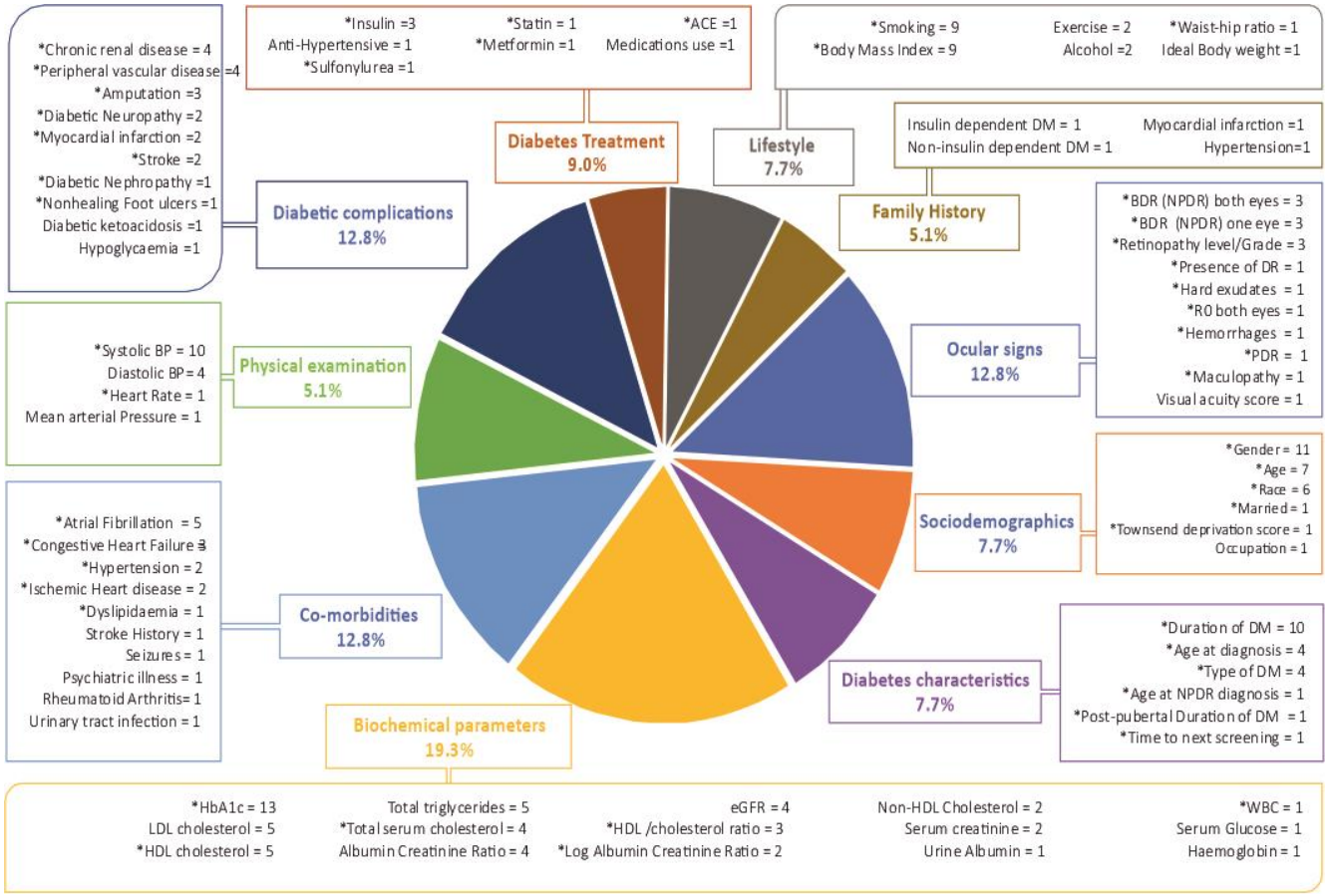


Figure 1: List of candidate predictors from systematic review model development studies.

Pie chart illustrates the percentage proportion of each category of predictors. Boxes indicate the type of candidate predictors in each category along with the number of modelling studies that considered them. Information on the full list of candidate predictors is in appendix 14.

* The predictor was used in at least one model development as a predictor

Table 1: List of predictors chosen by NGT by number of votes and voting frequency

Predictor		**Voting frequency**	**Proportion voted (total n = 44) (%)[a]**
1	Glycated haemoglobin (HbA1c)	31	70
2	Duration of diabetes	24	55
3	Retinopathy level	17	39
4	Townsend score	16	36
5	Smoking (Lifestyle)	12	27
6	Race	11	25
7	Proliferative Diabetic Retinopathy	11	25
8	Did not attend (DNA)[b]	10	23
9	Nephropathy	9	20
10	Hypertension	9	20
11	Maculopathy	8	18
12	Pregnancy [b]	7	16
13	Co-morbidities [b]	7	16
14	Exercise/physical activity	6	14
‘15	Type of Diabetes Mellitus	5	11
16	Body Mass Index	5	11
17	eGFR	4	9
18	Chronic renal disease	4	9
19	Rapid reduction of blood sugar (early worsening) [b]	4	9
20	Dyslipidaemia	3	7
21	Psychiatric illness	3	7
22	Visual acuity score	2	5
23	Diastolic Blood Pressure	2	5
24	Only eye [b]	2	5
25	Age at diagnosis	1	2

26	Chronic infection [b]	1	2
27	Pre-proliferative [b]	1	2
28	Neuropathy	1	2
29	Age	1	2
30	Statins	1	2
31	Insulin	1	2
32	Gender	1	2
33	Diet [b]	1	2

[a] Voting frequency in percentages % arranged in order of high to low. NGT selected 26 out of 78 candidate predictors. Eight new predictors (marked [b]) were added by NGT to give a total of 33 predictors.

4.6.2 Evidence evaluation

Following NGT, during the evidence review, two further predictors from the original list of 78 were added back in to give a total of 35 predictors (Table 2). **"**Systolic blood pressure" (SBP) was included because of the participation of patient expert panel in Individualised Screening for Diabetic Retinopathy (ISDR) model (35), was the third commonest predictor used in prognostic models and has a good evidence base. The Nominal group participants probably implicitly included it by selecting hypertension as well. Total serum cholesterol was added back as a proxy for dyslipidaemia as difficulties with reporting of this predictor have previously been reported (36).

Table 2: Predictors and their primary studies with confirmation and replication status.

Group	No	Predictors	Model mentioning Predictor	Status
Ocular features	1	Retinopathy level / DR grade	Lagani et al (14)	C (37), R (38)
	2	Proliferative diabetic retinopathy	*Hippisley – Cox et al (11)	C (37), R (39)
	3	Maculopathy	*Hippisley – Cox et al (11)	C (37), R (39)
	4	Visual acuity score	Lagani et al (14)	C (40)
Socio demographics	5	Age	Multiple studies (11, 41, 42)	C (43), R (44)
	6	Race	[L] Harris et al (11, 45)	E (45), (3) Lack of evidence
	7	Gender	Harris et al (45)	C (46), R (43)
	8	Social Deprivation score	*Hippisley – Cox et al (11)	C (47), R (48)
Diabetes characteristics	9	Type of DM	*Icelandic model (49)	C (50)
	10	Age at diagnosis	*UKPDS OM2 (51)	C (52), R (44)
	11	Duration of DM	Icelandic model (49)	C (37), R (38)
Biochemical parameters	12	HbA1c	UKPDS OM1 (53)	C (40), R (54)
	13	eGFR	*UKPDS OM2 (51) Stratton et al (55), ISDR[A] (35)	C (56), R (57) C (58), R (44)
	14	Total Serum Cholesterol	Soedamah-Muthu et al (41) UKPDS OM1 (53)	C (59), R (44) C (60), R (61)
Physical examination	15	Diastolic Blood Pressure (DBP)	*Icelandic model (49)	C (50)
	16	Systolic Blood pressure (SBP)	*UKPDS OM2 (53)	C (52), R (44)
Co-morbidities	17	Hypertension	Harris et al (45)	C (59)
	18	Dyslipidaemia	Harris et al (45)	C (58), R (62)
	29	Psychiatric illness	*[L]Lagani et al (14)	Absence of evidence
Diabetic complications	20	Chronic renal disease	*Hippisley – Cox et al (11)	C (63)
	21	Diabetic Nephropathy	Harris et al (45)	C (46), R (64)
	22	Diabetic Neuropathy	*[L] Harris et al (45)	Absence of evidence
Diabetes treatment	23	Statin	Harris et al (45)	C (65), R (62)
	24	Insulin	*Harris et al (45)	C (66), R (67)
Lifestyle	25	Smoking	[L]McEwan et al (68)	C (43, 69), Absence of evidence
	26	Body Mass Index (BMI)	*[L] McEwan et al and others (14, 42, 68)	C (70), Absence of evidence
	27	Exercise/physical activity	[L]Tanaka (13)	Absence of evidence

New from NGT	28	Only eye situation	[L] NA	Absence of evidence
	29	Early worsening	NA	C (71), R (72)
	30	Frequent DNA	[L] NA	C (73) Lack of evidence
	31	Pregnancy	NA	C (74), R (75)
	32	Diet	[L] NA	Absence of evidence
	33	Pre proliferative DR	NA	C (37), R (76)
	34	Chronic infection	[L] NA	Absence of evidence
	35	C0-Morbidities	[L] NA	Absence of evidence

*****Modelling Study did not clearly identify the primary study for the predictor. Details in text, [A] Individualised Screening for Diabetic Retinopathy (ISDR), E (exploration) C (confirmation) and R (replication), UKPDS (United Kingdom prospective Diabetes Study), NA Not applicable

[L] Lacking evidence

After NGT, residual duplication/overlap still remained, e.g; “diabetic nephropathy” and “chronic kidney disease”; “hypertension”, “diastolic blood pressure” and “systolic blood pressure”; “dyslipidaemia”, “cholesterol” and “statins”. “Diabetic Nephropathy” is the primary cause of “Chronic kidney disease” characterised by progressive decline of (64) “estimated glomerular filtration rate” (eGFR). Therefore “eGFR” was retained in preference to “chronic kidney disease”, and “diabetic nephropathy” as it is established as a predictor for diabetic retinopathy and is more sensitive than the earlier two predictors (57). In the case of the overlap between “hypertension”, “diastolic blood pressure” and “systolic blood pressure”, the latter two predictors are represented in “hypertension” but seem to have different prognostic values, so were retained (59, 60).

We did not find any primary studies supporting the association between diabetic retinopathy progression and “psychiatric illness” or “diabetic neuropathy”. Among the studies found, no proven association of diabetic retinopathy with “BMI” (70), “exercise / physical activity” (42) or “smoking” (43, 69) was seen but a weak association was seen with “serum cholesterol”. The last item was therefore retained, and the rest excluded.

Among the new predictors suggested by NGT, primary studies were found confirming the association for “early worsening“ (71), “pregnancy“ (74), frequent “DNA” (73), and “pre proliferative DR“ (37). Due to the lack of evidence, the remaining four predictors (“co-morbidities, “only eye situation”, “Diet” and “chronic infection”) were excluded.

The commonest primary studies quoted were UKPDS (37, 43, 54, 59, 60, 67), Diabetes Control and Complications Trial (DCCT) (40, 71, 75), Wisconsin Epidemiological Study of Diabetic Retinopathy (61, 77), Epidemiology and Prevention

of Diabetes Study (EURODIAB) (44), and Action to control cardiovascular risk in diabetes (ACCORD) (78). Table 2 tabulates primary studies and status of confirmation and replication for individual predictors. 14 predictors were excluded for reasons of no proven association, duplication / overlap. The remaining 21 predictors were condensed into 19 predictors (Table 3) by combining the DR categories together.

Only two out of 14 modelling studies (41, 42) reported primary studies for all of the prognostic factors/predictors used. One of the modelling studies used the items without giving reference to any primary study (35) but instead quoted literature review /expert patients' panel. A model partly relied on borrowing the items from other models (42), some could be traced from the reference list but were not in context (42, 79). It was not clear in 8 out of 35 predictors (23%), as to which primary studies confirmed them (given in Table 2).

Most of the primary studies had multiple publications. Out of 35 predictors, 25 (71%) had good supporting evidence of predictive value from the literature (Table 2). 4 predictors were excluded because of overlap / duplication. Out of 17 primary studies critically appraised in Appendix 24, the evidence base for only one primary study had a high risk of bias, mainly due to confounders issues and highly selective population (47). The factor involved did have another supporting study with low risk of bias. 3 predictor studies were judged to have moderate risk of bias and 15 predictor studies were low risk of bias (79% of primary predictors studies) on QUIP criteria.

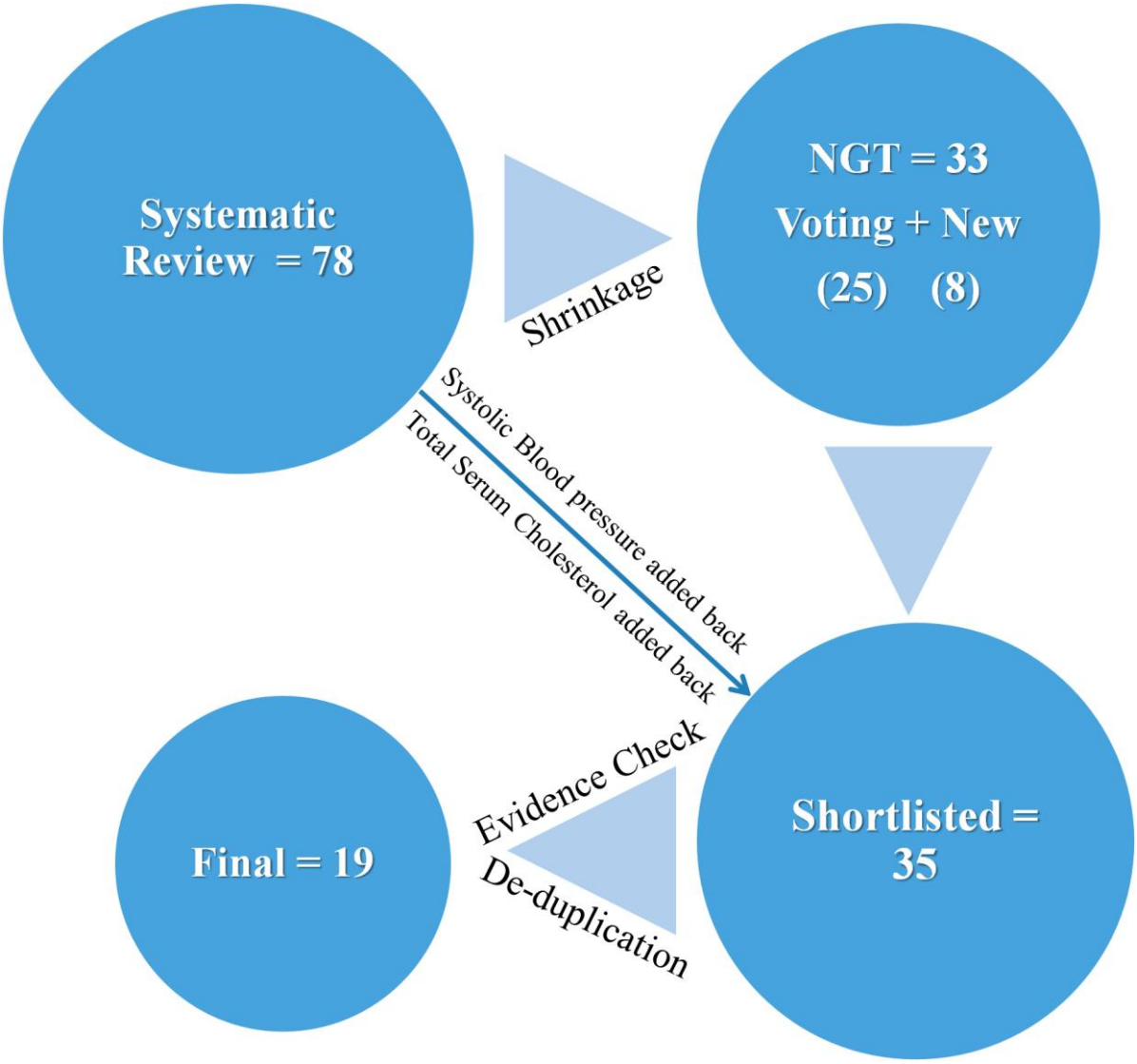


Figure 2: Summary of the sequence involved in reaching the final list of candidate predictors in Table 3.

Only two out of 14 modelling studies (42, 43) reported primary studies for all of the prognostic factors/predictors used. One of the modelling studies used the factors/predictors without giving reference to any primary study (34) but instead quoted a literature review or expert patients' panel. One model partly relied on borrowing items from other models (43), some of which could be traced from the reference list but were not in context (43, 78). It was not clear which primary studies confirmed 13 out of the 35 (37%) predictors, (marked * in Table 2).

Most of the primary studies had multiple publications. Out of 35 predictors, 26 (74%) had good supporting evidence of predictive value from the literature (Table 2). Out of 17 primary studies that were critically appraised (Appendix 24), the evidence base for only one primary study had a high risk of bias, mainly due to confounders issues and a highly selective population (50). The factor involved however did have another supporting study. Three predictor studies were judged to have a moderate risk of bias and 13 predictor studies had a low risk of bias (76% of primary predictors studies) on QUIP criteria.

Table 3: Final list of candidate predictors

	Group	Predictors
1	Ocular features	Presence and Diabetic Retinopathy grade a. Proliferative diabetic retinopathy b. *Pre-proliferative DR c. Maculopathy
2		Visual acuity score
3	Sociodemographic	Age at STR diagnosis
4		Age
5		Race
6		Gender
7		Social Deprivation score
8	Diabetes characteristics	Type of DM
9		Duration of DM > 10 years
10	Biochemical parameters	HbA1c
11		eGFR
12		Total Serum Cholesterol
13	Physical examination	Systolic Blood pressure
14		Diastolic Blood Pressure
15	Diabetes treatment	Statin
16		Insulin
17	*NGT	Pregnancy
18		Early worsening
19		Frequent DNA

***** Pre-proliferative diabetic retinopathy, new from NGT, is moved up to appear with the other DR categories, five condensed into two and thus 22 predictors condensed into 19. Also “age at diagnosis” was specified to “age at STR diagnosis” to conform with our target population of referable diabetic retinopathy.

4.7 Discussion

4.7.1 Statement of principal findings

We used clinical opinion in the NGT meetings to reduce a list of 78 previous candidate predictors to a list of 25 potential predictors. The study also suggested a further 8 potential predictors (Table). After evidence review, we added back another two predictors.

The new predictors suggested by the NGT made good clinical / biological sense, but four of them (“Co-morbidities”, ”only eye situation”, ”diet”, “chronic infections”) have not been explored for association with DR progression, although it is possible that the last two may be operational through other variables such as HBA1c rise due to uncontrolled blood sugar. The “only eye” situation clearly has a higher risk of blindness because of the absence of function in the affected eye, or higher risk because of the same pathology as the lost eye but has not been investigated for its association with DR progression. The other three predictors, ”early worsening”, ”pregnancy” and ”pre proliferative DR” have a proven association with our outcome of interest and were therefore included. While specific co-morbidities were considered as mentioned above, “co-morbidities” presence or numbers as a predictor had no supporting evidence. “Frequent DNA” has been proven to add prognostic value, so has been added to the final list but is in need for further confirmation, adjusted against all potential confounders.”

The top three ranked predictors from NGT (HBA1c, duration of diabetes and DR grade) were the same as the low risk of bias models’ predictor sets (Appendix 25) chosen by the systematic review (7). This was in spite of this information being

withheld from the participants in the NGT meetings. This shows the clinicians' intuitive thinking matches the findings of the systematic review as well. The approach of using retinopathy stages data alone to develop a risk stratification tool in DESP environment has been suggested (55). Individualized Screening for Diabetic Retinopathy (ISDR) model (35) in complete contrast has suggested using clinical predictors alone, esp. for higher-risk patients. We looked at risk estimates of the various predictors within various models. They tend to vary depending on the combination of predictors used (6, 14). However ocular predictors generally showed higher relative risk estimates. We suggest that any future model should contain a combination of ocular and systemic predictors for use in higher risk hospital population, as the values of for example SBP and HbA1c etc are likely to be higher and can thus add significantly to the predictive ability. A practical sized set is now available, with individual predictors ranked in importance as perceived by the NGT participants. A prediction model for higher risk diabetic retinopathy could be built based on them in appropriate population data.

Several predictors in previously reported models were associated with diabetic complications other than retinopathy - in multiple outcomes or composite outcome models (14, 41). The evidence for association of some of these predictors with diabetic retinopathy is unclear. For example, DR was a main covariate for severity of "diabetic polyneuropathy" (14, 59), but it still remains to be seen if the reverse is also true. "Smoking" as a candidate predictor was included in the majority of the models but made it to the final set only in one of them (68). The primary studies failed to show it as a predictor for diabetic retinopathy progression (43, 46). Harris et al included ethnicity in their final model but HR p value crossed the recommended

threshold of > 0.2, so was of doubtful statistical significance. It was included in 5 models as candidate predictor but not made to any of the final set on statistical testing. However, there is indirect evidence of higher risk of South Asians to develop STR (3) and blacks and South Asians among patients under diabetic eye screening programme having a higher prevalence of visual impairment (80). We therefore included ethnicity in the final candidate predictors set.

4.7.2 Strengths and weaknesses of the study

We based this study on the full list of predictors that had duplication, overlap and extrapolation of evidence from outcomes other than DR progression. In this situation with a large number of predictors already confirmed and replicated, incorporating clinical insight through a qualitative study design was invaluable to reach a shorter, more manageable list and also generate some new predictors. This is also the first study to evaluate the evidence base of potential predictors used in existing prognostic models for diabetic retinopathy progression and to follow the Prognosis Research strategy (PROGRESS) framework recommendations (9). We noted reporting deficiencies and have suggested possible preventive strategies for future.

4.7.3 Strengths and weaknesses in relation to other studies

The Individualised Screening for Diabetic Retinopathy (ISDR) model (35) mentioned using a patient expert panel for decision making during the predictor selection process. There is not enough detail in the study design, and we assume that it is only comparable to our work in that the study design was a qualitative one. We provide here the details of our methods and results with the interpretation. Our list also includes all the ISDR predictors, reflecting their expert patients' input. The

Standardisation of Uveitis Nomenclature (20) and Consensus on Outcome Measures for Glaucoma Effectiveness Trials (25) are prime examples of successful use of NGT to arrive at a decision with the help of expert panels in Ophthalmology. Another qualitative study looked into the patient-perceived risks of disease and benefits of treatment (81) but did not address predictors selection. We wanted to build on this approach seeking clinical experts' opinion with the help of NGT meetings.

4.7.4 Implications for clinicians and policymakers

This set of predictors derived will be useful to risk-stratify patients, optimise treatment strategies, inform patients about their personal risks and improve research strategies as well as providing the building blocks for future prognostic models (9). A set of predictors based on these 19 finalised predictors could be used to risk stratify the population received within the HES after referral from diabetic eye screening programme, to help with prioritisation of appointments and thus direct the resources more appropriately**.** Alternatively, a model could be constructed to estimate an individual patient's risk**.** This research will help the clinicians manage their patients according to their risk of progression.

While we have attempted to develop a list of predictors that are useful in predicting patients who progress from referable diabetic retinopathy to a stage of needing treatment or vision loss, the list is primarily from patients with diabetes under screening for incident diabetic retinopathy and referable diabetic retinopathy and as such are generalizable as markers for progression to any stages.

4.7.5 Unanswered questions and future research

Risk of bias assessment did not affect our decision to exclude any predictors as vast majority of the predictors had good evidence base. PROGRESS criteria was used to decide which predictors were not confirmed yet. That has helped decide where there is need for further research, e.g., “Race”, “diet” “exercise” etc.

We also found the reporting of the evidence base for the predictors selected in the modelling studies sub-optimal. Under-reporting has been mentioned by other observers before (9) and needs improving. The Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) statement asks for all predictors used to be clearly defined (82). One way of improving this situation may be for TRIPOD checklist to encourage investigators to report the primary study of origin for every candidate predictor used in a model. Out of 35, 10 predictors (29%) were not supported by the existing evidence base applying PROGRESS standards and require further research.

There have been recent useful reviews and studies on ocular predictors and their natural history use as predictors of diabetic retinopathy progression (5, 12, 13, 83) Ocular signs identified in OCT and FFA could also be potential predictors of diabetic retinopathy progression. While the ocular signs recommended are suitable for prospective research studies, existing retrospective longitudinal data most commonly used for prognostic research are unlikely to have sufficient information on these predictors. There is an ever-increasing interest in fundus images based detection, assisted by artificial intelligence (84). Prognostic factor research is a dynamic field and will benefit greatly with these newer technologies. Machine learning can handle the data from wider sources, can bring additional benefits from automation,

unsupervised clustering of a much larger number of predictors and can also add new phenotypes associated with the outcomes. However, there are ethical, governance, interpretability issues and the process of development of these techniques are at an early stage of development and application (85). It is likely that future research will identify further potentially important predictors so an update may be required in the future.

4.8 Conclusion

We have been able to identify 19 evidence-based predictors for diabetic retinopathy progression, using a novel method (NGT) and evidence review in line with PROGRESS recommendations. This smaller and more practical set provides a useful resource for a potential model to stratify patients for risk of DR progression, to aid clinical decision making and optimise clinical care pathways. This set is ranked in importance by the NGT.

Acknowledgements: The authors would like to thank four NHS trusts and their clinicians for their support and involvement in NGT. Also would like to express gratitude to Sophie Beese for kindly reviewing the manuscript for an internal peer review.

Contributors: SH, MT, HS and EL designed the study and participated in the NGT meetings, performed data extraction, were involved in drafting and revising the manuscript for important intellectual content. MT, SNS, MP, DM and KN made a substantial contribution to the interpretation of data and revision of the draft. All authors critically revised drafts of the work and approved the final version of the manuscript.

4.9 References:

1. Fong DS, Aiello L, Gardner TW, King GL, Blankenship G, Cavallerano JD, et al. Retinopathy in diabetes. Diabetes care. 2004;27 Suppl 1:S84-7.
2. Worldwide trends in diabetes since 1980: a pooled analysis of 751 population-based studies with 4.4 million participants. Lancet. 2016;387(10027):1513-30.
3. Mathur R, Bhaskaran K, Edwards E, Lee H, Chaturvedi N, Smeeth L, et al. Population trends in the 10-year incidence and prevalence of diabetic retinopathy in the UK: a cohort study in the Clinical Practice Research Datalink 2004-2014. BMJ open. 2017;7(2):e014444.
4. The Royal College of Ophthalmologists. The way forward age-related macular degeneration and diabetic retinopathy 2016

[ ]. Available from: https://www.rcophth.ac.uk/wp-content/uploads/2015/10/RCOphth-The-Way-Forward-AMD-300117.pdf, Accessed 2019, July 20.
5. Jyothi S, Elahi B, Srivastava A, Poole M, Nagi D, Sivaprasad S. Compliance with the quality standards of National Diabetic Retinopathy Screening Committee. Primary care diabetes. 2009;3(2):67-72.
6. Scanlon PH, Aldington SJ, Leal J, Luengo-Fernandez R, Oke J, Sivaprasad S, et al. Development of a cost-effectiveness model for optimisation of the screening interval in diabetic retinopathy screening. Health Technol Assess. 2015;19(74):1-116.
7. Haider S, Sadiq SN, Moore D, Price MJ, Nirantharakumar K. Prognostic prediction models for diabetic retinopathy progression: a systematic review. Eye. 2019;33(5):702-13.
8. Offord DR, Kraemer HC. Risk factors and prevention. Evidence Based Mental Health. 2000;3(3):70-1.
9. Riley RD, Hayden JA, Steyerberg EW, Moons KG, Abrams K, Kyzas PA, et al. Prognosis Research Strategy (PROGRESS) 2: prognostic factor research. PLoS medicine. 2013;10(2):e1001380.
10. Steyerberg EW, Moons KG, van der Windt DA, Hayden JA, Perel P, Schroter S, et al. Prognosis Research Strategy (PROGRESS) 3: prognostic model research. PLoS medicine. 2013;10(2):e1001381.
11. Hippisley-Cox J, Coupland C. Development and validation of risk prediction equations to estimate future risk of blindness and lower limb amputation in patients with diabetes: Cohort study. BMJ (Online). 2015;351 (no pagination)(h5441).
12. Sivaprasad S, Pearce E. The unmet need for better risk stratification of non-proliferative diabetic retinopathy. Diabetic medicine : a journal of the British Diabetic Association. 2019;36(4):424-33.
13. Jenkins AJ, Joglekar MV, Hardikar AA, Keech AC, O'Neal DN, Januszewski AS. Biomarkers in Diabetic Retinopathy. The Review of Diabetic Studies. 2015;12(1-2):159-95.
14. Lagani V, Chiarugi F, Thomson S, Fursse J, Lakasing E, Jones RW, et al. Development and validation of risk assessment models for diabetes-related complications based on the DCCT/EDIC data. Journal of diabetes and its complications. 2015;29(4):479-87.
15. Sivo SA, Willson VL. Is parsimony always desirable? Identifying the correct model for a longitudinal panel data set. The Journal of experimental education. 1998;66(3):249-55.
16. Thangaratinam S, Ismail K, Sharp S, Coomarasamy A, O'Mahony F, Khan K, et al. Prioritisation of tests for the prediction of preeclampsia complications: a Delphi survey. Hypertension in pregnancy. 2007;26(1):131-8.
17. Gallagher M, Hares T, Spencer J, Bradshaw C, Webb I. The nominal group technique: a research tool for general practice? Fam Pract. 1993;10(1):76-81.
18. Hayden JA, van der Windt DA, Cartwright JL, Cote P, Bombardier C. Assessing bias in studies of prognostic factors. Ann Intern Med. 2013;158(4):280-6.
19. Richard Poling. Nominal Group Technique. University fo Arkansas. 2009.

20. Trusko B, Thorne J, Jabs D, Belfort R, Dick A, Gangaputra S, et al. The Standardization of Uveitis Nomenclature (SUN) Project. Development of a clinical evidence base utilizing informatics tools and techniques. Methods of information in medicine. 2013;52(3):259-65, s1-6.
21. Angeles-Han ST, Lo MS, Henderson LA, Lerman MA, Abramson L, Cooper AM, et al. Childhood Arthritis and Rheumatology Research Alliance consensus treatment plans for juvenile idiopathic arthritis-associated and idiopathic chronic anterior uveitis. Arthritis care & research. 2018.
22. Avitabile T, Azzolini C, Bandello F, Boscia F, De Falco S, Fornasari D, et al. Aflibercept in the treatment of diabetic macular edema: a review and consensus paper. Eur J Ophthalmol. 2017:0.
23. Douglas RS, Tsirbas A, Gordon M, Lee D, Khadavi N, Garneau HC, et al. Development of criteria for evaluating clinical response in thyroid eye disease using a modified Delphi technique. Archives of ophthalmology (Chicago, Ill : 1960). 2009;127(9):1155-60.
24. Dreer LE, Girkin CA, Campbell L, Wood A, Gao L, Owsley C. Glaucoma medication adherence among African Americans: program development. Optometry and vision science : official publication of the American Academy of Optometry. 2013;90(8):883-97.
25. Ismail R, Azuara-Blanco A, Ramsay CR. Consensus on Outcome Measures for Glaucoma Effectiveness Trials: Results From a Delphi and Nominal Group Technique Approaches. J Glaucoma. 2016;25(6):539-46.
26. Mosca M, Tani C, Aringer M, Bombardieri S, Boumpas D, Brey R, et al. European League Against Rheumatism recommendations for monitoring patients with systemic lupus erythematosus in clinical practice and in observational studies. Annals of the rheumatic diseases. 2010;69(7):1269-74.
27. Radomski MV, Finkelstein M, Llanos I, Scheiman M, Wagener SG. Composition of a vision screen for servicemembers with traumatic brain injury: consensus using a modified nominal group technique. The American journal of occupational therapy : official publication of the American Occupational Therapy Association. 2014;68(4):422-9.
28. Suttle CM, Challinor KL, Thompson RE, Pesudovs K, Togher L, Chiavaroli N, et al. Attitudes and barriers to evidence-based practice in optometry educators. Optometry and vision science : official publication of the American Academy of Optometry. 2015;92(4):514-23.
29. McMillan SS, King M, Tully MP. How to use the nominal group and Delphi techniques. International journal of clinical pharmacy. 2016;38(3):655-62.
30. Fusch PI, & Ness, L. R. Are We There Yet? Data Saturation in Qualitative Research. The Qualitative Report. 2015;20(9):1408-16.
31. Saunders B, Sim J, Kingstone T, Baker S, Waterfield J, Bartlam B, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Quality & quantity. 2018;52(4):1893-907.
32. Palinkas LA, Horwitz SM, Green CA, Wisdom JP, Duan N, Hoagwood K. Purposeful Sampling for Qualitative Data Collection and Analysis in Mixed Method Implementation Research. Administration and policy in mental health. 2015;42(5):533-44.
33. British Medical Association. Doctors' titles: explained 2017 [Available from: file:///C:/Users/Admin/Downloads/PLG-doctors-titles-explained%20(1).pdf.
34. Berger R. Now I see it, now I don't: Researcher's position and reflexivity in qualitative research. Qualitative research. 2015;15(2):219-34.
35. Eleuteri A, Fisher AC, Broadbent DM, Garcia-Finana M, Cheyne CP, Wang A, et al. Individualised variable-interval risk-based screening for sight-threatening diabetic retinopathy: the Liverpool Risk Calculation Engine. Diabetologia. 2017.
36. Oake J, Aref-Eshghi E, Godwin M, Collins K, Aubrey-Bassler K, Duke P, et al. Using Electronic Medical Record to Identify Patients With Dyslipidemia in Primary Care Settings: International Classification of Disease Code Matters From One Region to a National Database. Biomedical informatics insights. 2017;9:1178222616685880.

37. Kohner EM, Stratton IM, Aldington SJ, Holman RR, Matthews DR, Group UKPDS. Relationship between the severity of retinopathy and progression to photocoagulation in patients with Type 2 diabetes mellitus in the UKPDS (UKPDS 52). Diabetic Medicine. 2001;18(3):178-84.
38. Mehlsen J, Erlandsen M, Poulsen PL, Bek T. Identification of independent risk factors for the development of diabetic retinopathy requiring treatment. Acta Ophthalmol. 2011;89(6):515-21.
39. Grauslund J, Green A, Sjolie AK. Blindness in a 25-year follow-up of a population-based cohort of Danish type 1 diabetic patients. Ophthalmology. 2009;116(11):2170-4.
40. DCCT. The effect of intensive diabetes treatment on the progression of diabetic retinopathy in insulin-dependent diabetes mellitus. The Diabetes Control and Complications Trial. Archives of ophthalmology (Chicago, Ill : 1960). 1995;113(1):36-51.
41. Soedamah-Muthu SS, Vergouwe Y, Costacou T, Miller RG, Zgibor J, Chaturvedi N, et al. Predicting major outcomes in type 1 diabetes: a model development and validation study. Diabetologia. 2014;57(11):2304-14.
42. Tanaka S, Tanaka S, Iimuro S, Yamashita H, Katayama S, Akanuma Y, et al. Predicting macro- and microvascular complications in type 2 diabetes: the Japan Diabetes Complications Study/the Japanese Elderly Diabetes Intervention Trial risk engine. Diabetes care. 2013;36(5):1193-9.
43. Stratton IM, Kohner EM, Aldington SJ, Turner RC, Holman RR, Manley SE, 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.
44. Sjolie AK, Stephenson J, Aldington S, Kohner E, Janka H, Stevens L, et al. Retinopathy and vision loss in insulin-dependent diabetes in Europe. The EURODIAB IDDM Complications Study. Ophthalmology. 1997;104(2):252-60.
45. Harris Nwanyanwu K, Talwar N, Gardner TW, Wrobel JS, Herman WH, Stein JD. Predicting development of proliferative diabetic retinopathy. Diabetes care. 2013;36(6):1562-8.
46. Klein R, Klein BE, Moss SE, Cruickshanks KJ. The Wisconsin Epidemiologic Study of Diabetic Retinopathy: XVII. The 14-year incidence and progression of diabetic retinopathy and associated risk factors in type 1 diabetes. Ophthalmology. 1998;105(10):1801-15.
47. Lane M, Mathewson PA, Sharma HE, Palmer H, Shah P, Nightingale P, et al. Social deprivation as a risk factor for late presentation of proliferative diabetic retinopathy. Clinical ophthalmology (Auckland, NZ). 2015;9:347-52.
48. Denniston AK, Lee AY, Lee CS, Crabb DP, Bailey C, Lip PL, et al. United Kingdom Diabetic Retinopathy Electronic Medical Record (UK DR EMR) Users Group: report 4, real-world data on the impact of deprivation on the presentation of diabetic eye disease at hospital services. The British journal of ophthalmology. 2018.
49. Aspelund T, Porisdottir O, Olafsdottir E, Gudmundsdottir A, Einarsdottir AB, Mehlsen J, et al. Individual risk assessment and information technology to optimise screening frequency for diabetic retinopathy. Diabetologia. 2011;54(10):2525-32.
50. Yau JW, Rogers SL, Kawasaki R, Lamoureux EL, Kowalski JW, Bek T, et al. Global prevalence and major risk factors of diabetic retinopathy. Diabetes care. 2012;35(3):556-64.
51. Hayes AJ, Leal J, Gray AM, Holman RR, Clarke PM. 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. 2013;56(9):1925-33.
52. Davis TM, Stratton IM, Fox CJ, Holman RR, Turner RC. U.K. Prospective Diabetes Study 22. Effect of age at diagnosis on diabetic tissue damage during the first 6 years of NIDDM. Diabetes care. 1997;20(9):1435-41.
53. Clarke PM, Gray AM, Briggs A, Farmer AJ, Fenn P, Stevens RJ, et al. 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. 2004;47(10):1747-59.

54. Stratton IM, Adler AI, Neil HA, Matthews DR, Manley SE, Cull CA, 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.
55. Stratton IM, Aldington SJ, Farmer AJ, Scanlon PH. Personalised risk estimation for progression to sight-threatening diabetic retinopathy: How much does clinical information add to screening data? Diabetic Medicine. 2014;31:23-4.
56. Man RE, Sasongko MB, Wang JJ, MacIsaac R, Wong TY, Sabanayagam C, et al. The Association of Estimated Glomerular Filtration Rate With Diabetic Retinopathy and Macular Edema. Invest Ophthalmol Vis Sci. 2015;56(8):4810-6.
57. Romero-Aroca P, Baget-Bernaldiz M, Navarro-Gil R, Moreno-Ribas A, Valls-Mateu A, Sagarra-Alamo R, et al. Glomerular Filtration Rate and/or Ratio of Urine Albumin to Creatinine as Markers for Diabetic Retinopathy: A Ten-Year Follow-Up Study. J Diabetes Res. 2018;2018:5637130.
58. Klein BE, Moss SE, Klein R, Surawicz TS. The Wisconsin Epidemiologic Study of Diabetic Retinopathy. XIII. Relationship of serum cholesterol to retinopathy and hard exudate. Ophthalmology. 1991;98(8):1261-5.
59. Matthews DR, Stratton IM, Aldington SJ, Holman RR, Kohner EM. Risks of progression of retinopathy and vision loss related to tight blood pressure control in type 2 diabetes mellitus: UKPDS 69. Archives of ophthalmology (Chicago, Ill : 1960). 2004;122(11):1631-40.
60. Adler AI, Stratton IM, Neil HA, Yudkin JS, Matthews DR, Cull CA, 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.
61. Klein R, Lee KE, Gangnon RE, Klein BE. The 25-year incidence of visual impairment in type 1 diabetes mellitus the wisconsin epidemiologic study of diabetic retinopathy. Ophthalmology. 2010;117(1):63-70.
62. Shi R, Zhao L, Wang F, Liu F, Chen Z, Li R, et al. Effects of lipid-lowering agents on diabetic retinopathy: a Meta-analysis and systematic review. International journal of ophthalmology. 2018;11(2):287-95.
63. Klein R, Klein BE, Moss SE, Davis MD, DeMets DL. The Wisconsin epidemiologic study of diabetic retinopathy. IV. Diabetic macular edema. Ophthalmology. 1984;91(12):1464-74.
64. Jeng CJ, Hsieh YT, Yang CM, Yang CH, Lin CL, Wang IJ. Diabetic Retinopathy in Patients with Diabetic Nephropathy: Development and Progression. PloS one. 2016;11(8):e0161897.
65. Gupta A, Gupta V, Thapar S, Bhansali A. Lipid-lowering drug atorvastatin as an adjunct in the management of diabetic macular edema. Am J Ophthalmol. 2004;137(4):675-82.
66. Ohkubo Y, Kishikawa H, Araki E, Miyata T, Isami S, Motoyoshi S, et al. Intensive insulin therapy prevents the progression of diabetic microvascular complications in Japanese patients with non-insulin-dependent diabetes mellitus: a randomized prospective 6-year study. Diabetes research and clinical practice. 1995;28(2):103-17.
67. UKPDS 33. 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.
68. McEwan P, Bennett H, Ward T, Bergenheim K. Refitting of the UKPDS 68 risk equations to contemporary routine clinical practice data in the UK. Pharmacoeconomics. 2015;33(2):149-61.
69. Klein R, Knudtson MD, Lee KE, Gangnon R, Klein BE. The Wisconsin Epidemiologic Study of Diabetic Retinopathy: XXII the twenty-five-year progression of retinopathy in persons with type 1 diabetes. Ophthalmology. 2008;115(11):1859-68.
70. Zhou Y, Zhang Y, Shi K, Wang C. Body mass index and risk of diabetic retinopathy: A meta-analysis and systematic review. Medicine (Baltimore). 2017;96(22):e6754.
71. DCCT. Early worsening of diabetic retinopathy in the Diabetes Control and Complications Trial.[Erratum appears in Arch Ophthalmol 1998 Nov;116(11):1469]. Archives of Ophthalmology. 1998;116(7):874-86.

72. Feldman-Billard S, Larger E, Massin P. Early worsening of diabetic retinopathy after rapid improvement of blood glucose control in patients with diabetes. Diabetes Metab. 2018;44(1):4-14.
73. Forster AS, Forbes A, Dodhia H, Connor C, Du Chemin A, Sivaprasad S, et al. Non-attendance at diabetic eye screening and risk of sight-threatening diabetic retinopathy: a population-based cohort study. Diabetologia. 2013;56(10):2187-93.
74. Klein BE, Moss SE, Klein R. Effect of pregnancy on progression of diabetic retinopathy. Diabetes care. 1990;13(1):34-40.
75. DCCT. Effect of pregnancy on microvascular complications in the diabetes control and complications trial. The Diabetes Control and Complications Trial Research Group. Diabetes care. 2000;23(8):1084-91.
76. Muqit K, Yang Y, Lipinski H, Chong V. Progression of pre-proliferative to proliferative diabetic retinopathy: a 3-year study in the Oxford population-based diabetic retinopathy screening programme. Diabetic Medicine. 2014;31(8):1018-9.
77. Klein R. Hyperglycemia and microvascular and macrovascular disease in diabetes. Diabetes care. 1995;18(2):258-68.
78. Accord Study Group, Accord Eye Study Group, Chew EY, Ambrosius WT, Davis MD, Danis RP, et al. Effects of medical therapies on retinopathy progression in type 2 diabetes.[Erratum appears in N Engl J Med. 2011 Jan 13;364(2):190], [Erratum appears in N Engl J Med. 2012 Dec 20;367(25):2458]. New England Journal of Medicine. 2010;363(3):233-44.
79. Kawasaki R, Tanaka S, Tanaka S, Yamamoto T, Sone H, Ohashi Y, et al. Incidence and progression of diabetic retinopathy in Japanese adults with type 2 diabetes: 8 year follow-up study of the Japan Diabetes Complications Study (JDCS). Diabetologia. 2011;54(9):2288-94.
80. Sivaprasad S, Gupta B, Gulliford MC, Dodhia H, Mohamed M, Nagi D, et al. Ethnic variations in the prevalence of diabetic retinopathy in people with diabetes attending screening in the United Kingdom (DRIVE UK). PloS one. 2012;7(3):e32182.
81. Meltzer D, Egleston B. How patients with diabetes perceive their risk for major complications. Effective clinical practice : ECP. 2000;3(1):7-15.
82. Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med. 2015;162(1):W1-73.
83. Lee CS, Lee AY, Baughman D, Sim D, Akelere T, Brand C, et al. The United Kingdom Diabetic Retinopathy Electronic Medical Record Users Group: Report 3: Baseline Retinopathy and Clinical Features Predict Progression of Diabetic Retinopathy. American Journal of Ophthalmology. 2017;180:64-71.
84. Gupta G, Kulasekaran S, Ram K, Joshi N, Sivaprakasam M, Gandhi R. Local characterization of neovascularization and identification of proliferative diabetic retinopathy in retinal fundus images. Computerized Medical Imaging & Graphics. 2017;55:124-32.
85. Riley RD, van der Windt D, Croft P, Moons KGM. Prognosis Research in Healthcare: Concepts, Methods, and Impact: OUP Oxford; 2019

4.10 Appendices for Chapter 4

5 Chapter 5: Development and Internal Validation of a Multivariable Prediction Model

This chapter is submitted as a manuscript for publication along with its appendices. The aim of this study was to develop and internally validate a prediction model for the progression of diabetic retinopathy to a treatment requiring stage or to vision loss in patients with referable diabetic retinopathy in a primary care database.

Development and Internal Validation of a Multivariable Prediction Model

Sajjad Haider[1], Anuradhaa Subramanian[3], Nicola J. Adderley[3], Kym I. E. Snell[4], Konstantinos A Toulis[3], David Moore[2], Krishnarajah Nirantharakumar[2], Malcolm James Price[2]

1. Doctoral researcher, Institute of Applied Health Research, University of Birmingham
2. *Joint Senior Authors Institute of Applied Health Research, University of Birmingham, Birmingham, United Kingdom
3. Institute of Applied Health Research, University of Birmingham
4. Centre for Prognosis Research, Keele University, Staffordshire, United Kingdom

Joint Corresponding authors:

Sajjad Haider, Doctoral Researcher, Institute of Applied Health Research, University of Birmingham, Birmingham, United Kingdom

E mail: [redacted]

Krishnarajah Nirantharakumar, Institute of Applied Health Research, University of Birmingham, Birmingham, United Kingdom

E mail: [redacted]

5.1 Abstract

5.1.1 Study background and aims

Patients with referable diabetic retinopathy are treated or regularly observed for the need and timing of treatment to prevent vision loss. A small but significant proportion progress to that stage, but presently there is no mechanism to risk stratify these patients. We aimed to develop and internally validate a prediction model to identify higher risk patients in this group.

5.1.2 Methods

Anonymized patient data was extracted from in the IQVIA Medical Research Data (a structured primary care routine practice database). Data on patients with diabetes joining the database from 2004 to 2018 were used for the model derivation and internal validation. We applied Cox regression with automated stepwise backward elimination using $p< 0.2$. We compared multiple models to test different scenarios and assessed discrimination and calibration. We then performed internal validation using bootstrap.

5.1.3 Results

The complete case analyses model was finally selected following a sensitivity analysis based on its stability with the predictors and overall satisfactory performance. This model included Diabetic Retinopathy stage, eGFR stage, HbA1C, Cholesterol, systolic blood pressure, Insulin and Statin use as final predictors. It identified patients with R3/ M1 retinopathy as those with the highest and patients with R2 retinopathy as those with the lowest risk. The models had a good calibration slope, a discrimination of 0.74 and optimism of .0039559 on internal validation.

5.1.4 Conclusion

This new algorithm can calculate the absolute risk of individual patient for progression from sight threatening retinopathy to treatment requiring stage or vision loss within a 5-year period based on their individual risk factors. External validation and decision curve analysis will be required before the model can be recommended for use in clinical practice.

Summary Box

What was already known on the subject of Model development and Internal Validation for diabetic retinopathy progression?

- There are quite a few models addressing this subject in the lower risk population under care of Diabetic Eye Screening Programme.
- None of these models start with the exposure of referable diabetic retinopathy and most of them deal with the outcome of patients reaching a stage of referral diabetic retinopathy.

What this study adds?

- This is the only study to date on the subject addressing the population of interest (diabetic population under care of hospital eye services).
- This is also the only study to date addressing the composite outcome of need for treatment and vision failure.
- Clinically relevant to an area not previously investigated, and is under severe stress with increased demand.

5.2 Introduction

There is a global and UK wide increase in the number of people with diabetes mellitus (1-3). Consequently, the burden of disease of diabetic retinopathy (DR), one of the major complications of diabetes has also increased (4). In the UK, diabetic retinopathy services are organised into the Diabetic Eye Screening Programme (DESP) for lower risk patients (defined as either no diabetic retinopathy (R0) or background diabetic retinopathy (R1)) and the Hospital Eye Services for higher risk patients needing closer observation and treatment, when they develop clinical signs of sight threatening retinopathy (STR) including the pre-proliferative diabetic retinopathy (R2), proliferative diabetic retinopathy (R3) and diabetic maculopathy (M1). The detection of retinopathy has also improved through wider population coverage (4), increasing the workload for both DESP and Hospital Eye Services. Approximately 50-70% of referrals do not require intervention and are observed in the Hospital Eye Services (5, 6), for a variable period of time. Difficulty in discriminating patients among them needing early intervention is a factor contributing to the increasing demand and capacity mismatch in hospital eye services. Risk stratification to estimate patients' individual risk for DR progression has already being used as a decision support instrument to establish screening intervals in DESP (7, 8), could also be of clinical value in the hospital setting.

In general, such prognostic prediction models combine the predictive performance of two or more predictor variables (9). Our recent systematic review (10) of the existing predictive models identified 14 distinct models, three of which had a moderate to low risk of bias and partial applicability to hospital eye service patients (8, 11, 12) assessed using PROBAST (13) and CHARMS (14) checklists. However, these three

models focused on the low risk patients, were modelled to detect incidence of referable retinopathy (as an outcome and not as exposure). Therefore, the clinical need to develop a similar predictive model, which would focus on patients with referable retinopathy and thereby would predict DR progression to the stage requiring therapy / or to vision loss, is conspicuous and imperative.

Aims:

Therefore we aimed to develop and internally validate a prognostic prediction model to predict diabetic retinopathy progression to treatment stage or vision loss within 5 years of diagnosis of referable diabetic retinopathy stage in patients with diabetes in a primary care database.

5.3 Methods

5.3.1 Study design

The study was a retrospective open cohort study in patients with type 1 or type 2 diabetes conducted between 2004 and 2018 in patients with referable diabetic retinopathy stage.

5.3.2 Data Source

Anonymised patient data were extracted from the IQVIA Medical Research Data (IMRD) database, which collects structured longitudinal data of patients registered in primary care. IMRD database has previously been used for a wide range of clinical research, including the development and validation of risk prediction models for diabetic retinopathy (15). Symptoms, diagnoses, referrals and screening data are recorded in the primary care system using Read codes, and prescriptions are

recorded according to the British National Formulary. Additionally, IMRD includes data on patient demographics collected by general practices when patients register.

5.3.3 Study population, inclusion, and exclusion criteria

Practices contributing to the IMRD database were eligible and included for data inclusion only when they were flagged as appropriate for research purposes i,e., a year after starting to use the Vision system for computerization of patient records and a year after the practice mortality rate became comparable to the national mortality rate.

Patients registered with an eligible general practice were included in our cohort if they were: 1) registered with an eligible practice at any time between 1st Jan 2004 to 31st December 2018; 2) aged 12 and above; 3) had a coded diagnosis of diabetes (identified as type 1 or type 2) and a diabetes duration of a minimum of 15 months before study entry, and thus providing before the study entry an adequate latency time or window of opportunity to undergo a retinopathy screening; 4) had a coded diagnosis of sight threatening retinopathy (R2, R3 and M1 or a combination thereof) after the diagnosis of diabetes; and 5) did not have a record of the outcome (a composite of vision loss, blindness, vitreous haemorrhage or procedures indicative of treatment such as laser and vitreous injections or surgery) at that time.

The exposure of referable retinopathy was categorized using Read code records as 1) R2 only as a reference category; 2) R3 only (without visual loss or vitreous haemorrhage as these are outcomes); 3) M1 only; 4) concurrent diagnosis of R2 and M1; (5) concurrent diagnosis of R3 and M1; and 5) generic record of referable retinopathy, not suitable to be classified into any of the above categories. While the primary model was designed with R2 only as a reference category, we also designed

a model with M1 as a reference category for exploratory analysis. A competing risk exploratory analysis was also carried out for the same purpose.

5.3.4 Follow-up period

Patient follow-up started at the time of diagnosis of referable diabetic retinopathy (R2, R3 or M1) (index date) and continued until the date of diagnosis of vision loss or the date of recording of the treatment procedures performed. Follow-up was censored at the earliest of, date of patient deregistering or transferring out of the GP surgery, last data collection date from the GP surgery, date of death, or the end of study period.

5.3.5 Selection of candidate predictors

A systematic review of existing prognostic models for diabetic retinopathy identified 78 candidate predictors (10). This list included significant overlap and duplication (e.g. diabetic nephropathy and chronic kidney disease); furthermore, it was not feasible to explore 78 predictors in this dataset. Therefore, to reduce the list of candidate predictors, a qualitative method called Nominal Group Technique (NGT) was performed by pooling opinions from clinical experts. Four separate NGT meetings were held in four different NHS trusts in the United Kingdom, with NHS hospital-based clinicians involved in caring for diabetic retinopathy patients. This was done to rank the predictors for importance by voting. This was followed by evidence evaluation for confirmation using QUIP tool (16). Feasibility of using each predictor was assessed at the time of data inspection of the IMRD database.

Demographic candidate predictors included age at diagnosis of sight threatening retinopathy, sex, socioeconomic status and ethnicity. Social deprivation was recorded as deprivation quintiles based on Townsend score graded from 1 to 5, 1

indicating the lowest deprivation and 5 indicating the highest deprivation. Recorded ethnicity was categorized as white, South Asian, Afro-Caribbean, and other ethnic group. Candidate predictors related to the diabetes diagnosis included diabetes type and duration (Diabetes Diagnosis to index date). Biochemical parameters included glycated haemoglobin A1c (HbA1c), glomerular filtration rate and serum cholesterol measured on or near to index date (diagnosis of referable diabetic retinopathy). eGFR was considered as a categorical variable as eGFR values >60 ml / min per 1.73m^2 are indicative of a mild disease. eGFR was categorized based on chronic kidney disease status: category 1 (stage 1 to 2 or > 60 ml / min per 1.73m^2), category 2 (stage 3 or 30-59) and category 3 (stage 4 or 15 to 29 ml / min) (17). Treatment with statins and insulin were also included in the model.

5.3.6 Statistical analysis

Sample size

Our sample included all eligible patients in the IMRD database. We assessed the adequacy of the sample size using recommendation by Riley et al and calculating events / per parameter considering three conditions. 1) The global shrinkage factor of > 0.9, 2) Small absolute difference of ≤ 0.05 in the model's apparent and adjusted Nagelkerke's R^2, 3) Precise estimation of the overall risk in the population (18)

Model development

We fitted a Cox proportional hazards regression model for vision loss or treatment (outcome). In the primary analysis only complete case data were used including only those participants with no missing data. Continuous predictors values were means centred. Predictor selection was performed using stepwise backward elimination with

a cut off of 0.2 for the p-value (19). For prediction, we estimated the baseline hazard by 5 years.

Exploratory Analysis

For the model to meet the clinical needs, in addition to the primary model (complete case analysis) for a full follow up period, model for 2 and 5 years, a multiple imputation model, a further model employing fractional polynomials (to mitigate the risk of bias from non-linearity) was also assessed. We also built models based on age groups, sex, systemic predictors only, ocular predictors only, M1 as a reference category and type 1 and 2 diabetes specific models. We also assessed models to test different clinical scenarios like systemic predictors only against ocular predictors, a model only for men against a model only for women, < 60 years old and > 60 years old and explored the need for competing risk model by looking into the effect of death as a competing outcome (20) using Fine and Gray model (21). All models were compared to the complete case analysis model using the criteria of stability of predictors, no of outcomes and model predictive performance.

5.3.6.1 Internal validation and Model performance:

Internal validation was carried out using bootstrapping (19) (Appendix 27). Predictive performance of the model was assessed using measures of goodness of fit, calibration and discrimination. Goodness of fit was assessed using the Martingale and the Cox-Snell residuals. Calibration measures included the expected/observed ratio (E/O) which is calculated as the average expected probability of the outcome (model prediction) divided by average observed outcome probability at 5 years (ideal value is 1) , a calibration slope (ideal value is 1) and a calibration plot showing observed versus expected probabilities for risk groups categorised using deciles of

predicted probabilities at 5 years (19). Discrimination was assessed using Harrell's C and (22). We also calculated the Heuristic shrinkage factor (23).

All analyses were performed using Stata IC version 14.2 (24). TRIPOD guidelines (25, 26) were followed for the transparent reporting of prediction model development and validation.

Assessment of model assumptions

Proportionality for categorical variables was assessed through inspection of log- log plots (18). For continuous variables we plotted the Schoenfeld residuals (19) to check for proportional hazards assumption. Potential interactions were identified from a previous prognostic model in the systematic review (27) and considered.

Missing data

Missing data on ethnicity, Townsend deprivation (21) quintile, and eGFR category were considered as a separate missing category within the corresponding variable. In the primary analysis only complete case data were used including only those participants with no missing data on serum cholesterol, HbA1c and blood pressure. In a subsequent analysis, missing values were imputed using multivariate imputation by chained equations (28, 29). We generated 10 imputations for HBA1C, Cholesterol, Systolic and diastolic blood pressure. We used Rubin's rule (30, 31) to combine estimates of the regression coefficients and to combine the predictive performance statistics.

5.4 Results

5.4.1 Study population

8,480,167 eligible patients aged 12 and above in 784 UK practices, during the during the study period were considered for inclusion. 25,820 patients had the exposure of referable diabetic retinopathy. Exclusions were: 9,307 patients with an outcome of interest on index date and 2,822 patients with less than 15 months of diabetes duration at study entry. 13,691 patients met our inclusion criteria so their data were used to develop the prediction model (Figure *1*).

5.4.2 Sample size adequacy

There were 2079 outcome events; 425 (17%) had a code for vision loss and 2143 (83%) for evidence of treatment. With 9964.8 years mean follow up time, outcome events with an overall event rate of 15.2%, and therefore an Event Per Parameter of 54.09, assuming a shrinkage factor of 0.9 and R^2 of .0866, a minimum sample size required for new model development was 2768 arrived by calculating using Stata IC version 14.2 (24).

5.4.3 Baseline characteristics / Predictor variables

The baseline characteristics of these patients are given in Table 1 below. Mean age at STR diagnosis was 60 years in men and 63 years in women. Fifty nine (59) % were female and 17% had type 1 diabetes. Thirty two (32)% had been diagnosed as having diabetes less than five years before cohort entry, 30% 5-10 years before, 19% 10-15 years before, and 19% > 15 years before. Missing-ness affected Ethnicity (48%), Townsend score (15.1%), eGFR (1.4%), HbA1C (6.1%), Cholesterol (2.3%),

and Systolic and Diastolic blood pressure (0.1%). The first three of these predictors had a missing category created. The next four were dealt with multiple imputation.

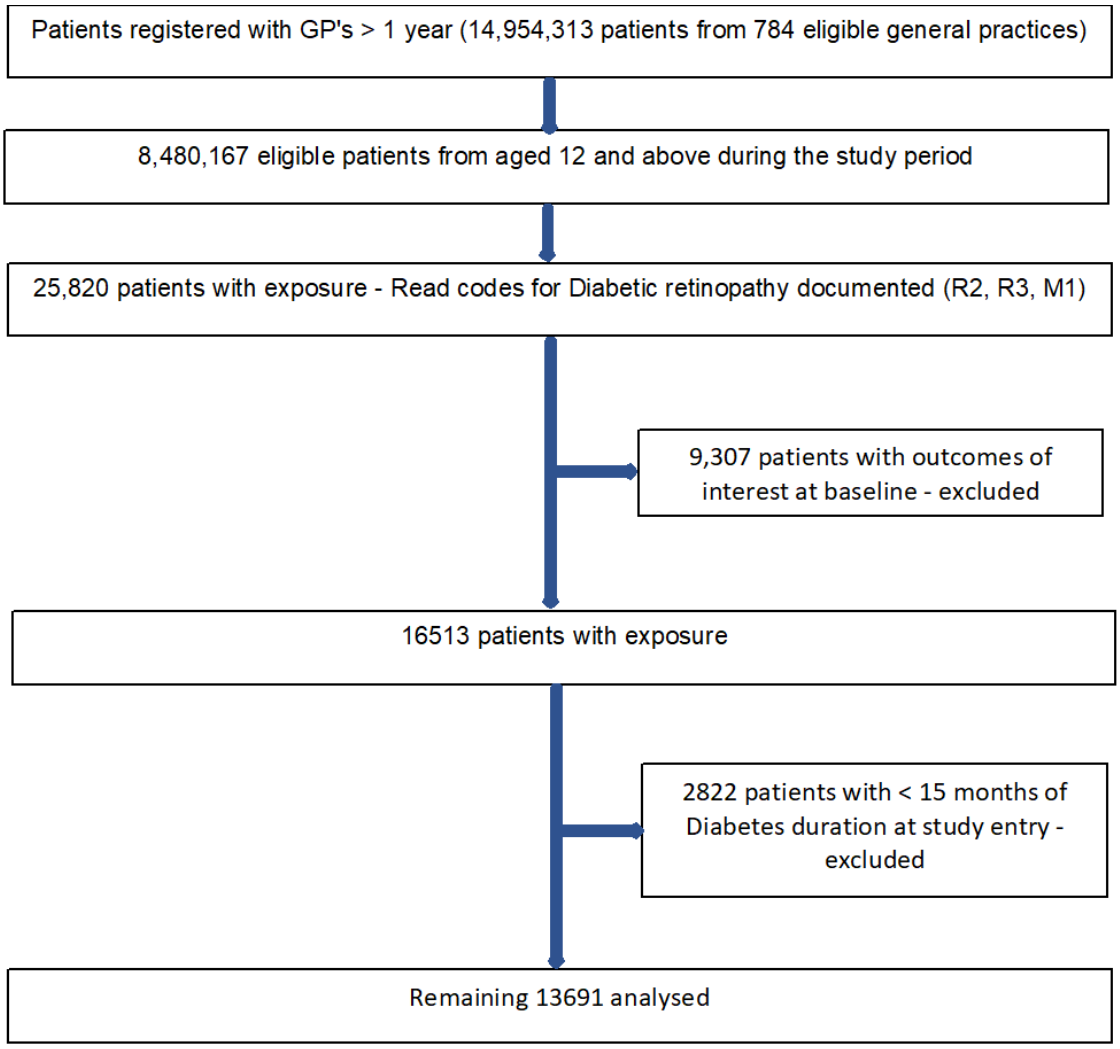


Figure 1: Patients Flow Diagram

Table 1: Population and Predictors

Variable (missing data)	Men (% / SD)	Women (% / SD)
Ocular features		
* £ Pre-proliferative DR (R2)	491 (6.1%)	296 (5.2%)
Proliferative diabetic retinopathy (R3)	1419 (17.7%)	848 (15%)
Pre-proliferative DR & Maculopathy (R2 and M1)	373 (4.6%)	207 (3.7%)
Proliferative diabetic retinopathy & Maculopathy (R3 and	822 (10.2%)	513 (9.1%)
Maculopathy (M1)	4755 (59.2%)	3647 (64.5%)
Unclassifiable sight threatening retinopathy	174 (2.2%)	146 (2.6%)
Socio demographics		
Age at STR diagnosis (years)	60.5 (14.5)	63.3 (16)
* Race - White	3670 (45.7%)	2478 (43.8%)
Race - Black	118 (1.5%)	114 (2%)
Race - Mixed	42 (1.1%)	42 (0.7%)
Race – Others	24 (0.3%)	23 (0.4%)
Race - South Asian	346 (4.3%)	256 (4.5%)
Race - Missing (48%)	3834 (47.7%)	2744 (48.5%)
Sex	8034 (58.7%)	5657 (41.3%)
*Townsend score 1	1398 (17.4%)	863 (15.3%)
Townsend score 2	1488 (18.5%)	927 (16.4%)
Townsend score 3	1528 (19%)	1038 (18.3%)
Townsend score 4	1376 (17.1%)	1137 (20.1%)
Townsend score 5	1033 (12.9%)	833 (14.7 %)
Townsend score 6 Missing (15.1%)	1211 (15.1%)	859 (15.2%)
Diabetes characteristics		
Type 1 DM	1416 (17.6%)	932 (16.5%)
Type 2 DM	6618 (82.4)%	4725 (83.5%)
Duration of DM (years)	9.5 (7.8)	9.8 (7.8)
Biochemical variables		
HbA1c mmol/mol (6.12%)	67.6 (20)	67.9 (21.1)
$* eGFR stage 1 and 2	6244 (77.7%)	3814 (67.4%)
eGFR stage 3	1441 (17.9%)	1517 (26.8%)
eGFR stage 4	232 (2.9%)	255 (4.5%)
eGFR Missing (1.4%)	117 (1.5%)	71 (1.3%)
Total Serum Cholesterol mmol / L (2.31%)	4.2 (1.1)	4.6 (1.1)
Physical examination		
Systolic Blood pressure (0.12%)	134.9 (16.2)	135.4 (17.9%)
Diastolic Blood Pressure (0.12%)	76 (10.2)	74.6 (10.1)
Drug treatment		
Statin	6424 (58.3%)	4516 (56.9%)
Insulin	4600 (41.7%)	3427 (43.1%)

$ eGFR, estimated glomerular filtration rate, * Reference categories, £ Belongs to NGT (new) category shifted to Ocular features for clinical reasons.

This table demonstrates the baseline characteristics in this IMRD model development cohort of referable diabetic retinopathy population. Values given are numbers (percentages) or, in the case of continuous predictors, the mean (SD).

5.4.4 Survival analysis

In the total sample (n=13,691 participants), median (IQR) follow-up was 3.23 (1.34-5.96) years. In participants who went on to develop the outcome (n=2079, 15.2%), median (IQR) follow-up to outcome was 1.3 (0.4-3.2) years. In participants who did not develop the outcome (n=11,612, 84.8%), median (IQR) follow-up was 3.9 (3.8 - 4) years. The crude incidence rate of the outcome of treatment or vision loss was 38.4 per 1000 person-years (in our population number of outcomes / divided by follow up time).

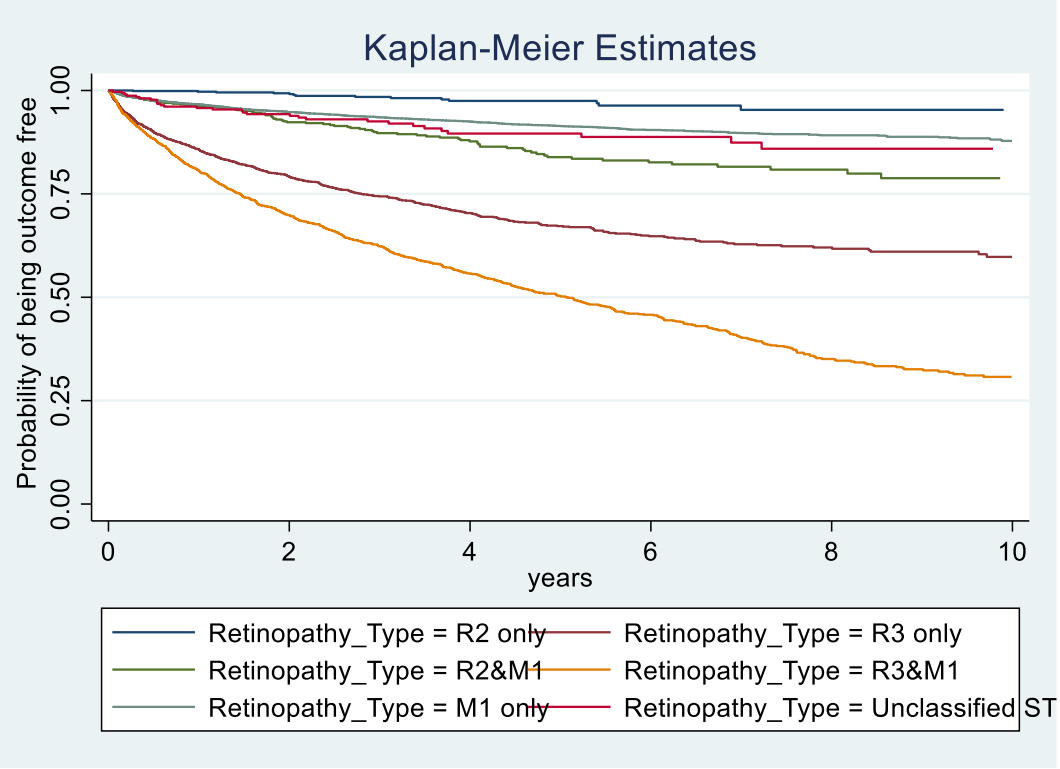


Figure 2: Kaplan-Meier Estimates by Retinopathy (Proportions without the outcome)

The probability of being outcome-free over time was highest in retinopathy type R2 closely followed by M1 (Figure 2). The categories with the lowest probability of being outcome-free were R3M1 followed by R3.

5.4.5 Statistical analysis

Model Development

In the primary model (complete case analysis) the final set of predictors included in the model were 1) retinopathy grade, 2) HbA1C, 3) eGFR categories, 4) SBP, 5) cholesterol, 6) insulin and 7) statins. Hazard ratios for the predictors in the final model are given in Table 3. The heuristic shrinkage factor was 0.99939707 and therefore shrinkage was not required. The baseline survival for 5 years was 0.9760651 and feeds into the prediction equation for an individual.

Exploratory Analysis

A comparative table between R2 as a reference and M1 as a reference is given in Appendix 28. In rest of the models the number of outcomes were smaller and the models' stability as well as discriminatory ability was suboptimal. The predictors in other analyses were mostly in common with those included in the primary model. DR Categories were the commonest predictors shared by every model. Rest of the models did not retain all of the predictors set in complete case analysis as they had p value > 0.2.. eGFR was the most frequent variables to be eliminated (6 out of 14 models). SBP, statins, Insulin and Cholesterol were not included in 4 of the models, while, HbA1C was not included in only one of the models intentionally. The C statistic (performance index) was lower in models with systemic factors only (0.62) and model for > 60 years of age (0.69) (Table 2).

In competing risk exploratory analysis, more events occurred for treatment or vision failure at earlier time points up to 2 years compared to death but then more deaths occurred by later time points from 3rd year onwards. The running total of deaths took over from 6th year onwards (Table 1:, Appendix 32). On Kaplan Meier failure analysis, the probability of the outcome isn't very different at the time points of interest, probably about 2% absolute difference at 5 years (figure 1, Appendix 32).

Table 2: Exploratory Analysis

Model	**Outcomes / Patients**	**Effect on predictors in the final model**	**Harrell's C**
Complete case analysis with R2 as reference category	1917 / 12584	DR Categories, SBP, HBA1C, Cholesterol, eGFR, Statin, Insulin	0.74
Complete case analysis with M1 as reference category	1917 / 12584	DR Categories, SBP, HBA1C, Cholesterol, eGFR, Statin, Insulin	0.74
Model for follow-up ≤ 2 year and outcome within 2 year	1171 / 12584	DR Categories, SBP, eGFR, HBA1C, Cholesterol, Statin, Insulin, age at STR diagnosis	0.73
Model for follow-up ≤ 5 year and outcome within 5 year	1687/12584	DR Categories, Diabetes duration, Age at STR diagnosis, SBP, HBA1C, Cholesterol, eGFR, Statin, Insulin	0.73
Cox - with multiple imputation	2079 / 13691	As above	0. 74
Cox - with multiple imputation and fractional polynomials (Appendix 31)	2079 / 13691	DR Categories, SBP, HBA1C, eGFR, Statin, Insulin	0.74
Ocular Predictors only	2079 / 13691	DR Categories only	0.72
Systemic Predictors only	1917 / 12584	SBP, HBA1C, Cholesterol, Statin, Insulin	0.62
Model for T1DM	446 / 2173	DR Categories, HBA1C, Sex	0.77
Model for T2DM	1,471 / 10411	DR Categories, Sex, SBP, HBA1C, Cholesterol, eGFR, Statin, Insulin	0.73

Model for men	1,159 / 7403	DR Categories, SBP, HBA1C, Cholesterol, eGFR, Statin	0.74
Model for women	758 / 5181	DR Categories, Diabetes type, HBA1C, Cholesterol, Insulin	0.74
Model for ≤ 60 years of age	997 / 5337	DR Categories, HBA1C, Cholesterol, eGFR, Age at STR diagnosis	0. 78
Model for > 60 years of age	920 / 7247	DR Categories, SBP, Diabetes type, HBA1C, Insulin, Statin	0.69

C statistic (concordance index) and Somers' D statistic are given and compared between different models.

* The p value for these variables crosses the cut off of > 0.2 with this option

Model performance:

On goodness of fit test, the model fitted with the data well overall (Appendix 27). The E/O ratio was 0.95, which means the model is slightly under-predicting the outcomes. The calibration plot displaying observed vs predicted risks is shown in Figure 3. The Harrell's C statistic was a value of 0.738 and a 95% confidence interval of 0.726-0.750. It showed a small amount of optimism, thus the adjusted figure for C statistics was 0.734 and the calibration slope was 0.98.

Measure	Apparent	Average optimism	Optimism Adjusted
C statistic	.7381313	.0039559	.73417539
Calibration slope	1	.0160065	.98399345

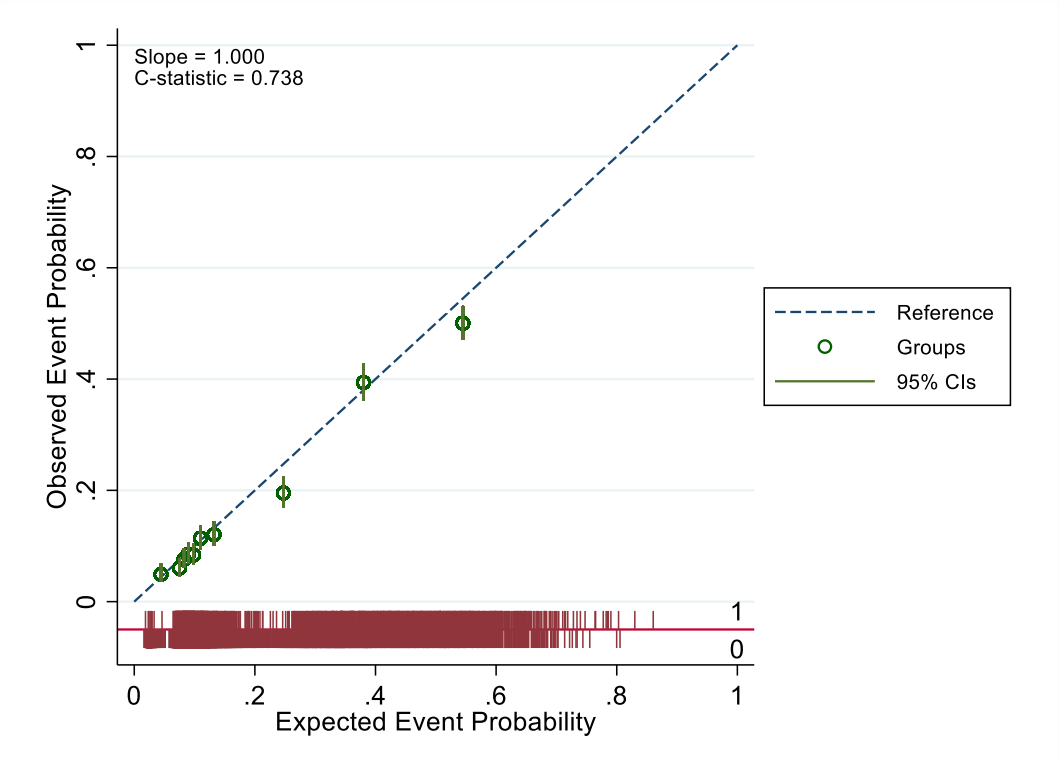


Figure 3: Calibration Plot

In this pmcal plot the predicted proportion is shown on the x-axis and the observed proportion on the y-axis.

Model equation

The baseline survival probability (intercept) is 0. 98

$h(t)=h_0(t)\exp(\beta_1X_1+\beta_2X_2+\ldots\ldots\ldots\ldots..\beta pXp)$

h(t)= .9760651 exp (1 R2 +2.72682 R3=1 + 1.822814 R2/M1=1 + 3.242615 R3/ M1=1 + 1.330706 M1=1 + 1.785901 Unclassified=1 + .0079924 (HbA1C - 67.7) + .116896 eGFR3=1 + . .4795509 eGFR4=1 +.0026764 (SBP- 135) + . .0407135 (Cholesterol- 4.4) + .1079604 Insulin=1 -.1269512 Statins=1)

Continuous predictor values were means centred and are shown underlined.

Assessment of model assumptions and missing data analysis

Proportionality assumption results are shown in Appendix 30Appendix 30. The proportional hazards assumption was valid. There was some suspicion of non-linearity of the predictor data in the three continuous predictors namely HBA1C, SBP and Cholesterol. We therefore also developed a fractional polynomial model and a model using multiple imputation to account for missing data. The final set of predictors was smaller (cholesterol excluded) in fractional polynomial model as compared to complete case analysis model but very similar in performance statistics.

Further analysis

Diabetic retinopathy grades gave the highest hazard values. The hazard for R3/ M1 and R3 had 26 and 15 times the hazard respectively as compared to the R2 group. The R2//M1 and Unclassified had a 6 times hazard. M1 had a 3.78 times hazard. The systemic predictors do offer a contribution to prognostic value of the model. A one unit increase in the HBA1C and SBP above the mean value increased the hazard by 1% and 0.3% respectively and same increase above the mean in the of eGFR 3 and eGFR 4 value was associated with a 12% and 62% increase in hazard respectively.

Every unit increase in Cholesterol value showed a change of 4% in hazard. Being on Insulin increased the risk by 11%, while being on Statin was protective. Age at STR diagnosis was neutral.

Table 3: Hazard ratios for the predictors

Predictor	Complete case analysis Model	95% Confidence intervals		P value
R2	1			
R3	15.28	8.817684	26.49302	<0. 001
R2/M1	6.19	3.432152	11.16116	<0.001
R3/ M1	25.6	14.78035	44.342	<0.001
M1	3.78	2.182933	6.558371	<0.001
Unclassified	5.96	3.099274	11.48031	<0.001
Age at STR diagnosis	0.99	0.99	1	0.002
HbA1C	1.01	1.005785	1.010269	<0.001
eGFR3	1.12	1.001647	1.270192	0.047
eGFR4	1.62	1.270192	2.054297	<0,001
SBP	1.003	1.000145	1.005221	0.038
Cholesterol	1.04	.9988176	1.086118	0.057
Insulin	1.11	.9994152	1.24173	0.051
Statins	0.88	.7915568	.9800529	0.02

5.5 Discussion

5.5.1 Statement of principal findings

The primary aim of the present study to develop a prediction model was fulfilled with a carefully calibrated and validated prediction model. Patients with the outcome spent on average 1/3 of follow-up time as compared to patients without the outcomes. Majority of patients ended up without the outcome, especially R2 and M1 (71% of the sample). These two groups of patients could offer an opportunity for optimisation of services by increasing the follow-up intervals for these patients less likley to develop the outcome of treatment or vision loss. This will help prioritise the care for R3M1 and R3 groups that developed the outcome several times more frequently.

We chose a combination of ocular and systemic predictors. It has recently been suggested that retinopathy data may be sufficient to develop a risk stratification tool, but clinical information is helpful (32). Majority of models in the systematic review used a combination. One model used only systemic predictors with good performance In high risk patients with diabetes, it is appealing to use both types. It will also help in populations with no access to photographic screening but may have systemic predictors available to risk stratify.

The exploratory analysis in Table 2 did show some differences in the performance of the models. Complete case analysis model was consistent with the imputation model in including the same variables that were in the complete case analysis model (29), the fractional polynomial model rejected cholesterol as it crossed the p value of significance. In the rest of the models some predictors crossed p value of 0.2. In the exploratory analysis, we performed a parallel analysis with M1 as a reference point

as opposed to R2 by using the same candidate predictors sets. The final predictors did not change.

Diabetic retinopathy is a microvascular complication in patients with diabetes, the risk increasing with longer duration of diabetes. These patients are also at risk of macrovascular complications (coronary artery disease, peripheral arterial disease, and stroke), where patients are also at risk of death. Death therefore can compete with the event of our outcome of interest which is treatment or vision failure. It can occur before the outcome of interest and thus can influence the measurement of the risk of outcome of interest. If the competing event is simply censored (Cox regression), then the outcome estimates and prediction risks are overestimated. Berry et al have put forward a rule of thumb for deriving absolute risk predictions. They proposed accounting for competing events if the follow up exceeds five years or when the proportion of subjects experiencing competing events is equal to or greater than the proportion experiencing primary outcome (33).

5 years competing risk Fine and Gray model (21) has E/O of 0.96 at 5 years. It is technically challenging to internally validate this model. So we decided to further process the Cox model for five years with the limitation that the death as a competing event may have inflated the risk estimates to a small extent.

The c statistic (performance index) was lower in models with systemic factors only and in model for > 60 years of age. Stratifying the models by gender, age groups or length of follow up, not only destabilised their predictors, but also lowered the number of outcomes and discriminatory performance. We chose complete case analyses option after a exploratory analysis because of its robustness, stability and simplicity.

It fits with the data well, with good calibration slope and a moderately good discrimination. Internal validation showed a small amount of optimism.

R3/ M1 had the highest hazard ratio with R2 being the lowest risk. Age at STR diagnosis did not contribute much risk. The systemic factors did contribute significantly. HBA1C was the commonest factor retained. Statin was protective.

5.5.2 Strengths and weaknesses

The predictor selection methodology incorporated a validated clinician input method (NGT), evidence evaluation and recommended statistical methods (19). We used IMRD database, which has been used for a wide range of clinical research, including the development and validation of risk prediction models (15). Our aim was to identify patients who may progress from referable retinopathy in a hospital outpatient setting to initiate treatment in a timely manner and prevent blindness. Therefore, vision loss and treatment initiation were chosen as outcome of interest. We performed comprehensive analysis to explore clinical scenarios and chose substantial period of follow up of 5 years. We used established methods to develop risk prediction equations for a combined outcome of treatments (laser, vitreous procedures) and vision loss / blindness in the derivation cohort. We tried to avoid overfitting and optimism with the following measures.

1. A large sample
2. Used a parsimonious predictors set, with least interaction / non-linear terms by using relevant tests
3. Shrinkage technique - Shrinkage was applied at model estimation (19).
4. Optimism corrected performance statistic (19)

This new model calculates the absolute risk of developing these complications and need for treatments taking account of individual patient's risk factors. We have reported easy to understand hazard ratios. We have also reported potential sources of misclassification of DR categories.

The study is restricted by the use of predictor variables measured in primary care data sets and the limitations of these measures. Due to the automated nature of the stepwise predictor selection, the functional form of continuous predictors could not be checked during the internal validation process. The limitations were missing data (mainly affecting Ethnicity, Townsend score, Type of diabetes, diabetes diagnosis) and miscoding possibly causing information bias. To mitigate that risk, an extra category (missing data) was created in Ethnicity and Townsend predictors. The missing data especially affected modelling for T1DM adversely as the sample size fell short of adequate size. The automated stepwise selection method does have some disadvantages like the selection being unstable, the estimated regression coefficients being extreme, and overestimation of the model performance (34). Outcomes may also be under-reported in the primary care records that we have used. However, our large sample size and a respectable number of outcomes, mitigates most of these risks.

Referable retinopathy patients downgraded to a lower category after clinical examination do not get recorded, as data is not collected any longer. A large proportion of participants fall in this category, except R3 downgraded to R2. This is an example of miscoding. False positives have been reported before to be between 50 to 70% (5, 6).

5.5.3 Strengths and weaknesses in relation to other studies, discussing particularly any differences in results

Our recent systematic review (10) included 14 models on diabetic retinopathy progression, but they were mostly limited by single type of diabetes (type 1 or 2 only) and were constructed having either incident retinopathy or incident referable retinopathy as an event (outcome of interest). A few of them had blindness as an event on its own. They were mostly focussed on the earlier lower risk part of the patient care pathway. Those with the outcome of progression of retinopathy (35-37) had high risk of bias and low applicability. Our target population for the present model was high risk diabetic retinopathy patients, exposure was referable diabetic retinopathy and the outcome chosen was an event of vision failure or any of the treatments. The existing modelling studies reported hazard ratios in 6 (12, 35, 38-41) out of 14 studies, 3 (15, 39, 42) reported regression coefficients (not that easy to interpret), another 4 reported risk ratios (8), screening intervals against risk thresholds (11), risk scores (36) or % risks (37), and one of them (32) did not report any risk estimates. We have reported easy to understand hazard ratios. We have also reported potential sources of misclassification of DR categories.

5.5.4 Implications for clinicians or policymakers

With a false positive referral rate of 50 – 70% (5, 6) or more the present service arrangements are under severe pressure. This model will help fulfil the need for risk stratification to help decide on follow-up intervals. This in turn will enable clinicians and managers to optimise the demand and capacity mismatch through prioritizing care for the high risk patients. The information can be used to empower clinicians and managers to direct resources where they are needed most. This information will

also lead to better treatment and prevention decisions for the management of modifiable risk factors. It may also be employed to help select appropriate patients for research. The model has the potential for an accurate individualized risk assessment to inform patients on risk benefit ratio in their own individual situations as well.

5.5.5 Unanswered questions and future research

Some of the predictors like the Early Worsening and Visual acuity were not available in the IMRD database. Pregnancy as a predictor was also not feasible. The external validation is essential and will also provide an opportunity for these predictors to be part of the list of candidate predictors. The model therefore may need updating in hospital data at the time of external validation and there will be an opportunity to consider the predictors not included in this model. Decision curve analysis for clinical decision by measuring net clinical benefit (improved outcomes) will also be needed to help make a clinical decision.

Risk groupings is based on linear predictor (also called risk score or prognostic index) calculation and then using cut offs to define categories of this score. A new individual can be assigned to a group based on their individual score. However, the individuals within the same group then all receive a standard care within the same group uniformly, and the difference between their individual risk scores is not taken into consideration. On the other hand, the model can predict an individual's survival probability or outcome risk. The model can thereby assist in providing a more accurate individual risk estimation and a simplified decision rule to guide interventions (19).

A recent review found the diagnostic performance of artificial intelligence models for imaging diagnostics to be as good as health-care professionals (43). These algorithms have recently been shown to improve the accuracy and confidence in DR diagnosis (44). Considering the importance of the ocular clinical features, reflected by the high hazard ratios of ocular predictors in our model, and judging by the recent developments in the field of image based artificial intelligence, it is envisaged that this discipline can have a hugely positive impact on future prediction modelling in the population with high risk diabetic retinopathy.

5.6 Conclusions

We have developed and internally validated new prognostic prediction model to assess the risk of vision loss and blindness or need for treatment in patients with referable diabetic retinopathy. The model shows good performance and has the potential to be used to identify patients with referable diabetic retinopathy at high risk of needing treatment or risk of vision loss for the purpose of triage and risk stratification. Further research is needed to perform external validation, evaluate any possible decision rules and net clinical benefit.

Acknowledgements / Contributions: SH designed, conducted, analysed, interpreted and wrote the first draft the study. KN and MP were the main supervisors but KS also contributed. AS performed data extraction and helped with internal validation. MP, KT, NA, KS and KN made a contribution verbally or by critical revision of the draft. All authors approved the final version of the manuscript.

Conflict of Interest: All authors declare no conflict of interest.

Funding: There was no funding / financial support for this research. KS is funded by the National Institute for Health Research School for Primary Care (NIHR SPCR Launching Fellowship). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

5.7 References:

1. (NCD-RisC) NRFC. Worldwide trends in diabetes since 1980: a pooled analysis of 751 population-based studies with 4.4 million participants. Lancet. 2016;387(10027):1513-30.
2. Gonzalez EL, Johansson S, Wallander MA, Rodriguez LA. Trends in the prevalence and incidence of diabetes in the UK: 1996-2005. Journal of epidemiology and community health. 2009;63(4):332-6.
3. Public Health England. Diabetes Prevalence Model 2016 [Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/612306/Diabetesprevalencemodelbriefing.pdf, Accessed 2019, July 12.
4. Mathur R, Bhaskaran K, Edwards E, Lee H, Chaturvedi N, Smeeth L, et al. Population trends in the 10-year incidence and prevalence of diabetic retinopathy in the UK: a cohort study in the Clinical Practice Research Datalink 2004-2014. BMJ open. 2017;7(2):e014444.
5. The Royal College of Ophthalmologists. The way forward age-related macular degeneration and diabetic retinopathy 2016

[ ]. Available from: https://www.rcophth.ac.uk/wp-content/uploads/2015/10/RCOphth-The-Way-Forward-AMD-300117.pdf, Accessed 2019, July 20.
6. Jyothi S, Elahi B, Srivastava A, Poole M, Nagi D, Sivaprasad S. Compliance with the quality standards of National Diabetic Retinopathy Screening Committee. Primary care diabetes. 2009;3(2):67-72.
7. Scanlon PH, Aldington SJ, Leal J, Luengo-Fernandez R, Oke J, Sivaprasad S, et al. Development of a cost-effectiveness model for optimisation of the screening interval in diabetic retinopathy screening. Health Technology Assessment. 2015;19(74):1-116.
8. Aspelund T, Porisdottir O, Olafsdottir E, Gudmundsdottir A, Einarsdottir AB, Mehlsen J, et al. Individual risk assessment and information technology to optimise screening frequency for diabetic retinopathy. Diabetologia. 2011;54(10):2525-32.
9. Steyerberg EW, Moons KG, van der Windt DA, Hayden JA, Perel P, Schroter S, et al. Prognosis Research Strategy (PROGRESS) 3: prognostic model research. PLoS medicine. 2013;10(2):e1001381.
10. Haider S, Sadiq SN, Moore D, Price MJ, Nirantharakumar K. Prognostic prediction models for diabetic retinopathy progression: a systematic review. Eye. 2019;33(5):702-13.
11. Eleuteri A, Fisher AC, Broadbent DM, Garcia-Finana M, Cheyne CP, Wang A, et al. Individualised variable-interval risk-based screening for sight-threatening diabetic retinopathy: the Liverpool Risk Calculation Engine. Diabetologia. 2017.
12. Scanlon PH, Aldington SJ, Leal J, Luengo-Fernandez R, Oke J, Sivaprasad S, et al. Development of a cost-effectiveness model for optimisation of the screening interval in diabetic retinopathy screening. Health Technol Assess. 2015;19(74):1-116.

13. Wolff R WP, Mallett S, Riley R, Westwood M, Kleijnen J, Moons K. PROBAST: a risk of bias tool for prediction modelling studies, 2015 3-7 Oct [
14. Moons KG, de Groot JA, Bouwmeester W, Vergouwe Y, Mallett S, Altman DG, et al. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. PLoS medicine. 2014;11(10):e1001744.
15. McEwan P, Bennett H, Ward T, Bergenheim K. Refitting of the UKPDS 68 risk equations to contemporary routine clinical practice data in the UK. Pharmacoeconomics. 2015;33(2):149-61.
16. Hayden JA, van der Windt DA, Cartwright JL, Cote P, Bombardier C. Assessing bias in studies of prognostic factors. Ann Intern Med. 2013;158(4):280-6.
17. The Renal Association. Classification of Chronic Kidney Disease [Available from: https://renal.org/information-resources/the-uk-eckd-guide/ckd-stages/.
18. Riley RD, Snell KI, Ensor J, Burke DL, Harrell FE, Jr., Moons KG, et al. Minimum sample size for developing a multivariable prediction model: PART II - binary and time-to-event outcomes. Stat Med. 2019;38(7):1276-96.
19. Riley RD, van der Windt D, Croft P, Moons KGM. Prognosis Research in Health Care: Concepts, Methods, and Impact: Oxford University Press; 2019.
20. Dignam JJ, Zhang Q, Kocherginsky M. The use and interpretation of competing risks regression models. Clin Cancer Res. 2012;18(8):2301-8.
21. Fine JP, Gray RJ. A Proportional Hazards Model for the Subdistribution of a Competing Risk. Journal of the American Statistical Association. 1999;94(446):496-509.
22. Guo C, Yo S, Jang W, editors. Evaluating predictive accuracy of survival models with PROC PHREG2018: SAS.
23. Van Houwelingen J, Le Cessie S. Predictive value of statistical models. Statistics in medicine. 1990;9(11):1303-25.
24. StataCorp. Stata Statistical Software: Release 14. College Station, TX: StataCorp LP 2015.
25. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): the TRIPOD Statement. The British journal of surgery. 2015;102(3):148-58.
26. Moons K. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): Explanation and Elaboration. Annals of Internal Medicine. 2015;162(1):W1-W73.
27. Hippisley-Cox J, Coupland C. Development and validation of risk prediction equations to estimate future risk of blindness and lower limb amputation in patients with diabetes: Cohort study. BMJ (Online). 2015;351 (no pagination)(h5441).
28. Royston P, White IR. Multiple imputation by chained equations (MICE): implementation in Stata. J Stat Softw. 2011;45(4):1-20.
29. Institute of Digital Research & Education UCLA. Multiple Imputation in Stata [Available from: https://stats.idre.ucla.edu/stata/seminars/mi_in_stata_pt1_new/ Accessed 2020, May 15.
30. Rubin DB. Multiple Imputation for Nonresponse in Surveys: Wiley; 2009.
31. Alpman A. Implementing Rubin's Alternative Multiple-imputation Method for Statistical Matching in Stata. The Stata Journal. 2016;16(3):717-39.
32. Stratton IM, Aldington SJ, Farmer AJ, Scanlon PH. Personalised risk estimation for progression to sight-threatening diabetic retinopathy: How much does clinical information add to screening data? Diabetic Medicine. 2014;31:23-4.
33. Berry SD, Ngo L, Samelson EJ, Kiel DP. Competing risk of death: an important consideration in studies of older adults. J Am Geriatr Soc. 2010;58(4):783-7.
34. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. European heart journal. 2014;35(29):1925-31.

35. Tanaka S, Tanaka S, Iimuro S, Yamashita H, Katayama S, Akanuma Y, et al. Predicting macro- and microvascular complications in type 2 diabetes: the Japan Diabetes Complications Study/the Japanese Elderly Diabetes Intervention Trial risk engine. Diabetes care. 2013;36(5):1193-9.
36. Harris Nwanyanwu K, Talwar N, Gardner TW, Wrobel JS, Herman WH, Stein JD. Predicting development of proliferative diabetic retinopathy. Diabetes care. 2013;36(6):1562-8.
37. Mehlsen J, Erlandsen M, Poulsen PL, Bek T. Individualized optimization of the screening interval for diabetic retinopathy: a new model. Acta Opthalmologica. 2012;90(2):109-14.
38. Lagani V, Chiarugi F, Thomson S, Fursse J, Lakasing E, Jones RW, et al. Development and validation of risk assessment models for diabetes-related complications based on the DCCT/EDIC data. Journal of diabetes and its complications. 2015;29(4):479-87.
39. Hayes AJ, Leal J, Gray AM, Holman RR, Clarke PM. 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. 2013;56(9):1925-33.
40. Soedamah-Muthu SS, Vergouwe Y, Costacou T, Miller RG, Zgibor J, Chaturvedi N, et al. Predicting major outcomes in type 1 diabetes: a model development and validation study. Diabetologia. 2014;57(11):2304-14.
41. Stratton IM, Aldington SJ, Taylor DJ, Adler AI, Scanlon PH. A simple risk stratification for time to development of sight-threatening diabetic retinopathy. Diabetes care. 2013;36(3):580-5.
42. Clarke PM, Gray AM, Briggs A, Farmer AJ, Fenn P, Stevens RJ, et al. 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. 2004;47(10):1747-59.
43. Liu X, Faes L, Kale AU, Wagner SK, Fu DJ, Bruynseels A, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. The lancet digital health. 2019;1(6):e271-e97.
44. Sayres R, Taly A, Rahimy E, Blumer K, Coz D, Hammel N, et al. Using a Deep Learning Algorithm and Integrated Gradients Explanation to Assist Grading for Diabetic Retinopathy. Ophthalmology. 2019;126(4):552-64.

5.8 Appendices for Chapter 5

6 Chapter 6: Discussion

The thesis is in alternate format and there is a detailed discussion section in each manuscript, therefore there will be some repetition during further elaboration of the subjects.

This PhD project aimed to determine how hospital diabetic retinopathy services could be optimised. The objectives were to 1) to determine the current and future demand on the diabetic retinopathy services and 2) to design a risk stratification model for hospital-based high risk diabetic retinopathy patients. The Table 1 below presents how the author attempted to meet the objectives described within Chapter 1. The process of external validation and beyond was outside the remit of this PhD project.

Table 1: Objectives and how this PhD research project met them.

	Objectives	How met?
1	To quantify and project the size of the diabetes and diabetic retinopathy disease and treatment burden	Carried out a series of cross-sectional surveys on the IQVIA Medical Research Data IMRD database and designed a model to project the future burden as presented in Chapter 2.
2	To develop an individual risk stratification model to identify high-risk diabetic retinopathy patients through the following steps:	
2 A	Summarize all existing prognostic models in the area of diabetic retinopathy progression and vision loss	This was achieved by undertaking a systematic review as presented in Chapter 3.
2 B	To identify the predictor variables and decide on a parsimonious set of candidate predictor variables.	This was achieved by undertaking a qualitative study (NGT) and Evidence Review Studies to reach a final set of 19 candidate predictors – an evidence-based and parsimonious set of variables as presented in Chapter 4.

2 C	Prognostic model development and validation for diabetic retinopathy progression to the stage of treatment or vision loss	This was achieved by undertaking an IMRD based cohort study and is presented in Chapter 5.

6.1 Disease Burden: Trends, future projections and current burden

6.1.1 Principal findings

This cross sectional IMRD data analysis was carried out in patients aged 12 and above. The study describes the last 21 years of disease burden trends, forecasts the same until the year 2030, and gives the details of the prevalence breakdown numbers in 2017. The trends in T2DM, STR and DR prevalence showed a higher rise in the decade from 1998 to 2007, with a relative slowing down between 2009 and 2018. T1DM showed the opposite trend in this respect, by accelerating in the second decade. This is in line with an increasing incidence of T1DM that has previously been reported in children in Europe (1) and the United States (2).

The overall prevalence of all categories of DR more than doubled, while STR nearly tripled in T2DM, and almost quadrupled in T1DM. The relatively smoother increments over the second decade (2009-2018) in every category except T1DM (Appendix 7), might reflect increased clinical awareness and understanding of DM / DR / STR, followed by implementation of early detection strategies and targeted management.

While the DR disease burden was greater with T2DM, STR was more common in T1DM (Manuscript Table 3, Chapter 2). T1DM acceleration during the second decade is a warning of the difficulties ahead in predicting a disproportionate increase in the more severe disease burden.

In 2017, one third of the participants had some DR and one eighth had STR. Participants needing treatment or losing sight together accounted for 5% of the population with diabetes, and 2.8% had blindness or vision loss. STR accounted for

52% of total retinopathy in T1DM and 34% in T2DM. DR severity was higher in T1DM, but the numbers affected are higher in T2DM.

This study went on to forecast future trends in both DR and STR as well as T1DM and T2DM prevalence until 2030, showing that in 10 years, approximately 1.6 million people will have some level of DR and about two thirds of a million will have STR (Table 2 in chapter 2). With rates of T1DM expected to increase further and more rapidly, it is likely that there will also be a corresponding higher rise in STR, forcing further increase in demand.

This prediction constitutes a significant increase in the disease burden of DR. The treatment burden is difficult to measure because of variations in local clinical practices and treatment protocols. If the rise in T1DM trend continues, this will pose a challenge as T1DM individuals are younger, and therefore need more life years of treatment and follow up, resulting in an increase in costs. The aforementioned greater severity of DR in T1DM patients adds to this complexity.

6.1.2 Strengths and weaknesses of the study

The prevalence figures are given in a clinically relevant form for the sake of clinicians and managers leading the professional services in the area of diabetic retinopathy care especially for patients under HES. This work is based on a cross-sectional analysis of primary care data, and thus is closer to routine practice. This is the second study to forecast the disease burden of DM and DR (3, 4) for the UK population.

The limitations are possible coding errors, missing data difficulties, and the potential risk of an overestimation of vision failure. The findings of this study should be

interpreted within that context. Firstly, the possible impact of coding errors, as well as subjectivity in documentation across a retrospective nationwide database involving several practices in different areas, cannot be precluded. This potential risk was minimised through a strict read code selection process. Importantly, missing data, especially with regards to ethnicity coding, prevented an interpretable analysis of ethnic origin on the estimates and severity of DR. Prevalence of severe DR was higher for South Asian and mixed ethnicity, therefore this could have implications for CCG to CCG variation in prevalence, and estimates could be different depending on the local ethnic mix (5). This can be mitigated by incorporating local factors when making local estimations (6). Finally, the potential impact of several concomitant medications on the course of DR was not captured in the present study design. Notwithstanding its limitations, this study provides rigorously tested prevalence estimates for DR, involving a large number of affected individuals, and a revised and pragmatic approach to the definition of STR.

As mentioned earlier, the estimation of the prevalence of diabetes and diabetic retinopathy, based on the IMRD database, is likely to be slightly different because of the number of undiagnosed cases of diabetes included in the Public Health England (PHE) figures and differences between age groups.

6.1.3 Strengths and weaknesses in relation to other studies, discussing particularly any differences in results

DM estimates

6.1.3.1.1 Past trends and future forecasts

Recent literature on DM (7, 8) reported an increasing prevalence of diabetes from 1996 to 2005 and between 2004 and 2014, but these increases were limited to

different age groups or type of disease. Direct comparisons are therefore not possible.

Bagust is a future forecast for the UK, but is limited to T2DM and is an underestimation (3). PHE predicted that diabetes prevalence in 2025 will be 5 million (9), as compared to the forecast from this study of just under 4.3 million. Their estimate includes undiagnosed diabetes, therefore both figures seem approximately in line with each other. However, an underestimation due to the misclassification of population with diabetes in this research cannot be ruled out.

6.1.3.1.2 Recent prevalence figures

The IDF (10) figure of 2.7 million cases of DM in 2017 seems to be an underestimation compared to the PHE prediction of 3.7 million cases with diagnosed DM, in those over 17 years of age, and the prediction of 3.4 million cases in those over 12 years of age in the present study. The latter two seem to be in line with each other. The PHE estimate was arrived at by quality and outcome framework figures (11). The figures arrived at in this study are based on a much larger database as compared to the health survey England 2016, on which the PHE figures for diabetes estimates are based. After implementation of the Quality and Outcomes Framework in 2004, the recording of diabetes diagnosis also improved in primary care, however the choice of codes can also influence disease estimates in the present study (12).

DR estimates

Mathur et al (5) reported an increasing trend in the prevalence of severe DR which is in agreement with this study. However, the STR figures in this study showed a 67%

increase in the corresponding years as compared to their data analysis which showed stabilization or a marginal decrease. The reason is most probably a difference in case definition. Li et al is the only study that has projected DR prevelance until 2050 (4), but is based on T2DM data only from 2003 and 2009 (13, 14). Case definitions and patient pathways have changed since.

This study's DR results were compared against Diabetic Eye Screening Programme (DESP) referral data (15). In the Midlands and East of England group, the incidence (annual referral rate) of 2.13% when converted into the prevalence was much higher (almost double) at 24.1%. This is most likely due to a high false-positive referral rate by the DESP of approximately 50-70% as already reported (16, 17). Welsh data (18, 19) up to 2009 are over a decade old. The care pathway has changed since and data is from lower-risk patients in the earlier, lower part of the care pathway. Keenen et al. (20) is the only study, that describes the annual incidence of diabetic retinopathy among higher risk hospital patients. However, the study is over a decade old and they measured DR prevalence in eyes rather than patients. The heterogeneity of methods in counting patients or eyes, differences in population characteristics, and different study periods due to changing practices, creates difficulties in interpreting and comparing results and trends. The study presented in this thesis is the only up to date research that gives long term trends, forecasts the future disease burden and provides a detailed breakdown of current estimates in clinically relevant categories of DR.

6.1.4 Meaning of the study: possible mechanisms and implications for clinicians or policymakers

This research study presented trends of the last 21 years for DM and DR figures, changes in trends and projected estimates of disease burden for the future to prepare the service providers for what to expect in terms of capacity needed in the future. The study also provides a novel forecasting model, which will be useful for clinicians and managers leading professional services in the area of diabetic retinopathy in order to plan the capacity needed to meet increasing demand. Unless there is an unforeseen breakthrough in the management of DM / DR soon, it is highly likely that increasing needs and resources required to cope with the impact of increasing DM / DR burden will be increasingly pressing in the ensuing years and that estimates presented will be of interest in public health policy strategy.

This study also presents up to date evidence on the prevalence estimates of various clinically relevant DR categories of the DR and DM disease burden. By presenting the detailed analysis of different forms and severity of DR, this thesis captures the information required by the public health system. Local demand can be calculated with the help of the national figures provided, to help clinicians and managers in their service planning.

6.1.5 Unanswered questions and future research

Drawing from the experience of the last 21 years, further disease burden studies will need to be carried out to capture the changing trends and any unforeseen cyclical changes (when data rises and falls but not for fixed periods). The estimates should be preferably presented as patient numbers rather than eye numbers, as a wealth of

information can then be transported from other studies for extrapolation of evidence or for comparison for reasons of verifiability. NHS trusts are increasingly converting from paper notes to electronic medical records, which offers an opportunity for relatively easier data collection. A combination of DESP and HES data will also help.

This study has not adjusted for the risk factors for the incidence of diabetes or diabetic retinopathy and I suggest that is best done locally. Public Health England has recently used a diabetes prevalence model to predict the local disease burden (21). The predictive factors they used were increasing age, ethnicity, and gender, and after adjusting for these, found additional benefit in the use of a deprivation index. Their model could be used for the calculation of the local disease burden. Further research in this area is required, preferably with combination of data from DESP and HES.

With all these challenges of increasing demand and the efforts needed to increase capacity to scale, the resources remain finite. One way is to identify higher risk patients and prioritise their care. The next section deals with that possible solution.

6.2 Disease progression and risk prediction

6.2.1 Principal findings

A systematic review summarized, and critically appraised, the existing models in predicting the progression of diabetic retinopathy and their applicability for patients with diabetic retinopathy under hospital care. The review included 14 predictive model development studies. Overall, the included studies were heterogeneous, principally due to differing disease classification, having few internal validations, even fewer external validations, and reporting deficiencies. Although the systematic review

identified three low risk of bias studies, they had only partial overlap on the outcome of choice, and therefore had only partial applicability (22-24). Since these models were primarily developed for lower-risk patients, their population samples did not fit the aims of this research project and exposure was mostly incident retinopathy or incident referable retinopathy. It was concluded that a model fully covering the target population, exposure and outcomes of this research project did not exist, and therefore a new primary modelling study for higher-risk patients under hospital eye services was needed.

The modelling studies in the systematic review did however yield a list of 78 candidate predictors. Based on the clinical advice in NGT meetings, this list of 78 predictors was reduced to a list of 25 potential predictors. The participants also suggested 8 additional new predictors (Chapter 4 manuscript, Table 1). After the evidence review, two further predictors were added back in, taking the list to 35 predictors. As a result of excluding predictors without any evidence base, predictors with duplication/overlap and condensing retinopathy grades, a final set of 19 predictors was generated.

This set included ocular and systemic predictors. The approach of using ocular predictors (DR stages) alone to develop a risk stratification tool in the DESP environment has been suggested (25). An Individualised Screening for Diabetic Retinopathy (ISDR) model (22), in complete contrast, has suggested using clinical predictors alone. The risk estimates of the various predictors within various models tend to vary depending on the combination of predictors used (23, 26), however, ocular predictors generally showed higher estimates. It was concluded therefore that the new model should be based on a combination of ocular and systemic predictors

for use in the higher risk hospital population, as the values of, for example, SBP and HbA1c etc are likely to be higher in the higher-risk hospital patients and can thus contribute to the performance of the model. A practical sized set was now available, with individual predictors ranked in importance, as perceived by the NGT participants. A model based on these predictors could be built in appropriate population data.

IMRD database was used for model development with a combined outcome of treatments (laser, vitreous procedures) and loss of vision/blindness. A model based on complete case analysis was derived and compared against a series of models for sensitivity and exploratory analysis of different clinical scenarios. The complete case analysis model was consistent with the multiple imputation model in that it included the same variables. In the other models some predictors crossed the p-value of 0.2, the c statistic (performance index) was suboptimal, and also had a lower number of outcomes. The complete case analyses option was chosen because of its robustness, stability and simplicity. It fitted the data well, has a good calibration slope, moderately good discrimination, and showed only a small amount of optimism on internal validation.

Although the hazard ratios for the systemic predictors are much lower than the ocular predictors, they still make a significant contribution to the predictive ability at higher measurement values (Table 3, chapter 5). R3/M1 had the highest hazard ratio with R3 being the second highest. These categories are also the ones that are more likely to result in the outcomes of interest. R2 and M1 were at a lower risk of getting the outcomes of interest. This also fits in with the evidence that the disagreement rate between DESP graders and Ophthalmology departments was high for M1, and as a

result, only 21% were treated in 2009 (17). I suspect that R2 is also similar. There is a shortage of contemporary survival study evidence on this subject. The cohort data survival analysis for model development showed a rather low combined outcome rate of 15.2% although the primary care data is likely to have underestimated hospital episodes of treatment.

6.2.2 Strengths and weaknesses of the study

The development of a disease prediction model followed a systematic approach starting with a systematic review, using a systematic search strategy, along with recommended search filters. The searches yielded several conference abstracts and posters without full journal articles. Extra precautions were taken to follow them up by searching for publications through the authors' names. This allowed us to identify two modelling studies not captured by conventional searches. This was fed back to the authors specialising in the field. PROBAST (27) and CHARMS (28) were used for risk of bias assessment, PRISMA (29) guidance for reporting and registered the protocol prospectively with PROSPERO (30). Authors were also contacted where needed and most of them responded to explain the information that was not reported. The comments on PROBAST (which was a pilot at the time) were fed back for the authors to use for any amendments in a future review of the guidance. Predictor selection is generally based on previous knowledge and clinical input (31). I started this process with the full list of predictors from previous modelling studies reported in the systematic review. The advantage was the likelihood of availability of the variables, and the evidence backing them, in the records. There was duplication, overlap, reporting deficiencies, and extrapolation of evidence from outcomes other than DR progression, for example macrovascular complications. With a large number of 78

predictors, the author decided to gain insight from clinician colleagues through a formal qualitative study design (NGT), then evaluating their evidence base, following the Prognosis Research framework recommendations rules (32) and by using Quality in Prognosis Studies (QUIP) tool (33). Predictors without evidence base were dropped to enhance the chances of generalisability.

The database used (IMRD) can give representative information about the UK population (7). The case definition was developed after consulting previous literature on the subject of diabetic retinopathy (5, 34). A rigorous selection process of read codes (Appendix 4) with high sensitivity and specificity was used to ensure inclusion of all cases and exclusion of patients without the diagnosis of diabetes. The assumptions were checked. Proportional hazards assumption was valid and there was a suspicion of non-linearity. This was dealt with by including models dealing with various clinical requirements for exploratory analysis (Table 2, Chapter 5, page 163).

Missingness and miscoding affected predictors like ethnicity, Townsend score, type of diabetes, and date of diabetes diagnosis. This was dealt with by excluding records without diabetes or diabetic retinopathy diagnosis to prevent consequent risk of bias. A missing data category was created for ethnicity and Townsend score to preserve the related data. Multiple imputation was also carried out. The missing data especially affected modelling for T1DM as the sample size fell short of adequate size. TRIPOD (35) reporting guidance was strictly followed to ensure high quality reporting of the developed model. Some of the covariates were not available in the IMRD database, namely “Early Worsening” and Visual acuity score, therefore these could not be used.

Diabetes duration is worth mentioning as a predictor in how it was handled and its results in multivariate modelling. This factor was second best voted in the NGT in chapter 4. The evidence available is the use of a cut off of 6 years in T2DM in UKPDS (36) and being an independent risk factor in affecting progression to treatment stage in T1DM patients in a Danish cohort (37). The risk factor was highly significant in T1DM patients with a 15 years cut off (38). Categorization of continuous variables is generally discouraged in multivariate modelling to avoid loss of information (39). I have carried out a sensitivity analysis by modelling for T1DM and T2DM separately and the factor was not significant in either type. I also attempted modelling for a 10 years' cut off and the factor was not significant on both sides of 10 years cut off. While it will remain part of the candidate predictor set to see how it performs in the external validation exercise, so far it has not been significant in the multivariate study analysis. For the outcome of referable DR, it was the second commonest variable used by the 14 models in our recent systematic review. It is possible that duration of diabetes may not be important in the context of our model which explored referable DR exposure with an outcome of treatment or vision loss.

6.2.3 Strengths and weaknesses in relation to other studies, discussing particularly any differences in results

The only other systematic review on the topic of prognostic models for diabetic retinopathy in progression of T2DM patients (40) was recently published and addressed the outcome of referable diabetic retinopathy but also included external validation of one of the models of their choice. Our systematic review (41) included 14 models on diabetic retinopathy progression but, apart from a few (22-25, 42-45), these models were limited by a single type of diabetes (type 1 or 2 only). Two

modelling studies did use the hospital population (24, 43), but had incident diabetic retinopathy as an outcome. Some models had blindness as an event on its own. Overall, these were mostly focussed on the earlier, lower-risk part of the patient care pathway. Those with the outcome of the progression of retinopathy (43, 44, 46) had a high risk of bias and low applicability. No prognostic model aimed at the target population, exposure or the outcome of interest for this research project existed.

There was evidence base available for predictors in the form of this systematic review and in the literature (47-49). This research project used the list of 78 predictors used by 14 models, for reasons explained in Chapter 4. The ISDR model (22) mentioned using a patient expert panel, presumably through a qualitative study design. The finalised predictor list includes all ISDR predictors to capture patients' perspective. The ocular predictors have a much higher hazard ratio value than the other predictors. It is suspected that this is due to the high-risk group of patients, and the more clinically relevant outcomes chosen.

6.2.4 Meaning of the study: possible mechanisms and implications for clinicians or policymakers

While conducting the systematic review, considering the difficulties in searches for prognostic models, conference abstracts without publication were followed up, and two extra records were found by searching by the author names. The two leading authors (50, 51) on the subject were requested to include this technique in any of their future publications. The pilot version of the PROBAST risk of bias tool was used along with the CHARM checklist. Feedback was sent back to the first author to reconsider in their future iteration of the tool, thus contributing to the methodology.

It was found that the reporting of the primary studies for the predictors selected in the modelling studies was sub-optimal (36% not reported). The Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) statement asks for all predictors used to be clearly defined (35). One way of improving this situation may be for the TRIPOD checklist to encourage investigators to report the primary study of origin for every candidate predictor used in the model. This is mentioned in the manuscript for publication and shall be recommended to the author of TRIPOD (35), once the paper is published, for the sake of their future checklist review.

The enthusiasm of the NGT participants was evident prior to the meetings (during the process of setting the meetings up and in the recruitment phase) and throughout the meeting. The participants wanted to get involved in the external validation project and get an update of the outcome of this work. They were also keen to know the evidence behind the research question and when the final tool will be available to use. The set of predictors chosen through NGT and the evidence evaluation could help with prioritisation of appointments and thus direct resources more appropriately**.** This research project has been able to derive a model based on this set of predictors to estimate an individual patient's risk, however if another researcher is interested, the set of predictors can be used in another data set for a modified model, or for the present model's external validation. Although the model was designed for UK where high risk patients are cared for in a care pathway separate from low risk patients. The model could also be applied outside of the UK, preferably after external validation in the local population. In the UK, baseline retinopathy grades and pregnancy are already being used to risk stratify patients in deciding their follow up intervals (52).

In the longitudinal cohort data used for modelling, outcomes of treatment or vision loss occurred in 15.2% of total cases within an average period of 1.34 years. Among the majority of participants (without the outcome), the average follow-up was 3.9 years. There is an opportunity to save some of this clinic time resource by prioritization of the first group for safety and efficiency through risk stratification. This model, after external validation, can help with this process.

This model has the potential for an accurate, individualized risk assessment to inform patients on the risk-benefit ratio in their special situations. The model can predict an individual's survival probability or outcome risk and can help the risk thresholds to be categorised for a simplified (preferably agreed by clinicians) risk cut off to guide interventions (39). It can also help patients improve their awareness of their modifiable risk factors, and hence prevention of further progression.

With the introduction of electronic health records in most eye clinics, a formal individual risk stratification model can help in many ways. Used prospectively to stratify patients according to their personal risk, this could rank newly diagnosed patients with referable diabetic retinopathy by the risk of the combined outcome (treatments / vision failure) so that those at highest risk could be allocated early assessment or additional interventions to manage their risk factors. The RCOphth has just embarked on a risk stratification initiative (53), which will be helped with this model in the long term. It could be implemented for high risk case finding and thus help optimise the demand and capacity mismatch in the service through a triage. The model will put the risk stratification efforts on a firm evidence-based footing and empower clinicians/managers to direct resources towards higher-risk patients.

6.2.5 Unanswered questions and future research

Further predictor research

During the predictor selection process (NGT and evidence review studies) one third of predictors selected by NGT (Chapter 4, Table 2) were not supported by the existing evidence base, therefore these will require further research.

A recent review found the diagnostic performance of artificial intelligence models for imaging diagnostics to be as good as healthcare professionals (54). These algorithms have recently been shown to improve the accuracy and confidence in DR diagnosis (55). Considering the prognostic importance of ocular features (can be all images in a virtual clinic situation or a reading centre) in the model, and judging by the recent developments in the field of ophthalmic imaging and artificial intelligence, it is envisaged that this discipline could have a hugely positive impact on future prediction modelling in the high-risk diabetic retinopathy population. There is an ever-increasing interest in fundus images based detection, assisted by artificial intelligence and the natural history of retinal lesions in prediction research (47, 56). Machine learning can handle data from wider sources, bring additional benefits from automation, handle unsupervised clustering of a much larger number of predictors, and can also add new phenotypes associated with the outcomes. Although there are ethical, governance, and interpretability issues, and the process of development of these techniques are at an early stage of development and application (57), they are quickly catching up.

While the present predictor set was finalized based on feasibility of which of the predictors from the systematic review list were likely to be recorded in clinical

practice, the future direction should be to use the evidence based ocular predictors that were recently suggested (47, 48). This could be encouraged by incorporating them in the newer technologies such as imaging modalities and the artificial intelligence platform. Modelling studies will have to be repeated as and when new or previously found evidence based predictors find their way into practice.

Survival analysis

Only a small fraction of referable diabetic retinopathy patients (15.2%) were found to reach the combined outcome of any of the treatments or loss of vision, while 71% did not, in a total median follow-up of 3.9 years. This finding goes further than the earlier reports of 50-70% false-positive rate (16, 17). The conclusions from the analysed primary care data may not truly reflect the situation though, and the measurement is best carried out on hospital data. It is fair to conclude though that M1 and R2 are the main source of false positives and should therefore be the area that is further investigated. It is quite clear that a service based on risk stratification (individual better than risk grouping), can improve system efficiency and safety of higher risk patients.

External validation

External validation is essential for a prognostic prediction model after development and internal validation. It essentially involves examining the model's predictive performance in a dataset not related to the original development set (39). The steps involved are 1) investigating the extent of relatedness between samples, 2) assessing model performance and 3) interpretation of the results - how well does the model perform in the new setting (transportability) and if poor, can it be improved by updating, re-calibrating or changing the predictors (58). While the same researcher

who developed the model can perform external validation (59), selective reporting and publication bias should be avoided by following the TRIPOD reporting guidelines for prediction models (35).

Impact on Clinical Practice

The model presented in this thesis can influence patient outcomes and improve cost effectiveness of the care provided based on predictions made. However, before suggesting or directing clinical decision making, the model will have to be assessed for its overall clinical benefit on outcomes and costs. For example, if the risks are above or below certain probability threshold then the follow-up interval could be longer or shorter than the routine current practice. Benefits, harms, and overall impact of this approach can be evaluated in three different ways. Convincing evidence can be obtained by carrying out comparative follow-up studies between usual practice and model directed practice. Randomised controlled trials can be very long and very expensive. Cost effectiveness modelling is another recognised method. Costs and outcomes of two approaches (with or without the model use) can be modelled for their impact on quality of life and costs. The third approach is decision curve analysis and net benefit assessment to quantify probability estimates for different scenarios before adoption for use in clinical practice. Decision curve analysis for clinical decision making by measuring net clinical benefit (improved outcomes) will be needed to help make a clinical decision, short of having to carry out a long and expensive impact study (usually a randomised clinical trial).

A proportion of referable diabetic retinopathy under the care of hospital eye service do not ever receive any treatment or lose vision. However, all patients need to be observed until they either are treated or show no progression up to a point when the

clinician is confident that they can be discharged back to the diabetic retinopathy screening service. Model performance measures I used (discrimination and calibration) don't capture the clinical consequences of the use of the model. Model performance tests alone cannot help decide on whether the model will be useful in clinical practice. Decision curve analysis considers the clinical consequences of using a model and thus attempts to solve this problem. To determine the model's clinical role, decision curve analysis will use validation data to measure the net benefit of overall consequences of the decisions made by using the model. This method requires weighing of the benefits of improved patient outcomes against harms of worse outcomes and increase in costs.

Therefore, if use of the model assists in the decision as to which patients are higher risk, it will reduce unnecessary visits for a large number of lower risk patients by increasing their follow up intervals. Unfortunately, it will miss some high-risk patients, who are then prone to harm. A degree of trade-off therefore needs to be considered, where a certain level of risk becomes acceptable for the benefit of reducing the burden of unnecessary follow-ups for the majority of patients. Net benefit compares benefits and harms by putting them on one scale, with a specified exchange rate mechanism (60). The model does not currently have thresholds for action (e.g. prescribe treatment or to change arranging follow-up interval routines based on a cut off threshold of risk) and may require expert input in the form of a qualitative study design based on decision curve analysis. This additional research in terms of decision curve analysis would be required to quantify probability estimates for different scenarios before adoption for use in clinical practice.

Once the net benefit is established, an early use of the model might be to rank newly diagnosed patients with referable diabetic retinopathy by the risk of the combined outcome (treatments/vision failure) so that those at the highest risk could be allocated early assessment or additional interventions to manage their risk factors. This risk model could be implemented for case finding in current electronic clinical systems (61) by the system suppliers, using the coded clinical data that they contain to populate the risk equations.

Once the model has had external validation and/or update if necessary, it may be considered for the clinic setting to 1) predict the risk of a patient with referable diabetic retinopathy reaching treatment requiring stage or vision failure, 2) provide a cut-off for recommending a follow-up interval or to assist clinicians in deciding on starting treatment, 3) help patients understand their individual risk of reaching the treatment stage or vision failure therefore motivating them to initiate lifestyle changes or to improve adherence to prescribed treatments, 4) find high-risk patients in electronic medical records (61) to prioritize their care and thus 5) allocate resources to those at highest risk.

Model presentation

When a clinical prediction model is developed, and is found suitable for use in clinical practice after external validation, a decision on the form of presentation has to be made, based on the intended users (clinicians), settings where it is intended to be used (secondary care), and the timing of use (clinic) is considered. I have included the full model equation as recommended, which will enable independent external validation and implementation by clinicians or use by other researchers (35). It is also best to involve stakeholders (clinicians) in the decision process. Patients may need

further tailoring of the model. The medium can be paper or electronic. The formats range from points score systems and nomograms to websites and mobile apps (62). Since this model is mainly for clinicians the most appropriate presentation form will be a user-friendly nomogram, incorporated within an electronic health record system.

External validation will lend an opportunity to carry out predictor research on hospital data and, if necessary, will update the model as well. Decision curve analysis, clinical benefit and model presentation can then be finalised. The model's integration into electronic medical records can help decision-making and needs to be the goal for future work on this model.

6.3 Conclusions

The studies presented within this thesis addressed two important topics in the area of diabetic retinopathy, (i) disease burden projections and (ii) a prediction model to determine individual prognosis prediction.

Most disease burden estimates suggested a trend of rising prevalence rates between type 1 and type 2 diabetes, but the T1DM prevalence rose significantly higher than the decade before. There was also a continuing rise in the prevalence of diabetic retinopathy, however the rate of increase slowed in the last decade. The rise in prevalence of all cases of DR was higher with T2DM, and STR was higher with T1DM. If the present trend continues, T1DM numbers are expected to grow faster, and are likely to contribute to a more severe DR caseload. Preventive strategies are needed, and care services need to look into increasing their capacity to prepare them to cope with this increasing demand.

However, against the background of the higher, and more severe, diabetic retinopathy on the horizon, a mechanism to predict which patients are likely to fall within that smaller higher risk group will help triage patients more effectively. A prognostic model is needed to determine their individual risk of progression to individualize their care. This will improve the safety of higher risk individuals, but will also direct resources where they are needed most.

The Systematic Review collated all the information on the existing models but found that they do not directly address the range of outcomes and target population of interest. The systematic review provided a comprehensive list of 78 candidate predictors though. A list of 19 evidence-based predictors for diabetic retinopathy progression was finalized with the help of expert clinical opinion through NGT and evidence evaluation. This set is ranked in importance by the NGT and can be used by any researcher to put together a model in an appropriate data set. Modifiable risk factors can be controlled to delay or eliminate risk.

A new Prognostic Prediction Model to assess the risk of vision loss and need for treatment in patients with referable diabetic retinopathy has been developed and internally validated. The model has shown good performance and can be used to identify patients with diabetes at high risk of these events for triage and risk stratification. Routinely collected primary care data can therefore be used to predict future risk of vision loss and/or the need for treatment in patients with referable diabetic retinopathy among the diabetic population.

During the model estimation, it became clear that hazard ratios were much higher for ocular predictors, as compared to systemic risk factors. There is a strong drive towards imaging-based artificial intelligence-driven algorithms. This upcoming

discipline, along with fast developing imaging technologies are likely to have a major impact on further modelling developments in this area.

Further research is needed to perform external validation, define any possible risk threshold, look into net clinical benefit, and to present as an easy to use nomogram before implementation. This may also compensate for the limitations of the study, and help the model to become more clinically useful.

6.4 References:

1. Patterson CC, Harjutsalo V, Rosenbauer J, Neu A, Cinek O, Skrivarhaug T, et al. Trends and cyclical variation in the incidence of childhood type 1 diabetes in 26 European centres in the 25 year period 1989-2013: a multicentre prospective registration study. Diabetologia. 2019;62(3):408-17.
2. Mayer-Davis EJ, Lawrence JM, Dabelea D, Divers J, Isom S, Dolan L, et al. Incidence Trends of Type 1 and Type 2 Diabetes among Youths, 2002-2012. The New England journal of medicine. 2017;376(15):1419-29.
3. Bagust A, Hopkinson PK, Maslove L, Currie CJ. The projected health care burden of Type 2 diabetes in the UK from 2000 to 2060. Diabetic medicine : a journal of the British Diabetic Association. 2002;19 Suppl 4:1-5.
4. Li JQ, Welchowski T, Schmid M, Letow J, Wolpers C, Pascual-Camps I, et al. Prevalence, incidence and future projection of diabetic eye disease in Europe: a systematic review and meta-analysis. Eur J Epidemiol. 2019.
5. Mathur R, Bhaskaran K, Edwards E, Lee H, Chaturvedi N, Smeeth L, et al. Population trends in the 10-year incidence and prevalence of diabetic retinopathy in the UK: a cohort study in the Clinical Practice Research Datalink 2004-2014. BMJ open. 2017;7(2):e014444.
6. Public Health England. Diabetes Prevalence Model 2016 [Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/612306/Diabetesprevalencemodelbriefing.pdf, Accessed 2019, July 12.
7. Gonzalez EL, Johansson S, Wallander MA, Rodriguez LA. Trends in the prevalence and incidence of diabetes in the UK: 1996-2005. Journal of epidemiology and community health. 2009;63(4):332-6.
8. Zghebi SS, Steinke DT, Carr MJ, Rutter MK, Emsley RA, Ashcroft DM. Examining trends in type 2 diabetes incidence, prevalence and mortality in the UK between 2004 and 2014. Diabetes Obes Metab. 2017;19(11):1537-45.
9. Diabetes UK. Facts and Figures 2019 [Available from: https://www.diabetes.org.uk/professionals/position-statements-reports/statistics, accessed 2020 May, 28.
10. Cho NH, Shaw JE, Karuranga S, Huang Y, da Rocha Fernandes JD, Ohlrogge AW, et al. IDF Diabetes Atlas: Global estimates of diabetes prevalence for 2017 and projections for 2045. Diabetes research and clinical practice. 2018;138:271-81.
11. Diabetes UK. Diabetes Prevalence 2017, Quality and Outcomes Framework 2018 [Available from: https://www.diabetes.org.uk/professionals/position-statements-reports/statistics/diabetes-prevalence-2017, Accessed 2019, July 12.
12. Tate AR, Dungey S, Glew S, Beloff N, Williams R, Williams T. Quality of recording of diabetes in the UK: how does the GP's method of coding clinical data affect incidence estimates? Cross-sectional study using the CPRD database. BMJ open. 2017;7(1):e012905.
13. Thomas RL, Dunstan F, Luzio SD, Roy Chowdury S, Hale SL, North RV, et al. Incidence of diabetic retinopathy in people with type 2 diabetes mellitus attending the Diabetic Retinopathy Screening Service for Wales: retrospective analysis. BMJ. 2012;344:e874.
14. Younis N, Broadbent DM, Harding SP, Vora JP. Incidence of sight-threatening retinopathy in Type 1 diabetes in a systematic screening programme. Diabetic Medicine. 2003;20(9):758-65.
15. Public Health England. Diabetic Eye Screening: 2016 to 2017 Data 2017 [Available from: https://www.gov.uk/government/publications/diabetic-eye-screening-2016-to-2017-data, Accessed 2019, July 12.
16. The Royal College of Ophthalmologists. The way forward age-related macular degeneration and diabetic retinopathy 2016

[ ]. Available from: https://www.rcophth.ac.uk/wp-content/uploads/2015/10/RCOphth-The-Way-Forward-AMD-300117.pdf, Accessed 2019, July 20.
17. Jyothi S, Elahi B, Srivastava A, Poole M, Nagi D, Sivaprasad S. Compliance with the quality standards of National Diabetic Retinopathy Screening Committee. Primary care diabetes. 2009;3(2):67-72.
18. Thomas RL, Dunstan FD, Luzio SD, Chowdhury SR, North RV, Hale SL, et al. Prevalence of diabetic retinopathy within a national diabetic retinopathy screening service. The British journal of ophthalmology. 2015;99(1):64-8.
19. Minassian DC, Owens DR, Reidy A. Prevalence of diabetic macular oedema and related health and social care resource use in England. The British journal of ophthalmology. 2012;96(3):345-9.
20. Keenan TD, Johnston RL, Donachie PH, Sparrow JM, Stratton IM, Scanlon P. United Kingdom National Ophthalmology Database Study: Diabetic Retinopathy; Report 1: prevalence of centre-involving diabetic macular oedema and other grades of maculopathy and retinopathy in hospital eye services. Eye (London, England). 2013;27(12):1397-404.
21. Public Health England. Diabetes Prevalnce Estimates for CCG's by GP registered populations 2016 [Available from: https://www.gov.uk/government/publications/diabetes-prevalence-estimates-for-local-populations, Accessed 2020, May 12.
22. Eleuteri A, Fisher AC, Broadbent DM, Garcia-Finana M, Cheyne CP, Wang A, et al. Individualised variable-interval risk-based screening for sight-threatening diabetic retinopathy: the Liverpool Risk Calculation Engine. Diabetologia. 2017.
23. Scanlon PH, Aldington SJ, Leal J, Luengo-Fernandez R, Oke J, Sivaprasad S, et al. Development of a cost-effectiveness model for optimisation of the screening interval in diabetic retinopathy screening. Health Technol Assess. 2015;19(74):1-116.
24. Aspelund T, Porisdottir O, Olafsdottir E, Gudmundsdottir A, Einarsdottir AB, Mehlsen J, et al. Individual risk assessment and information technology to optimise screening frequency for diabetic retinopathy. Diabetologia. 2011;54(10):2525-32.
25. Stratton IM, Aldington SJ, Farmer AJ, Scanlon PH. Personalised risk estimation for progression to sight-threatening diabetic retinopathy: How much does clinical information add to screening data? Diabetic Medicine. 2014;31:23-4.
26. Lagani V, Chiarugi F, Thomson S, Fursse J, Lakasing E, Jones RW, et al. Development and validation of risk assessment models for diabetes-related complications based on the DCCT/EDIC data. Journal of diabetes and its complications. 2015;29(4):479-87.
27. PROBAST. <PROBAST assessment of bias tool.pdf>.
28. Moons KG, de Groot JA, Bouwmeester W, Vergouwe Y, Mallett S, Altman DG, et al. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. PLoS medicine. 2014;11(10):e1001744.
29. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS medicine. 2009;6(7):e1000100.
30. Sajjad Haider KN, Salman Naveed, David Moore,. Prognostic prediction models for the progression of diabetic retinopathy (DR) and vision loss in patients with sight-threatening diabetic retinopathy (STDR): protocol for a systematic review. : PROSPERO 2017 [Available from: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42017057767 accessed 2020 May, 28.
31. Harrell FE. Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis: Springer; 2001.
32. Riley RD, Hayden JA, Steyerberg EW, Moons KG, Abrams K, Kyzas PA, et al. Prognosis Research Strategy (PROGRESS) 2: prognostic factor research. PLoS medicine. 2013;10(2):e1001380.
33. Hayden JA, van der Windt DA, Cartwright JL, Cote P, Bombardier C. Assessing bias in studies of prognostic factors. Ann Intern Med. 2013;158(4):280-6.

34. Martin-Merino E, Fortuny J, Rivero E, Garcia-Rodriguez LA. Validation of diabetic retinopathy and maculopathy diagnoses recorded in a U.K. primary care database. Diabetes care. 2012;35(4):762-7.
35. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): the TRIPOD Statement. The British journal of surgery. 2015;102(3):148-58.
36. Stratton IM, Kohner EM, Aldington SJ, Turner RC, Holman RR, Manley SE, 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.
37. Mehlsen J, Erlandsen M, Poulsen PL, Bek T. Identification of independent risk factors for the development of diabetic retinopathy requiring treatment. Acta Opthalmologica. 2011;89(6):515-21.
38. Romero-Aroca P, Baget-Bernaldiz M, Fernandez-Ballart J, Plana-Gil N, Soler-Lluis N, Mendez-Marin I, et al. Ten-year incidence of diabetic retinopathy and macular edema. Risk factors in a sample of people with type 1 diabetes. Diabetes research and clinical practice. 2011;94(1):126-32.
39. Riley RD, van der Windt D, Croft P, Moons KGM. Prognosis Research in Health Care: Concepts, Methods, and Impact: Oxford University Press; 2019.
40. van der Heijden AA, Nijpels G, Badloe F, Lovejoy HL, Peelen LM, Feenstra TL, et al. Prediction models for development of retinopathy in people with type 2 diabetes: systematic review and external validation in a Dutch primary care setting. Diabetologia. 2020;63(6):1110-9.
41. Haider S, Sadiq SN, Moore D, Price MJ, Nirantharakumar K. Prognostic prediction models for diabetic retinopathy progression: a systematic review. Eye. 2019;33(5):702-13.
42. Stratton IM, Aldington SJ, Taylor DJ, Adler AI, Scanlon PH. A simple risk stratification for time to development of sight-threatening diabetic retinopathy. Diabetes care. 2013;36(3):580-5.
43. Mehlsen J, Erlandsen M, Poulsen PL, Bek T. Individualized optimization of the screening interval for diabetic retinopathy: a new model. Acta Opthalmologica. 2012;90(2):109-14.
44. Harris Nwanyanwu K, Talwar N, Gardner TW, Wrobel JS, Herman WH, Stein JD. Predicting development of proliferative diabetic retinopathy. Diabetes care. 2013;36(6):1562-8.
45. Hippisley-Cox J, Coupland C. Development and validation of risk prediction equations to estimate future risk of blindness and lower limb amputation in patients with diabetes: Cohort study. BMJ (Online). 2015;351 (no pagination)(h5441).
46. Tanaka S, Tanaka S, Iimuro S, Yamashita H, Katayama S, Akanuma Y, et al. Predicting macro- and microvascular complications in type 2 diabetes: the Japan Diabetes Complications Study/the Japanese Elderly Diabetes Intervention Trial risk engine. Diabetes care. 2013;36(5):1193-9.
47. Sivaprasad S, Pearce E. The unmet need for better risk stratification of non-proliferative diabetic retinopathy. Diabetic medicine : a journal of the British Diabetic Association. 2019;36(4):424-33.
48. Lee CS, Lee AY, Baughman D, Sim D, Akelere T, Brand C, et al. The United Kingdom Diabetic Retinopathy Electronic Medical Record Users Group: Report 3: Baseline Retinopathy and Clinical Features Predict Progression of Diabetic Retinopathy. American Journal of Ophthalmology. 2017;180:64-71.
49. Jenkins AJ, Joglekar MV, Hardikar AA, Keech AC, O'Neal DN, Januszewski AS. Biomarkers in Diabetic Retinopathy. The Review of Diabetic Studies. 2015;12(1-2):159-95.
50. Geersing GJ, Bouwmeester W, Zuithoff P, Spijker R, Leeflang M, Moons KG. Search filters for finding prognostic and diagnostic prediction studies in Medline to enhance systematic reviews. PloS one. 2012;7(2):e32844.
51. Chatterley T, Dennett L. Utilisation of search filters in systematic reviews of prognosis questions. Health information and libraries journal. 2012;29(4):309-22.
52. Public Health England. NHS Diabetic Eye Screening Programme, Information for health professionals, 2016 [Available from:

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/505587/DES_07_GP_information_sheet_March_2016.pdf (Accessed 05/03/2020).
53. The Royal College of Ophthalmologists. Risk stratification coding framework March 2020 [updated 2020. Available from: https://www.rcophth.ac.uk/wp-content/uploads/2020/03/Measuring-follow-up-timeliness-and-risk-for-performance-reporting-improvement-actions-and-targeting-failsafe-procedures-in-England.pdf, accessed 2020, May 11.
54. Liu X, Faes L, Kale AU, Wagner SK, Fu DJ, Bruynseels A, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. The lancet digital health. 2019;1(6):e271-e97.
55. Sayres R, Taly A, Rahimy E, Blumer K, Coz D, Hammel N, et al. Using a Deep Learning Algorithm and Integrated Gradients Explanation to Assist Grading for Diabetic Retinopathy. Ophthalmology. 2019;126(4):552-64.
56. Raman R, Srinivasan S, Virmani S, Sivaprasad S, Rao C, Rajalakshmi R. Fundus photograph-based deep learning algorithms in detecting diabetic retinopathy. Eye (London, England). 2019;33(1):97-109.
57. Riley RD, van der Windt D, Croft P, Moons KGM. Prognosis Research in Healthcare: Concepts, Methods, and Impact: OUP Oxford; 2019.
58. Debray TP, Vergouwe Y, Koffijberg H, Nieboer D, Steyerberg EW, Moons KG. A new framework to enhance the interpretation of external validation studies of clinical prediction models. Journal of clinical epidemiology. 2015;68(3):279-89.
59. Steyerberg EW, Harrell FE, Jr. Prediction models need appropriate internal, internal-external, and external validation. Journal of clinical epidemiology. 2016;69:245-7.
60. Vickers AJ, Van Calster B, Steyerberg EW. Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests. Bmj. 2016;352:i6.
61. Marshall T, Caley M, Hemming K, Gill P, Gale N, Jolly K. Mixed methods evaluation of targeted case finding for cardiovascular disease prevention using a stepped wedged cluster RCT. BMC Public Health. 2012;12:908.
62. Bonnett LJ, Snell KIE, Collins GS, Riley RD. Guide to presenting clinical prediction models for use in clinical settings. Bmj. 2019;365:l737.

7 Appendices

7.1 Appendices for Chapter 2

7.1.1 Appendix 1: SRC approval

SRC Feedback

Researcher Name: Sajjad Haider
Organisation: University of Birmingham
SRC Reference Number: 19THIN028
Date: 18th April 2019
Study title: Diabetic retinopathy disease occurrence and prognostic research into diabetic retinopathy progression.

Committee opinion: Approved

The following feedback has been supplied by the SRC.

Notes from the Chair:

Advice (General advice for the researchers as information only)
A) The investigators could consider using multiple cross-sections (e.g. annual) to look at prevalence over time. Also, some thought is needed about the analysis of incidence from the period July to December 2017. Is this restricted to patients with diabetes? What is the rationale for the time period chosen?
B) Do the investigators have any information about the quality/validity of recording in primary care of their chosen outcomes for the prognistic modelling (i.e. laser/injection/vitrectomy treatment and vision loss)? Also, it could be worth assessing/noting (at least in the study limitations), the mentioned diabetic retinopathy services in the UK, especially Hospital Eye Service for higher risk patients, and the impact this could have on the identification of diabetic retinopathy in THIN.
C) The investigators are also advised to consider the best method for validation. The use of bootstrapping was not clear - wouldn't bootstrap samples overlap with the training data? Is the intention rather a cross-validation approach? As an alternative, would there be enough data to split the sample into separate training and validation sets?

Approved documents:

Approved document	Version	Date
SRC_Protocol_19THIN028_v1_03-04-2019	1	03/04/2019

We are pleased to inform that you can proceed with the study as this is now approved. IQVIA will let the relevant Research Ethics Committee (REC) know this study has been approved by the SRC.

Once the study has been completed and published, it is important for you to inform IQVIA for us to advise the SRC and your reference number to be closed.

References to all published studies are added to IQVIA's online bibliography enabling other researchers to become aware of your work. To identify your study as using the THIN database, this statement **must** be included in all publication (i.e. slides, posters, manuscripts, articles, abstracts etc):

"THIN is a registered trademark of Cegedim SA in the United Kingdom and other countries. Reference made to the THIN database is intended to be descriptive of the data asset licensed by IQVIA"

"This work uses de-identified data provided by patients as a part of their routine primary care"

In addition, studies that use THIN-HES linked data **must** include the following:

"Copyright © 2019, re-used with the permission of The Health & Social Care Information Centre. All rights reserved"

I wish you and your team all the best with the study progression.

Mustafa Dungarwalla
SRC Coordinator

7.1.2 Appendix 2: Ethical approval

From: Susan Cottam

Sent: 12 April 2018 14:49

To: Krishnarajah Nirantharakumar (Institute of Applied Health Research)

[redacted]

Cc: Malcolm James Price [redacted]

Subject: Application for Ethical Review ERN_18-0319

Dear Dr Nirantharakumar

Re: "Individualised Risk Assessment for Diabetic Retinopathy Progression to Optimize Management of Diabetic Retinopathy"

Application for Ethical Review ERN_18-0319

Thank you for your application for ethical review for the above project, which was reviewed by the Science, Technology, Engineering and Mathematics Ethical Review Committee.

On behalf of the Committee, I confirm that this study now has full ethical approval.

I would like to remind you that any substantive changes to the nature of the study as described in the Application for Ethical Review, and/or any adverse events occurring during the study should be promptly bought to the Committee's attention by the Principal Investigator and may necessitate further ethical review.

Please also ensure that the relevant requirements within the University's Code of Practice for Research and the information and guidance provided on the University's ethics webpages (available at https://intranet.birmingham.ac.uk/finance/accounting/Research-Support-

Group/Research-Ethics/Links-and-Resources.aspx) are adhered to and referred to in any future applications for ethical review. It is now a requirement on the revised application form (https://intranet.birmingham.ac.uk/finance/accounting/Research-Support-Group/Research-Ethics/Ethical-Review-Forms.aspx) to confirm that this guidance has been consulted and is understood, and that it has been taken into account when completing your application for ethical review.

Please be aware that whilst Health and Safety (H&S) issues may be considered during the ethical review process, you are still required to follow the University's guidance on H&S and to ensure that H&S risk assessments have been carried out as appropriate. For further information about this, please contact your School H&S representative or the University's H&S Unit at healthandsafety@contacts.bham.ac.uk.

Kind regards

Susan Cottam

Research Ethics Officer

Research Support Group

7.1.3 Appendix 3: Protocol for SRC

Diabetic retinopathy disease occurrence and prognostic research into diabetic retinopathy progression

Background, Rationale and Research question

There has been a global increase in the prevalence of diabetes (422 million in 2014) (1) and consequently, the burden of the disease of diabetic retinopathy (a complication of diabetes) has increased. Prevalence of diagnosed diabetes in the UK was reported to have risen by 50% from 1996 to 2005 (2). Around 4 million people aged 16 years and over in England have diabetes (diagnosed and undiagnosed). This is equal to 8.6% of the population of this age group (3). The same study projected the numbers for diabetes for years up to 2035. Recent figures and predictions for diabetic retinopathy are uncertain though (4). It will be useful to have these figures from a cross sectional study on up to date The Health Improvement Network **(**THIN) data for 2017 and design a future projection model to help clinicians and managers decide on resources needed to meet the demand.

In the UK, diabetic retinopathy services are organised into Diabetic Eye Screening Programme (DESP) for lower risk patients and Hospital Eye Service providing them a closer observation and treatment for higher risk patients. Diabetic retinopathy patients are referred to the hospital when they develop clinical signs of sight threatening diabetic retinopathy. The detection of diabetic retinopathy has also improved through better population coverage (5), increasing the workload for DESP and Hospital Eye Services. However, approximately 50% of referrals don't need intervention and are observed in the Hospital Eye Services, for a variable period of time. This has resulted in increasing demand and capacity mismatch. There are various service modifications being looked into, to improve this aspect (4).

Risk stratification with the help of prognostic prediction models has been used successfully in DESP by dividing patients into two groups: low risk, who are followed up by the Diabetic Eye Screening Programme; and high risk who are referred to the Hospital Eye Service (6, 7). A similar approach of stratifying patients will also help

hospital eye services optimisation. Such a model combines two or more predictors of diabetic retinopathy progression to a treatment requiring stage or loss of vision (8).

A recent systematic review (9) of the existing predictive models identified three models with moderate to low risk of bias (of the predictive characteristics and applicability to patients under hospitals care). They could potentially (after an update) serve the purpose in our population of interest (higher risk diabetic retinopathy patients in UK hospital). However, we envisage a new model will be needed considering the aforementioned models are designed for low risk part of the disease pathway and do not primarily predict progression of retinopathy from referable retinopathy stage to the stage needing therapy / or ending up with vision loss.

Aim of this doctoral research is to determine the prevalence and incidence of diabetic retinopathy, describe prognostic prediction factors in cross sectional data in patients with diabetes, and then develop and validate a prognostic model in longitudinal data to help stratify higher risk diabetic retinopathy patient populations according to their risk of outcomes (progression to treatment requiring stage or vision failure). This will enable clinicians and managers to more effectively direct resources toward the more needful higher risk patients.

Data Source

This study will use THIN general practice data (one of the commonest used longitudinal data available in the UK). THIN database covers 5.7% of the UK population (10) and has been shown to be generalizable to UK demographics (11).

Many of the variables included in the systematic review (10) are available in this database, as a previous model has been built using THIN database (12). A significant amount of literature related to our subject has been published based on THIN database in the subject of diabetic retinopathy (2, 12-16). Similarly, there has been some literature published out of another similar primary care database - clinical practice research datalink - CPRD (5, 17).

Methods

Study designs

Firstly a cross-sectional study will be carried out in THIN data to map the prevalence of diabetic retinopathy (both background and sight threatening) and risk factors for diabetic retinopathy on 1st of January 2017. We will also determine the incidence of sight threatening retinopathy from 1st of July 2017 to 31st December 2017. We shall then compare it with earlier studies and national screening programme figures to provide an insight into the changing trends and completeness of recording in THIN database. For example we will compare the incidence of sight threatening retinopathy in the THIN database to that observed in the national screening programme.

Thereafter we will extract demographic, predictor factor and event data from THIN longitudinal data to derive the model (development set) and then internally validate it by bootstrapping. The source population for this will be patients with referable sight threatening retinopathy. Clinical outcomes will be Laser / injection / vitrectomy treatment or vision loss. Data on the predictor variables will be collected from THIN. Patient leaving the practice, having an outcome of interest, their death or the termination of documentation in the THIN database will mark the end of recording for that participant.

Study population

- Patients with type 1 and 2 diabetes
- Over 12 years of age (age at when screening for diabetic retinopathy starts)
- Patients with diabetes registered in THIN database and active on 1st January 2017 for cross sectional study.
- Patients with newly diagnosed sight threatening retinopathy registered in THIN from 2000 until last data collection date for prognostic model

Inclusion Criteria

- Patients aged >12 years with either Type 1 or Type 2 diabetes will be included.
- A patient is eligible to take part in the study one year after registration with the practice or a year after the practice meets the criteria for inclusion.
- A practice will be eligible to participate a year after installation of the electronic medical record or a year after their Acceptable Mortality Recording date.

Follow-up period for the prognostic model study

- Index date will be the date of documentation of sight threatening retinopathy
- Patient leaving the practice, having an outcome of interest, their death or the termination of documentation in the THIN database will mark the end of recording for that participant (exit date).
- All patients will be followed up from index date to exit date

Covariates

- **Study Variables** will be the prognostic factors for the outcomes of progression to stage requiring treatment (laser, vitreous injections, vitreous surgery, and vitreous haemorrhage) or vision loss. The predictors identified in the systematic review are given in table 1.

- **Data Analysis**

Prevalence will be reported as a proportion and incidence per 1000 person years. We will report prevalence for any retinopathy and sight threatening retinopathy. If sufficiently well recorded we will also report for individual retinopathy stages.
Model development will be by fitting a survival model (cox regression) and then assessing its performance by ROC curves, calculating sensitivity / specificity / positive predictive value and negative predictive value; and calibration by constructing lowess calibration plots
Internal validation will be carried out using bootstrapping.

In the modelling data analysis, before applying multiple imputations for the missing data, a complete case analysis will be carried out on the assumption that data missing is missing at random.

Limitations

There will be some missing data for variables and outcomes. Multiple imputations technique may need to be used where possible to overcome bias induced by the missing data. Incomplete recording of stages of retinopathy is another issue. We will investigate this and report as a limitation.

Ethical approval

Application for Ethical Review (ERN_18-0319) was approved by our University.

References:

1. (NCD-RisC) NRFC. Worldwide trends in diabetes since 1980: a pooled analysis of 751 population-based studies with 4.4 million participants. Lancet. 2016;387(10027):1513-30.
2. Gonzalez EL, Johansson S, Wallander MA, Rodriguez LA. Trends in the prevalence and incidence of diabetes in the UK: 1996-2005. Journal of epidemiology and community health. 2009;63(4):332-6.
3. Public Health England. Diabetes Prevalence Model 2016 [Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/612306/Diabetesprevalencemodelbriefing.pdf, Accessed 2019, July 12.
4. The Royal College of Ophthalmologists. The way forward age-related macular degeneration and diabetic retinopathy 2016 [ ]. Available from: https://www.rcophth.ac.uk/wp-content/uploads/2015/10/RCOphth-The-Way-Forward-AMD-300117.pdf, Accessed 2019, July 20.
5. Mathur R, Bhaskaran K, Edwards E, Lee H, Chaturvedi N, Smeeth L, et al. Population trends in the 10-year incidence and prevalence of diabetic retinopathy in the UK: a cohort study in the Clinical Practice Research Datalink 2004-2014. BMJ open. 2017;7(2):e014444.
6. Scanlon PH, Aldington SJ, Leal J, Luengo-Fernandez R, Oke J, Sivaprasad S, et al. Development of a cost-effectiveness model for optimisation of the screening interval in diabetic retinopathy screening. Health Technology Assessment. 2015;19(74):1-116.
7. Aspelund T, Porisdottir O, Olafsdottir E, Gudmundsdottir A, Einarsdottir AB, Mehlsen J, et al. Individual risk assessment and information technology to optimise screening frequency for diabetic retinopathy. Diabetologia. 2011;54(10):2525-32.
8. Steyerberg EW, Moons KG, van der Windt DA, Hayden JA, Perel P, Schroter S, et al. Prognosis Research Strategy (PROGRESS) 3: prognostic model research. PLoS medicine. 2013;10(2):e1001381.
9. Haider S, Sadiq SN, Moore D, Price MJ, Nirantharakumar K. Prognostic prediction models for diabetic retinopathy progression: a systematic review. Eye. 2019.
10. Dhalwani NN, Tata LJ, Coleman T, Fleming KM, Szatkowski L. Completeness of maternal smoking status recording during pregnancy in United Kingdom primary care data. PloS one. 2013;8(9):e72218.
11. Blak BT, Thompson M, Dattani H, Bourke A. Generalisability of The Health Improvement Network (THIN) database: demographics, chronic disease prevalence and mortality rates. Informatics in primary care. 2011;19(4):251-5.
12. McEwan P, Bennett H, Ward T, Bergenheim K. Refitting of the UKPDS 68 risk equations to contemporary routine clinical practice data in the UK. Pharmacoeconomics. 2015;33(2):149-61.
13. Mamtani R, Haynes K, Finkelman BS, Scott FI, Lewis JD. Distinguishing incident and prevalent diabetes in an electronic medical records database. Pharmacoepidemiology and drug safety. 2014;23(2):111-8.
14. Martin-Merino E, Fortuny J, Rivero E, Garcia-Rodriguez LA. Validation of diabetic retinopathy and maculopathy diagnoses recorded in a U.K. primary care database. Diabetes care. 2012;35(4):762-7.
15. Martin-Merino E, Fortuny J, Rivero-Ferrer E, Garcia-Rodriguez LA. Incidence of retinal complications in a cohort of newly diagnosed diabetic patients. PloS one. 2014;9(6):e100283.
16. Martin-Merino E, Fortuny J, Rivero-Ferrer E, Lind M, Garcia-Rodriguez LA. Risk factors for diabetic retinopathy in people with Type 2 diabetes: A case-control study in a UK primary care setting. Primary care diabetes. 2016;10(4):300-8.
17. Hippisley-Cox J, Coupland C. Development and validation of risk prediction equations to estimate future risk of blindness and lower limb amputation in patients with diabetes: Cohort study. BMJ (Online). 2015;351 (no pagination)(h5441).

7.1.4 Appendix 4: 7 Step Process of Read codes selection methods

Read codes cover clinical features, diagnosis, procedures, some drugs and investigations (1). Ones used in IMRD consist of 7 characters. They have a hierarchy with more specific ones down the order. This was done in collaboration with Jhot Chandan, a fellow doctoral researcher and my supervisor Krishnarajah Nirantharakumar (Institute of Applied Health Research)

1. The Read code database (MsAcess, MsExcel) is divided into two main columns: A Medcode column with unique 8 character codes for each condition and a description column. Both were used, starting from the description column.
2. We developed a list of key search terms for the read codes of interest. These were searched for in the description column. Appendix 5 below provides a list of key search words.
3. Results from the key word search were used to identify the main stem codes where the Read codes of interest belong to.
4. The Next step involved searching the MedCode column for the main stem codes to pick out codes that were otherwise missed on searching the description column.
5. We then also conducted an online search of published articles that have published similar Read Codes (2, 3).
6. Once collected, they were split into possible, probable and definite. There was deliberation between clinicians in the THINking group to achieve these three lists.
7. They were then hand over to a group of data scientists within the THINking group who split them into various files following epidemiological principles and saved them in CSV files database.

References:

1. Health I. THIN Data Guide for Researchers. 2017(Version 3):108.
2. Mathur R, Bhaskaran K, Edwards E, Lee H, Chaturvedi N, Smeeth L, et al. Population trends in the 10-year incidence and prevalence of diabetic retinopathy in the UK: a cohort study in the Clinical Practice Research Datalink 2004-2014. BMJ open. 2017;7(2):e014444.

3. McEwan P, Bennett H, Ward T, Bergenheim K. Refitting of the UKPDS 68 risk equations to contemporary routine clinical practice data in the UK. Pharmacoeconomics. 2015;33(2):149-61.

7.1.5 Appendix 5: Search Terms for diabetic Retinopathy

Keywords for identifying diabetic retinopathy in the Read Codes Dictionary
O/E or *PHOTOGRAPHY* or *RETINAL* or *SCR* **and** *HAEMORRHAGES* or *EXUDATE* or *MICROANEURYSMS* or *INTRARETINAL MICROVASCULAR ANAOMALY* or *ABNORMALITY*
RETINA or *FUNDUS* or *MACULAR* or *VITREOUS* **and** *LASER" or *PHOTOCOAGULATION* or *INTRA-VITREAL INJECTIONS* or *INJECTIONS* or *RANIBIZUMAB* or *BIVACIZUMAB* or *AFLIBERCEPT* or *TRIAMCINOLON* or *ILEUVIEN* or *DEXAMETHOSON*
RETINOPATHY or *FUNDOSCOPY* or *SEEN or *RETINAL SCR* or *RETINOSCOPY* or *SLIT LAMP* or *DIABETIC EYE* or *EXAMINATION OF RETINA* or *RETINA and OTHER PARTS OF EYE OPERATIONS* or *VITRECTOMY* or *MACULOPATHY* or *BACKGROUND* or *PRE PROLIFERATIVE* or *PROLIFERATIVE*
*BLIND" or *PARTIAL SIGHTED" or **SIGHT IMPAIRMENT" or *VISUAL IMPAIRMENT" or *VISUAL FAILURE"

7.1.6 Appendix 6: Read Codes

Code	Description	Status
	No Retinoipathy (ROMO)	
2BBD.00	O/E - Right retina normal	Probable
2BBJ.00	O/E - no right diabetic retinopathy	Definite
2BB1.00	O/E - retina normal	Probable
2BBI.00	O/E - no retinopathy	Definite
3128000	Fundoscopy normal	Probable
3128200	Dilated fundoscopy normal	Probable
2BBM.00	O/E - diabetic maculopathy absent both eyes	Possible
	Background Retinopathy (R1)	
2BBP.00	O/E - right eye background diabetic retinopathy	Definite
2BBQ.00	O/E - left eye background diabetic retinopathy	Definite
F420000	Background diabetic retinopathy	Definite
F421.00	Other background retinopathy	Definite
F421000	Unspecified background retinopathy	Definite
F421z00	Other background retinopathy NOS	Definite
2BB4.00	O/E - retinal microaneurysms	Definite
2BBa.00	O/E- non-referable retinopathy	Probable
	Pre proliferative Diabetic Retinopathy (R2)	
F420200	Pre proliferative diabetic retinopathy	Definite
2BBR.00	O/E - right eye pre proliferative diabetic retinopathy	Definite
2BBS.00	O/E - left eye pre proliferative diabetic retinopathy	Definite
F420800	High risk non proliferative diabetic retinopathy	Definite
	Proliefartive Diabetic Retinoipathy (R3)	
2BBk.00	O/E - right eye stable treated prolif diabetic retinopathy	Definite
2BBl.00	O/E - left eye stable treated prolif diabetic retinopathy	Definite
F420100	Proliferative diabetic retinopathy	Definite
F420700	High risk proliferative diabetic retinopathy	Definite
F422z00	Proliferative retinopathy NOS	Definite
F422.00	Other proliferative retinopathy	Definite
FyuF700	[X]Other proliferative retinopathy	Definite
2BBT.00	O/E - right eye proliferative diabetic retinopathy	Definite
2BBV.00	O/E - left eye proliferative diabetic retinopathy	Definite
7272500	Panretinal laser photocoagulation to lesion of retina NEC	Definite
7272800	Panretinal laser photocoagulation to lesion of retina	Definite
2BB7.00	O/E - retinal vascular prolif.	Probable
2BB8.00	O/E - vitreous haemorrhages	Probable
7276	Pan retinal photocoagulation for diabetes	Definite
F420500	Advanced diabetic retinal disease	Possible
F422y00	Other specified other proliferative retinopathy	Definite

F4K2800	Vitreous haemorrhage	Probable
FyuH400	[X]Vitreous haemorrhage in diseases classified elsewhere	Probable
2BB8.00	O/E - vitreous haemorrhages	Probable
	Diabetic Maculopathy (M1)	
2BBL.00	O/E - Diabetic maculopathy present both eyes	Definite
2BBm.00	O/E - right eye clinically significant macular oedema	Definite
2BBn.00	O/E - left eye clinically significant macular oedema	Definite
2BBW.00	O/E - right eye diabetic maculopathy	Definite
2BBX.00	O/E - left eye diabetic maculopathy	Definite
F425900	Maculopathy	Definite
F42y900	Macular oedema	Definite
C10EP00	Type 1 diabetes mellitus with exudative maculopathy	Definite
C10EP11	Type I diabetes mellitus with exudative maculopathy	Definite
C10FQ00	Type 2 diabetes mellitus with exudative maculopathy	Definite
C10FQ11	Type II diabetes mellitus with exudative maculopathy	Definite
F420300	Advanced diabetic maculopathy	Definite
7272900	Focal laser photocoagulation of retina	Probable
F420400	Diabetic maculopathy	Definite
	Referrable Retinopathy (R2, R3, M1)	
2BBY.00	O/E - referable retinopathy	Definite
2BBo.00	O/E - sight threatening diabetic retinopathy	Definite
	Advanced diabetic retinal disease	
F420500	Advanced diabetic retinal disease	Definite

Code	Description	Status
	Laser Procedures	
7276	Pan retinal photocoagulation for diabetes	Definite
7272012	Photocoagulation of the retina NEC	Definite
7272013	Laser therapy lesion of retina	Definite
7272300	Laser destruction of lesion of retina	Definite
7272500	Pan retinal laser photocoagulation to lesion of retina NEC	Definite
7272600	Laser photocoagulation to lesion of retina NEC	Definite
7272800	Pan retinal laser photocoagulation to lesion of retina	Definite
7272900	Focal laser photocoagulation of retina	Definite
2BBk.00	O/E - right eye stable treated proliferative diabetic retinopathy	Definite
2BBl.00	O/E - left eye stable treated proliferative diabetic retinopathy	Definite
2BBO.00	O/E - Laser photocoagulation scars	Definite
5B4..11	Retinal laser therapy	Definite
Z6F..11	Laser therapy	Definite

5B42.00	Laser therapy - retinal lesion	Definite
	Vitreous/ Peribulbar procedures / haemorrhage	
7270D00	Injection of Ranibizumab into vitreous body	Definite
7270z00	Operation on vitreous body NOS	Definite
7270300	Injection into vitreous body NEC	Definite
7274800	Injection of therapeutic substance around the eye	Possible
727C200	Injection therapeutic substance posterior segment of eye NEC	Definite
7270D00	Injection of Ranibizumab into vitreous body	Definite
7L19E00	Injection of triamcinolone	Probable
727C100	Injection of steroid into posterior segment of eye	Definite
7270200	Injection of vitreous substitute into vitreous body	Definite
7277600	Injection of therapeutic substance into macula	Definite
7270C00	Injection of vitreous substitute into vitreous body NEC	Definite
727C100	Injection of steroid into posterior segment of eye	Definite
7270400	Pars plana vitrectomy	Definite
727Cy00	Other specified operations on posterior segment of eye	Probable
727Cz00	Operations on posterior segment of eye NOS	Probable
7273000	Epiretinal dissection	Possible
727C000	Insertion sustained release device posterior segment of eye	Definite
7270y00	Other specified operation on vitreous body	Definite
7270800	Internal tamponade of retina using liquid	Possible
7270900	Internal tamponade of retina using oil	Possible
7270A00	Removal of internal tamponade agent from vitreous body	Possible
7270411	Vitrectomy using pars plana approach	Probable
7270500	Air/gas exchange of vitreous	Possible
7270600	Internal tamponade of retina using gas	Probable
7270200	Injection of vitreous substitute into vitreous body	Probable
7270300	Injection into vitreous body NEC	Definite
7270400	Pars plana vitrectomy	Definite
7270	Operations on vitreous body	Probable
7270100	Extirpation of vitreous body NEC	Probable
F4K2800	Vitreous haemorrhage	Definite
FyuH400	[X]Vitreous haemorrhage in diseases classified elsewhere	Definite
2BB8.00	O/E - vitreous haemorrhages	Definite
	Vision loss / blindness	
ZV52200	[V]Fitting or adjustment of artificial eye	Probable
ZV43000	[V]Has artificial eye globe	Probable
ZV43100	[V]Has artificial eye lens	Possible
FyuL.00	[X]Visual disturbances and blindness	Definite
F49z.11	Acquired blindness	Definite
F490900	Acquired blindness, both eyes	Definite
F495A00	Acquired blindness, one eye	Definite

F491.00	Better eye: low vision, Lesser eye: profound VI	Definite
F491500	Better eye: moderate VI, Lesser eye: blind, unspecified	Definite
F492300	Better eye: moderate VI, Lesser eye: low vision unspecified	Definite
F492500	Better eye: moderate VI, Lesser eye: moderate VI	Definite
F491700	Better eye: moderate VI, Lesser eye: near total VI	Definite
F491800	Better eye: moderate VI, Lesser eye: profound VI	Definite
F492400	Better eye: moderate VI, Lesser eye: severe VI	Definite
F491600	Better eye: moderate VI, Lesser eye: total VI	Definite
F490400	Better eye: near total VI, Lesser eye: near total VI	Definite
F490300	Better eye: near total VI, Lesser eye: total VI	Definite
F490200	Better eye: near total VI, Lesser eye: unspecified	Definite
F490700	Better eye: profound VI, Lesser eye: near total VI	Definite
F490800	Better eye: profound VI, Lesser eye: profound VI	Definite
F490600	Better eye: profound VI, Lesser eye: total VI	Definite
F490500	Better eye: profound VI, Lesser eye: unspecified	Definite
F491100	Better eye: severe VI, Lesser eye: blind, unspecified	Definite
F492100	Better eye: severe VI, Lesser eye: low vision unspecified	Definite
F491300	Better eye: severe VI, Lesser eye: near total VI	Definite
F491400	Better eye: severe VI, Lesser eye: profound VI	Definite
F492200	Better eye: severe VI, Lesser eye: severe VI	Definite
F491200	Better eye: severe VI, Lesser eye: total VI	Definite
8F62.00	Blind lead dog rehabilitation	Definite
8F6..11	Blind rehabilitation	Definite
8F61.00	Blind rehabilitation	Definite
ZN56800	Blind telephone user	Definite
F49..00	Blindness and low vision	Definite
F490z00	Blindness both eyes NOS	Definite
F490.00	Blindness, both eyes	Definite
F49A.00	Blindness, monocular	Definite
F495000	Blindness, one eye, unspecified	Definite
F490100	Both eyes total visual impairment	Definite
668C.00	Certificate of vision impairment	Definite
Fy1..00	Combined visual and hearing impairment	Definite
Fy1..12	Deafblind	Definite
ZN56A00	Deaf-blind telephone user	Definite
Fy1..11	Dual sensory impairment - deafblind	Definite
9m08.00	Exclu diab ret screen as blind	Definite
2BBr.00	Impair vision due diab retinop	Definite
F49..11	Impaired vision	Definite
ZK74.00	Issue of local authority blind registration	Definite
F494.00	Legal blindness USA	Definite
F496500	Lesser eye: moderate VI, Better eye: near normal vision	Definite
F496600	Lesser eye: moderate VI, Better eye: normal vision	Definite

F496400	Lesser eye: moderate VI, Better eye: unspecified	Definite
F495500	Lesser eye: near total VI, Better eye: near normal vision	Definite
F495600	Lesser eye: near total VI, Better eye: normal vision	Definite
F495400	Lesser eye: near total VI, Better eye: unspecified	Definite
F495800	Lesser eye: profound VI, Better eye: near normal vision	Definite
F495900	Lesser eye: profound VI, Better eye: normal vision	Definite
F495700	Lesser eye: profound VI, Better eye: unspecified	Definite
F496200	Lesser eye: severe VI, Better eye: near normal vision	Definite
F496300	Lesser eye: severe VI, Better eye: normal vision	Definite
F496100	Lesser eye: severe VI, Better eye: unspecified	Definite
F495200	Lesser eye: total VI, Better eye: near normal vision	Definite
F495300	Lesser eye: total VI, Better eye: normal vision	Definite
F495100	Lesser eye: total visual impairment, Better eye: unspecified	Definite
F49..12	Low vision	Definite
F492.00	Low vision, both eyes	Definite
F492z00	Low vision, both eyes NOS	Definite
F492000	Low vision, both eyes unspecified	Definite
F496.00	Low vision, one eye	Definite
F496z00	Low vision, one eye NOS	Definite
F496000	Low vision, one eye, unspecified	Definite
F498.00	Moderate visual impairment, binocular	Definite
F49C.00	Moderate visual impairment, monocular	Definite
2B7A.11	O/E - blind L-eye	Definite
2B6A.11	O/E - blind R-eye	Definite
22E6.11	O/E - false eye	Definite
22E6.00	O/E - glass (prosthetic) eye	Definite
22E6.12	O/E - glass eye	Definite
22EF.00	O/E - has one eye	Definite
2B7B.00	O/E - L-eye completely blind	Definite
2B7C.00	O/E - L-eye sees hand movements	Definite
2B7T.00	O/E - L-eye visual acuity (corrected) 1/60	Definite
2B7V.00	O/E - L-eye visual acuity (corrected) 2/60	Definite
2B7W.00	O/E - L-eye visual acuity (corrected) 4/60	Definite
2B7X.00	O/E - L-eye visual acuity (corrected) 5/60	Definite
2B7S.00	O/E - pinhole L-eye completely blind	Definite
2B7Q.00	O/E - pinhole L-eye counts fingers only	Definite
2B7R.00	O/E - pinhole L-eye perceives light only	Definite
2B7P.00	O/E - pinhole L-eye sees hand movements	Definite
2B6S.00	O/E - pinhole R-eye completely blind	Definite
2B6Q.00	O/E - pinhole R-eye counts fingers only	Definite
2B6R.00	O/E - pinhole R-eye perceives light only	Definite
2B6P.00	O/E - pinhole R-eye sees hand movements	Definite
2B7L.00	O/E - pinhole visual acuity L-eye=6/60	Definite

2B6L.00	O/E - pinhole visual acuity R-eye=6/60	Definite
22E6.13	O/E - prosthetic eye	Definite
2B6B.00	O/E - R-eye completely blind	Definite
2B6C.00	O/E - R-eye sees hand movements	Definite
2B6T.00	O/E - R-eye visual acuity (corrected) 1/60	Definite
2B6V.00	O/E - R-eye visual acuity (corrected) 2/60	Definite
2B6W.00	O/E - R-eye visual acuity (corrected) 4/60	Definite
2B6X.00	O/E - R-eye visual acuity (corrected) 5/60	Definite
2B7E.00	O/E - visual acuity L-eye=3/60	Definite
2B78.00	O/E - visual acuity L-eye=6/60	Definite
2B6E.00	O/E - visual acuity R-eye=3/60	Definite
2B68.00	O/E - visual acuity R-eye=6/60	Definite
2B79.00	O/E -L-eye counts fingers only	Definite
2B69.00	O/E -R-eye counts fingers only	Definite
2B7A.00	O/E-L-eye perceives light only	Definite
2B6A.00	O/E-R-eye perceives light only	Definite
F491000	One eye blind, one eye low vision	Definite
F491z00	One eye blind, one eye low vision NOS	Definite
Z9E2.00	Optical low vision aid provision	Definite
F49..13	Partial sight	Definite
F495z00	Profound impairment one eye NOS	Definite
F495.00	Profound impairment, one eye	Definite
Z96..00	Provision for visual and hearing impairment	Definite
Z9E5400	Provision of ancillary low vision aid	Definite
Z9E1100	Provision of artificial eye	Definite
Z962.00	Provision of communicator for visual and hearing impairment	Definite
Z9E5100	Provision of electronic low vision aid	Definite
Z961.00	Provision of guide help for visual and hearing impairment	Definite
Z9E3200	Provision of low vision hand magnifier	Definite
Z9E3400	Provision of low vision headband magnifier	Definite
Z9E3300	Provision of low vision stand magnifier	Definite
Z9E3100	Provision of magnifier low vision aid - near	Definite
Z9E5.00	Provision of non-optical low vision aid	Definite
Z9E4.00	Provision of optical low vision aid - distance	Definite
Z9E3.00	Provision of optical low vision aid - near	Definite
Z9E1200	Provision of removable artificial eye	Definite
Z9E3500	Provision of spectacle low vision aid - near	Definite
8HIE.00	Referral to visual impairment multidisciplinary team	Definite
6689	Registered blind	Definite
6688.11	Registered partially blind	Definite
6688	Registered partially sighted	Definite
6689.11	Registered severely sight impaired	Definite
668D.00	Registered sight impaired	Definite

8D36.00	Removable artificial eye	Definite
9Nfb.00	Requires deafblind block alphabet interpreter	Definite
9NfB.00	Requires deafblind communicator guide	Definite
9Nfc.00	Requires deafblind haptic communication interpreter	Definite
9Nfa.00	Requires deafblind manual alphabet interpreter	Definite
F497.00	Severe visual impairment, binocular	Definite
F49B.00	Severe visual impairment, monocular	Definite
F49..14	Sight impaired	Definite
F490000	Unspecified blindness both eyes	Definite
1a00000	Uses guide dog for the blind	Definite
F49D.00	Visual impairment	Definite
F493.00	Visual loss, both eyes unqualified	Definite
F49y.00	Visual loss, one eye, unqualified	Definite
F404200	Blind hypertensive eye	Definite
F404100	Blind hypotensive eye	Definite
Z9E3900	Near low vision aid - clip-on spectacle magnifier	Definite
Z9E3C00	Near low vision aid - clip-on spectacle telescope	Definite
Z9E3D00	Near low vision aid - extra cap for telescope	Definite
Z9E3800	Near low vision aid - integral spectacle magnifier	Definite
Z9E3B00	Near low vision aid - integral spectacle telescope	Definite
9NlD.00	Seen by visual impairment teacher	Definite
1B75.00	Loss of vision, Severe visual loss	Definite
1B77.00	Deteriorating vision, Severe visual loss	Definite

Unclassifiable

Code	Description
2BB5.00	O/E - retinal haemorrhages
2BB6.00	O/E - retinal exudates
2BBF.00	Retinal abnormality-diabetes related
2BBr.00	Impaired vision due diab retinop
C105.00	Diabetes mellitus with ophthalmic manifestation
C105000	Diabetes mellitus, juvenile type, + ophthalmic manifestation
C105100	Diabetes mellitus, adult onset, + ophthalmic manifestation
C105y00	Other specified diabetes mellitus with ophthalmic complicatn
C105z00	Diabetes mellitus NOS with ophthalmic manifestation
C108100	Insulin-dependent diabetes mellitus with ophthalmic comps
C108111	Type I diabetes mellitus with ophthalmic complications
C108112	Type 1 diabetes mellitus with ophthalmic complications
C108700	Insulin dependent diabetes mellitus with retinopathy
C108711	Type I diabetes mellitus with retinopathy
C108712	Type 1 diabetes mellitus with retinopathy
C109100	Non-insulin-dependent diabetes mellitus with ophthalm comps

C109111	Type II diabetes mellitus with ophthalmic complications
C109112	Type 2 diabetes mellitus with ophthalmic complications
C109600	Non-insulin-dependent diabetes mellitus with retinopathy
C109611	Type II diabetes mellitus with retinopathy
C109612	Type 2 diabetes mellitus with retinopathy
C10A300	Malnutrit-related diabetes mellitus wth ophthalmic complicat
C10E100	Type 1 diabetes mellitus with ophthalmic complications
C10E111	Type I diabetes mellitus with ophthalmic complications
C10E112	Insulin-dependent diabetes mellitus with ophthalmic comps
C10E700	Type 1 diabetes mellitus with retinopathy
C10E711	Type I diabetes mellitus with retinopathy
C10E712	Insulin dependent diabetes mellitus with retinopathy
C10F600	Type 2 diabetes mellitus with retinopathy
C10F611	Type II diabetes mellitus with retinopathy
F420.00	Diabetic retinopathy
F420600	Non proliferative diabetic retinopathy
F420z00	Diabetic retinopathy NOS
F421.11	Microvascular retinal changes
2BB5.00	O/E - retinal haemorrhages
2BBM.00	O/E - diabetic maculopathy absent both eyes

7.1.7 Appendix 7: Summary of Prevalence Trends 1998 to 2018

Decade	Prevalence estimate at the start of the decade	Prevalence estimate at the end of the decade	Percentage increase in prevalence within the decade	Percentage increase in prevalence between the decades
STR in T1DM in two decades				
1998 to 2007	8.15	17.57	216%	
2009 to 2018	20.54	30.22	147%	371%
STR in T2DM in two decades				
1998 to 2007	4.36	8.1	186%	
2009 to 2018	9.01	11.15	124%	256%
DR in T1DM in two decades				
1998 to 2007	26.62	40.32	151%	
2009 to 2018	45.39	57.75	127%	217%
DR in T2DM in two decades				
1998 to 2007	11.53	20.06	174%	
2009 to 2018	23.7	32.64	138%	283%
STR in DM in two decades				
1998 to 2007	4.87	8.84	182%	
2009 to 2018	9.86	12.48	127%	256%
DR in DM in two decades				
1998 to 2007	13.57	21.64	159%	
2009 to 2018	25.3	34.39	136%	253%
T1DM in two decades				
1998 to 2007	0.31%	0.32%	104%	
2009 to 2018	0.33%	0.41%	123%	132%
T2DM in two decades				
1998 to 2007	1.91%	3.65%	191%	
2009 to 2018	4.01%	5.24%	131%	273%

7.1.8 Appendix 8: Future projections

In the four figures below, the grey area is the prediction band (95% confidence interval) and signifies the uncertainty of the estimates.

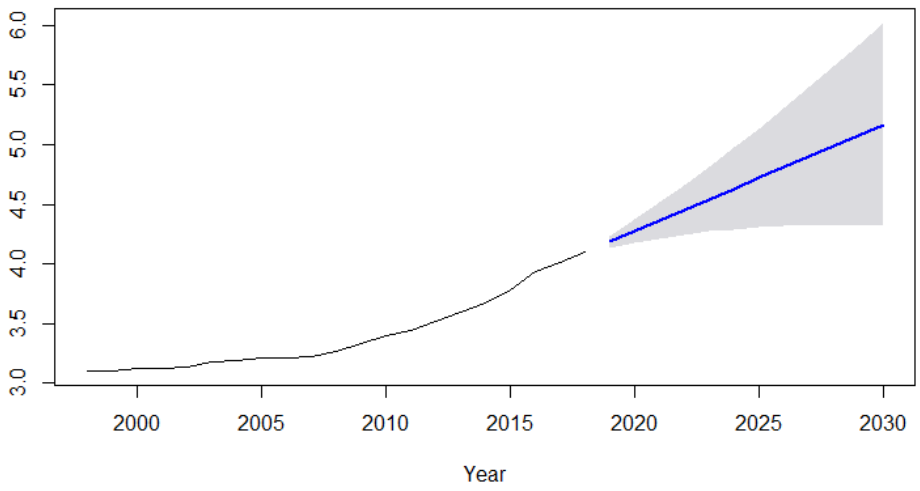


Figure 1: T1DM Projections / 1000 individuals

X axis is calendar years and Y axis is prevalence (cases / 1000 individuals general population), starts at 3.0

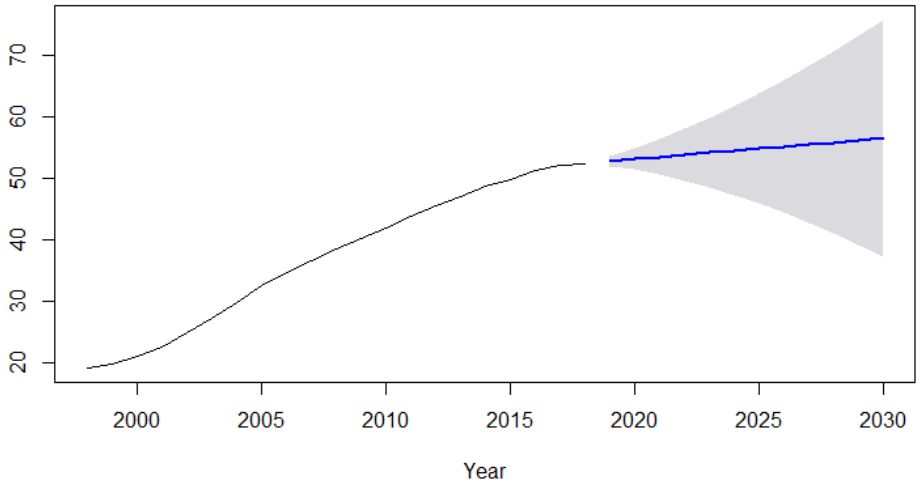


Figure 2: T2DM Projections / 1000 individuals

X axis is calendar years and Y axis is prevalence (cases / 1000 individuals general population) starts at 17

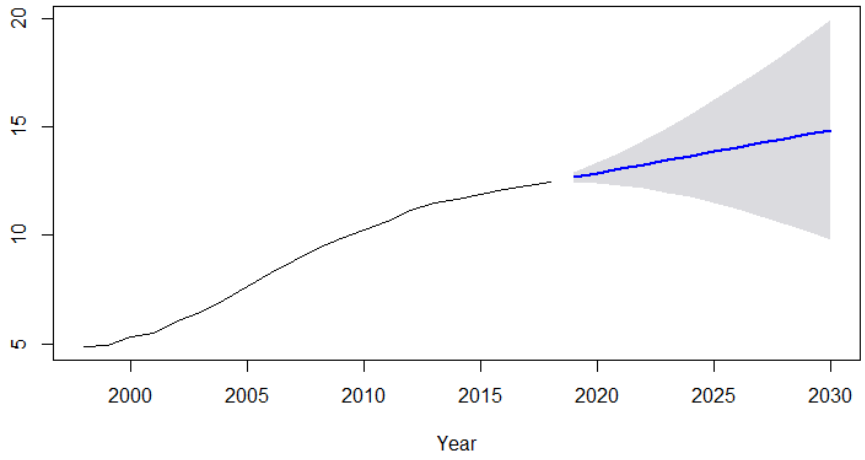


Figure 3: STR Projections (%)

X axis is calendar years and Y axis is prevalence (cases / 100 individuals with diabetes) starts at 4

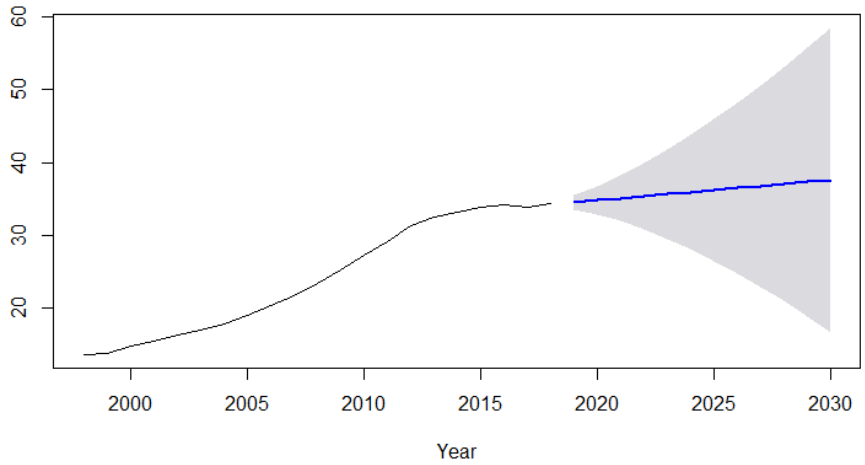


Figure 4: DR Projections (%)

X axis is calendar years and Y axis is prevalence (cases / 100 individuals with diabetes), starts at 10

Annual Prevalence Diabetes Mellitus / 1000 Population and retinopathy / 100 diabetic population (95% PI)

	T1DM Forecast			T2DM Forecast			DR Forecast			STR Forecast		
Year	Forecast	Low 95	High 95	Forecast	Low 95	High 95	Forecast	Low 95	High 95	Forecast	Low 95	High 95
2019	4.2	4.1	4.2	52.7	51.9	53.6	34.5	33.5	35.5	12.7	12.4	12.9
2020	4.3	4.2	4.4	53.1	51.4	54.8	34.8	32.9	36.7	12.9	12.4	13.3
2021	4.4	4.2	4.5	53.4	50.6	56.3	35.1	31.9	38.2	13.1	12.3	13.8
2022	4.5	4.2	4.7	53.8	49.6	57.9	35.4	30.8	39.9	13.3	12.2	14.4
2023	4.5	4.3	4.8	54.1	48.5	59.7	35.6	29.5	41.8	13.5	12.0	15.0
2024	4.6	4.3	5.0	54.4	47.2	61.6	35.9	28.1	43.8	13.7	11.8	15.6
2025	4.7	4.3	5.1	54.8	45.8	63.7	36.2	26.5	45.9	13.9	11.5	16.2
2026	4.8	4.3	5.3	55.1	44.3	65.9	36.5	24.7	48.2	14.1	11.2	16.9
2027	4.9	4.3	5.5	55.4	42.7	68.2	36.8	22.9	50.6	14.3	10.9	17.6
2028	5.0	4.3	5.7	55.8	41.0	70.6	37.0	20.9	53.2	14.5	10.5	18.4
2029	5.1	4.3	5.8	56.1	39.1	73.1	37.3	18.8	55.8	14.7	10.2	19.1
2030	5.2	4.3	6.0	56.4	37.2	75.7	37.6	16.7	58.5	14.9	9.8	19.9

7.1.9 Appendix 9: Workings behind the incidence

New referrals from DESP during one year into prevalence conversion

The English DESP referrals made were 2.1% (using this as incidence rate, equivalent to 24.1% of crude prevalence rate using the formula Average age at diagnosis (STR) = 66.4, Average age at death = 77.7 years and thus the formula used (Prevalence = (Incidence Rate) x Average Duration of Disease (11.3) (1).

Reference:

1. Tools E. Incidence into prevalence Modules/EP/EP713_DiseaseFrequency/EP713_DiseaseFrequency7html.

Wayne W. LaMorte. Relationship Among Prevalence, Incidence Rate, and Average Duration of Disease 2016 [Available from: http://sphweb.bumc.bu.edu/otlt/MPH-Modules/EP/EP713_DiseaseFrequency/EP713_DiseaseFrequency7.html.

7.1.10 Appendix 10: Previous prevalence studies compared with IMRD based analysis

Publication	Population	T1DM	T2DM	Any DM
Younis et al (1)	Liverpool diabetic retinopathy screening programme 1991 to 1999 – baseline prevalence at entry into the programme	Any DR 45.7% STED 16.4% PDR 3.7%	Any DR 25.3% STED 6.0% PDR 0.5%	
Misra et al (2)	Norwich Diabetic retinopathy screening programme 2006 with dynamic cohort design with repeated measures			Any DR 25.6% STDR 0.6% PPDR 4.6% PDR 0.08% Maculopathy 0.44% Referable (R2, R3, M1) retinopathy 4.7%
Thomas (3) and Minassian et al (4)	Welsh Diabetic retinopathy screening programme 2005 to 2009 and application to England	Any DR 56.3% STDR 11.2%	Any DR 30.9% STDR 2.9%	Any DR 32.4% STDR 3.4% Diabetic Macular Oedema 7.12%
Looker et al (5)	Newly diagnosed type 2 diabetes attending Scottish National screening programme 2005 to 2008. prevalence at first screening		Any DR 19.3% Referable DR 1.9% PPDR ± any maculopathy 0.4% PDR ± any maculopathy 0.3%	

Mathur et al (6)	CPRD based UK wide study 2014 - crude prevalence rate	Any DR 54.8% Severe DR 8.1%	Any DR 30.6% Severe DR 1.2%	Any DR 32.6% Severe DR 1.8%
The present study	IMRD based cross sectional study - 2017	Any DR 57.8% STR 30.2% Any maculopathy 19.62%	Any DR 32.6% STR 11.2% Any maculopathy 6.99%	Any DR 34.4% STR 12.3% Any maculopathy 7.86%

References:

1. Younis N, Broadbent DM, Harding SP, Vora JR. Prevalence of diabetic eye disease in patients entering a systematic primary care-based eye screening programme. Diabetic medicine : a journal of the British Diabetic Association. 2002;19(12):1014-21.
2. Misra A, Bachmann MO, Greenwood RH, Jenkins C, Shaw A, Barakat O, et al. Trends in yield and effects of screening intervals during 17 years of a large UK community-based diabetic retinopathy screening programme. Diabetic medicine : a journal of the British Diabetic Association. 2009;26(10):1040-7.

3. Thomas RL, Dunstan FD, Luzio SD, Chowdhury SR, North RV, Hale SL, et al. Prevalence of diabetic retinopathy within a national diabetic retinopathy screening service. The British journal of ophthalmology. 2015;99(1):64-8.

4. Minassian DC, Owens DR, Reidy A. Prevalence of diabetic macular oedema and related health and social care resource use in England. The British journal of ophthalmology. 2012;96(3):345-9.

5. Looker H, Nyangoma S, Cromie D, Olson J, Leese G, Black M, et al. Diabetic retinopathy at diagnosis of type 2 diabetes in Scotland. Diabetologia. 2012;55(9):2335-42.
6. Mathur R, Bhaskaran K, Edwards E, Lee H, Chaturvedi N, Smeeth L, et al. Population trends in the 10-year incidence and prevalence of diabetic retinopathy in the UK: a cohort study in the Clinical Practice Research Datalink 2004-2014. BMJ open. 2017;7(2):e014444.

7.1.11 Appendix 11: Previous publications reporting trends in prevalence rates of DR in the UK compared with IMRD based analysis

Publication	Population	T1DM	T2DM	Any DM
Misra et al (1)	Norwich Diabetic retinopathy screening programme 1990 to 2006 (Mostly Type 2) with dynamic cohort design with repeated measures			All DR prevalence increased from 23.2% to 25.3% Referable DR increased from 2 to 4.7%
Mathur et al (2)	CPRD based UK wide study population from 2004 to 2014	All DR remained stable at 55% Severe DR increased from 3.5% in 2004 to 8.0% in 2014	All DR reduced from 24.6% in 2004 to 23.1% in 2014 Severe DR increased from 0.3% in 2004 to 0.9% in 2014	All DR prevalence decreased from 2.6% to 2.2% Severe DR remained stable at 0.1%
This study	IMRD based serial cross-sectional studies 1998 to 2018			

References:

1. Misra A, Bachmann MO, Greenwood RH, Jenkins C, Shaw A, Barakat O, et al. Trends in yield and effects of screening intervals during 17 years of a large UK community-based diabetic retinopathy screening programme. Diabetic medicine : a journal of the British Diabetic Association. 2009;26(10):1040-7.

2. Mathur R, Bhaskaran K, Edwards E, Lee H, Chaturvedi N, Smeeth L, et al. Population trends in the 10-year incidence and prevalence of diabetic retinopathy in the UK: a cohort study in the Clinical Practice Research Datalink 2004-2014. BMJ open. 2017;7(2):e014444.

7.2 Appendices for Chapter 3

7.2.1 Appendix 12 : Supplementary material for the publication: Prognostic prediction models for diabetic retinopathy progression: A systematic review (1)

The classification table has been reproduced in chapter 1 (Table 1). Online supplementary material included appendices 13 to 15.

Reference:

1. Haider S, Sadiq SN, Moore D, Price MJ, Nirantharakumar K. Prognostic prediction models for diabetic retinopathy progression: a systematic review. Eye. 2019;33(5):702-13.

7.2.2 Appendix 13: Sample search strategy

1 diabetes mellitus, Type 1/ or diabetes.mp. or diabetes, gestational/ or diabetes complications/ or diabetes mellitus.ti,ab. (462309)

2 Blindness.ti,ab. (20537)
3 vision disorder$.mp. (24996)
4 (loss adj2 vision).ti,ab. (7657)
5 (visual adj2 loss).ti,ab. (11720)
6 or/2-5 (57618)

7 retinopathy.ti,ab. (31993)
8 diabetic retinopathy/ (21173)
9 diabetic retinopathy.ti,ab. (15966)
10 diabetic macular edema.ti,ab. (1840)
11 diabetic macular oedema.ti,ab. (363)
12 DMO.ti,ab. (499)
13 DME.ti,ab. (1484)
14 diabetic maculopathy.ti,ab. (369)
15 1 and 6 (2921)
16 1 and 7 (12829)
17 or/8-16 (30971)

18 (Stratification or ROC curve or discrimination or discriminate or c-statistic or c statistic or area or area under the curve or AUC or calibration or indices or algorithm$ or multivariable).ti,ab. (1118656)

19	((prognos$ or predict$) adj3 model$).ti,ab. (74085)
20	or/18-19 (1176150)
21	17 and 20 (1997)
22	limit 17 to "prognosis (maximizes specificity)" (1868)
23	limit 17 to "clinical prediction guides (maximizes specificity)" (221)

24	21 or 22 or 23 (3810)
25	limit 24 to humans (3653)

*****Repeated searches on 20/07/2017, did not yield any further records

7.2.3 Appendix 14: Candidate predictors

PREDICTOR GROUPS (n = type of candidate predictors in each group)	
Ocular signs (10)	Retinopathy level Presence of DR Hard exudates R0 both eyes BDR (NPDR) both eyes BDR (NPDR) one eye Haemorrhages PDR Maculopathy Visual acuity score
Socio-demographics (6)	Age Gender Race Married Townsend deprivation score Occupation
Diabetes characteristics (6)	Type of DM Age at diagnosis Age at NPDR diagnosis Duration of DM Post-pubertal Duration of DM Time to next screening
Biochemical parameters (15)	HbA1c Serum Glucose HDL /cholesterol ratio Total serum cholesterol LDL cholesterol HDL cholesterol Non-HDL Cholesterol Albumin /creatinine ratio Urine Albumin Serum creatinine Total triglycerides Log ACR

	Haemoglobin eGFR WBC
Physical examination (4)	SBP DBP Heart Rate Mean arterial pressure
Co-morbidities (10)	HTN Stroke History CHF history IHD history Dyslipidaemia Atrial Fibrillation Seizures Psychiatric illness Rheumatoid Arthritis Urinary tract infection
Diabetic complications (10)	Chronic renal disease Nephropathy Non-healing foot ulcer Amputation Stroke Neuropathy Peripheral vascular disease MI Diabetic ketoacidosis Severe Hypoglycaemia
Diabetes treatment (7)	Statin Metformin

	Sulfonylurea Insulin ACE Anti-Hypertensive Rx Medications use
Lifestyle (6)	Smoking Ideal Body weight BMI Exercise/physical activity Alcohol Waist-hip ratio
Family History (4)	IDDM NIDDM MI HTN
Total candidate predictors = 78	

7.2.4 Appendix 15: Abbreviations

AER	Albumin excretion rate
AIC	Corrected Akaike's information criterion
AUC	Areas under the receiver operating curves
CACTI	Coronary Artery Calcification in Type 1 Diabetes study
CHD	coronary heart disease
CHF	Congestive Heart Failure
CPRD	Clinical practice research datalink
CRF	Clinical risk factors
CKD	Chronic kidney disease
C-statistic	Concordance statistic
DESP	Diabetic eye screening programme
Dev	Model Development
EDC	Epidemiology of Diabetes Complications Study
EURODIAB PCS	EURODIAB Prospective Complications Study
EV	External Validation
FinnDiane	Finnish Diabetic Nephropathy Study
FPG	Fasting plasma glucose
GP2DRS	General Practice to Diabetic Retinopathy Screening
HDL	High-density lipoprotein
HTN	Hypertension
IHD	Ischaemic heart disease
IDDM	Insulin dependent diabetes mellitus
IV	Internal Validation
JDCS	Japan Diabetes Complications Study
J-EDIT	Japanese Elderly Diabetes Intervention Trial
LDS	Lipids in Diabetes Study
LOCF	Last observation carried forward
MAPE	Mean Absolute Percentage Error
MI	Myocardial infarction
NRI	Net Reclassification Improvement
NIDDM	Non-insulin dependent diabetes mellitus
NPDR	Non-proliferative diabetic retinopathy

PDR	Proliferative diabetic retinopathy
PTM	Post trial monitoring
PVD	Peripheral vascular disease
RCT	Randomised Controlled Trial
SE	Screening events
STDR	Sight threatening diabetic retinopathy
STR	Sight threatening retinopathy
THIN	The health improvement newtwork
T1 DM	Type 1 Diabetes Mellitus
T2 DM	Type 2 Diabetes Mellitus
WHR	Waist-hip ratio

7.2.5 Appendix 16: Systematic Review Protocol

This section of chapter 3 presents the protocol version published on PROSPERO website (https://www.crd.york.ac.uk/prospero/myprospero) reference given below

Prognostic prediction models for the progression of diabetic retinopathy (DR) and vision loss in patients with sight-threatening diabetic retinopathy (STDR) - Protocol for a systematic review

Reference:

Sajjad Haider KN, Salman Naveed, Krishnarajah Nirantharakumar, David Moore. Prognostic prediction models for the progression of diabetic retinopathy (DR) and vision loss in patients with sight-threatening diabetic retinopathy (STDR): protocol for a systematic review. : PROSPERO 2017 [Available from: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42017057767 accessed 2020 May, 28.

UNIVERSITY of York
Centre for Reviews and Dissemination

NHS
National Institute for Health Research

PROSPERO International prospective register of systematic reviews

Review title and timescale

1 Review title
Give the working title of the review. This must be in English. Ideally it should state succinctly the interventions or exposures being reviewed and the associated health or social problem being addressed in the review.
Prognostic prediction models for the progression of diabetic retinopathy (DR) and vision loss in patients with sight-threatening diabetic retinopathy (STDR) - Protocol for a systematic review

2 Original language title
For reviews in languages other than English, this field should be used to enter the title in the language of the review. This will be displayed together with the English language title.

3 Anticipated or actual start date
Give the date when the systematic review commenced, or is expected to commence.
14/03/2017

4 Anticipated completion date
Give the date by which the review is expected to be completed.
31/05/2017

5 Stage of review at time of this submission
Indicate the stage of progress of the review by ticking the relevant boxes. Reviews that have progressed beyond the point of completing data extraction at the time of initial registration are not eligible for inclusion in PROSPERO. This field should be updated when any amendments are made to a published record.

The review has not yet started ×

Review stage	Started	Completed
Preliminary searches	Yes	No
Piloting of the study selection process	No	No
Formal screening of search results against eligibility criteria	No	No
Data extraction	No	No
Risk of bias (quality) assessment	No	No
Data analysis	No	No

Provide any other relevant information about the stage of the review here.

Review team details

6 Named contact
The named contact acts as the guarantor for the accuracy of the information presented in the register record.
Sajjad Haider

7 Named contact email
Enter the electronic mail address of the named contact.
[redacted]

8 Named contact address
Enter the full postal address for the named contact.
Sajjad Haider C/O Krishnarajah Niranthanakumar, Institute of Applied Health Research, University of Birmingham, Edgbaston, Birmingham B15 2TT UK

9 Named contact phone number
Enter the telephone number for the named contact, including international dialing code.
[redacted]

10 Organisational affiliation of the review
Full title of the organisational affiliations for this review, and website address if available. This field may be completed

as 'None' if the review is not affiliated to any organisation.
Institute of Applied Health Research, University of Birmingham

Website address:
www.birmingham.ac.uk/research/activity/applied-health

11 Review team members and their organisational affiliations
Give the title, first name and last name of all members of the team working directly on the review. Give the organisational affiliations of each member of the review team.

Title	First name	Last name	Affiliation
Dr	Sajjad	Haider	University of Birmingham
Dr	Krishnarajah	Nirantharakumar	University of Birmingham
Dr	Malcolm	Price	University of Birmingham
Dr	David	Moore	University of Birmingham

12 Funding sources/sponsors
Give details of the individuals, organizations, groups or other legal entities who take responsibility for initiating, managing, sponsoring and/or financing the review. Any unique identification numbers assigned to the review by the individuals or bodies listed should be included.
N/A

13 Conflicts of interest
List any conditions that could lead to actual or perceived undue influence on judgements concerning the main topic investigated in the review.
Are there any actual or potential conflicts of interest?
None known

14 Collaborators
Give the name, affiliation and role of any individuals or organisations who are working on the review but who are not listed as review team members.

Title	First name	Last name	Organisation details

Review methods

15 Review question(s)
State the question(s) to be addressed / review objectives. Please complete a separate box for each question.
This systematic review will identify, summarize and assess the quality of studies of any design examining prediction / prognostic models utilizing combinations of predictors for the risk of progression of diabetic retinopathy reaching the treatment requiring stage or visual loss.

16 Searches
Give details of the sources to be searched, and any restrictions (e.g. language or publication period). The full search strategy is not required, but may be supplied as a link or attachment.
Bibliographic databases: MEDLINE, EMBASE (OVID), Cochrane, CENTRAL, using keywords - Prediction / prognostic / model / tool, progression, diabetic retinopathy and vision loss.and index terms text words / search string recommended in an update by Geersing et al. Filters for Prognostic model and prediction models will be added. No language or year limits will be applied. In addition proceedings of the key ophthalmic and diabetes conferences of The Royal College of Ophthalmologists, American Academy of Ophthalmology, European Society of Retina Specialists (EURETINA), European Society of Ophthalmology (SOE), Association for Research in Vision and Ophthalmology (ARVO), The American Diabetic Association (ADA), Diabetes UK and International Diabetes Foundation will be searched for the previous 3 years for conference abstracts etc

17 URL to search strategy
If you have one, give the link to your search strategy here. Alternatively you can e-mail this to PROSPERO and we will store and link to it.

I give permission for this file to be made publicly available

Yes

18 Condition or domain being studied
Give a short description of the disease, condition or healthcare domain being studied. This could include health and wellbeing outcomes.
Diabetic Retinopathy - a complication of diabetes mellitus affecting the eyes, one of the leading causes of vision loss through the development of new blood vessels and leaking blood vessels at the centre of the eye (macula).

19 Participants/population
Give summary criteria for the participants or populations being studied by the review. The preferred format includes details of both inclusion and exclusion criteria.
Patients with diabetes

20 Intervention(s), exposure(s)
Give full and clear descriptions of the nature of the interventions or the exposures to be reviewed
This review only considers prediction prognostic models

21 Comparator(s)/control
Where relevant, give details of the alternatives against which the main subject/topic of the review will be compared (e.g. another intervention or a non-exposed control group).
Nil

22 Types of study to be included
Give details of the study designs to be included in the review. If there are no restrictions on the types of study design eligible for inclusion, this should be stated.
This review will include studies of any design that develop, validate, update, compare, or evaluate a prediction / prognostic model / tool, utilizing multiple prediction factors to predict the risk of progression of diabetic retinopathy and vision loss.

23 Context
Give summary details of the setting and other relevant characteristics which help define the inclusion or exclusion criteria.
There has been some progress in the preparation of prediction models recently, but they have mostly focused on the incidence of diabetic retinopathy and optimization of the screening services rather than the progression of disease in patients under treatment in the hospital setting.

24 Primary outcome(s)
Give the most important outcomes.
The primary outcome for the review will be the predictive accuracy and applicability of prognostic prediction model for diabetic retinopathy progression needing treatment..

Give information on timing and effect measures, as appropriate.
The progression of diabetic retinopathy reaching treatment requiring stage.

25 Secondary outcomes
List any additional outcomes that will be addressed. If there are no secondary outcomes enter None.
Secondary outcomes will be the predictive accuracy and applicability of prognostic prediction model for diabetic retinopathy causing vision loss.

Give information on timing and effect measures, as appropriate.
The progression of diabetic retinopathy causing vision loss.

26 Data extraction (selection and coding)
Give the procedure for selecting studies for the review and extracting data, including the number of researchers involved and how discrepancies will be resolved. List the data to be extracted.
Study Selection: Titles (and abstracts where available) will initially be screened by two reviewers and a list of studies included developed by the two reviewers independently. The two lists from the above screening will be matched and consensus achieved by discussion. Full texts of any potentially relevant articles will then be obtained and two reviewers will independently apply the full inclusion criteria The study selection process will be documented using the PRISMA flow diagram Data Extraction Data extraction and critical appraisal will be independently carried out by two reviewers using CHARMS Checklist. Disagreements will be resolved through discussion or referral to a third reviewer.

27 Risk of bias (quality) assessment
State whether and how risk of bias will be assessed, how the quality of individual studies will be assessed, and whether and how this will influence the planned synthesis.
The included studies will be assessed using the Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. In addition, PROBAST tool will also be used for the same purpose and the utility of this tool will be commented upon. Any impact studies found (RCT, comparison of two cohorts, or a before and after study) will have to be looked at with another appropriate assessment tool as CHARM checklist will not be relevant.

28 Strategy for data synthesis
Give the planned general approach to be used, for example whether the data to be used will be aggregate or at the level of individual participants, and whether a quantitative or narrative (descriptive) synthesis is planned. Where appropriate a brief outline of analytic approach should be given.
A narrative summary will report the critical appraisal / risk of bias / interpretation and analysis of the included studies The reporting of the results will be carried out with the help of PRISMA check list. Results summary will include different studies in relation with their primary purpose, i.e. model development, validation, evaluation etc. The different studies on same or similar models will be grouped together in a tabulated form and will also be described in detail separately. Similarities and differences between the different models will be also be explained. Descriptive analysis will be based on predictor's selection, their relative importance, sample sizes used, modelling methods used, missing data handling, predictive horizon, models presentation, reporting quality, predictive performance, methods used for validation and assessment, updating etc. The various models' generalizability to patients reaching treatment requiring stage will be examined and described. A meta-analysis will be carried out if appropriate with the help of guidance. We shall also be looking into the impact of using the prediction model studied in case of evidence becoming available.

29 Analysis of subgroups or subsets
Give any planned exploration of subgroups or subsets within the review. 'None planned' is a valid response if no subgroup analyses are planned.
None planned

Review general information

30 Type and method of review
Select the type of review and the review method from the drop down list.
Systematic review

Eye disorders

31 Language
Select the language(s) in which the review is being written and will be made available, from the drop down list. Use the control key to select more than one language.
English

Will a summary/abstract be made available in English?
Yes

32 Country
Select the country in which the review is being carried out from the drop down list. For multi-national collaborations select all the countries involved. Use the control key to select more than one country.
England

33 Other registration details
Give the name of any organisation where the systematic review title or protocol is registered together with any unique identification number assigned. If extracted data will be stored and made available through a repository such as the Systematic Review Data Repository (SRDR), details and a link should be included here.

34 Reference and/or URL for published protocol
Give the citation for the published protocol, if there is one.
Give the link to the published protocol, if there is one. This may be to an external site or to a protocol deposited with CRD in pdf format.

I give permission for this file to be made publicly available
Yes

35 Dissemination plans
Give brief details of plans for communicating essential messages from the review to the appropriate audiences.
Do you intend to publish the review on completion?
Yes

36 Keywords
Give words or phrases that best describe the review. (One word per box, create a new box for each term)
Prognostic tools, diabetic retinopathy progression and visual loss

37 Details of any existing review of the same topic by the same authors
Give details of earlier versions of the systematic review if an update of an existing review is being registered, including full bibliographic reference if possible.

38 Current review status
Review status should be updated when the review is completed and when it is published.
Ongoing

39 Any additional information
Provide any further information the review team consider relevant to the registration of the review.

40 Details of final report/publication(s)
This field should be left empty until details of the completed review are available.
Give the full citation for the final report or publication of the systematic review.
Give the URL where available.

7.2.6 Appendix 17: 14 Models with their outcomes of interest

	Model	Outcomes
1	UKPDS 68 2004	One eye blindness
2	UKPDS OM-2 2013	One eye blindness
3	JJ risk engine 2013	Progression of retinopathy
4	Soedamah-Muthu, S.S. et al. 2014	Blindness
5	Lagani, Chiarugi et al. 2015	Retinopathy event
6	Harris Nwanyanwu, Talwar et al. 2013	Progression to PDR
7	Hippisley-Cox and Coupland 2015	Blindness in one or both eyes
8	UKPDS 68 risk equations 2015	Blindness
9	Aspelund T et al. 2011	STR
10	Mehlsen 2012	Treatment requiring DR
11	Stratton 2013	STR
12	Stratton, Aldington et al. 2014	STR
13	Scanlon PH 2015	STR
14	ISDR model 2017	Referable DR

7.2.7 Appendix 18: Models with their Candidate Predictors

	Candidate predictors	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Ocular signs	Retinopathy level/DR grade					1				1					0
	Presence of DR									1					
	Hard exudates										1				
	R0 both eyes											1			
	BDR (NPDR) both eyes											1	1	1	
	BDR (NPDR) one eye											1	1	1	
	Haemorrhages										1				
	PDR							1							
	Maculopathy							1							
	Visual acuity score					0									
Socio-demographics	Age		0	1	1	0		1	1					0	
	Race	0	0			0	1	0							0
	Gender	0	0	0	0	0	1		1	1	1			0	0
	Married					1									
	Townsend deprivation score							1							
	Occupation					0									
Diabetes characteristics	Type of DM							1		1				0	0
	Age at diagnosis	1	1								1				1
	Age at NPDR diagnosis						1								
	Duration of DM		0	1	0	0		1		1	1		1	1	1
	Post-pubertal Duration of DM					1									
	Time to next screening										1				
Biochemical predictors	HbA1c	1	1	1	1	1	1	1	1	1	1		1	1	1
	Serum Glucose					0									
	HDL /cholesterol ratio	0						1	1						
	Total serum cholesterol					0							1	0	1
	LDL cholesterol		0	0	0	0									0
	HDL cholesterol		0	0	1	0									0
	Non-HDL Cholesterol			0	0										
	Albumin /creatinine ratio		0			0								0	0
	Urine Albumin					0									
	Serum creatinine					0							0		
	Total triglycerides		0	0	0	0									0
	Log ACR			1	1										

	Haemoglobin		0												
	eGFR		0	0		0									0
	WBC		1												
Physical examination	DBP				0	0								0	0
	Mean arterial pressure					0									
	Heart Rate		1												
	SBP	0	1	0	0	0		1	1	1				0	1
Co-morbidities	HTN						1	0							
	Stroke History		0												
	CHF history		1					0	0						
	IHD history		1					0							
	Dyslipidaemia						1								
	Atrial Fibrillation	0	0	0				0	1						
	Seizures					0									
	Psychiatric illness					0									
	Rheumatoid Arthritis							0							
	Urinary tract infection					0									
Diabetic complications	Chronic renal disease	0	0					1	0						
	Diabetic Nephropathy						1								
	Nonhealing Foot ulcers						1								
	Amputation	0	0						1						
	Stroke	0							1						
	Diabetic Neuropathy					0	1								
	PVD	0	0					0	1						
	MI	0							1						
	Diabetic ketoacidosis					0									
	Hypoglycemia					0									
Diabetes treatment	Statin						1								
	Metformin						1								
	Sulfonylurea						1								
	Insulin					0	1								0
	ACE						1								
	Anti-Hypertensive Rx				0										
	Medications use														0
Lifestyle	Smoking	0	0	0	0	0		0	1					0	0
	Ideal Body weight					0									
	BMI	0	0	1	0	1		0	1					0	0
	Exercise/physical activity			0		0									

	Alcohol			0		0									
	Waist-hip ratio				1										
Family History	IDDM					0									
	NIDDM					0									
	MI					0									
	HTN					0									

0 part of candidate predictors list, 1 is part of final predictors list

7.3 Appendices for Chapter 4

7.3.1 Appendix 19: Participant information sheet

Prognostic Prediction Factors for Diabetic Retinopathy (DR) progression

These patients are referred to the hospital eye services (HES) when they develop clinical signs of sight-threatening diabetic retinopathy. However, approximately 50% of referrals don't need intervention and are observed in the HES for a variable period of time.

- As clinicians, you may have felt the necessity for a risk stratification tool supporting your decision in; 1) which patients to prioritise for more urgent appointments 2) whether to definitely be able to see them before they reach treatment stage or vision loss and need evidence to support your decision. Such a tool will need certain prognostic factors and we have put together a set (7.2.3) from our systematic review of published tools /models (mainly for screening patients).
- This study is a step to develop a prediction model for the progression of diabetic retinopathy with an aim to help predict the above outcomes safely and effectively.

- *Please help us select top prognostic factors which can be factors on this list or any additional factors that you can think of but is not on this list.*
- *To do that you may wish to write down without looking at the list given, ignore the order these factors are listed in or the groups they belong to and decide on the basis of your own personal knowledge and experience.*

Thanks for agreeing to take part in this study. Please mark on the candidate predictors list above (7.2.3). Please feel free to add at the end of the list any additional factors, name and sign the consent form and return at the end of the meeting.

7.3.2 Appendix 20: Participant consent form

		Please write your initials in the Box
1. I confirm that I have read and understand the information sheet dated 31/08/2018 (version 1) for the above study. I have had the opportunity to consider the information ask questions and have had these answered satisfactorily.		☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason.		☐
3. I understand that relevant sections of my notes and data collected during the study may be looked at by individuals above, from regulatory authorities or from the NHS Trust, where it is relevant to my taking part in this research. I give permission for these individuals to have access to my input.		☐
4. I consent to this material being used for research and publications.		☐
5. I freely agree to take part in this study		☐

Name of Participant	Date	Signature
Name of Researcher	Date	Signature

7.3.3 Appendix 21: The four-stage NGT Process*

1	Silent generation of ideas in writing (Brainstorming):	At the beginning of the meeting, participants were given an information sheet (Appendix 13) which includes a brief overview of the study, and the list of prognostic factors to select from and add to.	This was sent to participants before the meeting as well
2	Round-robin recording of ideas	Each participant was asked to propose one item that they believed is the most important predictor for DR progression. This happened over several rounds until there were no new predictors proposed.	Participants were asked to keep it simple and precise - One phrase only. Each prognostic factor mentioned was written on a flipchart.
3	Discussion of the list of ideas	Participants were encouraged to discuss their generated list of predictors and ask for any clarifications needed	An equal opportunity was provided for all
4	Voting	Using five stickers, each participant was asked to place stickers next to their top five predictors written on the flip chart from step 2.	Retina specialists were given a different coloured sticker, to be able to analyse two groups separately if needed

*Modified from: Gallagher M, Hares T, Spencer J, Bradshaw C, Webb I. The nominal group technique: a research tool for general practice? Fam Pract. 1993;10(1):76-81.

7.3.4 Appendix 22: Breakdown of participants

Grades	Midlands NHS Trust	North East NHS Trust	South of England NHS Trust	Midlands NHS Trust	Total
Medical Retina	2	1	4	4	11
Other Consultants	3	4	1	0	8
Associate specialists	1	3	0	1	5
Speciality Doctors	2	3	1	0	6
Registrars	7	1	3	1	12
Nurses	1	0	0	0	1
Optometrists	0	1	0	0	1
Total	16	13	9	6	44

7.3.5 Appendix 23: New factors introduced in the NGT meetings and saturation[a]

	Midlands	**North East**	**South England**	**Midlands**
Only eye situation	Y	-	-	-
Early Worsening	Y	-	-	-
Frequent DNA	-	Y	-	-
Pregnancy	-	Y	Y	Y
Diet	-	Y	-	-
Pre proliferative DR	-	-	Y	Y
Chronic infection	-	-	Y	-
Co-morbidities	Y	-	-	Y

[a] Please see text

7.3.6 Appendix 24: Primary studies Risk of Bias

STUDY	Reference Number	RISK OF BIAS (LOW, MODERATE, HIGH)
Non - Systematic Review		
Gupta et al	(1)	Moderate
UKPDS 50	(2)	Low
UKPDS 52	(3)	Low
DCCT	(4)	Low
Yau JW et al (Global)	(5)	Moderate
Lane et al	(6)	High
Ohkubo Y et al	(7)	Low
Grausland et al	(8)	Low
Jeng et al	(9)	Moderate
Zhou Y et al	(10)	Low
Systematic Review		
WESDR	(11)	Low
DCCT	(4)	Low
ACCORD	(12)	Low
EURODIAB	(13)	Low
UKPDS 68	(14)	Low
Mehlson et al	(15)	Low
Lagani et al	(16)	Low

References:

1. Gupta A, Gupta V, Thapar S, Bhansali A. Lipid-lowering drug atorvastatin as an adjunct in the management of diabetic macular edema. Am J Ophthalmol. 2004;137(4):675-82.
2. Stratton IM, Kohner EM, Aldington SJ, Turner RC, Holman RR, Manley SE, 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.
3. Kohner EM, Stratton IM, Aldington SJ, Holman RR, Matthews DR, Group UKPDS. Relationship between the severity of retinopathy and progression to photocoagulation in patients with Type 2 diabetes mellitus in the UKPDS (UKPDS 52). Diabetic Medicine. 2001;18(3):178-84.
4. DCCT. The effect of intensive diabetes treatment on the progression of diabetic retinopathy in insulin-dependent diabetes mellitus. The Diabetes Control and Complications Trial. Archives of ophthalmology (Chicago, Ill : 1960). 1995;113(1):36-51.
5. Yau JW, Rogers SL, Kawasaki R, Lamoureux EL, Kowalski JW, Bek T, et al. Global prevalence and major risk factors of diabetic retinopathy. Diabetes care. 2012;35(3):556-64.
6. Lane M, Mathewson PA, Sharma HE, Palmer H, Shah P, Nightingale P, et al. Social deprivation as a risk factor for late presentation of proliferative diabetic retinopathy. Clinical ophthalmology (Auckland, NZ). 2015;9:347-52.
7. Ohkubo Y, Kishikawa H, Araki E, Miyata T, Isami S, Motoyoshi S, et al. Intensive insulin therapy prevents the progression of diabetic microvascular complications in Japanese patients with non-insulin-dependent diabetes mellitus: a randomized prospective 6-year study. Diabetes research and clinical practice. 1995;28(2):103-17.
8. Grauslund J, Green A, Sjolie AK. Blindness in a 25-year follow-up of a population-based cohort of Danish type 1 diabetic patients. Ophthalmology. 2009;116(11):2170-4.
9. Jeng CJ, Hsieh YT, Yang CM, Yang CH, Lin CL, Wang IJ. Diabetic Retinopathy in Patients with Diabetic Nephropathy: Development and Progression. PloS one. 2016;11(8):e0161897.
10. Zhou Y, Zhang Y, Shi K, Wang C. Body mass index and risk of diabetic retinopathy: A meta-analysis and systematic review. Medicine (Baltimore). 2017;96(22):e6754.
11. Klein R, Klein BE, Moss SE, Davis MD, DeMets DL. The Wisconsin epidemiologic study of diabetic retinopathy. IV. Diabetic macular edema. Ophthalmology. 1984;91(12):1464-74.
12. Group AS, Group AES, Chew EY, Ambrosius WT, Davis MD, Danis RP, et al. Effects of medical therapies on retinopathy progression in type 2 diabetes.[Erratum appears in N Engl J Med. 2011 Jan 13;364(2):190], [Erratum appears in N Engl J Med. 2012 Dec 20;367(25):2458]. New England Journal of Medicine. 2010;363(3):233-44.
13. Sjolie AK, Stephenson J, Aldington S, Kohner E, Janka H, Stevens L, et al. Retinopathy and vision loss in insulin-dependent diabetes in Europe. The EURODIAB IDDM Complications Study. Ophthalmology. 1997;104(2):252-60.
14. Clarke PM, Gray AM, Briggs A, Farmer AJ, Fenn P, Stevens RJ, et al. 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. 2004;47(10):1747-59.
15. Mehlsen J, Erlandsen M, Poulsen PL, Bek T. Identification of independent risk factors for the development of diabetic retinopathy requiring treatment. Acta Ophthalmol. 2011;89(6):515-21.
16. Lagani V, Chiarugi F, Thomson S, Fursse J, Lakasing E, Jones RW, et al. Development and validation of risk assessment models for diabetes-related complications based on the DCCT/EDIC data. Journal of diabetes and its complications. 2015;29(4):479-87.

7.3.7 Appendix 25: Predictors distribution in models with low to moderate risk of bias and low concern of applicability*

	DR	Age	Gender	Age at diagnosis	Duration of diabetes	Type of diabetes	HBA1c	Total serum cholesterol	SBP
Aspelund T et al. 2011 (1)	Y	-	Y	-	Y	Y	Y	-	Y
Scanlon PH 2015 (2)	Y	-	-	-	Y	-	Y	-	-
ISDR model 2017 (3)	-	-	-	Y	Y	-	Y	Y	Y
7 common ones among 14 models	Y	Y	Y	Y	Y	-	Y	-	Y

***** Reproduced from Haider S, Sadiq SN, Moore D, Price MJ, Nirantharakumar K. Prognostic prediction models for diabetic retinopathy progression: a systematic review. Eye. 2019;33(5):702-13.

References:

1. Aspelund T, Porisdottir O, Olafsdottir E, Gudmundsdottir A, Einarsdottir AB, Mehlsen J, et al. Individual risk assessment and information technology to optimise screening frequency for diabetic retinopathy. Diabetologia. 2011;54(10):2525-32.
2. Scanlon PH, Aldington SJ, Leal J, Luengo-Fernandez R, Oke J, Sivaprasad S, et al. Development of a cost-effectiveness model for optimisation of the screening interval in diabetic retinopathy screening. Health Technol Assess. 2015;19(74):1-116.
3. Eleuteri A, Fisher AC, Broadbent DM, Garcia-Finana M, Cheyne CP, Wang A, et al. Individualised variable-interval risk-based screening for sight-threatening diabetic retinopathy: the Liverpool Risk Calculation Engine. Diabetologia. 2017.

7.3.8 Appendix 26: Poster presentation

Risk (Prognostic) factors for Diabetic Retinopathy Progression : Nominal Group Technique and a literature review

Sajjad Haider[1], Salman Naveed Sadiq[2], Harpreet Sihre[1], Eniya Lufumpa[1] and Mohammad Tallouzi[1] David Moore[3], Krishnarajah Nirantharakumar[3], Malcolm James Price[3]
Corresponding author : Sajjad Haider, e mail [redacted]
1 .Doctoral researchers, Institute of Applied Health Research, University of Birmingham , UK; 2. Royal Victoria Infirmary, Newcastle, UK; 3. Joint Senior Authors Institute of Applied Health Research, University of Birmingham, Birmingham, UK

Introduction:

Risk factors (1) are at the core of prevention and prediction of a clinical outcome and need be to present in the subjects before the outcome appears. A prognostic factor (2) is a risk factor defined as any measure used to predict in people with a certain health condition (diabetic retinopathy at referral - point zero) an outcome (treatment or vision loss as endpoint).
Prognostic models aiming to predict sight threatening diabetic retinopathy (DR) have recently been developed for diabetic eye screening. They are a combination of two or more predictors (2) of an outcome (DR progression in our case). A systematic review (SR) of the existing models (3) conducted recently identified 78 predictors (with significant duplication and overlap).

We aimed to :.

- To help clinicians to reduce the risk of progression by managing the modifiable risk factors
- To help the researchers build a model (for DR progression to treatment requiring stage or vision loss) to risk stratify higher risk hospital patients for better resource utilization.

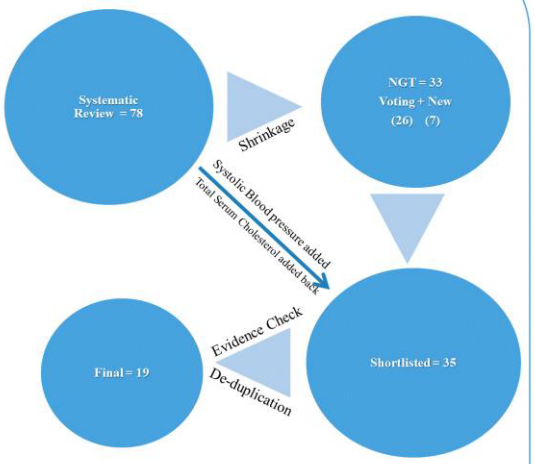
We did this by

- Shrinking this extensive list to be manageable
- Evaluated the evidence for confirmation of a prognostic factor association with the outcome of DR progression

Methods:

- A nominal group technique (NGT) study (4) was carried out to obtain clinicians input for de-duplication and shrinkage of the list
- A search of reference lists of modelling studies for primary studies of prognostic factors
- A critical appraisal of primary studies using QUIP tool (5) for evidence evaluation of confirmation
- A literature search for primary studies which were not clear from reference lists

Results:



- Only two modelling studies out of 14 in the SR, reported the primary studies for all their prognostic factors.
- The four NGT meetings arrived at 33 factors, including seven new ones.
- The primary studies for 36% of the prognostic factors could not be traced.

Conclusions:

- A parsimonious list of **19** predictors is finalised (Table below).
- De-duplication achieved
- Reporting deficiencies were found in modelling studies. It is suggested that future modelling studies should report the primary studies confirming each prognostic factor used. An amendment in next review of TRIPOD reporting guidance (6) is suggested

Final Set of Candidate Predictors

Group	Predictors	Group	Predictors	Group	Predictors
1	Retinopathy level / DR grade	6	Age	14 Biochemical parameters	HbA1c
2	Presence of DR	7	Race	15	eGFR
2a	Proliferative diabetic retinopathy	8 Socio demographics	Gender	16 Diabetes treatment	Statin
2b Ocular features	Maculopathy	9	Social Deprivation score	17	Insulin
2c	Pre-proliferative DR	10	Type of DM	18	Pregnancy
3	Visual acuity score (NGT)	11 Diabetes characteristics	Age at diagnosis	NGT New 19	Early worsening
4 Physical examination	Systolic Blood pressure	12	Duration of DM		
5	Diastolic Blood Pressure	13 Co-morbidities	Hypertension		

References:

1. Offord DR, Kraemer HC. Risk factors and prevention. Evidence Based Mental Health. 2000;3(3):70-1.
2. Riley RD, Hayden JA, Steyerberg EW, Moons KG, Abrams K, Kyzas PA, et al. Prognosis Research Strategy (PROGRESS) 2: prognostic factor research. PLoS medicine. 2013;10(2):e1001380.
3. Haider S, Sadiq SN, Moore D, Price MJ, Nirantharakumar K. Prognostic prediction models for diabetic retinopathy progression: a systematic review. Eye. 2019.
4. Gallagher M, Hares T, Spencer J, Bradshaw C, Webb I. The nominal group technique: a research tool for general practice? Fam Pract. 1993;10(1):76-81.
5. Hayden JA, van der Windt DA, Cartwright JL, Cote P, Bombardier C. Assessing bias in studies of prognostic factors. Ann Intern Med. 2013;158(4):280-6.
6. Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med. 2015;162(1):W1-73.

7.4 Appendices for Chapter 5

7.4.1 Appendix 27: Goodness of fit

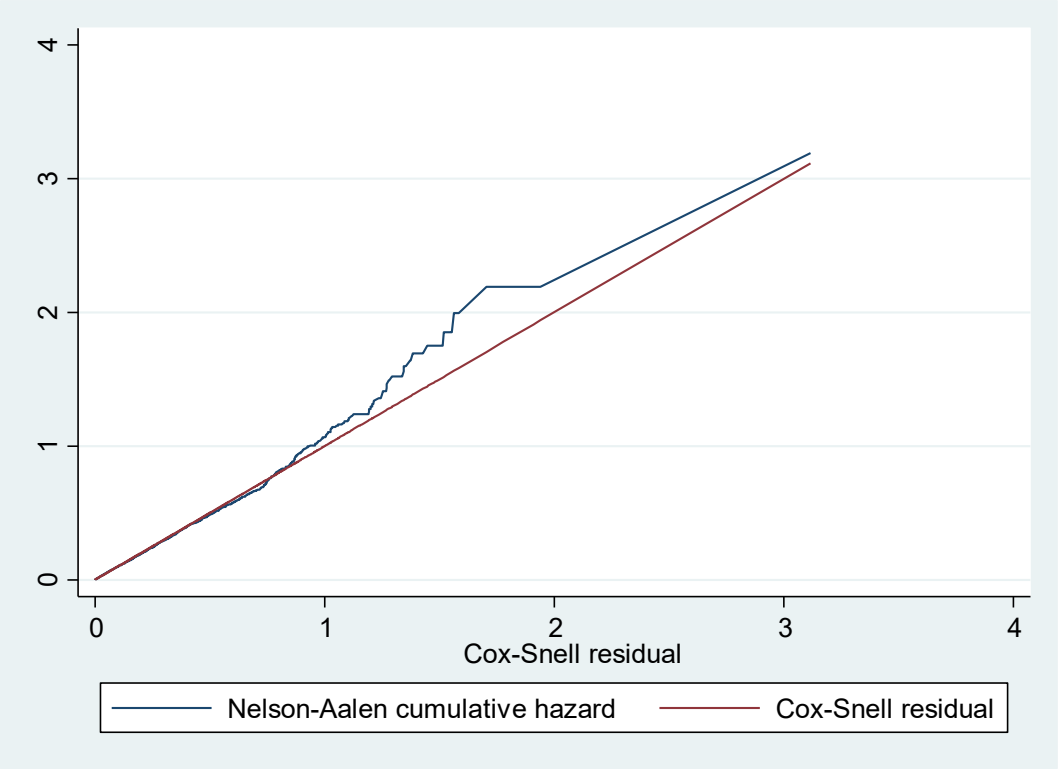


Figure 4: Goodness of fit

For goodness of fit**,** the Nelson-Aalen cumulative hazard function, plotted against the Cox-Snell residuals for the model are shown in

4 above, the hazard function roughly follows the 45 degree line (except for some medium to large values of time).

7.4.2 Appendix 28: Hazard Ratios Complete case Analysis R2 vs M1 as a reference:

Predictor	Complete case analysis Model (95% Confidence intervals) Reference R2	P value	Complete case analysis Model (95% Confidence intervals) Reference M1	P value
R2	1		.26 (.15- .46)	<0. 001
R3	15.28 (8.81 - 26.49)	<0. 001	4.04 (3.65-4.6)	<0. 001
R2/M1	6.19 (3.43 - 11.16)	<0.001	1.6 (1.3-2.1)	<0. 001
R3/ M1	25.6 (14.78 - 44.34)	<0.001	6.8 (6 - 7.6)	<0. 001
M1	3.78 (2.18- 6.56)	<0.001	1	
Unclassified	5.96 (3.1 - 11.48)	<0.001	1.6 (1.1- 2.3)	0.02
HbA1C	1.01 (1.01 - 1.01)	<0.001	1.01 (1 -1.01)	<0. 001
eGFR3	1.12 (1.00 - 1.27)	0.05	1.12 (1 -1.3)	0.05
eGFR4	1.62 (1.27 - 2.05)	<0,001	1.6 (1.3 - 2.1)	<0. 001
SBP	1.003 (1 - 1.005)	0.04	1 (1 - 1.005)	0.04
Cholesterol	1.04 (1 - 1.09)	0.06	1.04 (1 - 1.09)	0.06
Insulin	1.11 (1 - 1.24)	0.05	1.11 (1 - 1.24)	0.051
Statins	0.88 (.79 -.98)	0.02	.88 (.8 - .98)	0.02

7.4.3 Appendix 29: Internal Validation Steps *

I used the following steps for internal validation

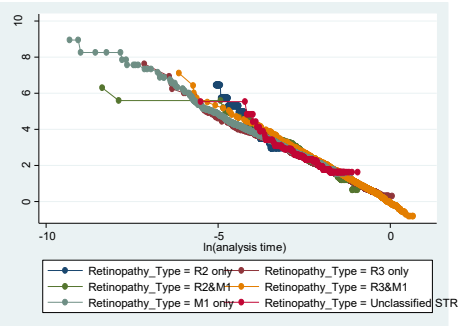
1. Produced bootstrap samples (800 samples) from our development dataset
2. In each bootstrap sample:
    a. A model was developed using automated backward elimination methods (same steps as complete case analysis methods).
    b. The apparent performance (C stat and C slope) of this model in the bootstrap sample and test performance of this model in the original dataset were calculated.
    c. The optimism statistics for both C stat and C slope were calculated as (Apparent - test).
3. The average optimism statistic across bootstrap samples was calculated.
4. Optimism adjusted C stat and C slope (apparent C in original development data – average optimism for C from bootstrap) were calculated in each case.

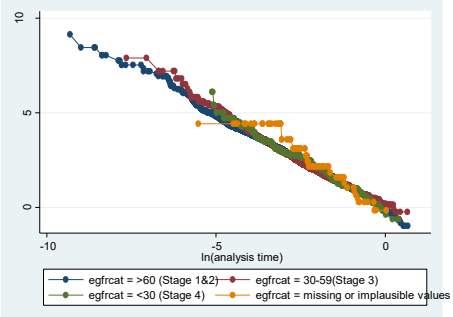
* Modified from Steyerberg EW. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating: Springer New York; 2008

7.4.4 Appendix 30: Assessment of Assumptions

During the process of model development, it was important to check that the model is able to give an appropriate representation of the data, and that various assumptions are valid to avoid misleading conclusions. This is done by checking for assumptions such as proportional hazards and linearity (1). Residuals were used for this purpose. In a regression, a residual is the difference between the observed value and the predicted value of dependent variable. The sum and the mean of residual are both equal to zero. A residual plot is a graph with residuals on the y axis and the independent variable on the horizontal axis (2). They have been used here in the diagnostics of the assumptions in various forms.

Proportionality is an inherent assumption of the Cox model, meaning that the relative risk for any two individuals with different covariate values is constant over time (1). This may not be true, for example, the covariate may have a larger effect on HR in the beginning, but wane off later (3). Proportional hazards assumptions were investigated with log log plots and in post-estimation phtest.





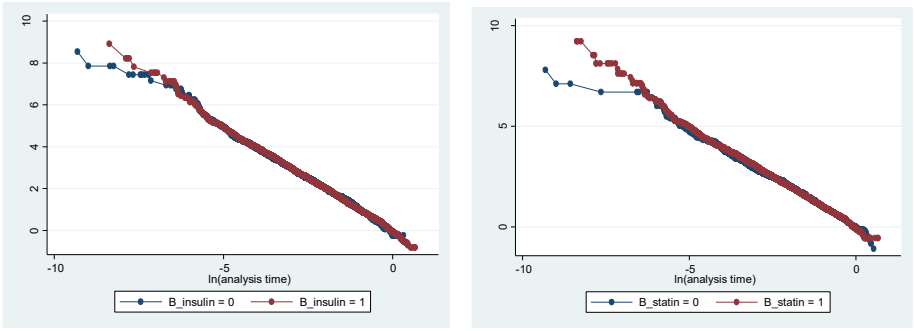


Figure 1: Proportional Hazards Assumption tests of categorical variables Produced using log log plots

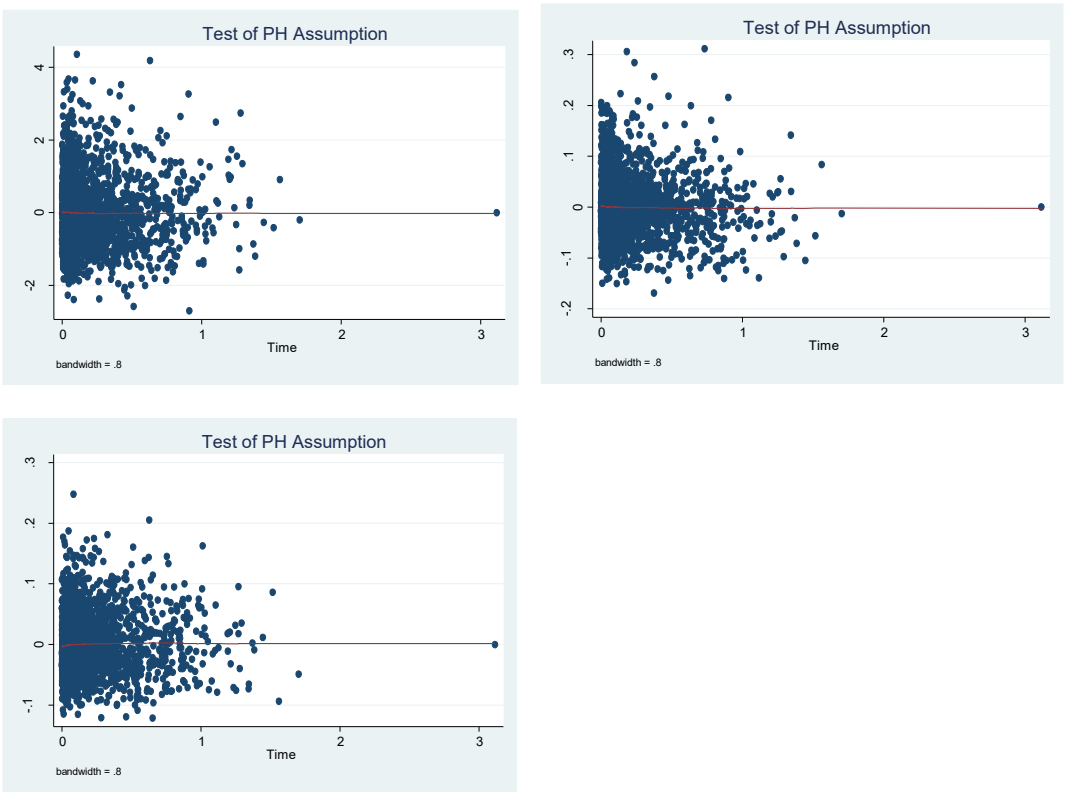


Figure 2: Proportional assumption tests for continuous variables

I used Schoenfeld Residuals and scaled Schoenfeld residuals. The log-log plots were mostly parallel, though a few of the lines do cross, but with very little separation. A

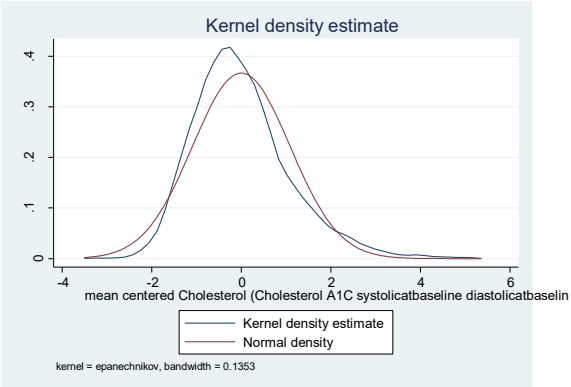
horizontal line in the graphs based on the Schoenfeld residuals also indicates non violation of the proportionality assumption (4, 5).

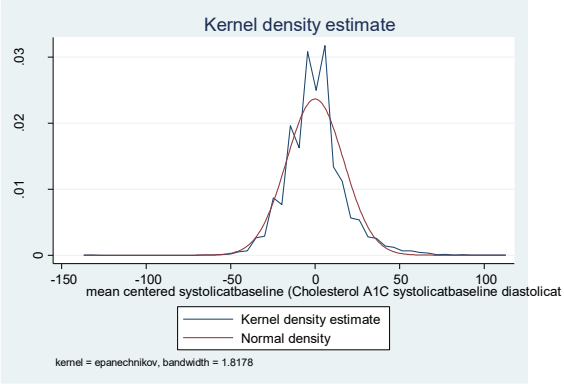
Another assumption in Cox modelling is linearity of the continuous predictors (1). While it is widely accepted that existing knowledge should guide predictor selection, there is some need for data-dependent selection of predictors and for choosing appropriate functions for continuous predictors. I checked for multicollinearity (with a variance inflation factor (Table). I checked skewness with kernal density plots.

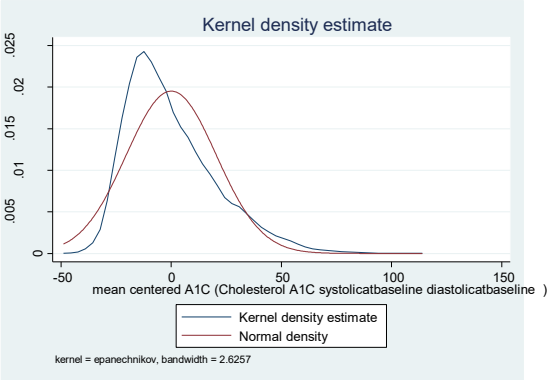
Table 1: Multi-collinearity Variance Inflation Factor (VIF)

Variable	VIF	1/VIF
HBA1C	1.07	0.934709
Cholesterol	1.07	0.932368
Systolic Blood Pressure	1.01	0.992536
Mean VIF	1.05	

Multi-collinearity is the occurrence of high inter-correlations among independent variables in a multiple regression and can give rise to skewed results. Our result was not significant (table 3). VIF score > 10 requires further investigations.







The k density plots for the three continuous variables to look for significant skewness. There was minimal skewness, not severe enough to need transformation.

References:

1. STHDA. Cox Model Assumptions [Available from: http://www.sthda.com/english/wiki/cox-model-assumptions.
2. Stat Trek. Residual Analysis in Regression 2019 [Available from: https://stattrek.com/regression/residual-analysis.aspx, Accessed 2020 May, 15.
3. Hernán MA. The hazards of hazard ratios. Epidemiology (Cambridge, Mass). 2010;21(1):13-5.
4. Riley RD, van der Windt D, Croft P, Moons KGM. Prognosis Research in Health Care: Concepts, Methods, and Impact: Oxford University Press; 2019.
5. Steyerberg EW. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating: Springer New York; 200

7.4.5 Appendix 31: Workings for Model with Fractional Polynomials / Multiple Imputation and its performance

I used automated Stata “mfpmi” fractional polynomial (FP) models to best predict the outcome variable from the list of candidate predictors. The data were multiply imputed and the programme selected an FP model based on a version of likelihood-ratio tests. The data was passed through two cycles of selection of non linear functions for all the continuous predictors. The procedure stopped when two consecutive cycles (cycles 2 and 3) contained the same covariates with the same FP transformations.

7.4.6 Appendix 32: Competing risk Analysis

Table 1: Running total comparison between outcome of interest vs death

0	Outcome of interest	Running total	Death	Running total
1	869	869	459	459
2	403	1272	396	855
3	246	1518	352	1207
4	178	1696	306	1513
5	133	1829	205	1718
6	88	1917	173	1891
7	67	1984	102	1993
8	44	2028	91	2084
9	20	2048	58	2142
10	15	2063	46	2188
11	9	2072	27	2215
12	1	2073	17	2232
13	3	2076	8	2240
14	3	2079	0	2240
15	0	2079	0	2240

More events occur for treatment or vision failure at earlier time points up to 2 years compared to death but then more deaths occur by later time points from 3rd year onwards. The running total of deaths takes over from 6th year onwards

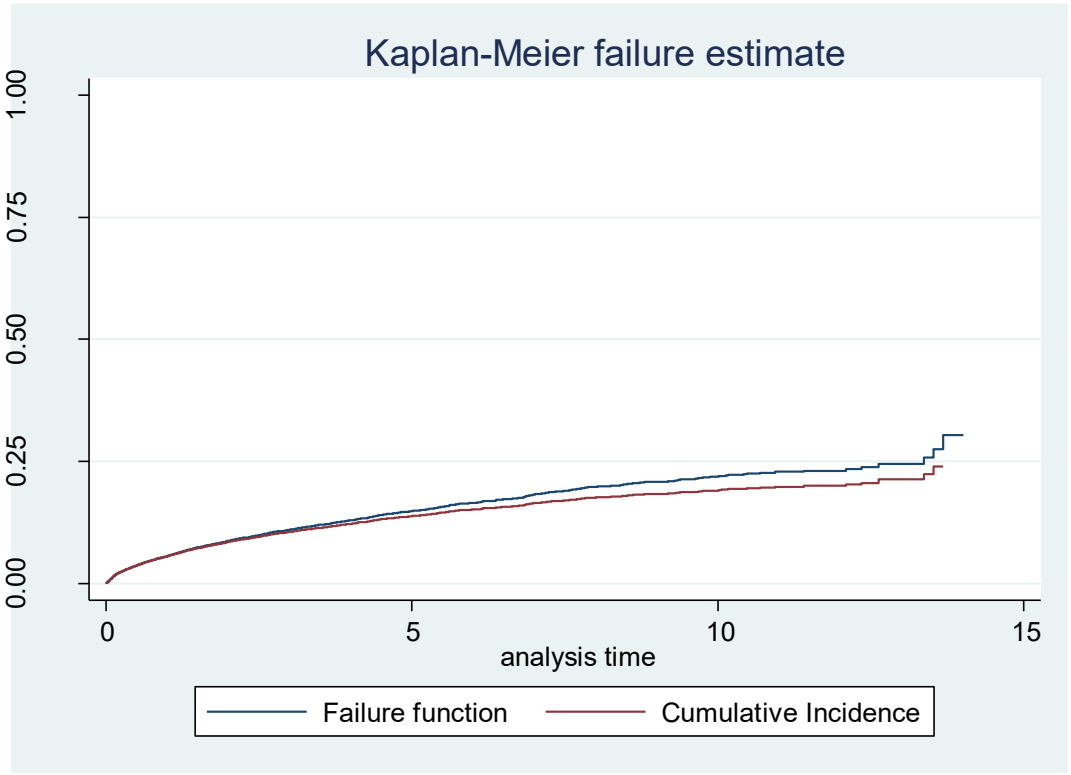


Figure : X axis – Survival, Y axis Analysis time - Comparing non-parametric cumulative incidence function at all time points up to 5 years with and without account for death

On average the probability of the outcome isn't very different at the time points of interest. Probably about 2% absolute difference at 5 years and even less for earlier time points (overestimation with cox as compared to Fine and Gray model).

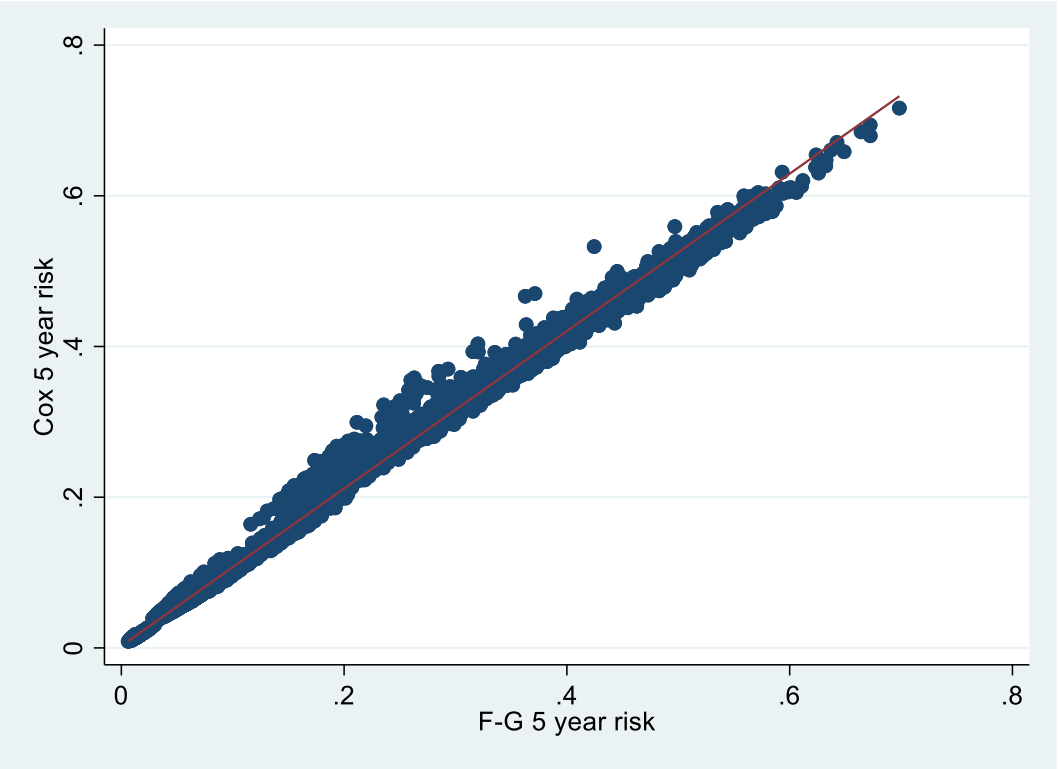


Figure : X axis - predictions from Cox model, Y axis – predictions from Fine and Gray model – Comparing cumulative incidence function for 5 years

E/O (5 years) = 0. 961489869

For the majority of individuals, there is very little difference but some larger differences are seen for a few individuals in the dataset.